Representation and the Direction of Innovation: Evidence from U.S. Patent Applications

by

Gauri S Subramani

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Toby Stuart, Co-chair
Professor Abhishek Nagaraj, Co-chair
Professor Mathijs De Vaan
Professor Lee Fleming

Summer 2021

Representation and the Direction of Innovation: Evidence from U.S. Patent Applications

Abstract

Representation and the Direction of Innovation: Evidence from U.S. Patent Applications

by

Gauri S Subramani

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Toby Stuart, Co-chair

Professor Abhishek Nagaraj, Co-chair

This dissertation is comprised of three chapters examining participation in and the direction of innovation in the United States. Chapter 1 details the creation of a novel data set constructed using publicly available data from the United States Patent and Trademark Office with which I am able to explore individual inventors' complete patent application histories. Chapter 2 examines the gender gap in innovation by studying male and female inventors' differential responsiveness to rejections in the patent process. Chapter 3 uses the data constructed in Chapter 1 to study variation in application behavior and the innovative trajectories of male and female patent applicants.

Chapter 1 considers how individual inventors' patenting activities have traditionally been measured in prior work and describes the construction of a more comprehensive data set. Prior research has documented repeat patenters as those who receive more than one patent. However, the lack of data that tracks all patenting activity, including ungranted and abandoned applications, has made it impossible to measure the effects of applications that do not convert to patents, but that still may represent meaningful innovative contributions. I detail the construction of a novel, open-access dataset that tracks inventors and identifies all patent applications and granted patents from individuals from 2001 onwards. I identify 2.4 million unique inventors across 4.6 million applications. I describe this data and outline future directions for research.

Chapter 2, coathored with Abhay Aneja and Oren Reshef, documents that over 86% of granted patents in the United States include no female inventors and asks: why are women underrepresented in innovation? We argue that differences in responses to early rejections between men and women are a significant contributor to the gender disparity in innovation. We evaluate the prosecution and outcomes of almost one million patent applications in the United States from 2001 through 2012 and leverage variation in patent examiners'

probabilities of rejecting applications to employ a quasi-experimental instrumental variables approach. Our results show that applications from women are less likely to continue in the patent process after receiving an early rejection. Roughly half of the overall gender gap in awarded patents during this period can be accounted for by the differential propensity of women to abandon applications.

We identify that teams with a significant female presence are 4.4-7.2 percentage points less likely to receive a patent if they receive a rejection early in the application process. The gender gap in responsiveness to rejection widens as the presence of women on inventor teams increases, indicating that the *intensive* margin of female representation affects how patenting teams respond to rejection. We explore why this may be the case and provide evidence that the gender gap in outcomes is reduced for applications that use attorneys and are affiliated with firms, consistent with a role for information and institutional support in mitigating gender disparities. These results suggest that female innovators' ideas are underleveraged, with negative implications for gender differences in entrepreneurship and the broader landscape of innovation.

In Chapter 3, I leverage the data constructed in Chapter 1 to study how male and female patent applicants' trajectories differ. Specifically, I evaluate the effect of prior experience on male and female inventors' subsequent patent applications. I identify that female inventors are significantly more likely than male inventors to explore different technology areas after success in the patent process. Interestingly, this effect is driven by experienced patent applicants as opposed to new patent applicants. I also find that experienced female inventors are more likely to patent across a broad set of technology areas than similarly experienced male inventors.

Through this work, I identify a significant piece of the puzzle of women's underrepresentation in patenting and implications for the landscape of innovation. I quantify the magnitude of gender differences in innovation contexts and identify the mechanisms by which these arise, along with opportunities for interventions to reduce performance gaps. This dissertation represents a step towards better understanding gender variation in participation in innovation, which is essential in order to develop solutions that promote gender equity in innovative fields.

Contents

List of Figures

List of Tables

Acknowledgments

Coauthor information: Chapter 2 of this dissertation is coauthored with Abhay Aneja, Assistant Professor of Law at UC Berkeley, and Oren Reshef, Assistant Professor of Strategy and Entrepreneurship at Washington University in St. Louis.

Acknowledgments: I am deeply grateful to the Macro MORS faculty for believing in me and welcoming me into their intellectual community. I am particularly thankful to Andreea Gorbatai and Abhishek Nagaraj for opening the door for me to join the group. I feel very lucky to have had the opportunity to learn from and work alongside such creative and insightful scholars.

I also want to extend my thanks to my brilliant committee members, Abhishek Nagaraj, Toby Stuart, Mathijs De Vaan, and Lee Fleming. Multiple people, upon hearing who is on my committee, have remarked to me something approximating, “Wow, what a group of all-stars.” I feel similarly! Thank you as well to Sameer Srivastava, Andreea Gorbatai, Heather Haveman, Solène Delecourt, and Douglas Guilbeault. I also had the good fortune of benefiting from the advice and guidance of Michael Luca and Ariel Stern.

My fellow PhD students helped me overcome innumerable bumps and challenges and stave off the inherent isolation of research. My office-mates and collaborators made spending hours on end in a windowless room fun. I have cherished being able to turn to so many wildly smart, interesting people to talk about anything from debugging code to elections in Brazil.

I would not have made it through the past six years without the women at The Dailey Method (TDM) in Berkeley and Piedmont, which served as my oasis. I am so appreciative of the TDM clients who trusted me to lead them through classes and help them connect with their bodies. I will miss this community more than I can put into words.

Another group of women from whom I have drawn inspiration is the East Bay Wellesley Book Group. This collection of amazingly intelligent, insightful, thoughtful, and kind women has been a beacon of hope for my future.

My friends and family have buoyed me throughout this process with encouragement and love. Thank you to my parents, Akka, Rohit, and Hadas for reliable and tireless cheerleading. Special thanks to Akka and Rohit for providing me with Vasantha and Prema, my favorite sources of distraction and joy.

Finally, to Ben; this accomplishment was only possible because of your unwavering support. I am always learning from your patience, curiosity, and compassion.

Chapter 1

Gender and Application Dynamics in U.S. Patent and Trademark Office Applications

1.1 Introduction

This paper details the construction of a novel dataset of applications to the U.S. Patent and Trademark Office (USPTO) from 2001 to 2014. I create a unique identifier that tracks patent applicants' repeat applications in this time period and identifies applicant gender. This data enables the study of gender-specific application trends, individual inventors' innovative trajectories, moves across firms, and other dynamics in application behavior.

Prior work (Fleming and Juda 2004; Singh 2005; Trajtenberg, Shiff, and Melamed 2006; Fleming, Mingo, and Chen 2007; Carayol and Cassi 2009; Pezzoni, Lissoni, and Tarasconi 2012; Li et al. 2014; Lai et al. 2015; Balsmeier et al. 2018) has focused on tracking the recipients of patents across time. Individuals do not disclose their gender when submitting a patent application, and researchers have used varying approaches to identify the gender of patent recipients (Raffo and Lhuillery 2009; Walsh and Nagaoka 2009; Jung and Ejermo 2014; Hoisl and Mariani 2016; Delgado, Mariani, and Murray 2019). This work provides insight into the experiences of successful inventors, but omits data about applications that were not granted. That is, it is not possible to identify all of the patent applications submitted by a single individual using these approaches, as some applications do not convert to patents and thus are missing from the set of granted patents.

I find that only 53.49% of patent applications are granted, which indicates that analyses that only evaluate inventors' granted applications miss over 46% of applications. Expanding the set of inventors to include unsuccessful patent applicants can meaningfully improve our understanding of the drivers of participation in innovation. For example, existing research has found that exposure to innovation in childhood has a causal effect on individuals' propensities to receive a patent in adulthood (Bell et al. 2019); however this work focuses only on patent recipients. In this work, I further contribute to an understanding of the trajectories of repeat patent applicants and captures previously unidentified interactions with the patent office, as well as make my code and data publicly available.

This paper is organized as follows. I begin by motivating the value of evaluating patent applications, not just granted patents, in innovation research. Next, I describe the data I leverage and discuss the methods I use to identify and track individuals over time. I show how my results compare to existing work and highlight opportunities for improvements. The last section concludes and identifies opportunities for research using this novel data.

1.2 Motivation

There are several reasons why it is useful to link individual inventors' patent applications as opposed to just granted patents, which represent a subset of all applications. First, patent applications are a valuable tool to evaluate trends in innovation. Much literature in economics and management has measured innovation by examining patents. "Patents provide a relatively objective measure of new knowledge" as they "are required to describe something novel and not obvious" (Katila 2000). To apply for a patent, an inventor must believe she has produced something with novel value and have the knowledge and resources to file an application. Filing a patent application requires time as well as a financial investment. Application fees vary depending on the applicant, type of patent application, the application's level of complexity, and other features of the application, but the total cost of fees paid to USPTO are a minimum of roughly $500 (excluding any attorney fees) (USPTO 2016). The assistance of a lawyer is resource intensive, as patent attorneys' hourly billing rates range from $300 to over $700 depending on attorney experience and geographic location (Quinn 2015). 96.99% of applications are filed using the services of an attorney, which suggests that frivolous attempts are uncommon and applications represent good-faith attempts to gain rights over a novel innovation.

While most research has focused on granted patents, even ungranted patent applications (or, equivalently, abandoned applications) can be valuable innovative contributions, as evidenced by their frequent reference as prior art by USPTO patent examiners. Abandoned patent applications are used by examiners to narrow the scope of subsequent patents and thus have implications for the patent system more broadly (Cotropia and Schwartz 2019). In fact, Cotropia and Schwartz (2019) find that patent examiners are more likely to cite abandoned applications as prior art when issuing rejections due to a lack of novelty and obviousness than they are to cite previously granted patents. This suggests two things: first, abandoned patent applications have value in terms of their innovative contributions, even if they did not translate to direct value in the form of legal rights for their inventors. Second, abandoned patent applications also play an important role in influencing *future* innovation by affecting what technologies USPTO grants patent rights to.

Between 2001 and 2014, only 53.49% of patent applications eventually converted to granted patents. Thus, analyses that focus only on the inventors and ideas represented in granted patents omit over 46% of filed applications. The costs of this omission are particularly striking when we consider the value of understanding the effects of failure in the patent application process on individuals' innovative paths as well as the broader landscape of innovation.

Further, granted patents are not representative of patent applications, particularly with respect to inventor characteristics. This indicates that only evaluating inventors who hold patents leads to sample bias. To begin, women are far less likely than men to apply for patents; from 2001 through 2014 only 14.31% of patent applicants were women. Only 20.76% of applications contain any female applicants, and this gender disparity only worsens through the patent review process. Conditional on application, women are less likely to receive a patent than men. Just 13.1% of patent recipients in this time period were women, and just 19.15% of granted patents include at least one female inventor. Thus, studying only granted patents disproportionately ignores the innovative contributions of female inventors.

In addition, just looking at granted patents elides gendered dynamics in representation and success in the patenting process (Jensen, Kovács, and Sorenson 2018). Examining the full set of patent applications gives researchers the opportunity to examine participation in the patent process and the demographic characteristics that affect repeat participation in patenting.

Finally, while granted patents have been the focus of much research, patent applications provide a more complete picture of inventors' careers and provide a valuable comparison set against which to study the effects of patent grants. Granted patents can be thought of as successful or "goals", while applications are akin to "shots on goal". Unsuccessful patent applications can represent significant learning opportunities for inventors (Cotropia and Schwartz 2019), and linking inventors' successful and unsuccessful applications opens up rich avenues for study.

1.2.1 Prior Work

Existing work has used many approaches to identify the gender of individuals who appear in patent data. Researchers have linked patent applications to administrative records (Walsh and Nagaoka 2009; Jung and Ejermo 2014; Hoisl and Mariani 2016) and most often, used the information provided by inventors' first names to impute gender. In early work, USPTO used a file of female-only names to identify female inventors, and characterized all other inventors as male (USPTO, Documentation Information, and Program 1999). Recent research has taken a more fine-grained approach and determined inventor gender using a combination of sources including the U.S. Census, WikiName, Wikipedia's given name list, and country-specific naming rules (Larivière et al. 2013). Another common approach has been to leverage U.S. Social Security Administration data to identify the gender of names of inventors in the United States (Jensen, Kovács, and Sorenson 2018; Delgado, Mariani, and Murray 2019).

Prior literature has examined patent applications and outcomes of patent applications (Frakes and Wasserman 2014; Jensen, Kovács, and Sorenson 2018; Sampat and Williams 2019; Farre-Mensa, Hegde, and Ljungqvist 2020). However, this work has examined patent outcomes at the level of the application as opposed to individual. One key challenge to identifying patent applicants and recipients across time in the USPTO data is the fact that there is no individual ID assigned to an inventor that can link independent applications. This data is not even available internally at USPTO; beyond name and address, there is no individual-specific information that is solicited to enable patent examiners to link a given application to an individual's prior patent applications, if any exist.

Researchers and the Office of the Chief Economist at USPTO have undertaken significant efforts to track individual patent recipients across time (Fleming and Juda 2004; Singh 2005; Trajtenberg, Shiff, and Melamed 2006; Raffo and Lhuillery 2009; Carayol and Cassi 2009; Pezzoni, Lissoni, and Tarasconi 2012; Li et al. 2014; Lai et al. 2015). The current gold standard in inventor disambiguation is the result of an Inventor Disambiguation Workshop hosted by USPTO on September 24, 2015. A research team from the University of Massachusetts Amherst led by Andrew McCallum and Nicholas Monath authored the successful algorithm that was integrated in the PatentsView data platform in March 2016. The algorithm uses discriminative hierarchical coreference as a new approach to increase the quality of PatentsView data (USPTO 2020a).

The only currently published work that evaluates patent applicants across time is by Melero, Palomeras, and Wehrheim (2020). This paper tracks individuals across time using proprietary data from PatentsView that applies the algorithms developed by Li et al. (2014) and Balsmeier et al. (2018) and evaluates the effect of receiving a patent on employee mobility. My hope is that the construction of the dataset described in this paper and its public availability will spur further research exploring the full population of potential inventors as measured by patent applications.

1.3 Methods

1.3.1 Data

The core of my work relies on detailed data on patent applications filed in the United States, including inventors' names, filing date, application correspondence address, and inventors' locations. I also merge in USPTO data on patent assignees, if any– that is, the organization or entity that will have the rights to the patent if granted.

I use several data sources to identify gender of patent applicants. First, I download data on baby names from 1880 through the present from the U.S. Social Security Administration. Names are included in a given year's dataset if they occur in the birth data more than 5 times that year. This data identifies each combination of first name×gender×year and identifies the number of occurrences of each first name×gender combination in a given year. I also take advantage of USPTO's gender identification for patent recipients by using PatentsView data on inventors.

1.3.2 Challenges

There are a number of challenges with regards to both identifying an inventor's gender and tracking the same individual across patent applications. Individuals do not self-identify their gender when applying for a patent, so there is no "ground truth" for an inventor's gender. There is also no ID with which one can link patent applications with any other administrative records that contain gender. While many first names can indicate an individual's gender, the gender association of a name can vary over time and by country, which complicates the process of gender identification.

As mentioned earlier, there is also no individual identifier elicited by USPTO that would allow for automatic tracking of individuals across patent applications. As a result, there is no perfect dataset against which to check this process. While name and location are useful pieces of information, both can vary in ways that make tracking a given person difficult. For example, an inventor may provide their middle initial in one application but not the following, or may use the name "Jon Smith" on one application and "Jonathan Smith" on the next. Additionally, individuals may move over time and apply for patents from different states. Even if name and location were fixed indicators for a single individual, there may be multiple people with the same name in the same geographic region. As a result, simply tracking people based on location and name is insufficient to accurately identify individuals who appear multiple times in the patent application data.

Table 1.1 shows the complexity of tracking a single individual across time by presenting data for a cross section of patent applications in my sample from inventors based in the United States named John M Adams. While one can discern that some of these entries (for example, rows 6 and 7) represent applicants from the same inventor, it is not straightforward to determine this for all entries.

1.3.3 Summary of Gender Identification Process

I begin by replicating the approach taken by Jensen, Kovács, and Sorenson (2018) and use data from the Social Security Administration on names given to babies born in the United States from 1880 through 2018. Using this data, I create a gender distribution for each name that identifies how often it is given to a female baby. If, for example, there are 1,000 babies born over the course of this timeframe named Julia, and 990 of these babies are female, the first name Julia would have a female value of 0.99. I then use a threshold of 0.90 or 0.10 respectively to determine whether names are female or male. That is, to consider a first name as indicating an individual is female or male, 90% of individuals given that name need to be of said gender.

The drawback of this approach is that it leads to the loss of a great deal of data, as many names fall below this 90% threshold. I am unable to identify the gender of 28.5% of inventors. Many applications have multiple authors, and to understand the gender composition and dynamics of applications, I limit the sample to those applications for which I know the gender of all inventors. 43.4% of all applications are dropped based on this process, and

28.5% of applications from U.S.-based individuals and teams are dropped. As a result, I investigate other approaches to identify gender that result in less data loss.

USPTO has invested significant effort in attributing gender to patent recipients Toole, Myers, et al. (2019). To do this, they leverage two different name dictionaries and make use of inventors' full names and countries of residence. Using this process, they are able to attribute gender to 92.08% of inventors. I extend USPTO's work and use their name-gender attribution to identify the gender of all patent applicants. For the subset of inventors whose applications are granted (53%), I can directly observe gender in the USPTO data. USPTO's gender identification is based on an inventor's first name, country, and the year of application filing. For some names, gender is fixed throughout the sample– for example, the first name Samantha is always identified as female, regardless of year or country.

To determine the gender of inventors on ungranted applicants, I start with inventors whose names are always either male or female in the USPTO data. This is the case for 47.95% of unique names. Due to the frequency of gender-specific names, this accounts for 88.86% of inventors in my sample.

Then, I turn to inventors whose first names appear as both male and female and identify those for whom first name is always one gender in their country of residence. If this is the case, I assign the same gender to the focal inventor. For example, there are inventors named Alex in France whose application are not granted, and USPTO identifies inventors named Alex as both male and female. I look to see whether Alex is consistently male or female in France. The name Alex is always coded as male in France, so I identify all French Alexes in the data as male. 0.73% of inventors' genders are identified in this manner.

Through this process, I am able to identify the gender of 89.59% of inventors in my sample, who make up 91.36% of inventor×application entries.

1.3.4 Summary of Individual Disambiguation Process

My next step is to identify repeat patent applicants. To do this, I use data from the USPTO's Patent Examination Research Dataset (Public PAIR) 2014 release. I leverage all publicly available information about inventors including full name, state, country, and detail about attorney and assignee, if any.

I adapt the algorithm used by the U.S. Patent Office to disambiguate patent *recipients* across time (Monath and McCallum 2015). This algorithm uses a hierarchical coreference model to identify records that come from the same individual. The key question here is: how can we accurately identify whether or not applications from applicants with the same name come from the same individual? This process begins by matching individuals based on name and then produces a similarity score based on information contained in patent records, like technology classification, location, and coauthors. This method was first developed and described by Wick, Singh, and McCallum (2012).

Due to my use of only publicly available patent application data and the fact that I am processing patent applications and not granted patents, my inputs to the algorithm differ from USPTO's work. First, I do not have the International Patent Classification

(IPCR), Cooperative Patent Classification (CPC), or NBER patent category or subcategory for patent applications. Second, I do not have the text of application claims. Third, I do not have detailed data on inventor location beyond state, as this is not made public to protect the privacy of patent applicants. Finally, I do not have data on citations to patent applications. My process makes use of more detailed information about attorneys affiliated with applications than USPTO's algorithm.

Broadly, my process can be explained as follows: first, I structure the data into a set of "inventor mentions" which encompass all the information related to an inventor – application name, patent classification, co-inventors, assignee, lawyer, and location. I then run a fast hierarchical method for disambiguation that clusters inventor mentions into disambiguated entities, each representing a single inventor. For efficiency, the algorithm is limited to disambiguation of inventors with the same last name and first initial. Finally, there is a post-processing step that cleans inventor name to standardize them. The updated code in its entirety is available for public use at https://github.berkeley.edu/gsubramani/inventor_disambig under the Apache license 2.0.

1.4 Results

From 2001 to 2014, I identify 4,635,541 patent applications from 2,483,491 unique applicants. On average, 53.49% of applications are granted, which indicates that this sample captures 2.1 million applications that are missing when researchers only study granted patents. Table 1.2 summarizes the disambiguated data.

Prior research has explored coauthor networks using disambiguated data on inventors on granted patents (Fleming and Juda 2004; Fleming, Mingo, and Chen 2007; Fleming, King, and Juda 2007). Figure 1.1 illustrates the additional information gained from examining patent applications as opposed to just granted patents. Panel A of Figure 1.1 shows the coauthor network for one inventor, Thomas Tufano, when I limit to the sample of granted applications. Panel B shows Thomas Tufano's coauthor network when considering all applications he has submitted. Each line indicates a shared patent or application between two people, respectively. The stark difference between these figures depicts the value added by studying the full set of patent applications for an individual.

Another important note is that in the granted patent coauthor network, all of Thomas Tufano's coauthors are male, whereas in the application network, he has two female coauthors on applications that are not granted (Jennifer Boettcher and Yu Jin Lee). This highlights another dimension on which granted patent data may lead to a biased understanding of patenting, particularly given that applications from female inventors have a lower conversion rate to granted patents than applications from male inventors. The information depicted in this figure suggests that previous work exploring coauthor networks using data on granted patents only captures part of the story.

I evaluate the accuracy of my disambiguated results in two ways: first, by comparing my tracking of individuals whose applications successfully convert to patents to USPTO's own repeat patent recipient tracking, and second, by manually checking a sample of 950 entries.

1.4.1 Validating Matches using USPTO Disambiguation

I adopt the approach taken by Monath, Jones, and Madhavan (2020) to test the accuracy of their tracking of entities across applications and focus on three metrics: 1) precision, 2) recall, and 3) F1 score, each defined in detail below. Together, these metrics allow me to see how well my individual ID compares to the best measure of truth available (USPTO's disambiguated patent recipients). To calculate these statistics, I evaluate the presence of true positives, false positives, and false negatives in my data. In order to do so, I link my disambiguated data on granted applications to USPTO's data by matching on on inventor name×application number to connect entries of inventorID (USPTO's individual identifier) to disambigID (my individual identifier). This creates a single dataset that allows me to check how often these individual identifiers agree on whether repeat observations belong to the same inventor.

For each unique inventorID in the data, I identify the most commonly occurring disambigID (ties are broken randomly) and consider this to be the "chosen truth". Then, I look at the converse and for each unique disambigID, identify the most commonly occurring inventorID and consider this to be the "chosen disambiguation" (again, ties are broken randomly). Observations in the data are at the inventor×application level; each row represents a single applicant on a patent application.

If for a given observation, the inventorID and disambigID are both equal to the "chosen truth" and "chosen disambiguation", I consider this entry to be a **true positive**.

If, for a given observation, the inventorID is not equal to the "chosen truth", I consider this to be a **false positive**. This indicates that an observation has been grouped with other observations by disambigID, but if we consider inventorID to be an accurate categorization, then this observation should not be treated as such. Thus, this entry is wrongly attributed as being in the same cluster as the others with the same disambigID.

Finally, if for a given observation, the inventorID is not equal to the "chosen disambiguation", I consider this to be a **false negative**. That is, this observation *should* be clustered with a different value of disambigID, and is incorrectly excluded from this cluster. Table 1.3 presents an example of this data.

Precision, recall, and F1 score are defined as follows (Monath, Jones, and Madhavan 2020):

Precision:

$$\frac{\#\text{ true positives}}{\#\text{ true positives} + \#\text{ false positives}}$$

Precision provides a measure of how precise disambigID is by indicating the proportion of entries in a cluster (i.e., for a given unique disambigID) that are correctly grouped together.

Recall:

$$\frac{\text{\# true positives}}{\text{\# true positives + \# false negatives}}$$

Recall measures what fraction of observations that should have been in a cluster were correctly predicted to be in that cluster.

F1 score:

$$2 * \frac{\text{precision * recall}}{\text{precision + recall}}$$

There is an inherent tension between recall and precision. High precision requires that entries are only included in a cluster if they actually belong there. High recall, on the other hand, requires that entries are not left out of the cluster in which they belong. I use the F1 score to measure the balance between precision and recall.

Table 1.4 presents the values of these statistics. Precision has a value of 0.979, recall is 0.933, and the F1 score for the data is 0.955. To put these values in context, in Monath, Jones, and Madhavan (2020), the authors calculate these same statistics for the PatentsView disambiguated data using the widely-used NBER patent assignee dataset (Bessen 2009) as a measure of ground truth. They find values for precision, recall, and F1 of 0.957, 0.928, and 0.942 respectively. Comparing my statistics to these values indicates that my disambiguation method is performing well and that my disambiguation ID effectively tracks the same individuals across time as USPTO's unique patent recipient ID.

1.4.2 Evaluating "Ground Truth"

For the above verification exercise, I assume that USPTO's inventor ID represents the truth and is the standard against which the matches I identify should be validated. However, through a manual examination of the data, I find multiple instances in which my unique identifier (disambigID) appears to perform better than USPTO's inventor ID. Table 1.5 presents several such examples. InventorID shows USPTO's unique inventor ID, while disambigID is the unique ID I have generated that tracks a given individual across applications. DisambigID identifies three unique inventors in this table, while inventorID identifies six unique inventors. A careful review of this data shows that disambigID seems to be more accurate than inventorID. This suggests that the statistics calculated in Table 1.4 may actually *underestimate* the accuracy of my algorithm.

1.4.3 Validating Matches Manually

One limitation of using USPTO's individual ID to validate my disambiguated ID is that I am only able to perform this verification exercise on granted patents for which USPTO's ID exists. To address this concern, I conduct a manual examination of 950 entries in my disambiguated dataset to check whether disambigID accurately tracks the same individual over time. This approach allows me to check if disambigID correctly identifies the same inventor, but it is less effective in determining whether or not disambigID *misses* entries that it should include. That is, if there is an entry that should be included in a disambigID cluster but is not, I am not able to identify it through this process.

I manually audit a sample of 950 entries, which represent all observations for 200 unique inventors (disambigID), each of whom is affiliated with more than one patent application. I identify false positives as cases in which the same ID wrongly identifies two people as one and false negatives as cases in which the same individual is wrongly identified as two different people with two different IDs. As noted above, this approach is more effective at identifying false positives than false negatives. I review each entry and confirm that a single disambigID appears to identify the same individual across observations based on inventor name, location, and the title and technology used in the patent application.

Based on this process, I identify 47 false positive entries and 0 false negative entries. That is, for 47 of the 950 entries, an individual observation is incorrectly clustered with other entries. These 47 false positives occur across 23 unique disambigIDs. I find no cases in which individuals with different names are grouped together under one disambigID; all of the false positive entries are the result of inconsistent locations and/or technologies within a given inventor. For example, I observe multiple patent applications from inventors named John Lai in the United States. 18 of these applications are identified by disambigID as being from the same individual. These applications cover the time period between March 2002 and May 2013. For 15 of these 18 applications, John Lai is listed as residing in Ohio, but for one application in 2002, one application in 2008, and one application in 2013, his residence is in other states. These three applications are also in technology areas that are distant from John Lai's other patent applications, which primarily deal with polymers. Based on this information, I infer that these three applications are false positives, and belong to other individual(s), and thus they should have had a different value of disambigID. Table 1.6 presents these results.

Together, these verification exercises provide evidence that my disambiguation process effectively groups applications that belong to the same inventor together.

1.4.4 Limitations

The data described in this paper has some limitations. First, as described earlier, I am unable to identify gender for about 10% of inventors. Regarding tracking inventors over time, as my manual validation shows, disambigID is not 100% accurate. The presence of false positives suggests that this ID sometimes incorrectly groups observations together, which creates an inaccurate picture of inventors' trajectories. Further, due to the limited public availability of patent application data, I am only able to capture patent applications beginning in 2001. While data on granted patents is available and included in my data before this time, I am unable to see ungranted patents. As a result, if an inventor has applied for and not received a patent prior to that year, I am unable to observe this. Thus, this dataset provides a complete picture of inventors' activities from 2001 through early 2014, but I am unable to conclusively identify first-time applicants due to the possibility that an individual applied for a patent which was not received prior to 2001.

1.5 Conclusion

Patent applications represent a tangible measure of innovation and are a useful tool for researchers to evaluate trends in innovation. In this work, I describe the construction of a novel dataset that tracks patent applicants across time and detail trends in application behavior. Unlike existing disambiguated patent data, the algorithm I describe uses only public available data inputs to enable replication. The code and resulting data is available for researchers to explore across-application trends and trajectories.

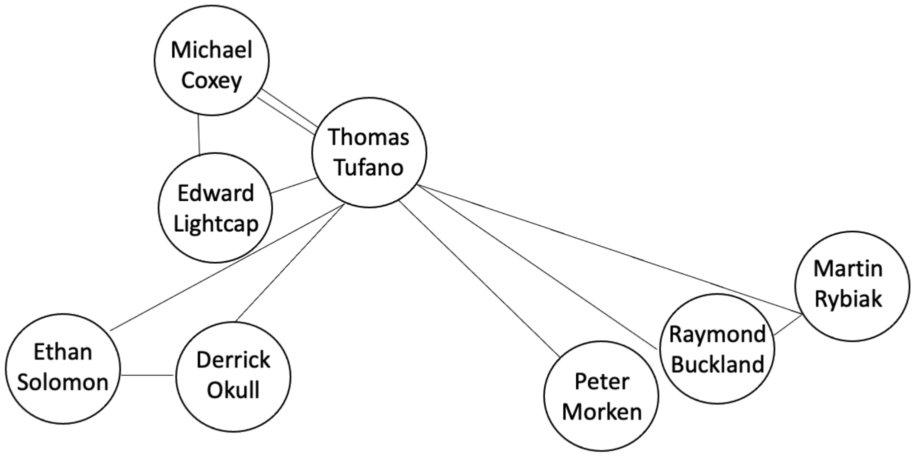
This dataset unlocks the potential to study inventors' intellectual careers and repeated interactions with the patent system, as well as variation in these outcomes by gender. This provides a context in which to track individuals' innovative contributions, including those that are not formally recognized through the legal protection of a patent. There are several promising avenues for future research using this data. First, it is possible to identify repeat co-inventors, as well as co-inventors' networks, and study how interpersonal ties affect both the frequency and direction of invention. Second, using this data, one can observe inventors' career moves between organizations and identify organization-specific effects on patent applications. Another broad set of questions that can be explored in this data relates to the role of experience on future innovative activity, and heterogeneity in the effects of experience and success or failure based on inventors' demographics or types of technology.

Using this data, I hope to continue to contribute to an understanding of the role of demographics, experience, networks, and prior success or failure on participation in and the process of innovation.

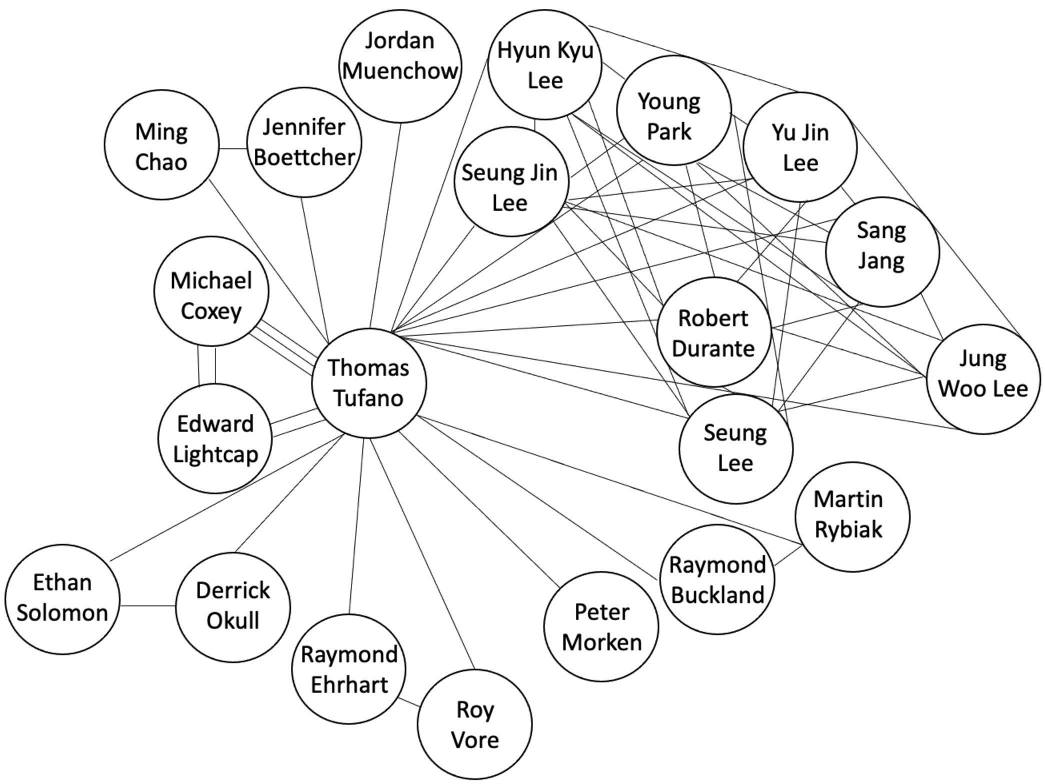
Figures

Figure 1.1: Evaluating Application and Patent Networks

Panel A: Full granted patent coauthor network



Panel B: Full patent application coauthor network



These two figures depict the differences between the coauthor network visible in the granted patent vs application data for one focal individual, Thomas Tufano. Each line represents a single joint patent (Panel A) or application (Panel B) between connected individuals. Note that inventors Seung Jin Lee and Seung Lee are distinct individuals.

Tables

Table 1.1: Challenges of Tracking Applicants Across Time

	Application #	First	MI	Last	Country	State	Filing Date	Art Unit	Invention title
1	11516467	John	M.	Adams	US	WA	6-Sep-06	3763	Disposable Infusion Device With Air Trapping ...
2	11528274	John	M.	Adams	US	VA	28-Sep-06	1791	Electrical Smoking System
3	11556722	John	M.	Adams	US	TX	6-Nov-06	2611	System And Method For Signal Phase Correction
4	11604166	John	M.	Adams	US	WA	22-Nov-06	3763	Disposable Infusion Device Filling Apparatus ...
5	11641596	John	M.	Adams	US	WA	18-Dec-06	3763	Cannula Delivery Apparatus And Method For A ...
6	12011472	John	M.	Adams	US	WA	25-Jan-08	3773	Transoral Endoscopic Gastroesophageal Flap ...
7	12011987	John	M.	Adams	US	WA	30-Jan-08	3773	Transoral Endoscopic Gastroesophageal Flap ...
8	12611997	John	M.	Adams	US	WA	4-Nov-09	3731	Shockwave Valvuloplasty Catheter System
9	12616582	John	M.	Adams	US	CA	11-Nov-09	3763	Wearable Infusion Device And System
10	12643983	John	M.	Adams	US	CA	21-Dec-09	2617	Distributed Architecure Wireless Rf Modem
11	12719758	John	M.	Adams	US	WA	8-Mar-10	3774	Device And Method For Modifying The Shape Of ...
12	13207381	John	M.	Adams	US	WA	10-Aug-11	3731	Shockwave Valvuloplasty Catheter System
13	61529564	John	M.	Adams	US	MS	31-Aug-11		Continuous Bladder Irrigation Alarm
14	13232730	John	M.	Adams	US	WA	14-Sep-11	3731	Shockwave Valvuloplasty Device With Guidewire ...

Table 1.2: Summary Statistics

	Mean	Std Dev	Min	Max
Applications (N=4,635,541)				
Inventors per application	2.39	1.76	1	98
Proportion of female inventors	0.10	0.24	0	1
Filing year	2006.9	3.53	2001	2014
Patent granted	0.53	0.50	0	1
Employer assigned	0.64	0.48	0	1
Using attorney	0.97	0.17	0	1
Inventors (N=2,483,491)				
Female	0.14	0.35	0	1
Applications per inventor	4.46	11.8	1	5083
Patents per inventor	2.38	7.92	0	4202
Patent grant rate per inventor	0.47	0.44	0	1
Proportion of coauthored applications	0.80	0.36	0	1
Mean number of coauthors	2.43	2.56	0	97

Table 1.3: Validating Disambiguation ID

Entry	inventor_ID (Truth)	inventor_ disambig	chosen_ disambig	chosen_disambig tiebreaker?	chosen_truth	chosen_truth tiebreaker?	True positive	False positive	False negative
1	A	100	100		A		1	0	0
2	B	200	200		B		1	0	0
3	C	300	300		C		1	0	0
4	A	100	100		A		1	0	0
5	B	200	200		B		1	0	0
6	B	200	200		B		1	0	0
7	D	100	100	X	A		0	1	0
8	E	400	400		E		1	0	0
9	F	500	500		F	X	1	0	0
10	C	600	300		J		0	1	1
11	C	300	300		C		1	0	0
12	D	700	100	X	D		0	0	1
13	E	800	400		E		0	0	1
14	G	500	500		F	X	0	1	0
15	H	900	900	X	H		1	0	0
16	H	1000	900	X	H		0	0	1
18	J	1100	1100		J		1	0	0
19	J	1100	1100		J		1	0	0
20	J	100	1100		A		0	1	1

This table shows the process for identifying a "chosen disabig" and "chosen truth" value for inventors across time so as to evaluate the accuracy of my disambiguation algorithm.

Table 1.4: Data Validation: Programmatic Comparison of Disambiguated Data vs. USPTO Data

	True Positives	False Positives	False Negatives
Number of entries	5349495	384949	114257
Proportion	0.916	0.066	0.019
	Recall	Precision	F1 Score
Score	0.93287	0.979	0.9554

Table 1.5: Disambig ID Compared to USPTO ID

inventor_id	disambigID	Application	First	MI	Last	Country	State	Title	Date
6906480-1	10097	10698342	Floyd	M.	Minks	US	FL	31-Oct-03	Regulator Control Circuit And Method
4024520-1	10097	11151690	Floyd	M.	Minks	US	FL	13-Jun-05	Regulator Control Circuit And Method
4024520-1	10097	11151690	Floyd	M.	Minks	US	FL	13-Jun-05	Regulator Control Circuit And Method
4024520-1	10097	11851621	Floyd	M.	Minks	US	GA	7-Sep-07	Power System For Producing Low . . .
4024520-1	10097	11851621	Floyd	M.	Minks	US	GA	7-Sep-07	Power System For Producing Low . . .
4118109-1	1224	10156151	Ian	D.	Crawford	US	FL	28-May-02	Laser Rangefinder Receiver
4118109-1	1224	10287665	Ian	D.	Crawford	US	FL	4-Nov-02	Pulse Discriminator
4118109-1	1224	10761126	Ian	D.	Crawford	US	FL	20-Jan-04	Single-Stage Power Factor . . .
4118109-1	1224	10889308	Ian	D.	Crawford	US	FL	12-Jul-04	Polyphase Diode Driver
4118109-1	1224	11054349	Ian	D.	Crawford	US	FL	9-Feb-05	Efficient Fast Pulsed Laser Or . . .
7773202-1	1224	11423367	Ian	D.	Crawford	US	FL	9-Jun-06	Laser Spot Tracker And Target Identifier
4118109-1	1224	12541915	Ian	D.	Crawford	US	FL	15-Aug-09	Biphase Laser Diode Driver And Method
4118109-1	1224	12835719	Ian	D.	Crawford	US	FL	13-Jul-10	Laser Spot Tracking With Off-Axis . . .
4118109-1	1224	12900453	Ian	D.	Crawford	US	FL	7-Oct-10	Smart Linear Pulsed Laser Diode . . .
4118109-1	1224	13015689	Ian	D.	Crawford	US	FL	28-Jan-11	Accuracy Of A Laser Rangefinder Receiver
4118109-1	1224	13531584	Ian	D.	Crawford	US	FL	25-Jun-12	Biphase Laser Diode Driver And Method
3978322-1	10150	10270799	Robert	C.	Dobkin	US	CA	11-Oct-02	Bidirectional Power Conversion . . .
3978322-1	10150	10324628	Robert	C.	Dobkin	US	CA	18-Dec-0	Circuits And Techniques For . . .
3978322-1	10150	10714825	Robert	C.	Dobkin	US	CA	14-Nov-03	Method For Manufacturing Sensing . . .
3978322-1	10150	10835693	Robert	C.	Dobkin	US	CA	29-Apr-04	Methods And Circuits For . . .
3978322-1	10150	11114267	Robert	C.	Dobkin	US	CA	25-Apr-05	Bidirectional Power Conversion . . .
7518351-2	10150	11131338	Robert	C.	Dobkin	US	CA	18-May-05	Switching Regulator Over Voltage . . .
3978322-1	10150	11493654	Robert	C.	Dobkin	US	CA	27-Jul-06	Class Ab Folded-Cascode Amplifier . . .
3978322-1	10150	11582443	Robert	C.	Dobkin	US	CA	18-Oct-06	Apparatus And Method For . . .
3978322-1	10150	11703718	Robert	C.	Dobkin	US	CA	8-Feb-07	Adaptive Output Current Control . . .
3978322-1	10150	11731279	Robert	C.	Dobkin	US	CA	30-Mar-07	Bandgap Voltage And Current Reference
3978322-1	10150	11827704	Robert	C.	Dobkin	US	CA	13-Jul-07	Paralleling Voltage Regulators
3978322-1	10150	12076302	Robert	C.	Dobkin	US	CA	17-Mar-08	Bidirectional Power Conversion . . .
3978322-1	10150	12174102	Robert	C.	Dobkin	US	CA	16-Jul-08	Class Ab Folded-Cascode Amplifier . . .
3978322-1	10150	12474938	Robert	C.	Dobkin	US	CA	29-May-09	Low Thermal Hysteresis Bandgap . . .
3978322-1	10150	12541557	Robert	C.	Dobkin	US	CA	14-Aug-09	Paralleling Voltage Regulators
3978322-1	10150	12610190	Robert	C.	Dobkin	US	CA	30-Oct-09	Voltage Regulator Compensating . . .
3978322-1	10150	13178302	Robert	C.	Dobkin	US	CA	7-Jul-11	Method For Clamping A . . .
3978322-1	10150	13424116	Robert	C.	Dobkin	US	CA	19-Mar-12	Circuitry To Prevent Overvoltage . . .

Table 1.6: Data Validation: Manual Audit of Disambiguated Data

	True Positives	False Positives	False Negatives
Number of entries	903	47	0
Proportion	95.05	4.96	0
	Accurate Match	>=1 False Positive	>=1 False Negative
Number of individuals	177	23	0
Proportion	88.5	11.5	0

Chapter 2

Try, try, try again? Differential Responses to Rejection & the Gender Innovation Gap

2.1 Introduction

By a variety of measures, women are underrepresented in innovation; in 2010, only 15.3% of all patents had at least one female inventor. This gap has narrowed over time as the number of female inventors has increased, but at the current rate, women would only achieve parity in patenting in 2092 (Milli et al. 2016). Women's differential participation in patenting is not driven solely by differences in education and occupation by gender; the gender gap in patents received is itself greater than the underrepresentation of women in STEM education and careers would suggest. Women comprise 28% of scientific and technical workers and only 12% of inventors on granted patents, despite making up over half of the workforce (Thébaud and Charles 2018). Indeed, women are even less likely to receive patents than they are to pursue entrepreneurship (Toole, Breschi, et al. 2019). These facts suggest that even when women are well-positioned to contribute to innovation, they are underrepresented in formal records of innovation.

Participating in innovation has implications both for individuals' careers in the form of compensation and career trajectories, as well as for firm growth. Independent inventors may commercialize or sell patent rights, and within organizations, individuals who receive patents can see direct wage increases (Kline, Petkova, et al. 2019). Having a patent can also affect future employment by affecting the likelihood that an inventor will switch jobs (Melero, Palomeras, and Wehrheim 2020). Patents are valuable for firms as well not only due to the intellectual property protection they provide but also because of the signal that receiving a patent sends to investors. Startups that hold patents have higher sales and employment growth, and receiving a patent increases access to external funding from venture capital firms and banks (Gaulé 2018; Farre-Mensa, Hegde, and Ljungqvist 2020). Gender disparities in innovation can thus exacerbate differences in labor market outcomes between men and women at both the individual and firm level.

If, conditional on having an innovative contribution, women do not patent at the same rates as men, women's inventions may be lost, with negative implications for potential inventors as well as the progress of innovation. This can have distributional consequences; innovation comes from expertise, but also from exposure to contexts that could benefit from innovation and situational awareness. For example, female-led patent teams and teams including women are significantly more likely to produce female-focused innovations (Koning, Samila, and Ferguson 2019). If women are less likely to participate in innovation, the female population may lose out differentially.

Work exploring the causes of the gender gap in innovation has primarily focused on the pipeline for innovators and bias in evaluation processes. The gender gap in patents could be, among other things, a result of the lack of gender diversity in science and technology and biases that women face in innovative fields with respect to career advancement and access to resources. Conditional on studying STEM subjects, female college graduates are less likely to transition to STEM jobs (Sassler, Michelmore, and Smith 2017), and female academic scientists are less likely to patent due to having more limited professional networks. (Ding, Murray, and Stuart 2006). Prior research has found that the small proportion of women who do participate in innovation are subsequently disadvantaged by biased evaluations of their accomplishments and capabilities as compared to similarly qualified men (Ridgeway and Correll 2004; Ding, Murray, and Stuart 2013). This has measurable implications for innovation and individuals' careers; one third of the gender gap in financing for entrepreneurs is due to venture capitalists' reliance on negative stereotypes in the face of uncertainty (Guzman and Kacperczyk 2019).

These findings provide important insight into the dynamics that affect who enters and succeeds in innovation and entrepreneurship. However, these explanations overlook another channel that arises when one considers the fact that innovation is a setting in which failure and rejection are a part of the process (Kline and Rosenberg 1986; Holmstrom 1989). In fact, tolerance for failure can even enable more significant innovations (Tian and Wang 2014). If women and men differ in the extent to which they continue to pursue innovative activities after facing rejection- whether due to support, resources, or other drivers- their innovative trajectories may also look quite different. In other words, if women who acquire the skills and resources to invent respond differently to rejection than male inventors, the set of successful inventors will be disproportionately male even conditional on the initial pool of ideas. Work in other domains suggests that women are less likely than men to consider a job opportunity at a firm that has previously rejected them (Brands and Fernandez-Mateo 2017), and women are less likely than men to create a crowdfunding campaign following an initial failed crowdfunding attempt (Kuppuswamy and Mollick 2016). Whether similar dynamics are at play when considering innovation remains an open question.

In this work, we focus on the role of gender differences in responses to rejection, which we describe using the term "persistence," in driving the gender gap in innovation in the context of patenting. There are a number of reasons why women may be less likely to follow up than men after receiving a rejection. For example, female inventors may have less familiarity with the patenting process and interpret a rejection as a signal that their idea is not worth pursuing further. It could be that women are interested in following up on their applications but lack the resources- time, money, or information- to do so (Ceci and Williams 2011). Or, it may be that women are more likely to interpret negative feedback as an appraisal of their quality and thus lose confidence in their inventions (Cech et al. 2011). Women might believe, correctly or not, that they face bias from evaluators and thus feel there is less value in responding to a rejection. Finally, it could be that women's ideas are of differing quality from men's ideas, and knowledge of their quality leads women to invest less in pursuing their innovations. We do not attempt to provide an exhaustive accounting of why this may be the case, nor do we make an argument about the innate differences between women and men. Rather, we explore potential drivers of this gap by leveraging our empirical strategy to remove concerns about quality and bias.

Specifically, we explore how resources and information affect patent outcomes by examining the effects of patent attorneys and firm affiliation. Many inventors elect to use patent attorneys to manage their applications. Lawyers can be an important source of information for inventors regarding how to write their applications and respond to comments from examiners so as to increase the chances of patent receipt. We also examine the outcomes of patent applications that are associated with firms. Firm assignment indicates that a firm will assume the ownership right of the patent, and is generally accompanied by financial support and guidance through the patenting process.

Patent prosecution is an excellent context in which to examine responses to rejection due to the large number of patent applications from both men and women, the availability of detailed data, and the iterative nature of patent review. Additionally, the patent process itself is an important setting in which to understand gender differences and identify potential policy interventions to increase the presence and success of women. We use newly available data from the United States Patent and Trademark Office (USPTO) on patent applications in the United States from 2001 through 2012. The final sample covers almost one million applications from individuals and teams based in the United States- both those that received patents and those that did not. The data include key information about a patent application, including the technology class, if application is associated with a firm, the outcome of the application (whether a patent is issued), and innovators' full names, which we use to elicit gender. We can also see each application's prosecution history, which details each step of the patent process. Most applications receive at least one rejection, coupled with feedback from a patent examiner, to which inventors must respond in order to continue their applications (Lemley and Sampat 2008). We consider innovators' tendencies to follow up on an application and amend their claims as a measure of persistence.

Estimating the causal effect of rejection on patent application is empirically challenging. The likelihood of receiving a rejection and the feedback that an inventor receives might be correlated with a host of unobservable application attributes. We leverage the quasi-random assignment of applications to examiners to isolate the role of rejection. The intuition behind this strategy is that different examiners have different propensities to approve patents. More lenient examiners are more likely to grant a patent than are harsher[1] examiners, holding the quality of the proposed invention constant. In order to identify the causal effect of rejection on patent continuation, we use examiner harshness across all *other* applications in the same technological specialization and year as an instrument for patent rejection.

Our estimates using this IV strategy indicate that majority-female teams are 3.3-7.3 percentage points less likely to continue the patent process after receiving an initial rejection compared to male inventors. This differential effect by innovators' gender is magnified when examining whether a patent is ultimately issued; we find that an initial rejection differentially reduces the probability that a patent is granted by 5.9-10.4 percentage points more for females as compared to their male counterparts. This disparity in the continuation of applications from men and women accounts for more than half of the gender gap in granted patents, conditional on application. When we restrict our attention only to applications filed by individuals, our estimates suggest that a rejection reduces the percentage of female applicants by 4.5 percentage points. This effect remains statistically significant when we examine the effect of the proportion of women inventors, or use indicators for whether the innovating team consists mostly or solely of female innovators.

We next explore the effect of using a lawyer on responses to rejection and ultimate patent outcomes. We find that lawyers are more likely to push an application forward following the receipt of a rejection, and that female inventors benefit more from using an attorney to represent their applications; applications from women are upwards of 4 percentage points more likely to proceed beyond an initial rejection if they are managed by an attorney. As such, attorneys can help shrink the gender gap in patent continuation and receipt. This result suggests that differences in information about the patenting process may be a key driver of the gender gap that we observe.

We then examine whether individuals and teams that submit an application with the support of a firm behave differently than unaffiliated inventors. We find that in the aggregate, firm-backed applications are considerably more likely to proceed beyond an initial rejection. This effect is even stronger for applications whose authorship is majority female; applications from half- and all-female applicant teams are 3.8 and 5.1 percentage points (respectively) more likely to follow up after receiving an initial rejection if they are affiliated with a firm than similarly situated male applicants. The positive effect of firm affiliation on female inventors' responses to rejection in pursuing patents suggests that the provision of resources to inventors and organizational management of patent applications can help shrink the gender gap in response to rejection, and thus in patenting outcomes.

Our findings contribute to the existing literature on the gender innovation gap. Understanding where in the process of innovation women fall out and why this happens is essential in order to develop solutions that address the gender gap in innovation (Delgado, Mariani, and Murray 2019; Cook 2020). We build on prior work identifying a gender gap in the conversion of applications to granted patents (Jensen, Kovács, and Sorenson 2018) by exam-

[1]Consistent with other work in this domain (Farre-Mensa, Hegde, and Ljungqvist 2020), we refer to more stringent examiners as being "harsh".

ining how a key feature of the patent process, the receipt of rejections and subsequent need for correspondence with patent examiners, drives differing outcomes for male and female inventors. Our results shed light on gender differences in entrepreneurship and innovation (Ding, Murray, and Stuart 2006; Ding, Murray, and Stuart 2013; Guzman and Kacperczyk 2019) and highlight a potential driver of broader gender disparities in participation in these fields.

Our results also speak to the literature studying gender differences in organizations (Fernandez-Mateo and Coh 2015; Brands and Fernandez-Mateo 2017) and the role that institutional investments can play in addressing and shrinking performance gaps between men and women (Blau et al. 2010; Srivastava 2015). Our findings suggest a channel through which organizations can effectively improve outcomes for women. Finally, we add to the literature studying gender differentials in response to rejection in other settings, such as politics (Wasserman 2018), and crowdfunding (Kuppuswamy and Mollick 2016) by showing how differential responses to rejection by gender contribute to variation in outcomes in the context of innovation.

2.2 Setting

To identify differences in responses to rejection as one channel that may contribute to gender disparities in rates of innovation, as well as underrepresentation in the science and technology labor market generally, we examine how men and women respond to first-time rejections within the patent process. Patents in the United States are granted by the the U.S. Patent and Trademark Office (USPTO). We begin with an overview of this setting, and how it provides the components of a natural experiment relevant to understanding gender differences in responses to rejection in this high-stakes setting.

Between 300,000 and 500,000 patent applications are filed at USPTO annually (Frakes and Wasserman 2017). Application fees vary depending on the applicant, type of patent application, the application's level of complexity, and other features of the application, but the total cost of fees paid to USPTO are a minimum of roughly $500 (excluding any attorney fees) (USPTO 2016). Each application makes a set of claims delineating the legal rights that an inventor is seeking. This includes disclosure of "prior art," or existing patents material to the patentability of the invention. An application submitted to USPTO is first directed to an art unit, which comprises patents of the same technological field. Within an art unit, a supervisory examiner quasi-randomly[2] assigns the application to a specific examiner, who oversees the application for the remainder of its existence. The assigned examiner first assesses the viability of an inventor's claim. She compares the prior art with the claims

[2]Based on qualitative and empirical evidence: for example, Frakes and Wasserman 2014 conducted a series of telephone interviews and confirms that supervisors do not make any substantive evaluation of an application before assigning it to an examiner. Additionally, Sampat and Williams 2019 (as well as our analysis) provide empirical evidence that, conditional on year and art unit, assignment of applications to patent examiners is plausibly random.

provided on the patent application, often conducting an interview with the applicant or the applicant's attorney. On the basis of the examiner's evaluation, she will write a "first office action" (FOA) letter that either accepts or rejects the claims. Some applications are accepted in their entirety during the examiner's initial examination. Most (over 80%) are not immediately accepted, with some or all of the claims failing to meet the requirements for patentability.

Patent applications are not categorically rejected by USPTO; rather, they are either implicitly or explicitly abandoned by applicants following what technically are appealable rejections issued by patent examiners (Lemley and Sampat 2008). This includes at the FOA stage; most patent applications are rejected in the first interaction with the examiner. An applicant will typically then respond to an initial rejection by amending the claims. The examiner will then again either allow the claims in the amendment or reject them. That second rejection is typically categorized as a "final" rejection; however, at times, an application receives more than one initial rejection or more than one final rejection. However, even when a "final rejection" is issued, an innovator can still attempt to convince the patent examiner that he/she is entitled to a patent via the submission of an amendment or appeal.

Receiving a rejection in the initial stage of the review process is not necessarily an indication that the innovation is of poor quality; in fact, over 60% of applications that are initially rejected eventually result in patents. As such, completing the patent prosecution process (i.e., the process of process of drafting, filing, and negotiating with the USPTO over patent rights) can require many months of work to remedy purported defects in an initial application and continue to try to convince examiners that a patent should be granted. Given the nature of the patent prosecution process, it can take several years from an application filing to the point at which a final patentability decision is made (Frakes and Wasserman 2014). Thus, the majority of patent applications require effort, patience, and persistence from the inventor in order to generate a "Notice of Allowance," which occurs when an examiner decides an applicant is entitled to patent protection for an invention. Our study examines whether, after receiving an initial rejection, women are more likely to exit the back-and-forth between applicant and examiner that is typically required for a patent to ultimately be granted. We measure responses to rejection by examining whether applicants follow up on early-stage administrative decisions by the USPTO.

Figure 2.1 provides a summary of the patent prosecution process and shows the raw proportions of applications in our sample, separated by gender of applicants, that proceed through each stage. Note that Figure 2.1 reports raw proportions without controls for application art units or time. "Initial" and "final" are used to describe the stages of rejection; however, as stated above, a single application can receive multiple initial rejections and final rejections to which the inventor(s) can respond with an amendment or appeal. Initial rejections are those that a patent application receives prior to an initial acceptance of some of the application's claims; some applications receive more than one initial rejection over the course of multiple rounds of review. Final rejections are those that the patent application receives following an initial acceptance of claims and prior to being granted the patent. At any point of time, the applicant can cease to respond to the examiner's communication.

As is evident in Figure 2.1, following both initial and final rejections, female applicants are far less likely to respond to the examiner's comments than male applicants. This difference affects which applications subsequently receive a patent; 63.6% of applications from male applicants that receive an initial rejection eventually go on to receive a patent, while for applications from women, that proportion is only 49.2%. Note that once controls are introduced in subsequent analyses (Table 2.3), the initial rejection rate for women disappears. For the purposes of this paper, we consider applications for which inventors did not respond to an office action to be those which the inventors are no longer pursuing. This is a broader set of applications than those that are formally identified as abandoned in the patent data, as it captures applications for which a lack of response followed examiner communication other than an action letter.

2.3 Empirical Framework

2.3.1 Data

To study differential responses to rejection between men and women in the patent prosecution process, we use individual patent application data over the course of 12 years from the Patent Office's Patent Application Information Retrieval (PAIR) database. The PAIR data cover all patent applications that were filed on or after March 2001. Our analysis focuses on applications that were filed from this date through 2012. USPTO application records include basic information on a patent application, including the technology class, art unit, firm assignment, and the outcome of the application (whether a patent is issued). Importantly for our analysis, the data also include complete prosecution histories, including the examiner assigned to review an application, which we use to implement our identification strategy. The PAIR data also includes the date and sequence of application filing, each examiner rejection, applicant amendments and appeals, and final patent allowances.

Note that the gender of the applicant is not an explicit field within the USPTO data. We follow the same process as Jensen, Kovács, and Sorenson (2018) to impute the gender of an applicant. We use publicly available data on the gender distributions of first names, in our case, downloaded from the U.S. Social Security Administration (SSA), to identify the frequency with which specific names are given to males and females born in the United States. There are 97,310 unique names in the SSA data, for which we construct a gender distribution by name. For example, if there are 10,000 people with the name Carol, 9,000 of whom are were women and 1,000 male Carols, then the name Carol would receive a female proportion of 0.9. We match these names to the 271,000 unique patent applicant names from the USPTO data to assign gender to patent applicants. We use a 90% cutoff threshold and drop applications for which any inventors' names are assigned either male or female less than 90% of the time (a proportion of less than 0.9). We only include applications for which we can assign gender to *all* inventors on the team. Using this method, we are able to identify the gender of all inventors for 71% of applications. Additionally, to ensure confidence in our gender identification of names, we limit our sample to applications for which all inventors are based in the United States. We focus on a few different measures of gender composition: (1) whether an application team is composed of 50% or more women (half-female), (2) whether it is composed of all women (all-female), and (3) the proportion of women on an application. We also conduct analyses in which we restrict attention to only applications submitted by individuals.

We use this same process to identify the gender of patent examiners. Each application is assigned to an examiner. Examiners are identified by unique IDs, and examiners' first and last names are also present in the data. In total, there are 7,700 unique examiners who appear in our sample. We are able to identify gender for 6,397 of these individuals. The majority of examiners are male, with just 27.9% of female examiners appearing in our data. The average examiner reviews 125 applications between 2001 and 2012. Male and female examiners see similar numbers of applications each year, with female examiners being assigned to slightly more applications than male examiners over our 12 year period (means of 138.6 and 123.3 for female and male examiners respectively).

We drop applications from teams composed of more than 10 individuals. In the remaining sample, we examine all applications, both those submitted by individuals and from inventor teams (composing 47% and 53% of our sample, respectively). We also use other information collected by USPTO to test for heterogeneity and examine mechanisms underlying our results. Each application is associated with a correspondent, who is most often the attorney managing the patent application but may also be the inventor(s) if they elect not to use an attorney. We leverage the presence of an attorney docket number to identify whether an lawyer is involved with a given application. Using natural language processing techniques, we track attorneys across time in the data. The majority of lawyers and law firms appear infrequently in the data; the median lawyer submits two applications and receives only one patent. We further identify top law firms in the correspondence data using company rankings from Vault, the leading source for law firm ratings.

Using USPTO's data on assignment of application rights, we can tell whether an inventor is connected to a firm or is applying solo. The majority of applications (63%) are associated with firms. A firm "assignment" indicates that a firm will assume the ownership right if granted, and is thus an indication of firm backing of an application.[3,4] Data on firm assignment is manually entered by applicants, and there is no post-processing standardization applied to firm names. As a result, employer names exhibit a great deal of variation and errors. For example, two applications that are assigned to the same company could have firm names that are "Apple, Inc" and "Apple Corporation of California," respectively. To accurately identify firms and evaluate effects by firm size, we use natural language processing

[3] Legally, the "original applicant is presumed to be the owner of an application for an original patent, and any patent that may issue therefrom, unless there is an assignment" (Graham, Marco, and Miller 2015).

[4] When an inventor or team of inventors apply for a patent, they are presumed to be the owner of the patent. Applicants can assign their idea to an organization or entity, generally the company for which they worked when they produced the idea.

techniques to standardize employer names. Using this approach, we reduce the number of unique firms in the application assignment data from 109,952 to 77,093.

The core of our analysis focuses on almost one million applications. Table 2.1 provides summary statistics for the sample. As is clear from this table, the majority of applications in our sample result in patents. By all measures, women are underrepresented in applications, both on teams and as solo applicants. Also evident is the frequency of rejections in the applications process; the mean number of initial rejections received is greater than one.

2.3.2 Empirical Strategy

The goal of our study is to cleanly identify heterogeneous responses from men and women to negative decisions made by patent examiners, particularly at the First Office Action (FOA) stage of the patent application process. In other words, we identify gender differences in the likelihood that an applicant "persists" after having initial patent claims rejected (an *initial rejection*). We operationalize "persistence" in this setting as the likelihood of continuing with the patent prosecution process following a rejection at the first stage. While a fraction of patents will be granted based only on the original claims, many more will involve adjustments to the application.

In other words, we examine whether, after receiving an initial rejection (the modal outcome in terms of an examiner's initial decision), women are differentially likely to submit an amendment responding to the points raised in the rejection. The ideal experiment to identify gender-specific responses to patent denials would be to randomize patent denials at the first instance across men and women, thus ensuring that successful applicants and their inventions do not differ systematically from unsuccessful ones ex-ante. Gender differences at the next stage would then indicate differential selection out of the patent prosecution process.

We model the relationship between receiving an initial rejection and continuing the patent prosecution process (including final receipt of a patent) as follows:

$$Y_a = \beta_1 Initial\ Rej_a + \beta_2 Female_a + \beta_3 [Female\ \times\ Initial\ Rej]_a + \mu_{ut} + \epsilon_{aut} \tag{2.1}$$

In this setup, a indexes a patent application, and ut the patent art unit*application year. Y_a is the outcome of interest, either whether inventors continued the application following a rejection or whether the application was approved. $Initial\ Rej$ is a dummy for whether the application received an initial rejection, and $Female$ is an indicator for the prevalence of females in the inventors team, as described in Section 2.3.1. We are interested in identifying the coefficient estimate for β_3, which tells us how likely women are to either amend an application or finally obtain a patent, conditional on receiving a rejection at phase one of the examination process, relative to men. In other words, β_3 captures the gender *differential* in response to rejection.

A simple comparison of how women compare to men in terms of continuing with the application process after an initial rejection will fail to cleanly isolate gender differences in responses to rejection independent of gender differences in application characteristics. That is, a comparison of gender-specific means may yield biased estimates because whether or not an application receives a rejection is likely correlated with application characteristics, many of which are largely unobservable. If, for example, men file applications for more incremental innovations that are inherently more easily obtained, they may receive rejections that have more straightforward, easy to address comments, and thus they may appear more persistent following a rejection. To estimate the effect that gender differences in response to rejection have on exit from the application process, we need variation in initial patent outcomes that is unrelated to other determinants of patents, such as the merit of a given application.

We use an instrumental variables (IV) strategy that leverages exogenous variation in the likelihood of receiving a rejection at the first step in the application review process. To obtain this type of exogenous variation, we use the quasi-random assignment of applications within USPTO's review process to get as close as possible to the ideal experiment. The intuition behind this approach is that examiner harshness directly affects the likelihood of receiving a rejection, and it is also not correlated (as we show) with application quality and potential outcomes. The USPTO assigns examiners to applications quasi-randomly, and examiners differ systematically in their propensity to approve patents at any stage. Some examiners are more lenient and are more likely to grant patents, while others are more stringent and give a greater proportion of rejections. Thus, the variation in examiner harshness allows us to isolate the effect of rejection on application continuation. Our design is similar in spirit to several recent studies about the patent prosecution process, such as Sampat and Williams (2019) and Farre-Mensa, Hegde, and Ljungqvist (2020).

We define examiner leniency as the leave-one-out initial rejection rate of examiner by art unit-year (i.e., the proportion of all *other* applications for which a decision of initial rejection is made by a given examiner in each art unit-year). A higher value of harshness indicates that an examiner gives more rejections.[5,6]

$$Harshness_{ae} = \left(\frac{1}{n_e}\right)\left(\sum_{k\neq a}^{n_e} ER_k\right)$$

In this expression, e indicates the examiner assigned to an application a, n_e is the total number of applications seen by examiner e in art unit- year, k indexes the applications seen by examiner e, and ER_k, an initial reject, is equal to one if the applicant did not receive a patent when the first response was given by examiner e for patent application k. We construct this measure for each application, so for a given application, the harshness measure captures how stringent the assigned examiner is based on all other applications she reviews. This measure of harshness avoids any bias from the current application. Figure 2.2

[5]We construct examiner harshness based on art unit-year because applications are assigned to examiners with art units in a given year.

[6]In Table A.1, we discuss an alternative definition of examiner leniency, using the leave-one-out patent rejection rates; that is, how harsh an examiner is when it comes to actually granting a patent. The main results are robust to using the alternative definition.

shows that examiner rejection rates at the initial decision stage vary substantially within year and art unit.

When we use this instrument to study the likelihood of patent continuation (responding to a rejection), one key insight is that examiners who are more likely to reject applications (those who are "harsher") are also those who give *more stringent* rejections. A useful analogy is to the process of academic publishing; referees who are more likely to reject submissions are also those who, when they write a review, are more likely to suggest substantial revisions. Similarly, when we compare the responses to rejections received from more lenient versus harsher patent examiners, we are examining how applicants react to incremental versus more significant comments.

Using examiner harshness as an instrument, we can then cleanly estimate heterogeneity in patent prosecution persistence:

$$Y_a = \beta_1 \widehat{Initial\,Rej}_a + \beta_2 Female_a + \beta_3 [Female \times \widehat{Initial\,Rej}]_a + \mu_{ut} + \epsilon_{aut} \tag{2.2}$$

where we instrument for $\widehat{Initial\,Rej}_a$ and $[Female \times \widehat{Initial\,Rej}]_a$ using $Harshness_e$ and $[Female_a \times Harshness_e]$, respectively.

Before proceeding, we provide evidence in favor of a strong first-stage relationship. We report results of the first stage of our IV specification and examine the effect of the assigned examiner's harshness on the likelihood of rejection in Table 2.2. An examiner's prior rate of rejection has a significant effect on the likelihood of receiving an initial rejection, robust to the inclusion of covariates that measure characteristics of the application and applicants. In each of these specifications, the effect of examiner harshness on rejection is statistically meaningful, with F-statistics well-above the threshold of 104.7 (Lee et al. 2020).

To satisfy the exclusion restriction, the only channel through which an examiner's stringency should affect outcomes is through the receipt of a rejection, and there should be no relationship between the assigned examiner's rejection rate for *other* applications and a patent application's potential outcomes. In our context, it is difficult to imagine how an examiner's harshness could affect patent outcomes other than through the likelihood that an application receives a rejection. Thus, any concern about identification would come from the harshness of the assigned examiner being correlated with omitted variables that affect the likelihood of following up after rejection or ultimately receiving a patent. Additionally, it is important to note that in our study, as we are interested specifically in the *differential* response of male and female applicants who are assigned to equivalently harsh examiners. That is, we do not examine the effect of patent rejection or receipt on potential outcomes like commercialization; instead, we identify the gender-specific effect of harshness on application continuation and receipt. As such, our design is even less vulnerable to these concerns regarding identification.

Following the advice of Righi and Simcoe (2017), we test our assumption of quasi-random assignment of applications to examiners by estimating our first-stage Ordinary Least Squares regression with and without subclass-fixed effects. This allows us to test whether the technology an application deals with is correlated with examiners' propensities towards rejection.

Table 2.2 presents these results. We find that adding subclass-year fixed effects in Column 3 has very little impact on the coefficient on the instrument, which varies from 0.709 to 0.695 after the inclusion of subclass controls. This represents a less than 2 percent change in magnitude and suggests that neither subclass nor our proxies for application quality and characteristics predict whether the application receives a rejection.

We further examine the relationship between an application's quality and how harsh the assigned examiner is in Figure 2.3, which relates examiner harshness to the *actual* initial rejection rate, shown in red, and the *predicted* rejection rate in yellow. We calculate the mean initial rejection rate for each examiner, residualized by Art Unit-by-application year fixed effects. When we relate this measure of examiner "leniency" to initial rejection, by construction, we observe a strong relationship. We predict rejection based on a number of application covariates that may be related to quality, like the number of inventors on an application, the proportion of female inventors, whether the application is assigned to an employer, and the employer's patenting track record. Together, these variables predict 40% of the variation in initial rejection rates. This figure provides indirect support of a strong first stage relationship, as well as of the exclusion restriction and suggests there is no visible relationship between examiner harshness and predicted rejection. Drawing a harsher examiner increases the probability of initial rejection, regardless of patent quality or other ex ante characteristics.

2.4 Results

2.4.1 Main Findings

We begin our analysis by presenting suggestive evidence that persistence has a central role in explaining the gender innovation gap. Panel A of Table 2.3 presents the results of a simple OLS regression of the impact of gender on initial rejection rates, controlling for application art unit-year. The definition of female changes across columns: Column 1 provides an estimate for the effect of the proportion of women on an application; Column 2 defines female as equal to one if 50% or more of the inventors on an application are women; in Column 3, female is equal to one if 100% of inventors on an application are women (includes applications from solo women), and Column 4 includes only solo applicants and female is equal to one if the sole applicant is a woman.

We find suggestive evidence that teams with more women are marginally more likely to receive initial rejection compared to majority male teams. While statistically significant, the effects are small in magnitude. For instance, in Column 4, which restricts attention to applications filed by individuals, we find that female inventors are 0.9 percentage points more likely to receive an initial rejection. This coefficient estimate thus suggests that women are only 1.16 percent more likely than men to have their initial claims rejected (evident from dividing 0.009 by the mean of 0.77). This effect is even smaller for all other specifications. Since innovator gender might be correlated with unobserved characteristics such as patent quality, we cannot rule out that examiners are discriminating against female innovators. Nevertheless, the above evidence suggests that discrimination, at least at the First Office Action stage, is not a major driver of the innovation gender gap.

In contrast, observing Panel B in Table 2.3, we see that the presence of women on inventor teams is correlated with significant reductions in the probability of a patent being granted. For instance, as is shown in in Column 3, applications filed by all-female teams are 7.2 percentage points less likely to receive a patent compared to applications filed by *all other* applicants. We observe similar magnitudes across all specifications. This suggests that there exists a gender gap in application conversion rates. More importantly, this gap is not driven by differences in initial rejection rates; women's higher rates of initial rejection do not alone explain the significantly different rates of patent receipt. Instead, this gap arises in subsequent stages of the patent application. This finding is consistent with our hypothesis that heterogeneous responses to persistence drive the innovation gender gap.

To further investigate the concern that examiners themselves may be biased and may apply more stringent requirements to female applicants, we investigate whether examiners exhibit gender bias and how this interacts with the examiners' own genders. To do so, we evaluate the interaction between examiner gender, applicant gender, and initial rejections. Results are presented in A.2. We find no evidence that the examiner's gender has any bearing on the persistence gap between men and women.

2.4.2 Differential Responses to Rejection by Gender

We now turn to our primary question, which studies heterogeneity by gender in innovators' responses to initial rejections. This approach allows us to directly compare male and female inventors' responses to rejection in a unified regression framework. We evaluate this using the full sample of patent applications.[7] Table 2.4 presents our primary results. Each column indicates an estimation using the instrumental variable strategy in Equation 2.2. The definition of Female changes across columns and is indicated in the last row of each column, similar to Table 2.3. We focus on two primary outcomes: 1) whether an applicant/team proceeds to the next step of the application, i.e. files an amendment, and 2) whether the application is eventually granted a patent. Collectively, the results demonstrate that women and majority-female teams are significantly less likely to continue in the patent process if they receive an initial rejection compared to their male counterparts. This finding is consistent all of our specifications.

We begin by focusing on Panel A. These regressions capture the effect of receiving an initial rejection on submitting an initial amendment, the next step to keep a rejected patent application alive. As expected, mechanically, receiving an initial rejection increases the

[7]In Table A.3, we conduct the same analysis but limit our sample to applications from individuals who apply completely independently; that is, applications that do not use attorneys and that do not indicate firm assignment. This sample represents applications for which we can reasonably assume 1) the inventor is handling communication with the examiner directly, and 2) there is no outside sponsorship of the patent application by an organization. Our results remain consistent.

likelihood that applicant(s) subsequently submit an initial amendment in response. Column 1 summarizes the negative relationship between the presence of women on applications and innovation across our entire sample: for every 10% increase in women on an application, the likelihood of abandoning an application following a rejection increases by up to 0.39 percentage points (p.p.). Recall that because we are leveraging random variation in likelihood of rejection at this stage, these estimates avoid potential bias from unobservable application characteristics. The coefficients on initial rejection in these analyses provide a baseline against which to compare the differential effect for women. Almost 85% of applications from men (in Columns 1-3) submit amendments after receiving initial rejections.

Columns 2 and 3 provide estimates of the primary specification using different measures of the female presence on inventor teams. These estimates again indicate that women are less likely than men to continue in the patent process following an initial rejection in the patent examination process. In Column 2, we observe that when patents whose authorship is primarily female (patents in which 50% or more of the inventors on an application are women) receive an initial rejection, these applications are over 3 p.p. less likely to continue with the application process than applications whose authorship is majority male.[8] Finally, all-female teams are over 7 p.p. less likely to continue with a patent application after an initial rejection. Given the coefficient on Initial Rejection which indicates a baseline response rate of 85%, these estimates represent a 3.8-8.5% reduction in the likelihood of submitting a response for majority and all-female teams.

In Column 4, we limit our sample further to only applications submitted by individual inventors. This approach, while reducing our sample size, provides arguably the cleanest estimate of gender persistence disparities in this innovation setting. Examining individual inventors allows us to avoid issues related to selection in team composition. The estimates in Column 4 are similar in magnitude to our other specifications, and suggest that a woman who applies for a patent is 4.5 p.p. less likely than a man who is facing an equally stringent examiner to amend her application if her initial set of patentability claims is rejected at the initial stage.

Given that the proportion of female applicants is low to begin, with just 15% of applications having at least one female inventor and only 4.3% of applications coming from all-female teams or solo women, these estimates represent a significant reduction in the number of women who remain in the patent application process. The monotonic negative relationship between the fraction of woman on an application and persistence provides compelling evidence that women are differentially deterred from continuing in the patent process after an initial rejection. This also indicates that the *intensive* margin of female representation matters when it comes to patent outcomes.

In Panel B, we focus on whether this mechanism explains the overall gender disparities in patenting. We examine whether women are differentially deterred from ultimately *completing* patent applications after initial rejections by their assigned patent examiners. Successful patent grants may involve several examiner rejections of specific claims, followed by applicant

[8]In Table A.4, we limit our sample to only mixed-gender teams and find consistent results.

amendments, before a patent is finally awarded. As is expected, receiving an initial rejection reduces the likelihood of receiving a patent, for both male and female applications. We start by examining the differential effect of a rejection on applications in which over 50% of inventors are women, as presented in Column 2. These applications are 5.9 p.p. less likely to receive patents than similarly situated patent applications with majority male inventors. Given that these teams were 3.3 p.p. more likely to drop out immediately (Panel A) – i.e., without refiling an amended application – our results suggest that 55% of the overall gender patent granting gap is explained by women's differential deterrence when an examiner makes her initial determination.

The effect of rejection on patent receipt is magnified when we look at outcomes for all-female teams and solo female inventors in Columns 3 and 4. These inventors are 10.4 p.p. and 7.5 p.p. less likely to be granted patents than all other applications and applications from solo male inventors, respectively. We observe consistent magnitudes of differences between immediate (as measured by initial amendment submissions) and downstream (measured by patent receipt) gender differences across our measures of gender application composition. In all cases, the differential deterrence of women after initial rejection is larger for final patent completion than for completing the immediate next step of the process, and the differential response to initial rejections accounts for 55% to 70% of the gender gap in granted patents.

We find that the effect of the *first* rejection on gender differentials in patent continuation is greater than the the effects of subsequent rejections. That is, after receiving one rejection and subsequently continuing the patent process by responding with an amendment, female teams are less likely to drop out following another rejection in the process. These results are shown in Table A.5.

This is interesting from a policy perspective; it may be that one way to address the gap in responsiveness to rejection is by providing resources to applicants about how to respond to an initial rejection, or information regarding the frequency of initial rejections. If female applicants are less likely to respond after receiving a rejection, it may be because they interpret an initial rejection as a signal of their quality. In this context, knowing that more than 80% of applications receive an initial rejection could help these applicants set their expectations and better interpret feedback from examiners. Bol, De Vaan, and Rijt (2018) find a similar (gender aspecific) effect among early-career grant recipients and identify the provision of information as a potential solution.

2.4.3 Drivers of Differential Responses to Rejection

We turn now to sources of heterogeneity in women's differential responses to rejection in the patent prosecution process relative to men. First, we evaluate whether using a lawyer to file a patent application has an effect on gender differentials in responses to rejection. 95.4% of applications use lawyers to manage their patent applications. When an application is prosecuted by a law firm, this means that the inventor(s) pay a lawyer to manage the application process. The lawyer typically helps draft the patent application, handles all communication with the patent examiner at the USPTO, and solicits input and information from the inventor as needed (USPTO 2020b). The assistance of a lawyer is resource intensive; patent attorneys' hourly billing rates range from $300 to over $700 depending on attorney experience and geographic location (Quinn 2015). Even just filing a patent application can be costly, with attorney fees alone amounting to between $5,000 to $16,000 – before getting to the stage of amendment and reapplication (Fechner 2019).

Patent attorneys are particularly specialized professionals who are required to be registered "patent practitioners" to represent patent applicants before the USPTO. Registered patent practitioners are "individuals who have passed the USPTO's registration exam and met the qualifications to represent patent applicants before the USPTO" (USPTO 2020b). The process of obtaining a patent is a complex legal process that requires properly drafting and filing an initial application and then potentially negotiating claims with government agents within the USPTO in order to finally receive legal protection for an invention.

Table 2.5 presents our results. We begin by evaluating the net effect of using a lawyer on patent outcomes; that is, are applications prosecuted by lawyers differentially likely to respond to rejections, and does using a lawyer affect the gender gap in persistence? Columns 1-2 display our aggregate results using differing definitions of female applications. When comparing applications using a lawyer to those that do not, on average, applications prosecuted by lawyers are more likely to respond after receiving initial rejections (Initial Rejection×Lawyer). Female applicants who use lawyers are still less than male applicants likely to proceed with their applications following the receipt of a rejection (Female×Initial Rejection), but the added benefit of using a lawyer is greater for female applicants than for male applicants. That is, using an attorney helps to close the gender gap in responsiveness to rejection.

We next explore the variation in these effects by law firm quality. Not all attorneys are equally effective, and we leverage differences in the ranking of law firms to evaluate if these effects are different based on whether an application is managed by an attorney from a more or less prestigious firm. We use Vault's Law 100 rankings in addition to Vault's list of the 30 best law firms for intellectual property to identify which applications are using attorneys from top firms. These firms (114 out of a total of over 63,000 firms in our sample) represent a disproportionate number of applications; they represent 11.4% of all applications and 11.2% of all granted patents.

When we break down the aggregate effects by looking separately at the top 100 patent attorneys and the top 50 attorneys, we find that more highly ranked lawyers are not more effective than less prestigious attorneys at pushing applications through or at closing the gap between male and female patent applicants. As indicated in Columns 3-6, attorneys from the top 100 and above firms have a positive effect on the likelihood of patent continuation (Initial Rejection×Lawyer) when applications using attorneys in this group are compared to applications that do not use an attorney. This effect is positive and significant, with attorneys boosting the likelihood of response to rejection by roughly 24 percentage points. However, this is the same magnitude of the effect we find when we examine the net benefit of having an attorney, regardless of ranking, and there is no significant difference between the coefficients on top 100 and top 50 law firms and lawyers overall.

When we focus on the top 100 law firms, the benefit of using an attorney is greater for applications that come from teams composed of 50% or more women. However, this effect disappears when we focus on applications from all-female applicant teams and solo female inventors (Column 4). In columns 5-6 we see that the effect of using a top 50 attorney shrinks the gap between male and female applicants for both applications whose authorship is at least 50% female and 100% female, the latter at the 10% significance level. Still, the differential effect for women in all of these specifications (the summation of four coefficients, Female×Initial Rejection×Lawyer + Female×Initial Rejection + Female×Lawyer + Female, displayed at the bottom of the table) remains negative.

Next, we consider whether applying for a patent with the support of an employer (rather than independently) affects women's greater tendency to exit the patent process after an initial rejection. When an inventor or team of inventors apply for a patent, they are presumed to be the owner of the patent. Applicants can assign their idea to an organization or entity, generally the company for which they worked when they produced the idea. Firm assignment indicates that a firm will assume the ownership right if granted, and is thus an indication of firm backing of an inventor team. Firm support of an application is often accompanied by financial support as well, with the firm paying for patent counsel. Additionally, many firms have internal processes to manage patent applications, and the individual inventor is less likely to be driving decisions regarding whether or not to continue in the application process. These analyses compare outcomes for applications that are associated with firms with applications from unaffiliated inventors, who are likely not receiving any financial or other support from organizations. Examining applications that are affiliated with firms compared with those that are not allows us to examine how access to resources and the handling of patent applications by professional experts affects gender gaps in innovation outcomes

Table 2.6 explores the effect of firm assignment on patent outcomes for majority and all-female applications. Columns 1-2 demonstrates that in the aggregate, working within firms does offset women's differential deterrence after an initial rejection. First, we observe that firm-backed patent applications are considerably more likely to proceed beyond an initial rejection (as indicated by Initial Rejection×Firm), and that this effect is greater for women (as seen in the coefficients on Female×Initial Rejection× Firm). That is, female applicants benefit *more* from firm affiliation than male applicants.

In Columns 3-6, we examine differential effects by firms' patenting track records. We sort firms based on the number of patents they have received to understand the effects of firm experience on outcomes. Given that the patent application process requires an understanding of the norms and techniques to successfully navigate both the initial application and subsequent communication with an examiner, experienced firms may differ substantially from novice firms in the support they provide to inventors. A small number of firms account for the majority of patent applications and granted patents in our data. In all, we have about 77,000 unique firms; the 770 companies that make up the top 1% of patent recipients account for 69.2% of all applications and 69% of all granted patents. The median firm receives just one patent, with the 75th percentile firm having just two patents.

Columns 3-4 evaluate outcomes for applications that are assigned to firms from the 1%-75% in terms of the number of patents received, relative to applications that are unaffiliated with firms. Columns 5-6 do the same for applications affiliated with firms in the top quartile. We find that there is a consistently positive effect of employer affiliation on the responses to rejection, regardless of firm size. Both male and female inventors are more likely to submit initial amendments when their application is assigned to a firm (Initial Rejection×Firm). This effect is of similar magnitudes for both sets of firms (11.3-14.5 percentage points).

For both types of firms, there is an additional positive effect for teams that are majority female, and this differential effect is greater for all-female teams. Relative to women and majority-female teams applying independently, female innovators from firms are significantly more likely to continue the patent application process after a receiving a rejection. Moreover, examination of the coefficients on Female×Initial Rejection×Firm suggests that the added benefit for women applying as part of a firm exceeds that of men applying in firms. That is, being affiliated with a firm improves outcomes for female innovators more than for male innovators.

However, depending on the firm, majority-women application teams backed by firms are still less likely to either appeal or amend their applications after a rejection compared to majority-men teams backed by firms. This result requires the summation of four coefficients, Female×Initial Rejection×Firm + Female×Initial Rejection + Female×Firm + Female, and is displayed below each column. Interestingly, applying for a patent while working at one of the the bottom 75% of firms appears to eliminate the gap between women and men-led teams, while being at a firm that has received more patents has a relatively smaller effect and does not erase the gender differential completely.

These results suggest that firms can play a role in ameliorating the gender gap in innovation, potentially by providing information and resources to inventors. It may also be that the benefit of firm affiliation is actually that the inventor ceases to be the decision maker regarding whether or not to proceed with an application, as that judgment now lies with the employer. However, the fact that there still remains a gap in persistence between men and women at some firms indicates that firm support is not a panacea, and further work is required to understand the dynamics at play in this setting.

2.5 Discussion and Conclusion

In this study, we seek to identify the extent to which gender differences in deterrence after early setbacks contribute to the underrepresentation of women in innovation. To do so, we study how male and female inventors respond to rejection in the patent application process. We use an instrumental variables strategy that takes advantage of the quasi-random assignment of applications to patent examiners. We identify that gender differentials in responses to rejection contribute significantly to differential outcomes in patenting for women. Female inventors who receive rejections early in the application process are less likely to submit amendments in response to examiner feedback, and this results in the abandonment of their applications. Female inventors' differential responses to rejection account for more than 50% of the gender gap in patent receipt, conditional on application. The gender gap in responsiveness to rejection widens as the presence of women on inventor teams increases, indicating that the *intensive* margin of female representation affects how patenting teams respond to rejection.

To examine how the gender gap in patent receipt can be mediated, we explore the effects of access to information on women's differential persistence. We do this by evaluating how using a lawyer to prosecute a patent application and having the support of a firm affects the gender gap in response to rejection. First, we find that using a lawyer increases the likelihood of submitting an amendment following a rejection for applications from both male and female inventors. We also see that using a lawyer shrinks but does not eliminate the gender differential in response to rejection, which suggests that having the guidance of a lawyer and the information that comes along with patent counsel can help address the problem we identify.

When we consider the effect of having patent rights assigned to a firm on outcomes, we see that firm assignment increases responsiveness to rejection for all applications. This increase is greater for applications that come from female inventors and teams of 50% or more women. That is, having the support of a firm, which generally encompasses financing the application and any associated costs and managing the application process, helps shrink the gap between how male and female inventors respond following an initial rejection. These findings provide suggestive evidence that the provision of resources and potentially distancing inventors from the decision of whether or not to continue an application can play a key role in increasing women's success in the patent application process.

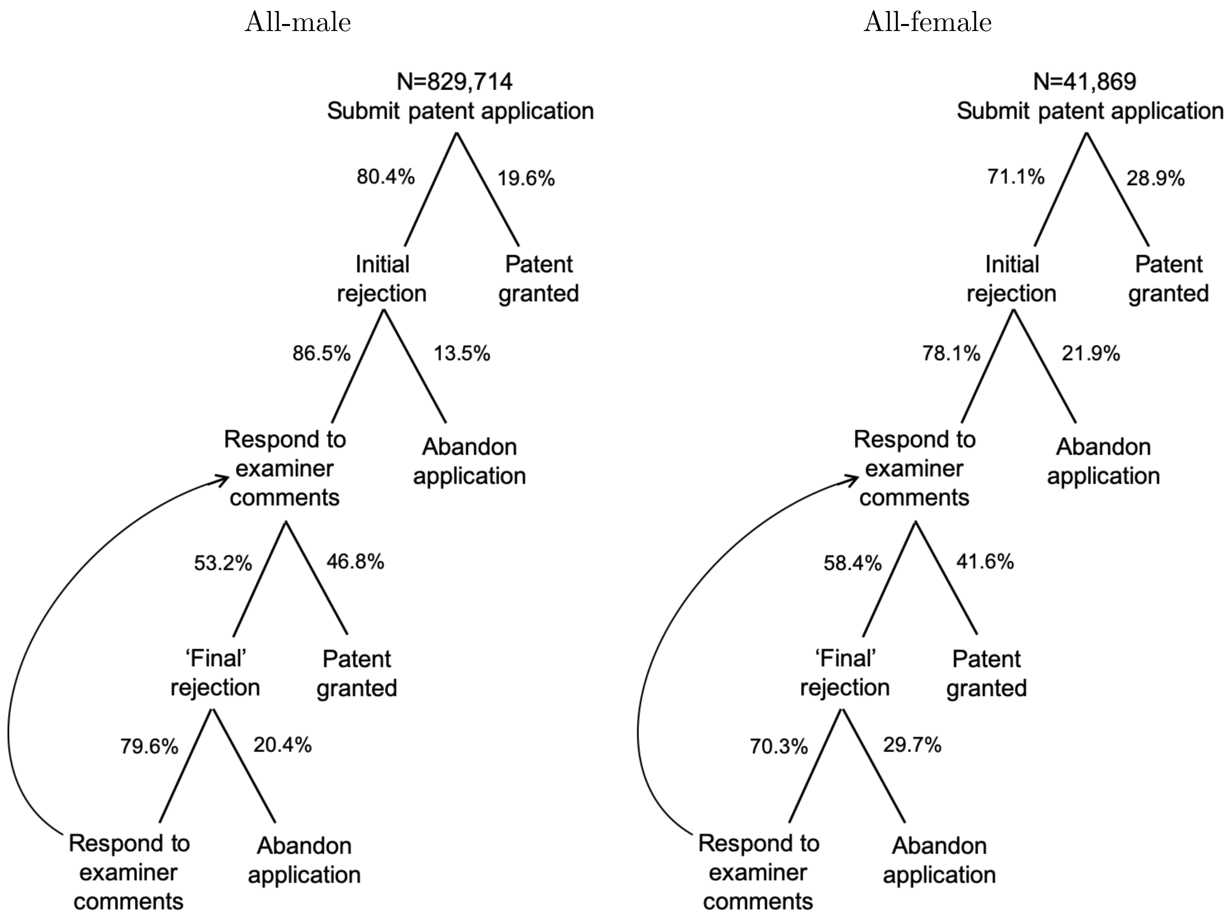
Our findings have implications for work on innovation and gender gaps more broadly. First, we identify a potential driver of the underrepresentation of women in entrepreneurship. Women's differential rates of patent receipt can contribute to the gender gap in access to financial resources as well as a disparity in opportunities for women to commercialize their inventions. We also identify a novel dimension of differences in responsiveness to rejection by gender; women are not only less likely to follow up on rejections than men, but they are also less likely to follow up on harsher rejections than more lenient rejections.

One limitation of this study is that we do not have access to communications between examiners and patent applicants. We are also not able to directly observe the mechanisms that drive the differential outcomes we observe. However, we leverage the available data to try to, at a minimum, rule out potential explanations.

This work makes an important contribution by opening the black box of gender disparities in innovation and identifying a key reason for this gap. We focus on how differential responses to rejection by gender, a previously unexplored mechanism, contributes to the underrepresentation of women in innovation. Our findings highlight the potential implications of women's decreased likelihood to convert applications into patents for both the direction and landscape of innovation as well as individuals' career trajectories and firm outcomes. We suggest potential interventions that policymakers can consider to begin addressing the gender gap in patenting, so that future innovations can better serve the needs of a diverse and varied world.

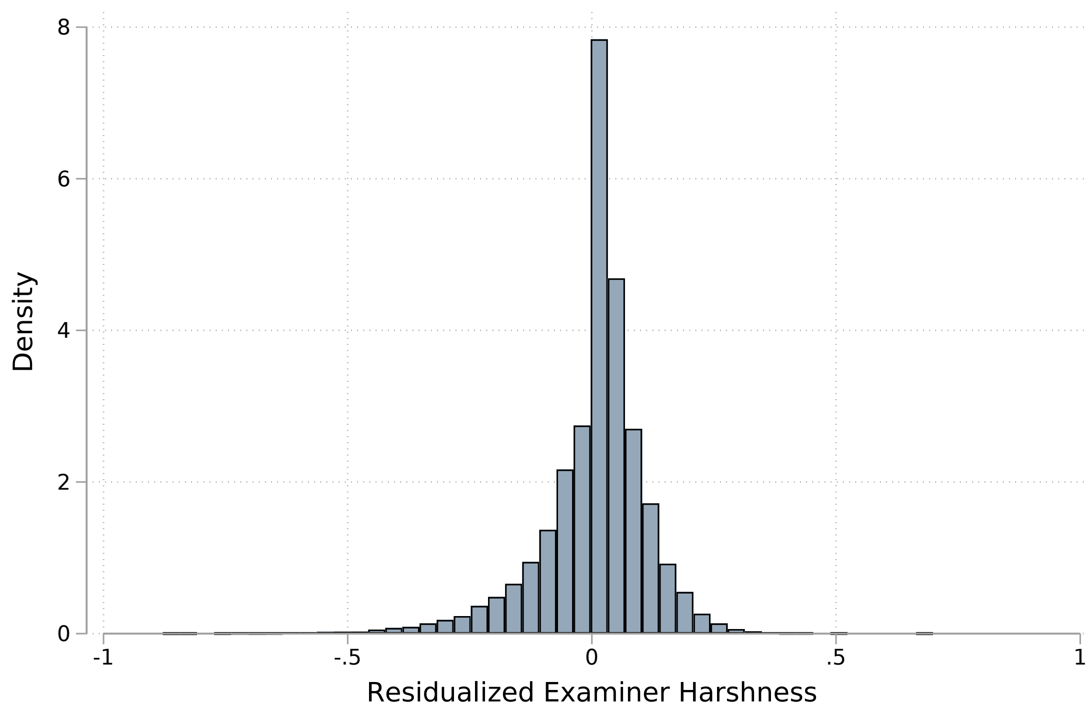
Figures

Figure 2.1: Evaluative Trajectory of Patent Applications



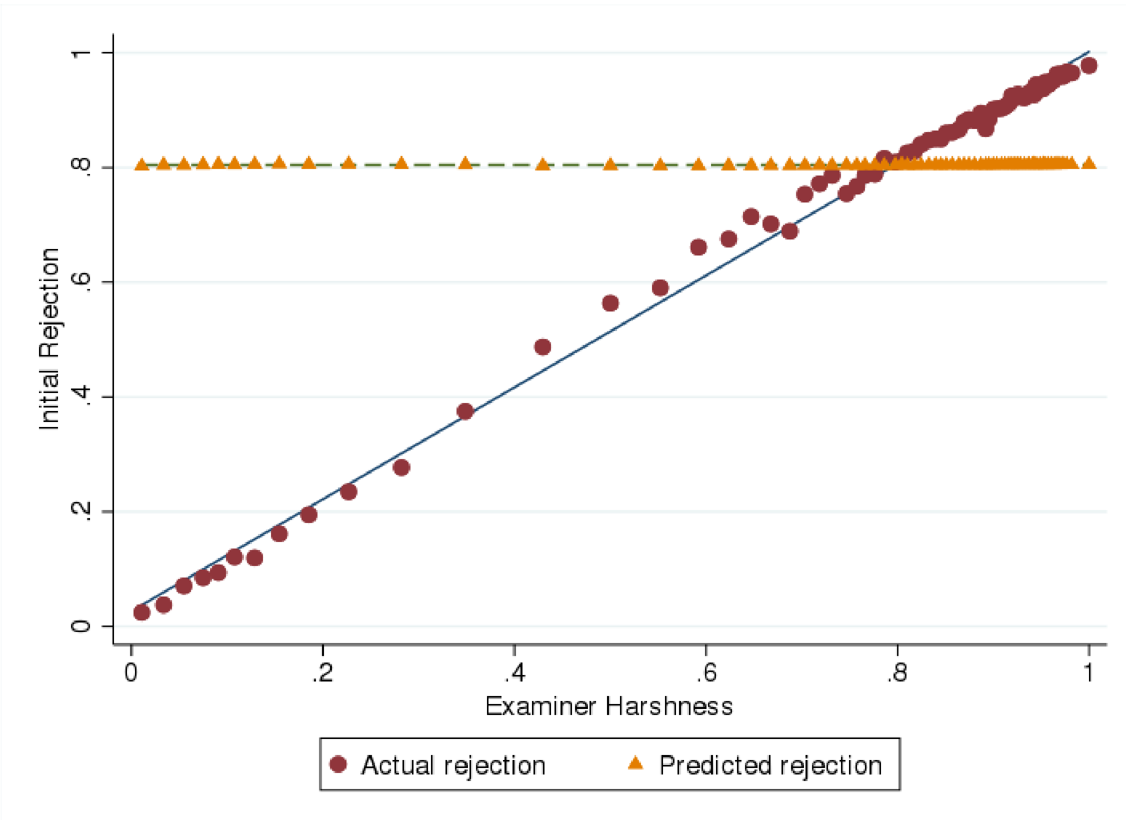
This figure shows the raw proportions of applications that progress through each stage of the patent process for applications from all-male and all-female inventors or teams. Applications from single-gender teams or solo individuals account for almost 90% of all applications in our sample.

Figure 2.2: Distribution of Initial Rejection by Examiner Harshness



This figure shows the distribution of patent initial rejection rates, residualizing by the full set of art-unit-by-application-year fixed effects.

Figure 2.3: Probability of Initial Rejection by Examiner Harshness



This figure relates examiner harshness to two variables: the actual initial rejection rate, shown in red, and the predicted rejection rate in yellow. By construction, the initial rejection rate is perfectly correlated with examiner harshness. Predicted rejection is based on observables that proxy for quality (the number of inventors on an application, the proportion of female inventors, whether the application is assigned to an employer and who that employer is). Together, these variables explain 40% of the likelihood of initial rejection (R^2=.40).

Tables

Table 2.1: Summary Statistics

	Mean	Std Dev	Min	Max
Applications (N=971,547)				
All-US Inventors	1	0	1	1
Patent Granted	0.70	0.46	0	1
Employer Assignment	0.63	0.48	0	1
Using Attorney	0.95	0.21	0	1
Number of Team Members	2.04	1.37	1	10
Solo Inventors	0.47	0.50	0	1
Individual female inventor	0.038	0.19	0	1
Proportion of Female Team Members	0.083	0.23	0	1
>=1 woman on team	0.15	0.35	0	1
>=50% women on team	0.089	0.28	0	1
All-female team	0.043	0.20	0	1
Number of Initial Rejections	1.15	0.92	0	12
Number of Initial Appeals	1.10	1.13	0	19
Number of Final Rejections	0.52	0.78	0	12
Number of Final Appeals	0.56	1.03	0	23
Proportion of Applications that receive Initial Rejections	0.80	0.40	0	1
Proportion of Applications that submit Initial Appeals	0.69	0.46	0	1
Proportion of Applications that receive Final Rejections	0.38	0.49	0	1
Proportion of Applications that submit Final Appeals	0.32	0.47	0	1

Table 2.2: First-Stage Results

	Application receives initial rejection		
	(1)	(2)	(3)
IV: Patent examiner harshness	0.694***	0.709***	0.695***
	(175.08)	(192.17)	(175.09)
Proportion of female inventors		0.00209	0.00133
		(1.42)	(0.89)
Using attorney		-0.0119***	-0.0112***
		(-7.75)	(-6.99)
Affiliated with employer		-0.00535***	-0.00504***
		(-7.03)	(-6.43)
Art Unit x Year FE	X	X	X
Subclass x Year FE			X
R^2	0.475	0.442	0.475
F-stat	30651.9	9265.1	7680.6
Observations	945373	971547	945373
# of Clusters	36851	36851	36851

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table reports the results of two different versions of the first-stage equation of our IV (2SLS) analysis. We use the initial rejection rate (for applications in the same art unit and year) of the assigned patent examiner to predict whether the focal application will receive an initial rejection. Column 1 includes only art unit-year fixed effects, Column 2 adds characteristics of the application and applicants that may proxy for quality, and Column 3 adds subclass-year fixed effects. We use the Kleibergen-Paap *rk* Wald F statistic and identify that examiner harshness is a good instrument for rejection.

Table 2.3: Motivating Evidence - Effect of Gender on Patent Application Outcomes (OLS)

	(1)	(2)	(3)	(4)
	Panel A: Effect of Gender on Initial Rejection			
Female	0.007***	0.005***	0.007***	0.009***
	(0.002)	(0.001)	(0.002)	(0.002)
Dependent Var. Mean	0.80	0.80	0.80	0.77
	Panel B: Effect of Gender on Patent Granted			
Female	-0.055***	-0.044***	-0.072***	-0.053***
	(0.002)	(0.002)	(0.002)	(0.003)
Art Unit x Year FE	X	X	X	X
Dependent Var. Mean	0.70	0.70	0.70	0.69
Observations	971547	971547	971547	461147
# of Clusters	36851	36851	36851	36727
Female Definition	Proportion	Half Female	All Female	Solo

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable mean shows the mean value of initial rejection and patent granted in the sample, respectively. All regressions include art unit-year fixed effects and are clustered at the examiner-year level.

Table 2.4: Effect of Initial Rejection on Patent Application Continuation (IV)

	(1)	(2)	(3)	(4)
Panel A: Effect of Initial Rejection on Initial Amendment				
Female X Initial Rejection	-0.039***	-0.033***	-0.073***	-0.045***
	(0.003)	(0.003)	(0.004)	(0.004)
Initial Rejection	0.849***	0.848***	0.849***	0.811***
	(0.005)	(0.005)	(0.005)	(0.007)
Panel B: Effect of Initial Rejection on Patent Granted				
Female X Initial Rejection	-0.074***	-0.059***	-0.104***	-0.075***
	(0.005)	(0.004)	(0.006)	(0.006)
Initial Rejection	-0.663***	-0.664***	-0.665***	-0.686***
	(0.011)	(0.011)	(0.011)	(0.013)
Art Unit x Year FE	X	X	X	X
Observations	971547	971547	971547	461147
# of Clusters	36851	36851	36851	36727
Female Definition	Proportion	Half Female	All Female	Solo

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Definitions of the Female variable are denoted below each column and are described in the text. We instrument for initial rejection using examiners' leave-out mean initial rejection rate for all other applications within art unit-year. All regressions include art unit-year fixed effects and are clustered at the examiner-year level.

Table 2.5: Effect of Lawyer on Initial Amendment Submission (IV)

	Attorney		Top 100 Attorney		Top 50 Attorney	
	(1)	(2)	(3)	(4)	(5)	(6)
Female × Initial Rejection × Lawyer	0.051***	0.045***	0.053***	0.027	0.050***	0.040*
	(0.013)	(0.016)	(0.015)	(0.020)	(0.017)	(0.023)
Initial Rejection × Lawyer	0.243***	0.245***	0.244***	0.246***	0.238***	0.240***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)
Female × Initial Rejection	-0.074***	-0.102***	-0.070***	-0.093***	-0.071***	-0.095***
	(0.013)	(0.016)	(0.013)	(0.016)	(0.013)	(0.016)
Female × Lawyer	-0.013	-0.018*	-0.014	-0.004	-0.011	-0.005
	(0.008)	(0.010)	(0.009)	(0.011)	(0.010)	(0.013)
Observations	971547	971547	156740	156740	111680	111680
# of Clusters	36851	36851	33493	33493	30296	30296
Net Female Effect	-0.02	-0.05	-0.02	-0.05	-0.02	-0.04
Female Definition	Half Female	All Female	Half Female	All Female	Half Female	All Female

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Specifications 2 and 3 group attorneys based on the ranking of the ranking of the law firm they are a part of, where Top 100 represents the top 100 overall and top IP firms and Top 50 represents the top 50 overall firms and top IP firms within that set. Definitions of the Female variable are denoted below each column and are described in the text. We instrument for initial rejection using examiners' leave-out mean initial rejection rate for all other applications within art unit-year. All regressions include art unit-year fixed effects and are clustered at the examiner-year level.

Table 2.6: Effect of Firm Affiliation on Initial Amendment Submission (IV)

	Employer Assigned		Emp 1-75		Emp 75+	
	(1)	(2)	(3)	(4)	(5)	(6)
Female × Initial Rejection × Firm	0.038***	0.051***	0.067***	0.097***	0.037***	0.055***
	(0.005)	(0.007)	(0.011)	(0.014)	(0.005)	(0.007)
Initial Rejection × Firm	0.120***	0.121***	0.113***	0.114***	0.144***	0.145***
	(0.002)	(0.002)	(0.004)	(0.004)	(0.002)	(0.002)
Female × Initial Rejection	-0.043***	-0.070***	-0.044***	-0.068***	-0.043***	-0.072***
	(0.004)	(0.006)	(0.005)	(0.006)	(0.004)	(0.006)
Female × Firm	-0.007**	-0.010***	-0.015***	-0.016***	-0.005*	-0.009***
	(0.003)	(0.003)	(0.006)	(0.006)	(0.003)	(0.003)
Observations	971547	971547	413757	413757	899387	899387
# of Clusters	36851	36851	36245	36245	36851	36851
Net Female Effect	-0.01	-0.02	0.02	0.02	-0.01	-0.02
Female Definition	Half Female	All Female	Half Female	All Female	Half Female	All Female

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Specifications 2 and 3 group employers based on the number of patents the firm has received, where 1-75 represents the bottom 75% of firms by number of patents granted and 75+ represents the top 25% of patent-receiving firms. Definitions of the Female variable are denoted below each column and are described in the text. We instrument for initial rejection using examiners' leave-out mean initial rejection rate for all other applications within art unit-year. All regressions include art unit-year fixed effects and are clustered at the examiner-year level.

Chapter 3

Exploration or Exploitation: The Effects of Past Success on the Direction of Innovation

3.1 Introduction

Exploration in innovation is essential to discover and broach new frontiers. The discovery of new ideas and processes is at the core of innovation, which we can think of as the production of knowledge through experimentation (Arrow 1969). However, there is an inherent tension for inventors between exploration and exploitation (March 1991); that is, the decision to pursue new ideas outside of one's area of expertise as opposed to continuing to focus on a field in which one already has experience.

Understanding the determinants of exploration in innovation is important for several reasons. The decision of whether to invest in exploration or exploitation can have meaningful implications on both firm and individual-level outcomes. Innovative firms that focus on exploration see higher short-term market performance than firms with exploitation strategies (Fang and Levinthal 2009; Fitzgerald et al. 2020). Individuals who are outsiders in fields are more likely to bring novel insights and solutions to problems and thus see greater rewards for participation in new areas than established participants (Jeppesen and Lakhani 2010). Researchers with diversified knowledge are also more likely to be able to combine different inputs to create meaningful innovations (Nagle and Teodoridis 2020). Still, it is not clear that exploration is always necessarily the best approach. While integrating novel inputs can be particularly beneficial in new technology areas, relying on internal firm knowledge is more helpful when innovating in established technologies (Chatterji and Fabrizio 2014; Eggers and Kaul 2018). Innovations using new technological components and new combinations of these components have less impact on average, but can on occasion lead to technological breakthroughs (Fleming 2001). Thus, it is important to understand what organizational or individual features motivate exploration in innovation.

Prior research has attempted to understand the drivers of exploration in innovation by examining the effects of incentives (Manso 2011; Ederer and Manso 2013) and organizational environment (Manso 2017) on individuals' innovative decisions. At the individual level, inventors' personal histories play a role in their decisions whether to participate in innovative fields (Bell et al. 2019), and workers' innovative outputs are affected by their exposure to recessions (Babina, Bernstein, and Mezzanotti 2020; Bernstein, Mcquade, and Townsend 2021). This work highlights the potential for individuals' past experiences to affect innovative trajectories. Personal background affects not only an individual's decision of whether or not to innovate, but also the *content* of their innovations. For example, female inventors are more likely to produce inventions that address the health concerns of women (Koning, Samila, and Ferguson 2019; Einiö, Feng, and Jaravel 2019). Similarly, minority status can affect innovators' propensities to develop novel insights (Jeppesen and Lakhani 2010; Hofstra et al. 2020).

In this work, I examine how gender and prior innovative experience affect inventors' propensities to explore in innovation contexts. I leverage a novel dataset that tracks individual inventors to study this question in the context of applications to the U.S. Patent and Trademark Office. Using this data, I am able to analyze inventors' comprehensive application trajectories, including both successfully patented and unsuccessful innovations. I start by examining the effect of receiving a patent on the area of technological focus for the same individual's *next* patent application. I then explore variation in this effect by gender to identify how male and female inventors' paths are differentially affected by past success. In order to reliably identify causal estimates of the effect of success on the technology in an inventor's subsequent application, I use an instrumental variables strategy that takes advantage of the quasi-random assignment of patent examiners to applications. Next, I evaluate the relationship between experience and exploration across different types of technologies, and specifically explore how this differs by gender.

My results show that female inventors are significantly more likely than male inventors to explore–that is, apply for a patent on a different type of technology–following the receipt of a granted patent. This effect is sizeable and statistically significant, with female applicants between 4.94-7.50 percentage points more likely to subsequently apply for a patent in a different technology center. This effect is larger in magnitude and present across more fine-grained classifications of technology for female applicants who have previously applied for a patent as compared to new applicants, which suggests that this effect is not driven by individuals who are unfamiliar with the patent process. Congruently, I find that more experienced female patent applicants are also more likely to apply for patents on a variety of different technologies than similarly experienced male inventors.

This work contributes to the growing literature on gender disparities in innovation by identifying an important driver of differences in male and female inventors' innovative trajectories. Understanding how and why male and female innovators' paths differ is necessary in order to address the gender innovation gap (Ding, Murray, and Stuart 2006; Ding, Murray, and Stuart 2013; Delgado, Mariani, and Murray 2019). Additionally, my findings speak to a broader literature on the role of individual characteristics and experiences on exploration in innovation (Chai 2017; Nagle and Teodoridis 2020).

3.2 Data

I use a novel dataset constructed using inputs from the U.S. Patent and Trademark Office's Patent Examination Research Dataset (Public PAIR) that tracks patent applicants across time. To construct my dataset, which includes a unique individual identifier for each applicant with which I can identify all of an individual's applications, I develop an algorithm adapted from Monath, Jones, and Madhavan (2020). This disambiguation process relies on detailed data on patent applications filed in the United States, including inventors' names, filing date, application correspondence address, assignees, and inventors' locations. I identify applicant gender by leveraging USPTO's own gender attribution dataset. Data on all applications is available from 2001 through 2014, while data on granted patents is available beginning in 1977. More details about this data are available in Subramani (2021).

3.2.1 Sample Construction and Description

I identify 2.4 million unique inventors across 6.6 million applications throughout the full sample. Due to the fact that data on ungranted applications is available beginning in 2001, I am only able to observe the complete set of applications for inventors after this time. Because I am interested in evaluating application behavior and not just granted patents, I drop all inventors who only appear in the data prior to 2001. These are inventors who were active before 2001 and for whom I am unable to observe any ungranted applications. To identify inventors whose full histories are captured in the data, I examine a sample of applicants who never appear in the data prior to 2001. To conservatively identify first-time applicants, I start by examining the gap in time between an individuals' first and second patent applicants. The average lag between the first and second application in the data is almost a year (359 days), and the 75th percentile of lag time is over two years (784 days). In order to reliably posit that an individual is a first-time patent applicant, I limit my sample to individuals who appear in the patent data only after February 24, 2003, which represents 784 days after January 1, 2001. That is, these are individuals who never appear in the available application data prior to this date and have not yet received a patent.

Table 3.1 summarizes both the full dataset of granted applications from 1977-2014 and ungranted applications from 2001-2014 as well as this limited sample of ungranted and granted applications from 2001-2014 at both the application and applicant levels. As is evident from this table, female inventors are underrepresented throughout the time period and in both samples, with the average proportion of female inventors on applications lower than 14%. Female inventors make up roughly 15% of inventors in both samples. Table 3.2 shows that on average, female patent applicants apply for fewer patents and are less likely to receive a patent conditional on application. Female patent applicants also apply to fewer unique technology centers and art units.

I measure technological classification in three ways, each of increasing specificity. I examine application behavior in different technology centers, art unit groups, and art units. First, I look at whether the second application is in the same patent technology center as the first application. Technology centers are broad classifications of patent applications (ex: "Biotechnology and Organic Fields"), and there are only nine such categorizations. Technology centers are composed of art unit groups, or clusters of art units that deal with related technology. There are 82 unique art unit groups (ex: "Fermentation, Microbiology, Isolated and Recombinant Proteins/Enzymes"), and I examine whether the second application is in the same art unit group as the first application. Finally, I examine the effect of success on the likelihood of subsequently submitting an application to the same art unit. Art units are the unit at which patent examiners are organized and group together patent applications that deal with the same technologies.

3.3 Empirical Strategy

In my initial analysis, I am interested in the effect of the outcome of one application on an individual's next patent application. I begin by focusing on the first two applications from individuals in my sample. Given that my goal is to isolate the response to the outcome of the first application, I limit my sample to individuals whose next application was only filed after a final determination was made on the first application that appears in the sample. These are individuals who submitted applications only after they knew the final outcome of their previous application, so I can reasonably infer that their decision making was informed by their experience. Additionally, my analysis is estimated only for individuals who have submitted more than one patent application. I start by estimating an OLS regression of the effect of success in the first patent application on the technological classification of the second application.

I evaluate the following:

$$\begin{aligned} Second\ App\ Same\ Tech_{(a+1)i} &= \beta_1 Prior\ Patent\ Issued_{ai} + \beta_2 Female_i \\ &+ \beta_3[Female\ \times\ Patent\ Issued]_{ai} + \mu_{aut} + \epsilon_{aut} \end{aligned} \tag{3.1}$$

where a indexes a patent application, i indexes a unique inventor, and ut the patent art unit*application year. $Patent\ Issued$ is a dummy for whether the first application was issued, and $Female$ is an indicator if the focal inventor is a woman. $SecondAppSameTech$ is a dummy variable equal to 1 if the second application is in the same technology classification as the first application, and 0 if not. However, it may be that individuals who receive rejections are different from those who do not in ways that also affect their likelihood of subsequently applying for patents in the same or different technological areas. To address this concern, I employ an instrumental variables strategy that leverages variation in examiners' propensities to grant patents.

I instrument for an application's grant using the leave-out mean of the assigned examiner's patent grant rate in the given art unit and year. This allows me to isolate the effect of an application's allowance on its inventors' subsequent application behavior. I refer to the patent examiner's grant rate as examiner leniency. This is the proportion of all other applications that a focal examiner sees in an art unit-filing year for which they grant a patent. A higher value of leniency indicates that an examiner is more likely to grant an application. This measure is calculated based on all *other* applications an examiner reviews, and as such, avoids any bias from the outcome of a given application.

$$Leniency_{ae} = \left(\frac{1}{n_e}\right)\left(\sum_{k \neq a}^{n_e} ER_k\right)$$

In this expression, e indicates the examiner assigned to an application a, n_e is the total number of applications seen by examiner e in art unit-year, k indexes the applications seen by examiner e, and P_k, a granted patent, is equal to one if the applicant at some point receives a patent from examiner e for patent application k.

For leniency to effectively instrument for patent grant, two things should be true. First, leniency should be related to the likelihood of a patent being granted– it should be relevant. This is directly observable; in a simple regression of leniency on patent issuance, the coefficient on leniency is 0.92. Second, an examiner's leniency should only affect the outcome of an application through whether a patent is granted at all; that is, there should be no relationship between an application's quality or likelihood of being granted and the leniency of the assigned examiner. Prior research (Frakes and Wasserman 2014; Sampat and Williams 2019) indicates that conditional on year and art unit, assignment of applications to patent examiners is plausibly random.

Using this instrument, my IV regression is as follows:

$$\begin{aligned} Second\ App\ Same\ Tech_{(a+1)i} &= \beta_1 Prior\ \widehat{Patent}\ Issued_{ai} + \beta_2 Female_i \\ &+ \beta_3 [Female \times \widehat{Patent}\ Issued]_{ai} + \mu_{aut} + \epsilon_{aut} \end{aligned} \tag{3.2}$$

where I instrument for $Prior\ \widehat{Patent}\ Issued_{ai}$ and $[Female \times \widehat{Patent}\ Issued]_{ai}$ using $Leniency_{ae}$ and $[Female_a \times Leniency_{ae}]$, respectively.

I next evaluate the relationship between experience and the number of unique technology areas in which an inventor applies for patents. That is, I am interested in understanding if inventors who are more active are also those who are more prone to exploration, and whether this relationship varies by gender of inventor. Mechanically, experience will be related to an increase in the number of unique technology classifications in which an individual has submitted applications, as a person who has submitted more applications has more opportunities to apply in different art units and thus is more likely to have applications in a greater number of unique technologies. However, my interest in this analysis is specifically on the interaction between gender and experience, as this isolates the relationship between experience and the number of unique technology areas for women as compared to men. I estimate the following:

$$\begin{aligned} Unique\,Tech\,Classifications &= \beta_1 Experience_i + \beta_2 Female_i \\ &+ \beta_3 [Female \times Experience]_i + \mu_i + \epsilon_i \end{aligned} \tag{3.3}$$

where i indexes a unique inventor. I measure experience in two ways; first, by counting the number of applications an individual has submitted, and then by the number of patents an individual has received. *Female* is an indicator if the focal inventor is a woman. The dependent variable of interest is the number of unique technological classifications (technology centers, art unit groups, or art units) that the inventor has submitted applications in.

Results

3.3.1 Effect of Success on Subsequent Applications

I start by estimating an OLS regression of the effect of success in the first patent application on the technological classification of the second application.

Table 3.3 presents the results of Equation 3.1, evaluating the linear relationship between success of the first patent application that appears in my data and the technological classification of the focal inventor's subsequent application. I run this analysis on three samples: 1) the full sample of individuals who appear in the patent application data following 2001 and who only file a second application after a decision is rendered on the first, 2) further restricting 1) by limiting to individuals who have never received a patent prior to 2001, and 3) limiting the sample described in 2) only to inventors who have never applied for or received a patent prior to February 24, 2003 (the 75th percentile for gap between first and second application, as discussed earlier). I reference these as samples 1, 2, and 3 respectively.

My results are remarkably consistent across these three samples. First, I find that if their first application that appears in the data is granted, inventors are far more likely to file their second applications in the same area of technological specialization. This is the case across technology centers, art unit groups, and art units. A successful male applicant is 8.47-11.11 percentage points more likely to subsequently apply in the same technology center and 1.2-2.44 percentage points more likely to apply within the same art unit. Additionally, female inventors are in general more likely to stay within a given technology specification in their second application. However, successful female inventors are *less* likely than successful male inventors to stay within the same technology in their subsequent applications. That is, receiving a patent increases the likelihood that female inventors subsequently apply for a patent on a different type of technology.

As discussed earlier, these estimates may not accurately capture the effect of success itself. The grant of the first application can be related to an inventor's innovative capabilities, which could in turn affect an individual's propensity to explore new technologies. Table 3.4 presents my results instrumenting for rejection using the assigned patent examiner's rejection rate for all other applications in the same art unit in the same year. This table shows results for the same three samples described above.

These results differ somewhat from my OLS results. The effect of patent issuance on continuing to innovate in the same technology remains significant and positive at the level of technology centers, but not art unit groups or art units. That is, a male inventor whose first application in the sample is successful is more likely to submit a second application in the same technology center than in a different technology center. This effect is insignificant with respect to the subsequent patent being in the same art unit group or art unit. The coefficient on female remains positive and significant for all technology classifications in samples 1 and 2 and at the level of technology centers in sample 3.

When examining the differential effect of prior success for female inventors, I find that for all three classifications of technology similarity, the coefficient on Female×Patent Issued is negative. This effect is significant at the technology center level for all samples. At the art unit group and art unit levels, the effect is more inconsistent. Art unit group is significant at a 5% level for sample 1, at a 10% level for sample 2 and insignificant for sample 3. Art unit is significant at a 10% level for sample 1 and insignificant for all other samples. Overall, these results indicate that successful female inventors are differentially likely to explore new technology areas after experiencing success in the patent process. Female inventors are more likely than successful male inventors to explore new technology centers across all samples, while the evidence is more mixed at the art unit group and art unit levels.

There are several things to note about these findings. First, the fact that the IV results differ from the OLS results suggests that some of the effects captured in my OLS regressions are the result of the relationship between the characteristics of applications that are successful and the gendered propensity to explore across technology types. For example, it may be that the quality distribution of applications from female applicants is different from the distribution of applications from male applicants. Thus, successful female applicants respond differently when it comes to making decisions about their subsequent work than successful male applicants because they evaluate their options differently than similarly situated male applicants, not because they are responding directly to the success of their initial applications. Using the instrumental variable approach removes this endogeneity, and therefore one can consider the IV estimates to be true measures of the effect of success on the content of the next patent application. The point estimates on the Patent Issued×Female interaction on subsequently applying for a patent in the same technology center are *greater* in the IV specifications. Successful female applicants are 4.94-7.5 percentage points less likely to reapply in the same technology center than successful male applicants.

Finally, the fact that female inventors are more likely to apply for a patent in a different technology area in my analyses of the full sample than in samples 2 and 3, which limit to new applicants, suggests that prior experience may in fact be driving female inventors' decisions. This effect runs counter to a story of new patent applicants being deterred from pursuing a given technology by an early failure and thus being more likely to subsequently apply for a patent on a different type of technology. Instead, it indicates that more experienced, successful applicants are driving the net effect of exploration.

3.3.2 Effect of Experience on Inventors' Technological Diversity

Next, I turn to my analyses examining the relationship between experience as measured by number of submitted applications and granted patents and an inventor's exploration across different technologies. Table 3.5 presents these results. I evaluate the full sample of inventors who appear in the data at least once after 2000 (that is, individuals for whom data on applications, not just granted patents, is available). In Panel 1, experience is measured by the number of applications an inventor has submitted, and in Panel 2, by the number of patents an inventor has received. The coefficient of interest in this context is the interaction between female and experience.

In both sets of specifications, I see that the effect of experience on the number of unique technology classifications explored is greater for female applicants than for men. That is, experienced female applicants are more likely to explore than experienced male applicants. This is true across all three measures of technology classifications. To interpret the magnitude of these coefficients, consider the following: take two inventors, one man and one woman, both of whom have submitted 20 patent applications. The female inventor's experience is predicted to increase the number of art units she has applied to by 0.848 relative to the male inventor. Given that the average female inventor applies to 2.13 art units, this difference represents a significant increase in the number of unique art units. While the magnitude of these effects are small, particularly at the technology center level, it highlights that experienced female inventors have a greater propensity to explore across different types of technologies than similar male inventors.

In order to ensure that my results are not driven by outliers, I run the same analyses with the outcome variable (number of unique technology centers) logged. These results remain consistent and are presented in Table 3.6.

3.4 Discussion and Conclusion

In this paper, I evaluate inventors' innovative trajectories using a novel data set in which I observe both successful and unsuccessful patent applications. I examine the effect of prior experience on male and female inventors' subsequent applications and find that female inventors are more likely than male inventors to explore different technology areas after success in the patent process. Interestingly, this effect is not driven by new patent applicants but rather by inventors with prior experience applying for a patent. I identify that experienced female inventors are more likely to patent across a broad set of technology areas than similarly experienced male inventors.

These findings shed light on prior experience as an important, and due to data limitations, previously difficult to evaluate, dimension on which to explore gender differences in innovative behavior. Prior work has found that women are less likely to participate in contests and innovative processes after receiving a rejection (Kuppuswamy and Mollick 2016; Brands and

Fernandez-Mateo 2017; Wasserman 2018); in this paper, I evaluate how success differentially affects the *way* in which women continue to participate.

In ongoing work, I am examining variation in individuals' decisions to explore new-to-them technology areas longitudinally (for multiple applications) as well as to examine the implications of exploration and exploitation on individuals' outcomes. I plan to leverage data on patent citations to evaluate the quality of applications and identify whether individuals appear to benefit from these decisions. Additionally, I am interested in what factors drive the decision to specialize versus explore, and am currently measuring the effect of the gender composition of the technology classification (as measured by the proportion of applicants who are female and male) on inventors' propensities to stay within the same technology groupings.

More broadly, given that gender is relevant to the types of innovations that individuals develop (Koning, Samila, and Ferguson 2019; Einiö, Feng, and Jaravel 2019), this work has implications for our understanding of the effect of differential success by gender on the landscape of innovation. That is, given that successful female inventors are more likely to pursue innovation in varied technology areas, and that female inventors are also more likely to pursue innovations that meet the needs of women, this gender-specific trend may be one explanation for differences that we observe in the products and ideas that come to market.

Tables

Table 3.1: Summary Statistics

	(1) Full Sample	(2) 2001 Onwards
Applications		
Inventors per application	2.213	2.362
	(1.648)	(1.736)
Female	0.0870	0.138
	(0.282)	(0.345)
Proportion of Female Inventors	0.0870	0.127
	(0.226)	(0.269)
Filing Year	2002.7	2007.4
	(7.693)	(3.394)
Patent Granted	0.627	0.451
	(0.484)	(0.498)
Employer Assignment	0.633	0.604
	(0.482)	(0.489)
Using Attorney	0.958	0.960
	(0.200)	(0.196)
Observations	6,662,840	2,414,784
Inventors		
Female	0.143	0.163
	(0.350)	(0.369)
Applications per inventor	5.938	3.014
	(15.58)	(6.227)
Patents per inventor	3.602	1.364
	(11.27)	(3.806)
Patent grant rate per inventor (2001-2014)	0.466	0.412
	(0.440)	(0.446)
Proportion of coauthored applications	0.785	0.797
	(0.360)	(0.377)
Mean number of coauthors	2.330	2.414
	(2.486)	(2.687)
Unique technology centers per inventor	1.652	1.158
	(1.577)	(0.749)
Unique art unit groups per inventor	2.305	1.447
	(3.029)	(1.389)
Unique art units per inventor	2.992	1.721
	(5.093)	(2.245)
Observations	2,483,108	1,892,120

Table 3.2: Application Statistics by Gender

	(1) Female	(2) Male
Applications per inventor	4.078	6.249
	(14.90)	(15.67)
Patents per inventor	2.118	3.850
	(10.11)	(11.43)
Patent grant rate per inventor (2001-2014)	0.431	0.472
	(0.445)	(0.439)
Unique technology centers per inventor	1.275	1.715
	(0.972)	(1.648)
Unique art unit groups per inventor	1.689	2.407
	(1.995)	(3.157)
Unique art units per inventor	2.127	3.137
	(3.459)	(5.304)
Observations	355,338	2,127,770

Table 3.3: Effect of Patent Issuance on Subsequent Technology (OLS)

	Same Technology Center	Same Art Unit Group	Same Art Unit
Sample 1: All Inventors			
Patent Issued	0.0847***	0.0362***	0.0120***
	(0.0111)	(0.00543)	(0.00317)
Female	0.0549***	0.0471***	0.0418***
	(0.00603)	(0.00778)	(0.00699)
Female x Patent Issued	-0.0374***	-0.0306**	-0.0294***
	(0.0103)	(0.0101)	(0.00757)
Art Unit FE	X	X	X
N	172017	172017	172017
N_clusters	13	13	13
Sample 2: New Inventors 2001-2014			
	Same Technology Center	Same Art Unit Group	Same Art Unit
Patent Issued	0.111***	0.0515***	0.0244***
	(0.00903)	(0.00496)	(0.00363)
Female	0.0465***	0.0402***	0.0381***
	(0.00656)	(0.00711)	(0.00715)
Female x Patent Issued	-0.0448***	-0.0358***	-0.0323***
	(0.00892)	(0.00826)	(0.00629)
Art Unit FE	X	X	X
N	101438	101438	101438
N_clusters	13	13	13
Sample 3: New Inventors Feb 2003-2014 (75% + lag)			
	Same Technology Center	Same Art Unit Group	Same Art Unit
Patent Issued	0.0950***	0.0469***	0.0235***
	(0.00918)	(0.00788)	(0.00604)
Female	0.0387***	0.0345***	0.0320***
	(0.00933)	(0.00958)	(0.00855)
Female x Patent Issued	-0.0369**	-0.0316**	-0.0273***
	(0.0148)	(0.0124)	(0.00819)
Art Unit FE	X	X	X
N	59487	59487	59487
N_clusters	11	11	11

Table 3.4: Effect of Patent Issuance on Subsequent Technology (IV)

	Same Technology Center	Same Art Unit Group	Same Art Unit
Sample 1: All Inventors			
Patent Issued	0.0474*	0.00126	-0.0218
	(0.0220)	(0.0267)	(0.0178)
Female	0.0790***	0.0662***	0.0480***
	(0.0168)	(0.0166)	(0.0141)
Female x Patent Issued	-0.0750***	-0.0589**	-0.0381*
	(0.0231)	(0.0249)	(0.0198)
Art Unit FE	X	X	X
N	170734	170734	170734
N_clusters	13	13	13
Sample 2: New Inventors 2001-2014			
	Same Technology Center	Same Art Unit Group	Same Art Unit
Patent Issued	0.0384	-0.0138	-0.0283
	(0.0216)	(0.0306)	(0.0208)
Female	0.0544***	0.0487***	0.0363**
	(0.0136)	(0.0149)	(0.0135)
Female x Patent Issued	-0.0577**	-0.0470*	-0.0271
	(0.0199)	(0.0240)	(0.0194)
Art Unit FE	X	X	X
N	100550	100550	100550
N_clusters	13	13	13
Sample 3: New Inventors Feb 2003-2014 (75%+ lag)			
	Same Technology Center	Same Art Unit Group	Same Art Unit
Patent Issued	0.0520**	0.00129	-0.0168
	(0.0176)	(0.0304)	(0.0237)
Female	0.0445**	0.0308*	0.0196
	(0.0141)	(0.0159)	(0.0135)
Female x Patent Issued	-0.0494*	-0.0262	-0.00705
	(0.0230)	(0.0279)	(0.0213)
Art Unit FE	X	X	X
N	58958	58958	58958
N_clusters	11	11	11

Table 3.5: Relationship Between Experience and Exploration

Panel 1: Application Experience	Same Technology Center	Same Art Unit Group	Same Art Unit
Female x Experience	0.00645***	0.0193***	0.0424***
	(0.000239)	(0.000446)	(0.000652)
N	1,583,296	1,583,296	1,583,296

Panel 2: Granted Patent Experience	Same Technology Center	Same Art Unit Group	Same Art Unit
Female x Experience	0.00820***	0.0258***	0.0617***
	(0.000282)	(0.000551)	(0.000862)
N	1,583,296	1,583,296	1,583,296

Table 3.6: Relationship Between Experience and Exploration (Log DV)

Panel 1: Application Experience	Same Technology Center	Same Art Unit Group	Same Art Unit
Female x Experience	0.00221***	0.00546***	0.00946***
	(0.0000773)	(0.000101)	(0.000113)
N	1,583,296	1,583,296	1,583,296
Panel 2: Granted Patent Experience	**Same Technology Center**	**Same Art Unit Group**	**Same Art Unit**
Female x Experience	0.00268***	0.00630***	0.0113***
	(0.0000910)	(0.000121)	(0.000139)
N	1,583,296	1,583,296	1,583,296

Bibliography

[1] Kenneth J. Arrow. "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-Market Allocation". In: *In The Analysis and Evaluation of Public Expenditures: The PBB-System, Joint Economic Committee, 91st Cong., 1st sess.* 1969.

[2] Tania Babina, Asaf Bernstein, and Filippo Mezzanotti. *Crisis Innovation.* en. SSRN Scholarly Paper ID 3567425. Rochester, NY: Social Science Research Network, Nov. 2020. DOI: 10.2139/ssrn.3567425. URL: https://papers.ssrn.com/abstract=3567425 (visited on 04/08/2021).

[3] Benjamin Balsmeier et al. "Machine learning and natural language processing on the patent corpus: Data, tools, and new measures". en. In: *Journal of Economics & Management Strategy* 27.3 (Sept. 2018), pp. 535–553. ISSN: 10586407. DOI: 10.1111/jems.12259. URL: http://doi.wiley.com/10.1111/jems.12259 (visited on 03/17/2021).

[4] Alex Bell et al. "Who Becomes an Inventor in America? The Importance of Exposure to Innovation". en. In: *The Quarterly Journal of Economics* 134.2 (May 2019), pp. 647–713. ISSN: 0033-5533. DOI: 10.1093/qje/qjy028. URL: https://academic.oup.com/qje/article/134/2/647/5218522 (visited on 02/21/2020).

[5] Shai Bernstein, Timothy Mcquade, and Richard R. Townsend. "Do Household Wealth Shocks Affect Productivity? Evidence from Innovative Workers During the Great Recession". en. In: *The Journal of Finance* 76.1 (2021), pp. 57–111. ISSN: 1540-6261. DOI: https://doi.org/10.1111/jofi.12976. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12976 (visited on 04/14/2021).

[6] James Bessen. *NBER PDP Project User Documentation: Matching Patent Data to Compustat Firms.* Tech. rep. May 2009. URL: https://users.nber.org/~jbessen/matchdoc.pdf.

[7] Francine D. Blau et al. "Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial". en. In: *American Economic Review* 100.2 (May 2010), pp. 348–352. ISSN: 0002-8282. DOI: 10.1257/aer.100.2.348. URL: https://www.aeaweb.org/articles?id=10.1257/aer.100.2.348 (visited on 04/24/2020).

[8] Thijs Bol, Mathijs De Vaan, and Arnout van de Rijt. "The Matthew effect in science funding". en. In: *Proceedings of the National Academy of Sciences* 115.19 (May

2018). Publisher: National Academy of Sciences Section: Social Sciences, pp. 4887–4890. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1719557115. URL: https://www.pnas.org/content/115/19/4887 (visited on 08/22/2020).

[9] Raina A. Brands and Isabel Fernandez-Mateo. “Leaning Out: How Negative Recruitment Experiences Shape Women’s Decisions to Compete for Executive Roles”. en. In: *Administrative Science Quarterly* 62.3 (Sept. 2017), pp. 405–442. ISSN: 0001-8392. DOI: 10.1177/0001839216682728. URL: https://doi.org/10.1177/0001839216682728 (visited on 11/13/2019).

[10] Nicolas Carayol and Lorenzo Cassi. *Who’s Who in Patents. A Bayesian approach.* en. Tech. rep. 2009-07. Publication Title: Cahiers du GREThA (2007-2019). Groupe de Recherche en Economie Théorique et Appliquée (GREThA), 2009. URL: https://ideas.repec.org/p/grt/wpegrt/2009-07.html (visited on 03/17/2021).

[11] Erin Cech et al. “Professional Role Confidence and Gendered Persistence in Engineering”. en. In: *American Sociological Review* 76.5 (Oct. 2011). Publisher: SAGE Publications Inc, pp. 641–666. ISSN: 0003-1224. DOI: 10.1177/0003122411420815. URL: https://doi.org/10.1177/0003122411420815 (visited on 10/12/2020).

[12] Stephen J. Ceci and Wendy M. Williams. “Understanding current causes of women’s underrepresentation in science”. en. In: *Proceedings of the National Academy of Sciences* 108.8 (Feb. 2011). Publisher: National Academy of Sciences Section: Social Sciences, pp. 3157–3162. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1014871108. URL: https://www.pnas.org/content/108/8/3157 (visited on 08/13/2020).

[13] Sen Chai. “Near Misses in the Breakthrough Discovery Process”. In: *Organization Science* 28.3 (June 2017). Publisher: INFORMS, pp. 411–428. ISSN: 1047-7039. DOI: 10.1287/orsc.2017.1134. URL: https://pubsonline.informs.org/doi/abs/10.1287/orsc.2017.1134 (visited on 06/08/2021).

[14] Aaron K. Chatterji and Kira R. Fabrizio. “Using users: When does external knowledge enhance corporate product innovation?” en. In: *Strategic Management Journal* 35.10 (2014). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.2168, pp. 1427–1445. ISSN: 1097-0266. DOI: 10.1002/smj.2168. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/smj.2168 (visited on 06/09/2021).

[15] Lisa D Cook. *Policies to Broaden Participation in the Innovation Process.* en. Tech. rep. Brookings Institution, Aug. 2020, p. 27.

[16] Christopher Anthony Cotropia and David L. Schwartz. “The Hidden Value of Abandoned Applications to the Patent System”. en. In: *SSRN Electronic Journal* (2019). ISSN: 1556-5068. DOI: 10.2139/ssrn.3465737. URL: https://www.ssrn.com/abstract=3465737 (visited on 03/10/2021).

[17] Mercedes Delgado, Myriam Mariani, and Fiona E. Murray. “The Role of Location on the Inventor Gender Gap: Women are Geographically Constrained”. English. In:

(2019). URL: https://research.cbs.dk/en/publications/the-role-of-location-on-the-inventor-gender-gap-women-are-geograp (visited on 11/12/2019).

[18] Waverly W. Ding, Fiona Murray, and Toby E. Stuart. "From Bench to Board: Gender Differences in University Scientists' Participation in Corporate Scientific Advisory Boards". en. In: *Academy of Management Journal* 56.5 (Oct. 2013), pp. 1443–1464. ISSN: 0001-4273, 1948-0989. DOI: 10.5465/amj.2011.0020. URL: http://journals.aom.org/doi/10.5465/amj.2011.0020 (visited on 08/11/2020).

[19] Waverly W. Ding, Fiona Murray, and Toby E. Stuart. "Gender Differences in Patenting in the Academic Life Sciences". en. In: *Science* 313.5787 (Aug. 2006), pp. 665–667. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1124832. URL: https://science.sciencemag.org/content/313/5787/665 (visited on 02/21/2020).

[20] Florian Ederer and Gustavo Manso. "Is Pay for Performance Detrimental to Innovation?" In: *Management Science* 59.7 (Feb. 2013). Publisher: INFORMS, pp. 1496–1513. ISSN: 0025-1909. DOI: 10.1287/mnsc.1120.1683. URL: https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1120.1683 (visited on 04/14/2021).

[21] J. P. Eggers and Aseem Kaul. "Motivation and Ability? A Behavioral Perspective on the Pursuit of Radical Invention in Multi-Technology Incumbents". In: *Academy of Management Journal* 61.1 (Feb. 2018). Publisher: Academy of Management, pp. 67–93. ISSN: 0001-4273. DOI: 10.5465/amj.2015.1123. URL: https://journals.aom.org/doi/10.5465/amj.2015.1123 (visited on 06/08/2021).

[22] Elias Einiö, Josh Feng, and Xavier Jaravel. *Social Push and the Direction of Innovation*. en. SSRN Scholarly Paper ID 3383703. Rochester, NY: Social Science Research Network, May 2019. DOI: 10.2139/ssrn.3383703. URL: https://papers.ssrn.com/abstract=3383703 (visited on 04/15/2021).

[23] Christina Fang and Daniel Levinthal. "Near-Term Liability of Exploitation: Exploration and Exploitation in Multistage Problems". en. In: *Organization Science* 20.3 (June 2009), pp. 538–551. ISSN: 1047-7039, 1526-5455. DOI: 10.1287/orsc.1080.0376. URL: http://pubsonline.informs.org/doi/abs/10.1287/orsc.1080.0376 (visited on 04/15/2021).

[24] Joan Farre-Mensa, Deepak Hegde, and Alexander Ljungqvist. "What Is a Patent Worth? Evidence from the U.S. Patent "Lottery"". en. In: *The Journal of Finance* 75.2 (2020). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12867, pp. 639–682. ISSN: 1540-6261. DOI: 10.1111/jofi.12867. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12867 (visited on 04/24/2020).

[25] Holly Fechner. *The Study of Underrepresented Classes Chasing Engineering and Science Success (SUCCESS) Act*. Detroit, Michigan, June 2019. URL: https://www.uspto.gov/sites/default/files/documents/SUCCESSAct-Fechner.pdf.

[26] Isabel Fernandez-Mateo and Marko Coh. "Coming with Baggage: Past Rejections and the Evolution of Market Relationships". In: *Organization Science* 26.5 (Aug.

2015), pp. 1381–1399. ISSN: 1047-7039. DOI: 10.1287/orsc.2015.0995. URL: https://pubsonline.informs.org/doi/10.1287/orsc.2015.0995 (visited on 11/13/2019).

[27] Tristan Fitzgerald et al. "Innovation Search Strategy and Predictable Returns". In: *Management Science* 67.2 (Sept. 2020). Publisher: INFORMS, pp. 1109–1137. ISSN: 0025-1909. DOI: 10.1287/mnsc.2019.3480. URL: https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2019.3480 (visited on 04/13/2021).

[28] Lee Fleming. "Recombinant Uncertainty in Technological Search". In: *Management Science* 47.1 (Jan. 2001). Publisher: INFORMS, pp. 117–132. ISSN: 0025-1909. DOI: 10.1287/mnsc.47.1.117.10671. URL: https://pubsonline.informs.org/doi/10.1287/mnsc.47.1.117.10671 (visited on 06/23/2021).

[29] Lee Fleming and Adam Juda. "A Network of Invention". In: *Harvard Business Review* 82 (Apr. 2004).

[30] Lee Fleming, Charles King, and Adam I. Juda. "Small Worlds and Regional Innovation". In: *Organization Science* 18.6 (Dec. 2007). Publisher: INFORMS, pp. 938–954. ISSN: 1047-7039. DOI: 10.1287/orsc.1070.0289. URL: https://pubsonline.informs.org/doi/abs/10.1287/orsc.1070.0289 (visited on 06/23/2021).

[31] Lee Fleming, Santiago Mingo, and David Chen. "Collaborative Brokerage, Generative Creativity, and Creative Success". en. In: *Administrative Science Quarterly* 52.3 (Sept. 2007). Publisher: SAGE Publications Inc, pp. 443–475. ISSN: 0001-8392. DOI: 10.2189/asqu.52.3.443. URL: https://doi.org/10.2189/asqu.52.3.443 (visited on 06/22/2021).

[32] Michael Frakes and Melissa F. Wasserman. *Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents?: Evidence from Micro-Level Application Data*. en. SSRN Scholarly Paper ID 2467262. Rochester, NY: Social Science Research Network, July 2014. URL: https://papers.ssrn.com/abstract=2467262 (visited on 11/12/2019).

[33] Michael D Frakes and Melissa F Wasserman. *Knowledge Spillovers and Learning in the Workplace: Evidence from the U.S. Patent Office*. Working Paper 24159. Series: Working Paper Series. National Bureau of Economic Research, Dec. 2017. DOI: 10.3386/w24159. URL: http://www.nber.org/papers/w24159 (visited on 08/13/2020).

[34] Patrick Gaulé. "Patents and the Success of Venture-Capital Backed Startups: Using Examiner Assignment to Estimate Causal Effects". en. In: *The Journal of Industrial Economics* 66.2 (2018), pp. 350–376. ISSN: 1467-6451. DOI: 10.1111/joie.12168. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/joie.12168 (visited on 10/05/2020).

[35] Stuart J. H. Graham, Alan C. Marco, and Richard Miller. *The USPTO Patent Examination Research Dataset: A Window on the Process of Patent Examination*. en. SSRN Scholarly Paper ID 2702637. Rochester, NY: Social Science Research Network, Nov. 2015. URL: https://papers.ssrn.com/abstract=2702637 (visited on 11/18/2019).

[36] Jorge Guzman and Aleksandra (Olenka) Kacperczyk. "Gender gap in entrepreneurship". en. In: *Research Policy* 48.7 (Sept. 2019), pp. 1666–1680. ISSN: 0048-7333. DOI: 10.1016/j.respol.2019.03.012. URL: http://www.sciencedirect.com/science/article/pii/S0048733319300757 (visited on 08/13/2020).

[37] Bas Hofstra et al. "The Diversity–Innovation Paradox in Science". en. In: *Proceedings of the National Academy of Sciences* 117.17 (Apr. 2020). ISBN: 9781915378118 Publisher: National Academy of Sciences Section: Social Sciences, pp. 9284–9291. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1915378117. URL: https://www.pnas.org/content/117/17/9284 (visited on 05/05/2021).

[38] Karin Hoisl and Myriam Mariani. "It's a Man's Job: Income and the Gender Gap in Industrial Research". In: *Management Science* 63.3 (Apr. 2016). Publisher: INFORMS, pp. 766–790. ISSN: 0025-1909. DOI: 10.1287/mnsc.2015.2357. URL: https://pubsonline.informs.org/doi/10.1287/mnsc.2015.2357 (visited on 05/08/2021).

[39] Bengt Holmstrom. "Agency costs and innovation". en. In: *Journal of Economic Behavior & Organization* 12.3 (Dec. 1989), pp. 305–327. ISSN: 0167-2681. DOI: 10.1016/0167-2681(89)90025-5. URL: http://www.sciencedirect.com/science/article/pii/0167268189900255 (visited on 08/11/2020).

[40] Kyle Jensen, Balázs Kovács, and Olav Sorenson. "Gender differences in obtaining and maintaining patent rights". en. In: *Nature Biotechnology* 36.4 (Apr. 2018), pp. 307–309. ISSN: 1546-1696. DOI: 10.1038/nbt.4120. URL: https://www.nature.com/articles/nbt.4120 (visited on 11/18/2019).

[41] Lars Bo Jeppesen and Karim R. Lakhani. "Marginality and Problem-Solving Effectiveness in Broadcast Search". In: *Organization Science* 21.5 (Feb. 2010). Publisher: INFORMS, pp. 1016–1033. ISSN: 1047-7039. DOI: 10.1287/orsc.1090.0491. URL: https://pubsonline.informs.org/doi/10.1287/orsc.1090.0491 (visited on 05/17/2021).

[42] Taehyun Jung and Olof Ejermo. "Demographic patterns and trends in patenting: Gender, age, and education of inventors". en. In: *Technological Forecasting and Social Change* 86 (July 2014), pp. 110–124. ISSN: 0040-1625. DOI: 10.1016/j.techfore.2013.08.023. URL: https://www.sciencedirect.com/science/article/pii/S0040162513001911 (visited on 05/08/2021).

[43] Riitta Katila. "Using patent data to measure innovation performance". en. In: *International Journal of Business Performance Management* 2.1/2/3 (2000). Publisher: Inderscience Enterprises Ltd, pp. 180–193. URL: https://ideas.repec.org/a/ids/ijbpma/v2y2000i1-2-3p180-193.html (visited on 03/30/2020).

[44] Patrick Kline, Neviana Petkova, et al. "Who Profits from Patents? Rent-Sharing at Innovative Firms". en. In: *The Quarterly Journal of Economics* 134.3 (Aug. 2019). Publisher: Oxford Academic, pp. 1343–1404. ISSN: 0033-5533. DOI: 10.1093/qje/qjz011. URL: https://academic.oup.com/qje/article/134/3/1343/5420483 (visited on 05/05/2020).

[45] Stephen Kline and Nathan Rosenberg. “An Overview of Innovation”. en. In: *The Positive Sum Strategy. Harnessing Technology for Economic Growth.* National Academy of Sciences, 1986, pp. 173–203. ISBN: 978-981-4273-58-9 978-981-4273-59-6. DOI: 10.1142/9789814273596_0009. URL: http://www.worldscientific.com/doi/abs/10.1142/9789814273596_0009 (visited on 08/13/2020).

[46] Rembrand Koning, Sampsa Samila, and John-Paul Ferguson. *Female Inventors and Inventions*. en. SSRN Scholarly Paper ID 3401889. Rochester, NY: Social Science Research Network, June 2019. URL: https://papers.ssrn.com/abstract=3401889 (visited on 11/18/2019).

[47] Venkat Kuppuswamy and Ethan R. Mollick. “Second Thoughts About Second Acts: Gender Differences in Serial Founding Rates”. en. In: *SSRN Electronic Journal* (2016). ISSN: 1556-5068. DOI: 10.2139/ssrn.2752689. URL: http://www.ssrn.com/abstract=2752689 (visited on 01/09/2020).

[48] Ronald Lai et al. “Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010)”. en. In: (Dec. 2015). Publisher: Harvard Dataverse type: dataset. DOI: 10.7910/DVN/5F1RRI. URL: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/5F1RRI (visited on 05/08/2021).

[49] Vincent Larivière et al. “Bibliometrics: Global gender disparities in science”. en. In: *Nature News* 504.7479 (Dec. 2013). Section: Comment, p. 211. DOI: 10.1038/504211a. URL: http://www.nature.com/news/bibliometrics-global-gender-disparities-in-science-1.14321 (visited on 03/18/2021).

[50] David S. Lee et al. “Valid t-ratio Inference for IV”. en. In: *arXiv:2010.05058 [econ, q-fin]* (Oct. 2020). arXiv: 2010.05058. URL: http://arxiv.org/abs/2010.05058 (visited on 10/20/2020).

[51] Mark A. Lemley and Bhaven N. Sampat. *Is the Patent Office a Rubber Stamp?* en. SSRN Scholarly Paper ID 999098. Rochester, NY: Social Science Research Network, Oct. 2008. URL: https://papers.ssrn.com/abstract=999098 (visited on 08/13/2020).

[52] Guan-Cheng Li et al. “Disambiguation and co-authorship networks of the U.S. patent inventor database (1975–2010)”. en. In: *Research Policy* 43.6 (July 2014), pp. 941–955. ISSN: 0048-7333. DOI: 10.1016/j.respol.2014.01.012. URL: https://www.sciencedirect.com/science/article/pii/S0048733314000225 (visited on 03/17/2021).

[53] Gustavo Manso. “Creating Incentives for Innovation”. en. In: *California Management Review* 60.1 (Nov. 2017). Publisher: SAGE Publications Inc, pp. 18–32. ISSN: 0008-1256. DOI: 10.1177/0008125617725287. URL: https://doi.org/10.1177/0008125617725287 (visited on 04/15/2021).

[54] Gustavo Manso. “Motivating Innovation”. en. In: *The Journal of Finance* 66.5 (2011). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2011.01688.x, pp. 1823–1860. ISSN: 1540-6261. DOI: https://doi.org/10.1111/j.1540-6261.2011.01688.x. URL:

https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2011.01688.x (visited on 04/15/2021).

[55] James G. March. “Exploration and Exploitation in Organizational Learning”. In: *Organization Science* 2.1 (1991). Publisher: INFORMS, pp. 71–87. ISSN: 1047-7039. URL: https://www.jstor.org/stable/2634940 (visited on 04/15/2021).

[56] Eduardo Melero, Neus Palomeras, and David Wehrheim. “The Effect of Patent Protection on Inventor Mobility”. In: *Management Science* 66.12 (May 2020). Publisher: INFORMS, pp. 5485–5504. ISSN: 0025-1909. DOI: 10.1287/mnsc.2019.3500. URL: https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2019.3500 (visited on 03/11/2021).

[57] Jessica Milli et al. *Equity in Innovation: Women Inventors and Patents*. Tech. rep. Institute for Women’s Policy Research, Dec. 2016, p. 47. URL: https://iwpr.org/publications/equity-in-innovation-women-inventors-and-patents/ (visited on 02/21/2020).

[58] Nicholas Monath, Christina Jones, and Sarvo Madhavan. “Disambiguating Inventors, Assignees, and Locations”. en. In: *American Institutes for Research* (Aug. 2020), p. 28.

[59] Nicholas Monath and Andrew McCallum. “Discriminative Hierarchical Coreference for Inventor Disambiguation”. en. In: *Machine Learning* (Aug. 2015), p. 124.

[60] Frank Nagle and Florenta Teodoridis. “Jack of all trades and master of knowledge: The role of diversification in new distant knowledge integration”. en. In: *Strategic Management Journal* 41.1 (2020). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.3091, pp. 55–85. ISSN: 1097-0266. DOI: https://doi.org/10.1002/smj.3091. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/smj.3091 (visited on 05/17/2021).

[61] Michele Pezzoni, Francesco Lissoni, and Gianluca Tarasconi. *How To Kill Inventors: Testing The Massacrator© Algorithm For Inventor Disambiguation*. en. Tech. rep. 2012-29. Publication Title: Cahiers du GREThA (2007-2019). Groupe de Recherche en Economie Théorique et Appliquée (GREThA), 2012. URL: https://ideas.repec.org/p/grt/wpegrt/2012-29.html (visited on 05/08/2021).

[62] Gene Quinn. *Patent Cost: Understanding Patent Attorney Fees*. en-US. Section: Inventors Information. Apr. 2015. URL: https://www.ipwatchdog.com/2015/04/18/patent-cost-understanding-patent-attorney-fees/id=56970/ (visited on 09/11/2020).

[63] Julio Raffo and Stéphane Lhuillery. “How to play the “Names Game”: Patent retrieval comparing different heuristics”. en. In: *Research Policy* 38.10 (Dec. 2009), pp. 1617–1627. ISSN: 0048-7333. DOI: 10.1016/j.respol.2009.08.001. URL: https://www.sciencedirect.com/science/article/pii/S0048733309001528 (visited on 03/17/2021).

[64] Cecilia L. Ridgeway and Shelley J. Correll. “Unpacking the Gender System: A Theoretical Perspective on Gender Beliefs and Social Relations”. en. In: *Gender & Society* 18.4 (Aug. 2004), pp. 510–531. ISSN: 0891-2432, 1552-3977. DOI: 10.1177/

0891243204265269. URL: http://journals.sagepub.com/doi/10.1177/0891243204265269 (visited on 08/11/2020).

[65] Cesare Righi and Timothy Simcoe. *Patent Examiner Specialization*. en. SSRN Scholarly Paper ID 2951107. Rochester, NY: Social Science Research Network, Sept. 2017. URL: https://papers.ssrn.com/abstract=2951107 (visited on 02/21/2020).

[66] Bhaven Sampat and Heidi L. Williams. "How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome". en. In: *American Economic Review* 109.1 (Jan. 2019), pp. 203–236. ISSN: 0002-8282. DOI: 10.1257/aer.20151398. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20151398 (visited on 11/18/2019).

[67] Sharon Sassler, Katherine Michelmore, and Kristin Smith. "A Tale of Two Majors: Explaining the Gender Gap in STEM Employment among Computer Science and Engineering Degree Holders". en. In: *Social Sciences* 6.3 (2017). Publisher: MDPI, Open Access Journal, pp. 1–26. URL: https://ideas.repec.org/a/gam/jscscx/v6y2017i3p69-d103441.html (visited on 08/12/2020).

[68] Jasjit Singh. "Collaborative Networks as Determinants of Knowledge Diffusion Patterns". en. In: *Management Science* 51.5 (May 2005), pp. 756–770. ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.1040.0349. URL: http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1040.0349 (visited on 03/17/2021).

[69] Sameer B. Srivastava. "Network Intervention: Assessing the Effects of Formal Mentoring on Workplace Networks". en. In: *Social Forces* 94.1 (Sept. 2015), pp. 427–452. ISSN: 0037-7732. DOI: 10.1093/sf/sov041. URL: https://academic.oup.com/sf/article/94/1/427/1753648 (visited on 10/05/2019).

[70] Gauri Subramani. "Gender and Application Dynamics in U.S. Patent and Trademark Office Applications". In: (2021).

[71] Sarah Thébaud and Maria Charles. "Segregation, Stereotypes, and STEM". en. In: *Social Sciences* 7.7 (2018). Publisher: MDPI, Open Access Journal, pp. 1–18. URL: https://ideas.repec.org/a/gam/jscscx/v7y2018i7p111-d156990.html (visited on 08/12/2020).

[72] Xuan Tian and Tracy Yue Wang. "Tolerance for Failure and Corporate Innovation". en. In: *The Review of Financial Studies* 27.1 (Jan. 2014). Publisher: Oxford Academic, pp. 211–255. ISSN: 0893-9454. DOI: 10.1093/rfs/hhr130. URL: https://academic.oup.com/rfs/article/27/1/211/1571535 (visited on 08/11/2020).

[73] Andrew Toole, Stefano Breschi, et al. *Progress and Potential: A profile of women inventors on U.S. patents*. en. Tech. rep. 1. U.S. Patent and Trademark Office, Feb. 2019, p. 20.

[74] Andrew A Toole, Amanda Myers, et al. "On-line appendixes to "PROGRESS AND POTENTIAL: A profile of women inventors on U.S. patents"". en. In: (Feb. 2019), p. 19.

[75] Manuel Trajtenberg, Gil Shiff, and Ran Melamed. "NBER WORKING PAPER SERIES". en. In: (Aug. 2006), p. 75.

[76] USPTO. *Disambiguation Overview — PatentsView*. 2020. URL: https://patentsview.org/disambiguation (visited on 05/08/2021).

[77] USPTO. *General Requirements Bulletin for Admission to the Examination for Registration to Practice in Patent Cases before the United States Patent and Trademark Office*. July 2020. URL: https://www.uspto.gov/sites/default/files/documents/OED_GRB.pdf.

[78] USPTO. *Table of Patent Fees- Current, Proposed and Unit Cost*. Feb. 2016. URL: https://www.uspto.gov/sites/default/files/documents/Table_of_Patent_Fees-Current_Proposed_and_Unit_Cost.pdf.

[79] USPTO, Office of Documentation Information, and Technology Assessment and Forecast Program. *Buttons to Biotech: 1996 Update Report with Supplemental Data Thru 1998 : U.S. Patenting by Women, 1977 to 1996*. U.S. Department of Commerce, Patent et al., 1999. URL: https://books.google.com/books?id=lWyvI_JyqNwC.

[80] John P. Walsh and Sadao Nagaoka. *Who Invents?: Evidence from the Japan-U.S. inventor survey*. Discussion paper. Research Institute of Economy, Trade and Industry (RIETI), July 2009. URL: https://econpapers.repec.org/paper/etidpaper/09034.htm (visited on 05/08/2021).

[81] Melanie Wasserman. "Gender Differences in Politician Persistence". en. In: *SSRN Electronic Journal* (2018). ISSN: 1556-5068. DOI: 10.2139/ssrn.3370587. URL: https://www.ssrn.com/abstract=3370587 (visited on 11/12/2019).

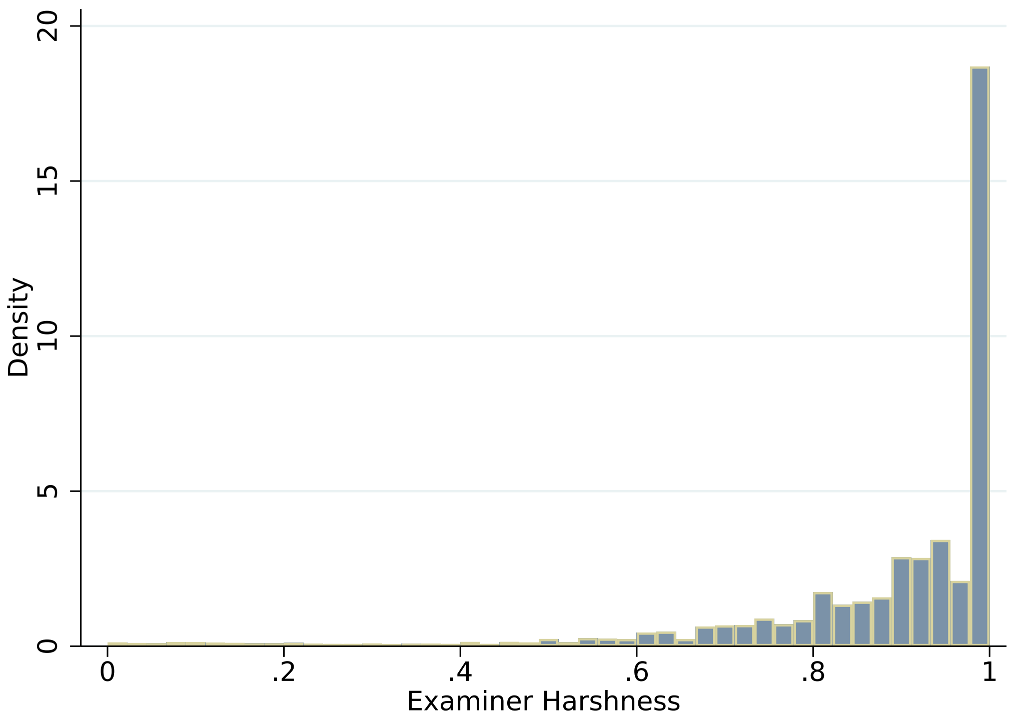
[82] Michael Wick, Sameer Singh, and Andrew McCallum. "A Discriminative Hierarchical Model for Fast Coreference at Large Scale". en. In: *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics* (July 2012), pp. 379–388. URL: https://www.aclweb.org/anthology/P12-1040.

Appendix A

Try, try, try again? Differential Responses to Rejection & the Gender Innovation Gap

A.1 Appendix Figures and Tables

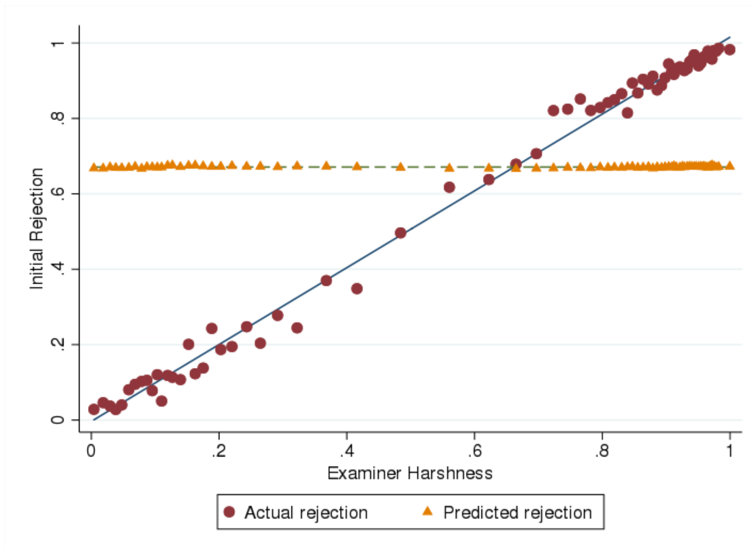
Figure A.1: Distribution of Examiner Harshness by Initial Rejection (Nonresidualized)



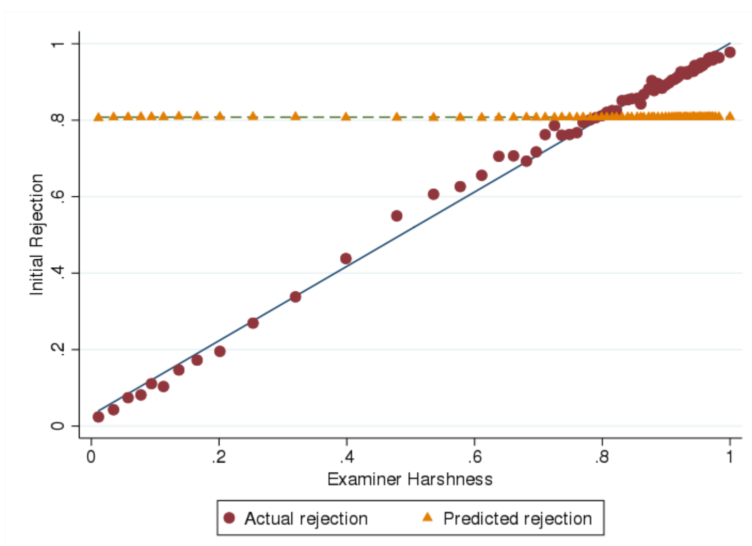
This figure shows the distribution of patent examiners' initial rejection rates.

Figure A.2: Probability of Initial Rejection by Examiner Harshness: Figures by Gender

All-female inventors



All-male inventors



These figures relate examiner harshness to two variables: the actual initial rejection rate, shown in red, and the predicted rejection rate in yellow. By construction, the initial rejection rate is perfectly correlated with examiner harshness. Predicted rejection is based on observables that proxy for quality. These figures illustrate that examiner harshness is an effective instrument regardless of applicant gender; for both male and female applicants, there is no relationship between examiner harshness and these application characteristics.

Table A.1: Effect of Initial Rejection on Patent Application Continuation: Alternate Definition of Harshness IV

	(1)	(2)	(3)	(4)
Panel A: Effect of Initial Rejection on Initial Amendment				
Female X Initial Rejection	-0.055***	-0.046***	-0.095***	-0.058***
	(0.008)	(0.007)	(0.009)	(0.010)
Initial Rejection	0.380***	0.379***	0.381***	0.274***
	(0.015)	(0.015)	(0.015)	(0.021)
Panel B: Effect of Initial Rejection on Patent Granted				
Female X Initial Rejection	-0.116***	-0.086***	-0.109***	-0.099***
	(0.019)	(0.017)	(0.020)	(0.020)
Initial Rejection	-2.892***	-2.894***	-2.895***	-2.830***
	(0.041)	(0.041)	(0.041)	(0.048)
Observations	971547	971547	971547	461147
# of Clusters	36851	36851	36851	36727
Female Definition	Proportion	Half Female	All Female	Solo

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the eight separate regressions displayed in this table, we instrument for initial rejection using examiners' leave-out mean **overall** rejection rate for all other applications within art unit-year. This alternate definition of harshness allows us to check that our results are robust and not reliant on exclusively defining harshness in terms of the rate of giving initial rejections. Definitions of the Female variable are denoted below each column and are described in the text. All regressions include art unit-year fixed effects and are clustered at the examiner-year level.

Table A.2: Heterogeneity by Examiner Gender (IV)

	Initial Amendment		Patent Received	
	(1)	(2)	(3)	(4)
Female × Initial Rejection × Examiner Female	-0.003	0.002	-0.003	0.011
	(0.006)	(0.009)	(0.010)	(0.013)
Initial Rejection × Examiner Female	-0.002	-0.003	-0.036***	-0.037***
	(0.002)	(0.002)	(0.006)	(0.006)
Female × Initial Rejection	-0.034***	-0.078***	-0.059***	-0.109***
	(0.004)	(0.006)	(0.007)	(0.009)
Female × Examiner Female	0.003	-0.001	0.006	-0.001
	(0.003)	(0.004)	(0.007)	(0.009)
Observations	816168	816168	816168	816168
# of Clusters	30300	30300	30300	30300
Female Definition	Half Female	All Female	Half Female	All Female

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table examines whether examiner gender affects determinations of initial rejection, and if there is any interaction between examiner and applicant gender. Each column reports coefficients from a separate regression. An observation is a patent application. We instrument for initial rejection using an examiner's leave-out mean initial rejection rate for all other applications within art unit-year. We find that female examiners are no more or less likely than male examiners to lead to differential outcomes for male vs female applicants. It does appear that applications reviewed by female examiners are less likely to convert to granted patents, but this does not vary based on the gender of the applicants. Examiner Female is a dummy variable for whether patent examiner is a female. All regressions include art unit-year fixed effects and are clustered at the examiner-year level.

Table A.3: Effect of Initial Rejection on Patent Application Continuation: Limited Sample (IV)

	(1)	(2)	(3)	(4)
Panel A: Effect of Initial Rejection on Initial Amendment				
Female X Initial Rejection	-0.044**	-0.032**	-0.043**	-0.040**
	(0.017)	(0.016)	(0.018)	(0.019)
Initial Rejection	0.555***	0.555***	0.554***	0.559***
	(0.038)	(0.038)	(0.038)	(0.042)
Panel B: Effect of Initial Rejection on Patent Granted				
Female X Initial Rejection	-0.090***	-0.074***	-0.092***	-0.099***
	(0.024)	(0.022)	(0.025)	(0.026)
Initial Rejection	-1.005***	-1.005***	-1.007***	-0.980***
	(0.044)	(0.044)	(0.045)	(0.048)
Observations	40310	40310	40310	32931
# of Clusters	15305	15305	15305	13320
Female Definition	Proportion	Half Female	All Female	Solo

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table captures the effect of receiving an initial rejection on continuing in the patent process for the sample of applications that do not use an attorney and are not associated with a firm. This sample represents applications for which we can reasonably assume 1) the inventor is handling communication with the examiner directly, and 2) there is no outside sponsorship of the patent application by an organization. Each column reports coefficients from a separate regression. An observation is a patent application. We instrument for initial rejection using examiners' leave-out mean initial rejection rate for all other applications within art unit-year.

Table A.4: Effect of Initial Rejection on Patent Outcomes (Mixed-Gender Teams only)

	Initial Amendment	Patent Granted
Female X Initial Rejection	-0.034***	-0.051***
	(0.004)	(0.009)
Initial Rejection	0.886***	-0.679***
	(0.013)	(0.023)
Observations	99,731	99,731
# of Clusters	29,758	29,758
Female Definition	Half Female	Half Female

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In this table, we limit our analyses only to mixed-gender teams and evaluate how teams composed of $\geq 50\%$ of women differ from teams whose members are majority male. Our effect sizes are very similar to what we find when we do not limit the sample only to teams composed of both men and women as in our main analyses.

Table A.5: Variation in Persistence by Number of Rejections

	(1)	(2)	(3)	(4)
Female X Rejection	-0.038***	-0.033***	-0.078***	-0.043***
	(0.003)	(0.003)	(0.004)	(0.004)
Female X Rejection X Round 2	0.112***	0.074***	0.093***	0.084**
	(0.032)	(0.024)	(0.035)	(0.038)
Female X Rejection X Round 3	0.073	0.013	0.094	0.060
	(0.063)	(0.045)	(0.082)	(0.090)
Female X Rejection X Round 4	0.244*	0.139	0.226	0.151
	(0.126)	(0.098)	(0.139)	(0.138)
Female X Rejection X Round 5	0.054	-0.043	0.119	0.131
	(0.250)	(0.157)	(0.391)	(0.533)
Observations	1751944	1751944	1751944	785307
# of Clusters	36851	36851	36851	36733
Female Definition	Half Female	Proportion	All Female	Solo

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table examines variation in persistence across multiple rejections. We instrument for initial rejection using examiners' leave-out mean initial rejection rate for all other applications within art unit-year. We find that female examiners are no more or less likely than male examiners to lead to differential outcomes for male vs female applicants. It does appear that applications reviewed by female examiners are less likely to convert to granted patents, but this does not vary based on the gender of the applicants. All regressions include art unit-year fixed effects and are clustered at the examiner-year level.