Try, Try Again? Gender Differences in Self-Assessment and Evaluation of Innovation in the NASA SBIR program

Isabel Hanewicz, Andrea Belz[1], and Alexandra Graddy-Reed

University of Southern California

Richard Terrile

California Institute of Technology Jet Propulsion Laboratory

November 29, 2020

Author Note

[1] Corresponding author. Andrea Belz served previously as co-Principal Investigator and Research PI on the awards acknowledged below. She currently serves as Division Director of Industrial Innovation and Partnerships at the National Science Foundation. To manage the potential conflicts of interest she has resigned from all roles associated with the NSF awards that funded this research and is recused from all matters related to the awards named herein. Her current research activities follow NSF policies.

This research was funded in part by the National Science Foundation (NSF) I-Corps awards 1440080 and 1740721); the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004); and in part by the University of Southern California's Undergraduate Research Associates Program. The authors were provided access to the data by the NASA SBIR program. Any opinions, findings, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the aforementioned organizations. We thank members of the Management of INnovation, Entrepreneurial Research, and Venture Analysis (MINERVA) lab for their constructive comments on earlier versions of the draft.

Abstract

Gender differences in innovation financing are well documented but poorly understood. We study applicants to the National Aeronautics and Space Administration Small Business Innovation Research program, an initiative seeking to increase access for women. Suggesting that underlying confidence disparities affect the selection process for female proposers, we find that: (1) prior application experience and a patenting history predict selection, but only for males; (2) females develop earlier-stage technologies, and highly qualified women underestimate their projects' technical maturity; and (3) highly qualified female applicants receive a boost in a selection indicator in the first funding phase, but the average pool of women scores lower in the second phase - which we name a "Mattea" effect. These results impact diversity in innovation.

Keywords: Innovation, R &D Funding, Gender, SBIR, NASA, Technology Readiness

JEL codes: H5, J16, O3, O32

Try, Try Again? Gender Differences in Self-Assessment and Evaluation of Innovation in the NASA SBIR program

Introduction

The ability and drive to discern opportunities, acquire scarce resources, and act strategically represent the heart of innovation and entrepreneurship. Crucially, this activity requires self-efficacy (Bandura, 1989), a sense of capability and agency. This generalized notion of confidence contributes to venture emergence and growth (Baum & Locke, 2004; Dimov, 2010; Vilanova & Vitanova, 2020). The importance of such psychological traits in the process of entrepreneurship introduces the possibility of gender effects through disparities in confidence and risk tolerance (Caliendo, Fossen, Kritikos, & Wetter, 2015; Fisk, 2016; Zeffane, 2015), and indeed, the well-known gender gap in entrepreneurship may be explained in part by confidence differentials (Koellinger, Minniti, & Schade, 2013).

Gender disparities may be exacerbated in technology commercialization, wherein women show a lower tendency to invent (W. W. Ding, Murray, & Stuart, 2006; Sugimoto, Ni, West, & Larivière, 2015; Whittington & Smith-Doerr, 2005) and hold fewer commercialized patents (Hunt, Garant, Herman, & Munroe, 2013). Some of these differences may be partially attributed to risk aversion (Stephan & El-Ganainy, 2007) and ambiguous attitudes toward commercialization (Murray & Graham, 2007).

In addition, women may suffer from difficulties in obtaining resources. While grant selection typically does not reveal direct gender bias (Ceci & Williams, 2011), examples exist of indirect effects. For instance, in advanced career stages, where external funding is both crucial for research development and professional growth, some blind application review processes inadvertently preference male applicants due to gendered language variation in proposal writing (Kolev, Fuentes-Medel, & Murray, 2019). Similarly, women's patent applications are more likely to be rejected (Jensen, Kovács, & Sorenson, 2018).

Confidence underlies risk preferences and self-assessment, and these factors may impact resource acquisition in realizing a woman's innovative potential, especially with respect to scientific commercialization. In this light, we explore the impact of this behavior in a competitive government innovation program. As a laboratory, we use the Small Business Innovation Research (SBIR) program, created in 1982 to stimulate innovation as well as increase participation by women and other historically disadvantaged groups (U.S. Small Business Administration, 2020). As of 2020, the SBIR program awards $3 billion annually to businesses with fewer than 500 employees through 11 federal agencies over two competitive stages, known as Phases (U.S. Small Business Administration, 2020). SBIR aims to increase research and development (R&D) for technology offering benefits to federal agencies. This program has been linked to many positive entrepreneurial and innovation outcomes, such as firm growth and subsequent patent activity (c.f. Giga, Graddy-Reed, Belz, Terrile, and Zapatero (2019); Howell (2017); Lerner (1999)). This is a staged program in which only the recipients of Phase 1 awards are eligible to apply for Phase 2. The first Phase generally represents a feasibility study, whereas the second Phase focuses on prototype development.

Unfortunately, after roughly forty years, the fraction of female principal investigators (FPIs) in the SBIR program across all agencies was 13.2% between 2011 and 2018 (Servo et al., 2020). While select studies have examined positive selection and commercialization outcomes associated with FPIs in this program (Bednar, Gicheva, & Link, 2019; Link & Wright, 2015), less is known about gender disparities within the application and review process, due in part to the paucity of data.

We use a unique data set comprising applications to the SBIR program at the National Aeronautics and Space Administration (NASA) to address the underlying issues that drive gender differences in entrepreneurial financing. We look first at the role of experience in selection and assess how gender disparities may manifest. Second, we examine gender differences in how principal investigators (PIs) assess their technology and subsequent impacts on the selection process. For the latter, we use Technology Readiness Level (TRL), an ordinal metric ranging from one to nine. It is used extensively throughout the aerospace industry to track a technology's progress toward deployment (Mankins, 2009). Because NASA requires that the TRL must reach level 6 to be used in a flight project (National Aeronautics and Space Administration, 2007), this is an important indicator in NASA's efforts to identify potential technologies for later use. In principle, this metric should be largely independent of gender effects. However, because the PI estimates the technology's TRL in the proposal, the possibility for behavioral impact exists.

Viewing confidence and risk tolerance through the lens of observable past experience, we find that both the applicant's prior history of applying for the program and filing patents impact selection, but only for males; this may be a byproduct of the sharply reduced histories of the FPI population. We extend this perspective to the TRL reported by the PI. The analysis shows that in Phase 2, FPIs report lower TRLs because they work on earlier stage technologies; furthermore, the most competitive applicants also underestimate the TRL, as determined by NASA's own assessment. Finally, we find that the most competitive women receive a small increase in their scores at Phase 1. In Phase 2, women on average receive lower scores in the second Phase of the program, but this effect is smaller for the most competitive women and is mediated completely by TRL. We term this combined set of results a "Mattea" effect, in that the strongest applicants are not affected.

We contribute to several research streams. First, we add to the human capital discussion regarding gendered behaviors in the context of scientific innovation and entrepreneurship, a finding with important implications beyond federally subsidized R&D. In addition, our findings add to the literature of systems engineering by illuminating a previously less studied aspect of technology assessments and the role of behavior therein. Finally, we make recommendations to the field of public policy to support gender diversity in the technology-based entrepreneurship and financing.

Context: The NASA SBIR Program

The SBIR program aims to help small businesses with fewer than 500 employees complete technological innovation deemed critical by federal agencies (U.S. Small Business Administration, 2020). The goals of the SBIR program are fourfold: (1) increase innovation; (2) support federal R&D goals; (3) fund and expand involvement of women, minorities, and other disadvantaged groups; and (4) increase likelihood that federally-funded R&D will be commercialized in the private sector. While each participating agency operates its own SBIR application process, the program always consists of two consecutive Phases. Within NASA, the SBIR program provides roughly $150 B in funding each year. The program is generally viewed as successful; about two-thirds of SBIR projects reported “significant research value”, and a similar fraction reported that they likely would not have completed their proposed research without the funding (National Research Council, 2009).

Since 2009, the NASA SBIR program has required PIs to provide TRL estimates in their proposals. TRL is commonly used in the aerospace and defense industries to evaluate and express a technology’s maturity from 1 (idea) to 9 (fully tested) (Héder, 2017; Mankins, 2009) (Table 1). The scale is designed to be technology-agnostic for broad use. The TRL metric is used extensively throughout most NASA funding programs, and technologies must achieve a level of 6 for infusion into flight projects (National Aeronautics and Space Administration, 2007). At more advanced levels (i.e., above 5), the protocols for advancing are more formalized via the use of specific validation tests. However, earlier stages may be more difficult to distinguish, such as the difference between “proof of concept” (TRL 3) and “validation in laboratory environment” (TRL 4). Consequently, these distinctions at lower TRLs can produce subjective and inconsistent evaluations (Hay, Reeves, Gresham, Williams-Byrd, & Hinds, 2013; Héder, 2017; Kujawski, 2013).

Insert Table 1 here.

As mentioned above, one of the primary goals of the SBIR program is specifically to increase participation of women. As a result, the program encourages applications from women-owned small businesses (WOSBs). These types of firms have been previously studied in the SBIR context and found to be less likely to receive Phase 2 funding (Joshi, Inouye, & Robinson, 2018) and in some cases, private investment (Gicheva & Link, 2015). They are also associated with fewer patent applications (Link & van Hasselt, 2020). However, analysis of WOSBs can be problematic as the status is not well enforced, yet carries significant benefit (Layman, 2016; National Research Council, 2005; United States Government Accountability Office, 2014).

Due to these challenges of correctly identifying WOSBs, we focus instead on the gender of the PI to understand differences. Beyond conferring the benefit of more accurate identification for analysis, the PI authors the proposal and serves as the key individual driving the technology development. The required proposal contents include a complete description of the technology, the envisioned applications, as well as the potential risks and associated mitigation plan. Thus we expect to see any impacts of gender behavior operate through the PI rather than the ownership structure, which may be fairly separate from the development of the technology, even in small firms with up to 500 employees. Literature on this specific group is sparse but prior work has found that proposals led by FPIs have a lower chance of failure (Link & Wright, 2015) and Phase 2 projects led by FPIs experience a higher likelihood of commercialization (Bednar et al., 2019).

Theory & Hypotheses

The role of experience

Experience is an important dimension of generating strong innovation outcomes as it shapes basic cognitive functions beyond simple domain-related skills, thereby supporting the identification of opportunities and the ability to exploit them (Baron & Henry, 2010; Dencker & Gruber, 2015). In translational science, experience mediates the relationship between an invention and opportunities (Shane, 2016). Entrepreneurial experience is a major driver for predicting the transition to launching a new business (Rotefoss & Kolvereid, 2005). Once the business is launched, founding team members with prior experience in entrepreneurship contribute to venture growth (Colombo & Grilli, 2005), particularly in the earliest stages of innovative ventures (Samuelsson & Davidsson, 2009). Founding teams with any prior startup experience are more likely to survive, but those with a higher number of previous startups report higher sales (Delmar & Shane, 2006). Broad experience helps teams commercialize new products (Furr, 2019), and in the SBIR context, prior work with the subject technology was linked to better outcomes (Link & Wright, 2015).

Beyond its role in predicting the likelihood of venture launch, experience has been positively linked to entrepreneurial outcomes. We extend this line of reasoning and suggest that it will play the same role in a competitive grant application process related to discovery of new technology opportunities, particularly because setbacks in a grant application process do enhance long-term performance in a federal program (Wang, Jones, & Wang, 2019). Specifically, we hypothesize that:

Hypothesis 1A (H1A). Experience is positively linked to the likelihood of proposal selection.

Gender effects may emerge as related disparities in risk perception (c.f. Zeffane (2015)). This lower tolerance may stem from a greater fear of failure (Koellinger et al., 2013), but these measured differences may be more subtle than previously appreciated (Filippin & Crosetto, 2016; Nelson, 2016). Cost-benefit analyses may shift if women estimate different social costs (Fisk, 2016), particularly with their tendencies to overestimate negative outcomes while underestimating utility from success (Harris, Jenkins, & Glaser, 2006). These assessments can impact interest in entrepreneurship, with lower rates of female participation partially attributed to higher rates of risk aversion among women (Caliendo et al., 2015); this difference may be heightened by a tendency to hold themselves to higher standards (Thébaud, 2010). Decreased entrepreneurial participation by women may also stem from lower interest in competition (Bönte & Piegeler, 2013), in contrast to over-confidence that may motivate men to compete more frequently (Niederle & Vesterlund, 2007).

Entrepreneurial self-efficacy is linked to venture emergence (Vilanova & Vitanova, 2020) and growth (Baum & Locke, 2004); in addition, it moderates entrepreneurial experience (Dimov, 2010). However, women report lower confidence levels in their entrepreneurial skills (Smith, Hamilton, & Fabian, 2019) and appear to need higher levels of confidence to pursue the important task of financing their startups (Coleman & Kariv, 2014). Similar dynamics appear in academic science as well. When awardees of a prestigious early-career National Institutes of Health (NIH) grant were asked to read a mock grant rejection, women translated the written feedback into a lower application score and reported less motivation to reapply (Biernat, Carnes, Filut, & Kaatz, 2020). Male postdoctoral fellows at NIH were more confident in both their ability to secure a principal investigator role and their subsequent ability to earn tenure (Martinez et al., 2007).

In other words, women may experience lower self-efficacy for reasons that include higher perception of risk, stricter self-assessments, and reduced confidence. We posit that these psychological traits may deter women from obtaining important experience in the program, and that this will create a gender effect in selection. This leads to the second proposition:

Hypothesis 1B (H1B). Gender moderates the effect of experience on selection.

The advancement of innovative technologies.

Gender can impact innovation through several avenues, including commercial propensity and patenting tendencies. Women represent a smaller fraction of disciplines active in technology transfer (Link, Siegel, & Bozeman, 2007) and engage in less commercial work, but when they do, the quality and impact are equal to or exceed those of their counterparts (Whittington & Smith-Doerr, 2005). Among university faculty members, males are significantly more likely to disclose than females, even with similar publication activity (Thursby & Thursby, 2005). In medical faculties, while women have the same likelihood of reporting any inventions, they disclose fewer (Colyvas, Snellman, Bercovitz, & Feldman, 2012). Ultimately, academic females are less likely to found companies than their counterparts (W. Ding & Choi, 2011). Similar to the trends in general entrepreneurship, lower participation by women in commercializing science may be linked to risk aversion and a lower desire for competition (Stephan & El-Ganainy, 2007).

An important measure of innovation is patent activity. In 2013, women represented 10.8 percent of inventorships in the United States Patent and Trademark Office (USPTO) database (Sugimoto et al., 2015). Despite publishing work with the same impact, American academic female scientists hold fewer patents (Whittington & Smith-Doerr, 2005), patenting at 40 percent of their counterparts' rate (W. W. Ding et al., 2006).

Due to these lower rates of commercialization, we theorize that women will self-report lower TRLs than men, driven in part by two factors. First, we posit that the confidence gap between men and women will lead them to assess the stage of their technologies differently. As previously discussed, lower levels of TRL lack formal indicators and are subjective (Hay et al., 2013; Héder, 2017; Kujawski, 2013). Thus if a woman is less confident, she may assign a lower TRL than is accurate. Furthermore, women's under-representation in advanced engineering has been linked to their small representation of commercialized patents (Hunt et al., 2013). Therefore, we suggest that lower rates of commercialization among women could also be attributed to working on earlier stage technologies.

Both of these factors - lower confidence, and the choice to work on less advanced research - could contribute to women reporting technologies at early stages. We propose the following hypothesis:

Hypothesis H2A (H2A). Female gender will be negatively associated with TRL.

However, prior work has found that women with higher capabilities have confidence issues that show stronger effects (Fisk, 2018). It is therefore conceivable that higher performing women will be more risk averse with respect to failure. Thus, we further conjecture that this will negatively impact the self-reported TRL of FPIs. Specifically, we suggest that:

Hypothesis H2B (H2B). A larger negative effect will be seen within highly competitive FPI applicants.

The evaluation of innovative technologies.

While experience and technology are critical elements in selection for financing, the possibility of underlying gender bias against women also exists. Multiple examples identify this concern in the existing rich discussion of entrepreneurial financing. Specifically, it is well documented that women experience less success in securing funding, especially from external sources (Bigelow, Lundmark, McLean Parks, & Wuebker, 2014; Coleman & Kariv, 2014; Wheadon & Duval-Couetil, 2019). Additionally, investors may use an evaluation heuristic favoring the dominant (male) gender (Brush & Gatewood, 2008; Kaatz et al., 2016; Malmström, Johansson, & Wincent, 2017) or penalize women due to explicit concerns that they are not capable of taking necessary entrepreneurial risk (Marlow & Swail, 2014). Angel investors can stereotype female applicants as dependent and needy, disregarding their education and professional expertise (Edelman, Donnelly, Manolova, & Brush, 2018).

Similar dynamics may operate in the grant review community. Many programs use blinded review processes concealing an applicant's personal characteristics to eliminate biases in evaluation processes and increase diversity (Goldin & Rouse, 2000). However, even in double-blind processes, other gendered behavior may impact outcomes; for instance, women's tendencies to choose more specific words may lead to lower application scores (Kolev et al., 2019). Similarly, female-authored NIH applications generated lower scores but contained more linguistically positive elements than responses to men, suggesting women needed to prove their worthiness but needed protection from disappointment (Kaatz et al., 2016). In the innovation arena, women's patent applications were more likely to be rejected, whereas granted ones had fewer claims allowed and more words added to them (Jensen et al., 2018).

These specific observations can be reflections of how evaluators may unconsciously create criteria favoring dominant stereotypes, creating barriers for women in so-called "male" fields (Uhlmann & Cohen, 2005). Women may not be viewed as legitimate or qualified in domains when the evaluators are mostly males (Brush & Greene, 2020). We suggest that in the NASA SBIR environment, these forces may operate, but that this will be mitigated by technical sophistication, as evidenced by TRL. Therefore, we formulate our final evaluation hypothesis:

Hypothesis 3 (H3). TRL mediates the negative effect of female gender associated with the selection process.

Together, these hypotheses are illustrated in Figure 1, a path diagram of the gender effects on selection.

Insert Figure 1 here.

Methods & Data

To test these hypotheses, we estimate a series of regressions using a sample of applicants to NASA's SBIR program. After a two-stage gender identification process, we employ entropy balancing on relevant covariates in the entire sample. We then estimate the impact of gender separately on innovation and selection.

Data

As illustrated in Figure 2, SBIR funding is awarded in two tranches: a Phase 1 feasibility analysis, and a larger Phase 2 project. At NASA, only Phase 1 awardees are eligible to apply for Phase 2. Our dataset comprises the complete set of SBIR proposals in the years 2009 to 2016. In the first two years of our study, the agency offered up to $100,000[2] over the course of six months for Phase 1 funding; this amount increased to $125,000 starting in 2011. Similarly, NASA offered up to $600,000 over the course of 24 months for Phase 2 funding; this amount increased to $750,000 starting in 2010. We use data collected from the Electronic Handbook (EHB), an internal grant management tool. These EHB data are searchable across years of recorded program history for restricted use. To protect the procurement-sensitive nature of this compilation, we aggregated data from 2009 to 2016 from all internal units, or Mission Directorates. The unit of observation is the proposal, with a Phase 2 proposal considered independently of its associated Phase 1 award. The set of candidate records consists of 10,502 observations.

Insert Figure 2 here.

Gender Classification

The application reports the PI name but not the gender, and so we follow the prior literature and extract the first name to classify gender. Early examples of this research relied on manually coding the data (Kurichi, Kelz, & Sonnad, 2005; Robinson, McKay, Katayama, & Fan, 1998), but more recent studies utilize automated processes (Dion, Sumner, & Mitchell, 2018; Hart, Frangou, & Perlis, 2019). In the latter case, associated errors may include incorrect assignments or misclassifications, as well as incomplete assignments or non-classifications (Santamaría & Mihaljević, 2018).

We predict PI gender using a two-step automated process[3]. First, because this is a US federal program restricted to firms with a majority citizen-ownership, we use birth record data from the Social Security Administration (SSA) (West, Jacquet, King, Correll,

[2] The cost information contained in this document is of a budgetary and planning nature and is intended for informational purposes only. It does not constitute a commitment on the part of the Jet Propulsion Laboratory and/or the California Institute of Technology.

[3] In the first step, we used the *gender* package; for the remaining unclassified observations, we used the *genderize* package.

& Bergstrom, 2013). By applying an appropriate birth-year range to the SSA data, we can assess the frequency with which names were given to male or female children. The potential limitations of this approach are two-fold. First, gender-neutral names or names given to the uncommon gender will likely result in incorrect assignment. Second, the SSA dataset will not accurately capture names that are not commonly used for US-born children, nor those names used commonly in other countries.

To augment this, we employ a second step that uses a dynamic search of the internet using a global database of names obtained through social media (Wais, 2006). By late 2020, it had over 115 million records in its database. This approach lacks the historical component of the SSA data and thus will fail to capture how name usage changes in time (Blevins and Mullen (2015) describe this as the so-called “Leslie problem”). Another source of misclassification is a lack of specificity for names such as “Andrea”, which is male in Italian and female in Spanish; however, as we do not have further information on the PI’s background, this cannot be fully resolved. Additionally, as women are more likely to have gender-neutral names, prediction methods tend to create a downward bias in their representation (Lieberson, Dumais, & Baumann, 2016; West et al., 2013) and in general, these algorithms tend to be less accurate in predicting women’s names (Blevins & Mullen, 2015; King, Bergstrom, Correll, Jacquet, & West, 2017; Szymkowiak & Rhodes-Reese, 2020).

All automated routines pose the challenge of defining a minimum probability optimizing the trade-off between statistical significance and a definitive prediction. Scholars do not yet agree on how to specify the gender probability threshold (GPT) for coding the gender consistently even with one algorithm. For instance, when the commercially available *genderize* routine is used, Hart (2019) and Dion (2018) use a GPT of 60% and 70% respectively, whereas Symkowiak (2020) and Kaji (2019) use 90%. Shen (2018) does not specify the GPT, but tests the robustness of their results with 90%. With these variations in mind, we use a GPT of 75% for our primary analysis and later test a more stringent GPT of 90% for robustness.

We first use the SSA data (Mullen, 2018) with a birth year range of 1945 to 1980, consistent with Department of Defense estimates that 75% of applicant ages fall between ages 35 and 64 (National Research Council, Policy and Global Affairs, Board on Science Technology and Economic Policy, & Committee on Capitalizing on Science Technology and Innovation, 2014), with an average of 53. Second, the unclassified observations are then inputted into the *genderize* platform (Wais, 2006). This combined process led to an estimated 5% loss in the sample, with 9,970 observations identified for analysis.

Previous Patents

We developed a variable to estimate the PI's level of technical sophistication and history within the field by analyzing the patent history. Patents serve as quality signals in external financing of small, early-stage firms (Conti, Thursby, & Thursby, 2013; Hottenrott, Hall, & Czarnitzki, 2016) and thus may be interpreted as relevant indicators of the PI's experience of commercializing technologies. Though not included in the EHB, applicants may cite relevant patents in their application. To integrate prior patenting activity, we used software designed for the USPTO database (Belz & Zapatero, 2020; Giga et al., 2019), extracting all patents associated with the PI's name in the five-year period prior to the proposal submission date. A subsequent automated disambiguation process proceeded by confirming matches between (1) the patent assignee and the SBIR firm name; and (2) the state of the inventor and the firm. Observations were retained if they were absent entirely from the database with no disambiguation concerns. We then created a binary variable *Previous Patent(s)*, set equal to 1 if any past patents were verified for that PI, and 0 otherwise. This process created a set of 8,293 observations comprising 6,903 Phase 1 applicants and 1,390 Phase 2 applicants.

Table 2 details the sample selection and construction. Women serve as PIs on 658 (9.53%) proposals at Phase 1, and 128 (9.21%) at Phase 2; of these, applications with female PIs represent 140 (8.99%) of the awards at Phase 1 and 49 (8.51%) at Phase 2.

Insert Table 2 here.

Descriptive Statistics

Initial TRL. As per Figure 2, TRLs are measured chronologically as follows: (1) The PI estimates the technology's initial TRL in the Phase 1 proposal; (2) NASA evaluates the technology's final TRL at the end of Phase 1; (3) The PI again estimates the initial TRL in the Phase 2 proposal; and (4) NASA evaluates the final TRL at the end of Phase 2. In principle, the NASA measurement of step (2) above should be equal to the PI's measurement of step (3); any differences may reveal the confidence-driven effects of interest in this work. The measurement of step (4) was not used in this analysis because it does not contribute to the proposal nor selection process. Although this is an ordinal scale, we treat it as a linear variable to facilitate analysis. We estimate a series of robustness analyses treating it as a categorical variable and find consistent results.

Technical Score. Multiple reviewers both internal and external to NASA independently provide appraisals of the scientific merit, technical feasibility, team capability, and the proposed plan of a SBIR proposal. These scores are averaged in an aggregated Technical Score ranging from 1 to 100, representing a key factor in selection. The minimum Technical Score for consideration in the selection process is 85. A Technical Score is evaluated at each Phase and is thus termed "TS1" and "TS2".

Center Rank. Applicants submit proposals to one of ten NASA Centers, each of which manages specific technology topics by reviewing its own proposals and making preliminary selections for ultimate review and approval by SBIR program managers. During the years of this study, program management began to infuse a ranking system such that each Center would prioritize each proposal in terms of its projected ability to meet NASA's objectives. This ranking system was introduced at different times in each Center but when present, a rank serves as a strong indicator of quality and relevance.

Therefore, it was used as a binary variable indicating quality, employed only in forming a sub-sample described further below.

Competitive History. To estimate the role of experience on selection, we create a binary variable to indicate the PI's specific experience in the NASA SBIR program at a given Phase. *Competitive History* assigns a value of 1 if the PI was previously a competitive applicant at that Phase, defined as having a prior application with a non-zero Center rank and a Technical Score greater than or equal to 85.

The descriptive statistics (Table 3) show that at Phase 1, 10% of female PIs had previous patents in the preceding five-year window, compared with 35% of male PIs, a disparity in line with the prior literature (Sugimoto et al., 2015; Whittington & Smith-Doerr, 2005) and consistent with the estimated correlation coefficient (-0.16, $p<0.01$) between patents and gender. Further, FPIs are less likely to have a Competitive History with the NASA SBIR program; in Phase 1, 30% of women show this experience, compared to 38% of men. However, at Phase 2 this disparity increases to just 16% of female PIs, versus 25% of male PIs. Interestingly, at Phase 1 the initial TRL and Technical Score do not show gender differences; however, the Technical Score in Phase 2 shows a lower mean value for FPI in the entire applicant population, but not the awardees.

Insert Table 3 here.

Estimation Approach

Stratifying the sample. To determine if the disparity in Competitive History and Previous Patents lead to differences in selection rates, we stratify the sample and begin with male Principal Investigators (MPI) (Table 4), with models 1-4 (5-8) estimating a logit on the probability of selection in Phase 1 (2). Each estimation includes a year fixed effect for the proposal submission date and a Center fixed effect to control for the technology type or discipline[4].

Model 1 suggests that Competitive History is important for selection in Phase 1, whereas model 5 suggests that Previous Patents drive selection in Phase 2. In other words, two distinct forms of experience are linked to selection, supporting H1A. Next, we find that Technical Score is used as a strong indicator of selection in models 2 and 6 for the two Phases. In our third test (models 3 and 7), we see that the initial TRL is an indicator of selection in each Phase, although the coefficient in Phase 2 is three times larger than that of Phase 1. Finally, in the combined model, we see that the Technical Score may mediate all these effects, although Previous Patents is still weakly significant. In summary, we show that H1A is supported and that Technical Score is an important mediator.

We perform the same estimation procedure for the total sample of FPI (Table 5) and find that in contrast, neither form of experience is linked to selection in any Phase. The positive impact of Competitive History and Previous Patents for men but not women supports Hypothesis H1B, such that gender moderates the effect of experience on selection with men receiving a boost and women receiving no similar benefit. In other words, experience positively affects selection, but only for male PIs. The importance of the Technical Score as an important potential mediator motivates further investigation.

Insert Table 4 here.

Insert Table 5 here.

Entropy Balancing. Given the differences in experience and its impacts on selection, we seek to improve the validity of our estimations by balancing our sample. As gender is clearly not a randomly assigned treatment in this context, we use observational

[4] The ten NASA Centers are: Armstrong Flight Research Center (AFRC), Ames Research Center (ARC), Glenn Research Center (GRC), Goddard Space Flight Center (GSFC), Jet Propulsion Laboratory (JPL), Johnson Space Center (JSC), Kennedy Space Center (KSC), Langley Research Center (LaRC), Marshall Space Flight Center (MSFC), Stennis Space Center (SSC). These Centers all work on R&D in a specific area of interest to NASA. Applicants submit proposals to a participating Center.

data to balance the applicant pool and obviate the selection bias. We balance our sample on observable characteristics that vary by gender. Specifically, we use entropy balancing (Hainmueller, 2012), which evaluates the moments of covariates and estimates weights for an untreated sample such that the first two moments (mean and variance) of the covariates are equivalent to those of the treated observations. It has been used successfully in economics research (Huang & Yeh, 2014; Marcus, 2013). Entropy balancing offers a key advantage in this analysis compared to other matching techniques by preserving the full sample of observations. Given our limited sample of female applicants, we are inclined to preserve the full set for improved statistical power.

We use entropy balancing to create a weighted distribution of untreated (male PI) observations to form a control group. As we lack access to other pre-proposal traits potentially useful as covariates, we use the Competitive History and Previous Patents as indicators of gender-specific differences in the applicant population. As discussed previously, the gender differences in patenting rates are consistent with prior work (W. W. Ding et al., 2006; Sugimoto et al., 2015; Whittington & Smith-Doerr, 2005). In principle, patenting can serve both as an indicator for technology development and for a PI's experience in the field. T-tests can be used to compare the results of the weighting procedure by comparing the distribution of weighted MPI observations with the female (unweighted) counterparts. As expected, the associated p-values are quite low prior to weighting and quite high subsequently (Table 6), suggesting that the observations are weighted appropriately along those covariates.

Insert Table 6 here.

Results

Gender and Innovation

Using the balanced sample of applicants, we turn to test hypotheses H2A and H2B examining gender effects on innovation. Our outcome of interest is the self-reported initial

TRL of Phase 2. We conduct an OLS regression on the TRL with the FPI as a predictor (Table 7). As discussed previously, we expect women to report lower TRLs driven by the dual factors of working on earlier stage technologies and reporting lower confidence in their innovations. We attempt to disaggregate the two issues by including the NASA measurement at the end of Phase 1 as a control. Model 1 shows a modest negative link between FPI and initial TRL, validating hypothesis H2A. However, in Model 2, with the addition of the NASA TRL measure, we see that the negative relationship between women and TRL is fully mediated by the NASA measurement, suggesting that FPI's assessment is indeed systematically lower because they work on earlier stage technologies.

Models 3 and 4 assess the second part of hypothesis 2, H2B, which posits that more competitive female applicants will have a larger negative correlation with TRL, due to disparities in confidence. In an extended analogy to the Competitive History variable, we constructed a sub-sample classified as "Competitive", with qualifying observations required to have both a non-zero Center Rank and a Technical Score greater than or equal to 85 (the selection threshold). As with model 1, model 3 shows a negative effect of FPI on self-reported TRL, with a marginal effect roughly twice as large, providing initial support for H2B. Model 4 then adds the NASA measurement as in model 2. In this estimation the larger effect seen in model 3 is partially mitigated by the NASA measurement, in agreement with the main result. However, a difference remains after accounting for the actual technology maturity, which may then be attributed to mis-estimation. This error has a negative sign, indicating that this population (highly competitive women) systematically underestimate technology maturity. This can be interpreted as a reflection of confidence or other determinants of self-efficacy. Together, the results of Table 7 support the gender differences proposed in H2A and H2B: women work on earlier stage technologies, but the most competitive women also underestimate their technologies' level of advancement.

Insert Table 7 here.

Gender and Evaluation

Next, we examine the impact of gender on the selection process to test hypothesis H3, which suggests that TRL mediates a negative effect of female gender in the selection process. We examine this with the balanced sample, conducting regressions on both technical score and selection in each Phase. We first estimate the impact of gender and TRL on the outcome of Technical Score (Table 8) using OLS regressions. The results for the full sample in Phase 1 (models 1 and 2) indicate no effect of gender on the Technical Score in Phase 1. However, models 3 and 4 suggest that there is indeed an effect on Technical Score in Phase 2 that is not mediated by TRL. The gender difference is approximately 1.2 points lower for female PIs compared to similar male PIs.

Models 5 through 8 re-estimate the regressions on the sub-sample of competitive applicants to see if there is a differential effect by quality as there was with TRL. The results are quite different. In the competitive sample, Phase 1 results show a positive gender effect that is not mediated by TRL. Thus competitive women appear to get approximately a half-point boost in Technical Score, compared to their male counterparts. Models 7 and 8 repeat this analysis at Phase 2 and show a smaller negative effect that is mitigated by TRL. Across these estimations we find mixed support for H3. Female applicants do receive lower technical scores in Phase 2 than their equivalent counterparts, but for competitive applicants, this is mediated by TRL. In addition, competitive female applicants at Phase 1 experience a positive effect on Technical Score that is not mediated by TRL.

Insert Table 8 here.

Finally, we also estimated the model on the outcome of selection to assess underlying gender differences beyond the factors previously examined (Table 9). Controlling for TRL and Technical Score, we find no gender difference in award selection at either Phase. Thus it appears that the negative effects materialize through two mechanisms: first, through the self-reported TRL; and second, through the Technical Score assigned in the evaluation process.

Insert Table 9 here.

Robustness Analyses

We address the analysis methods with several robustness checks.

Gender prediction threshold. As discussed earlier, the gender prediction algorithm can be subject to misclassification, and therefore we repeat the analysis with a more stringent GPT of 90%. The results are highly consistent and robust for both estimations of gender on TRL (Table A1) and Technical Score (Table A2).

Classification of the TRL as a multinomial outcome variable. As previously discussed, TRL is a categorical variable that we treat as linear to ease analysis. As a robustness check, we classify it as a categorical variable and aggregate higher-order TRLs given their low frequency (four and above in Phase 1 and five and above in Phase 2). We re-estimate a multinomial logit model, allowing the effect to vary across levels. Tables A3 and A4, report results for the main sample and the competitive subsample, respectively, with the first three columns labeled "Model 1" representing the estimation without controlling for the NASA measurement, and the second three columns labeled "Model 2" including this control. The multinomial logit estimations are generally consistent but not robust, which may reflect statistical differences in the frequencies at higher levels.

Classification of the TRL as a categorical independent variable. We use the same reasoning to reanalyze the impact of gender on Technical Score when controlling for TRL, re-estimating the OLS functional form for Technical Score in each of the two Phases with TRL serving as a categorical control variable with level one as the referent group. The results are consistent and robust (Table A5).

Mediation by the Technical Score in Phase 1. To understand the potential role of the mediator in the estimation of the Technical Score in Phase 2 (termed "TS2"), we repeat it using the Technical Score in Phase 1 (“TS1”) as an alternative mediator. This has two advantages: (1) TS1 is a linear scale ranging from 1-100 and is thus both finer and truly linear; and (2) TS1 is defined by NASA, and the results of H2A indicate that TRL is potentially an inaccurate measure of the technology. The results (Table A6) indicate that roughly 25% of the effect of the reduced value of TS2 can be attributed to underlying quality as measured by TS1, and that the FPI still is penalized by 1 point of TS on average in Phase 2. For competitive applicants, the gender difference is reduced to approximately one-third of a point lower; however, it is not moderated by TS1 and a systematic negative impact of gender remains.

Alternate method to address endogeneity. We conducted an additional check on the entropy balancing algorithm by using a different method to address endogeneity concerns. A two-step Heckman method may be used for truncated samples in which both the independent and dependent variables are not observed in the sample (Heckman, 1977); to do so, treatment must be estimated with a probit function treating instrumental variables (IV) as regressors, estimating the outcome with a tobit function. We consider Phase 1 selection as the treatment with Competitive History, Previous Patents, Phase 1 Technical Score, Initial TRL, and gender as predictors. We then use TS2 as the continuous outcome variable in the second stage with predictors of gender and Initial TRL at Phase 2. The results (Table A7) show consistent and robust results for the impact of gender, but suggest that TRL is also important. The coefficient for a 1-level increase in TRL is approximately equal to half that of the FPI, suggesting that on average, a FPI must have roughly 2 levels higher of TRL to compensate for the associated gender decrease in the score.

Discussion

This study generates important insights about drivers of success and gender on the financing of innovation. First, we show that experience positively impacts applicants, as expected - but only for males. The staged importance of Competitive History and then Previous Patents suggests that so-called "grantsmanship" may improve selection in Phase 1, whereas selection in the larger Phase 2 prize is linked both to a technical history, as evidenced by patents; and by the promise of the technology itself, as indicated by TRL. To our knowledge, this study is unique in explicitly documenting first, the role of experience in the application; and second, the strong experience disparity between male and female applicants to the SBIR program. A related experience base could potentially even include prior service as a reviewer and not just as an applicant. These results add to the field of human capital by providing another innovation-rich context in which experience confers significant benefits.

Despite four decades of effort, the NASA SBIR program does not show significant participation by women; they represent roughly 10% of the applicant pool. In Phase 2, only 16% of female PIs have submitted to the program before, compared with 25% of male PIs. Unfortunately, the second Phase is in many ways more important - certainly in terms of dollars available to the innovators. The failure to return denies applicants the chance to benefit from setbacks in grant applications (Wang et al., 2019) and preparing for future opportunities (Cope, 2011). If women are less likely to reapply as previously reported for other grant programs (Biernat et al., 2020), they will disproportionately not benefit. Indeed, the irrelevance of experience in selection of females could be a simple reflection of the absence thereof; applying the same framework would result in an award pool even more depleted of women. The deficiency of women in the proposal pool has important implications for science policy relatied to equity, as this extension of the "leaky pipeline" has not previously been explored.

While much attention has been directed on the outcomes of the SBIR program, interest in improving the incoming proposal pool has focused on general business education for scientists, such as the National Science Foundation I-Corps program (Belz & Zapatero, 2020; Huang-Saad, Fay, & Sheridan, 2017). This study motivates a new type of mentoring to support PIs with promising ability and technologies of interest, but less fluency in the language and standards for a federal agency. A simple orientation for proposing PIs with examples of TRL levels could help narrow this estimation gap. Offering coaching to female PIs to make them more successful in the second round could be a powerful method to increase diversity in the Phase 2 awardee pool and would be consistent with other indications that women may benefit more from a stronger entrepreneurial support system (Alonso-Galicia, Fernández-Pérez, Rodríguez-Ariza, & Fuentes-Fuentes, 2015; Coleman & Kariv, 2014; Smith et al., 2019).

By controlling for both the proposer's experience and the "real" technology status, we show that the most competitive females tend to underestimate the maturity of the technology. Because TRL is the effective currency of the NASA community, this disparity is potentially costly in terms of participation. This lack of confidence has impact with respect to both entrepreneurship and innovation. Overconfidence is a major indicator of entrepreneurial intention (Koellinger, Minniti, & Schade, 2007) and the ability to exploit innovation opportunities (Hirshleifer, Low, & Teoh, 2012). It is obvious that women are far from showing overconfidence if even accurate measures lag. Therefore, this work informs the intersection of systems engineering and human capital, as the gender implications of inconsistency in TRL evaluations have not previously been appreciated (Hay et al., 2013; Héder, 2017; Kujawski, 2013).

We have found new evidence that female PIs work on more basic research. The historically lower commercialization rate of women has partially been attributed to lower representation in advanced design and development (Hunt et al., 2013). However, the question of how women advance technologies from basic to applied research - a necessary prerequisite for patenting - has not been thoroughly explored in the literature. Women's participation in mathematics has outpaced that in physics and engineering (National Science Foundation, 2019). Is migration toward more theoretical work indicative of a different type of risk aversion? This fascinating question has not previously been thoroughly explored in the innovation policy community.

Finally, we illustrate the complexity of the evaluation process. This study reports robust results that the most competitive women benefit from the evaluation in Phase 1, with an estimated half-point boost; and any negative effects in Phase 2 are mediated by TRL. This could be a reflection of a different type of experience if the competitive women have also previously served as reviewers for similar funding programs.

On the other hand, the entire pool of female PIs experiences an average one-point lower score in a scoring rubric ranging from 1 to 100. This yields an estimated 1% negative effect, which may be even larger if the full dynamic range is not utilized - for instance, if scores typically range only from 40 to 100 (as suggested by the mean values of the descriptive statistics), then a one-point differential may represent an even larger impact. In addition, because this population includes the most competitive women, the impact may be even larger for less competitive FPIs. This may be a signal of a Matthew effect, wherein the strongest candidates receive a benefit in Phase 1 but the average female PI pays an additional cost in Phase 2. This so-called "Mattea" effect can inhibit ongoing participation by women in federally sponsored research.

A blinded review process that conceals applicant characteristics is often used to address gender disparities (Goldin & Rouse, 2000). This solution may seem attractive but could create new challenges, particularly for the competitive women who might no longer enjoy the gain of Phase I. Instead, this study motivates the development of a two-fold education and mentoring program. Recent trends in entrepreneurship support mechanisms, such as lean startup models, use business model training to simulate the mindset of startup experience Blank (2013). An analogous program for this population could encompass topics such as accurate TRL evaluation, interpreting NASA needs, effective proposal-writing techniques, and others to generate successful proposals. While the training could be offered broadly, outreach could focus on female PIs to improve their success rate.

Our robustness check reveals that when the accuracy of the gender identification increases, the margin for competitive women increases in magnitude, whereas the negative effect for the entire population does not. In other words, identifying a meritorious proposal authored by "Jennifer" rather than the more ambiguous "Leslie" increases the size of the benefit, whereas a moderate-quality proposal by "Jennifer" does not get a lower score. This suggests that the review process is indeed trying to reward exemplary proposals affirmatively authored by women, but an issue remains with the less competitive ones. As we explicitly control for technical quality (via TRL or technical score in Phase 1) and proposal quality (competitive history and patent experience), this difference remains unexplained and could be generated during the evaluation process. Again, we suggest development of a training program, but this time for reviewers. For instance, reviewers may not be aware of the gender differences in proposal word choice (Kolev et al., 2019). Guidance for reviewers could have a tremendous influence if it examined factors such as: (1) proposal language may vary by gender; (2) women tend to work on earlier-stage research; and (3) the tendency to under-report the TRL is linked to application quality for women. Such a training program might have a significant impact in diversifying the portfolio PI population.

In addition, it could be impactful to actively enhance diversity among evaluators, particularly from the prospective PI population. This could lead to two benefits: 1) diversity on a review panel could help mitigate the reaction to the language differentials and related effects; and 2) potential PIs could obtain valuable insight in submitting successful proposals. A separate outreach effort to recruit reviewers could generate long-term outcomes for the program.

Conclusions

We demonstrate that both programmatic and technical experience help male applicants to the NASA SBIR program. However, females show behaviors consistent with lower confidence as they are less likely to have applied previously. They work on technologies at earlier stages, but highly qualified women also under-report their innovation's maturity. Evaluation of the most competitive female applicants generally produces a higher score, but the complete pool experiences a negative effect, creating a so-called "Mattea" effect for highly qualified female candidates. Training for applicants and reviewers alike can potentially address some of these differences. These results contribute to human capital scholarship, systems engineering studies, and innovation policy.

References

Alonso-Galicia, P. E., Fernández-Pérez, V., Rodríguez-Ariza, L., & Fuentes-Fuentes, M. d. M. (2015). Entrepreneurial cognitions in academia: Exploring gender differences. *Journal of Managerial Psychology*, *30*(6), 630–644. doi: 10.1108/JMP-08-2013-0262

Bandura, A. (1989). Human Agency in Social Cognitive Theory. *American Psychologist*, *44*(9), 1175–1184. doi: 10.1109/SCORED.2002.1033040

Baron, R. A., & Henry, R. A. (2010). How entrepreneurs acquire the capacity to excel: insights from research on expert performance. *Strategic Entrepreneurship Journal*, *4*(1), 49–65. doi: 10.1002/sej.82

Baum, J. R., & Locke, E. A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology*, *89*(4), 587–598. doi: 10.1037/0021-9010.89.4.587

Bednar, S., Gicheva, D., & Link, A. N. (2019). Innovative activity and gender dynamics. *Small Business Economics*(September). Retrieved from `http://dx.doi.org/10.1007/s11187-019-00282-2` doi: 10.1007/s11187-019-00282-2

Belz, A., & Zapatero, F. (2020). Impact of Government Grants on Venture Capital Funding of Deep Technology University Spinoffs. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3493973

Biernat, M., Carnes, M., Filut, A., & Kaatz, A. (2020). Gender, Race, and Grant Reviews: Translating and Responding to Research Feedback. *Personality and Social Psychology Bulletin*, *46*(1), 140–154. doi: 10.1177/0146167219845921

Bigelow, L., Lundmark, L., McLean Parks, J., & Wuebker, R. (2014). Skirting the Issues: Experimental Evidence of Gender Bias in IPO Prospectus Evaluations. *Journal of Management*, *40*(6), 1732–1759. doi: 10.1177/0149206312441624

Blank, S. (2013). *Why the lean start-up changes everything.*

Blevins, C., & Mullen, L. (2015). Jane , John ... Leslie ? A Historical Method for Algorithmic Gender Prediction. *Digitial Humanities Quarterly*, *9*(3), 1–18. Retrieved from `http://www.digitalhumanities.org/dhq/vol/9/3/000223/000223.html`

Bönte, W., & Piegeler, M. (2013). Gender gap in latent and nascent entrepreneurship: Driven by competitiveness. *Small Business Economics*, *41*(4), 961–987. doi: 10.1007/s11187-012-9459-3

Brush, C. G., & Gatewood, E. J. (2008). Women growing businesses: Clearing the hurdles. *Business Horizons*, *51*(3), 175–179. doi: 10.1016/j.bushor.2008.01.007

Brush, C. G., & Greene, P. G. (2020). *Catalyzing Change in Equity Investing: Disruptive Models for Financing Women's Entrepreneurship* (Tech. Rep. No. January). Wellesley: Babson College.

Caliendo, M., Fossen, F. M., Kritikos, A., & Wetter, M. (2015). The gender gap in entrepreneurship: Not just a matter of personality. *CESifo Economic Studies*, *61*(1), 202–238. doi: 10.1093/cesifo/ifu023

Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(8), 3157–3162. doi: 10.1073/pnas.1014871108

Coleman, S., & Kariv, D. (2014). *'Deconstructing' entrepreneurial self-efficacy: A gendered perspective on the impact of ESE and community entrepreneurial culture on the financial strategies and performance of new firms* (Vol. 16) (No. 2). Taylor & Francis. Retrieved from http://dx.doi.org/10.1080/13691066.2013.863063 doi: 10.1080/13691066.2013.863063

Colombo, M. G., & Grilli, L. (2005). Founders' human capital and the growth of new technology-based firms: A competence-based view. *Research Policy*, *34*(6), 795–816. doi: 10.1016/j.respol.2005.03.010

Colyvas, J. A., Snellman, K., Bercovitz, J., & Feldman, M. (2012). Disentangling effort and performance: A renewed look at gender differences in commercializing medical school research. *Journal of Technology Transfer*, *37*(4), 478–489. doi: 10.1007/s10961-011-9235-6

Conti, A., Thursby, J., & Thursby, M. (2013). Patents as signals for startup financing. *Journal of Industrial Economics*, *61*(3), 592–622. doi: 10.1111/joie.12025

Cope, J. (2011). Entrepreneurial learning from failure: An interpretative phenomenological analysis. *Journal of Business Venturing*, *26*(6), 604–623. Retrieved from http://dx.doi.org/10.1016/j.jbusvent.2010.06.002 doi: 10.1016/j.jbusvent.2010.06.002

Delmar, F., & Shane, S. (2006). Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization*, *4*(3), 215–247. doi: 10.1177/1476127006066596

Dencker, J. C., & Gruber, M. (2015). The Effects of Opportunities and Founder Experience on New Firm Performance. *Strategic Management Journal*, *36*, 1035–1052. doi: 10.1002/smj

Dimov, D. (2010). Nascent entrepreneurs and venture emergence: Opportunity confidence, human capital, and early planning. *Journal of Management Studies*, *47*(6), 1123–1153. doi: 10.1111/j.1467-6486.2009.00874.x

Ding, W., & Choi, E. (2011). Divergent paths to commercial science: A comparison of scientists' founding and advising activities. *Research Policy*, *40*(1), 69–80. Retrieved from http://dx.doi.org/10.1016/j.respol.2010.09.011 doi: 10.1016/j.respol.2010.09.011

Ding, W. W., Murray, F., & Stuart, T. E. (2006). Gender differences in patenting in the academic life sciences. *Science*, *313*(5787), 665–667. doi: 10.1126/science.1124832

Dion, M. L., Sumner, J. L., & Mitchell, S. M. (2018). Gendered Citation Patterns across Political Science and Social Science Methodology Fields. *Political Analysis*, *26*, 312–327. doi: 10.7910/DVN/R7AQT1

Edelman, L. F., Donnelly, R., Manolova, T., & Brush, C. G. (2018). Gender stereotypes in the angel investment process. *International Journal of Gender and Entrepreneurship*, *10*(2), 134–157. doi: 10.1108/IJGE-12-2017-0078

Filippin, A., & Crosetto, P. (2016). A reconsideration of gender differences in risk attitudes. *Management Science*, *62*(11), 3138–3160. doi: 10.1287/mnsc.2015.2294

Fisk, S. R. (2016). Gender stereotypes, risk-taking, and gendered mobility. *Advances in Group Processes*, *33*, 179–210. doi: 10.1108/S0882-614520160000033007

Fisk, S. R. (2018). Who's on Top? Gender Differences in Risk-Taking Produce Unequal Outcomes for High-Ability Women and Men. *Social Psychology Quarterly*, *81*(3), 185–206. doi: 10.1177/0190272518796512

Furr, N. R. (2019). Product adaptation during new industry emergence: The role of start-up team preentry experience. *Organization Science*, *30*(5), 1076–1096. doi: 10.1287/orsc.2018.1278

Gicheva, D., & Link, A. N. (2015). The gender gap in federal and private support for entrepreneurship. *Small Business Economics*, *45*(4), 729–733. doi: 10.1007/s11187-015-9664-y

Giga, A., Graddy-Reed, A., Belz, A., Terrile, R. J., & Zapatero, F. (2019). Helping the Little Guy: The Impact of Government Grants on Small Technology Firms. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3054809

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*. doi: 10.1257/aer.90.4.715

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, *20*(1), 25–46. doi: 10.1093/pan/mpr025

Harris, C. R., Jenkins, M., & Glaser, D. (2006). Gender Differences in Risk Assessment : Why do Women Take Fewer Risks than Men ? *Judgment and Decision Making*.

Hart, K. L., Frangou, S., & Perlis, R. H. (2019). Gender Trends in Authorship in Psychiatry Journals From 2008 to 2018. *Biological Psychiatry*, *86*(8), 639–646. Retrieved from `https://doi.org/10.1016/j.biopsych.2019.02.010` doi: 10.1016/j.biopsych.2019.02.010

Hay, J., Reeves, J. D., Gresham, E., Williams-Byrd, J., & Hinds, E. (2013). Evidence for Predictive Trends in TRL Transition Metrics.. doi: 10.2514/6.2013-5369

Heckman, J. J. (1977). Sample Selection Bias As A Specification Error With An Application To The Estimation Of Labor Supply Functions. *NBER Working Paper Series*. doi: 10.1515/9781400856992.206

Héder, M. (2017). From NASA to EU: The evolution of the TRL scale in Public Sector Innovation. *Innovation Journal*, *22*(2), 1–24.

Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are Overconfident CEOs Better Innovators ? *The Journal of Finance*, *LXVII*(4), 1457–1498.

Hottenrott, H., Hall, B. H., & Czarnitzki, D. (2016). Patents as quality signals? The implications for financing constraints on R&D. *Economics of Innovation and New Technology*, *25*(3), 197–217. doi: 10.1080/10438599.2015.1076200

Howell, S. T. (2017). Financing innovation: Evidence from R&D grants. *American Economic Review*, *107*(4), 1136–1164. doi: 10.1257/aer.20150808

Huang, H. C., & Yeh, C. C. (2014). Inflation targeting on unemployment rates: A quantile treatment effect approach. *Applied Economics Letters*, *21*(7), 453–458. Retrieved from `http://dx.doi.org/10.1080/13504851.2013.866198` doi: 10.1080/13504851.2013.866198

Huang-Saad, A., Fay, J., & Sheridan, L. (2017). Closing the divide: accelerating technology commercialization by catalyzing the university entrepreneurial ecosystem with I-Corps^{TM}. *Journal of Technology Transfer*, *42*(6), 1466–1486. doi: 10.1007/s10961-016-9531-2

Hunt, J., Garant, J. P., Herman, H., & Munroe, D. J. (2013). Why are women underrepresented amongst patentees? *Research Policy*, *42*(4), 831–843. doi: 10.1016/j.respol.2012.11.004

Jensen, K., Kovács, B., & Sorenson, O. (2018). Gender differences in obtaining and maintaining patent rights. *Nature Biotechnology*, *36*(4), 307–309. Retrieved from `http://dx.doi.org/10.1038/nbt.4120` doi: 10.1038/nbt.4120

Joshi, A. M., Inouye, T. M., & Robinson, J. A. (2018). How does agency workforce diversity influence Federal R&D funding of minority and women technology entrepreneurs? An analysis of the SBIR and STTR programs, 2001–2011. *Small Business Economics*, *50*(3), 499–519. doi: 10.1007/s11187-017-9882-6

Kaatz, A., Lee, Y. G., Potvien, A., Magua, W., Filut, A., Bhattacharya, A., … Carnes, M. (2016). Analysis of national institutes of health r01 application critiques, impact, and criteria scores: Does the sex of the principal investigator make a difference'. *Academic Medicine*, *91*(8), 1080–1088. doi: 10.1097/ACM.0000000000001272

Kaji, A. H., Meurer, W. J., Napper, T., Nigrovic, L. E., Mower, W. R., Schriger, D. L., … Schenkel, S. (2019). State of the Journal: Women First Authors, Peer Reviewers, and Editorial Board Members at Annals of Emergency Medicine. *Annals of Emergency Medicine*, *74*(6), 731–735. doi: 10.1016/j.annemergmed.2019.05.011

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., & West, J. D. (2017). Men Set Their Own Cites High: Gender and Self-citation across Fields and over Time. *Socius: Sociological Research for a Dynamic World*, *3*, 237802311773890. doi: 10.1177/2378023117738903

Koellinger, P., Minniti, M., & Schade, C. (2007). "I think I can , I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology*, *28*, 502–527. doi: 10.1016/j.joep.2006.11.002

Koellinger, P., Minniti, M., & Schade, C. (2013). Gender Differences in Entrepreneurial Propensity. *Oxford Bulletin of Economics and Statistics*, *75*(2), 213–234. doi: 10.1111/j.1468-0084.2011.00689.x

Kolev, J., Fuentes-Medel, Y., & Murray, F. (2019). *Is Blinded Review Enough? How Gendered Outcomes Arise Under Anonymous Evaluation.* doi: 10.5465/ambpp.2019.15210abstract

Kujawski, E. (2013). Analysis and critique of the system readiness level. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, *43*(4), 979–987. doi: 10.1109/TSMCA.2012.2209868

Kurichi, J. E., Kelz, R. R., & Sonnad, S. S. (2005). Women authors of surgical research. *Archives of Surgery*, *140*(11), 1074–1077. doi: 10.1001/archsurg.140.11.1074

Layman, M. (2016). Mistake, Fraud, and Waste : Mass Inefficiencies in Small Business Programs That Allow Self-Certification. *Public Contract Law Journal*, *46*(1), 167–187.

Lerner, J. (1999). The government as venture capitalist: The long-run impact of the SBIR program. *Journal of Business*, *72*(3), 285–318. doi: 10.1086/209616

Lieberson, S., Dumais, S., & Baumann, S. (2016). The Instability of Androgynous Names : The Symbolic Maintenance of Gender Boundaries Author (s): Stanley Lieberson , Susan Dumais and Shyon Baumann Source : American Journal of Sociology , Vol . 105 , No . 5 (Mar ., 2000), pp . 1249-1287 Published by. *American Journal of Sociology*, *105*(5), 1249–1287.

Link, A. N., Siegel, D. S., & Bozeman, B. (2007). An empirical analysis of the propensity of academics to engage in informal university technology transfer. *Industrial and Corporate Change*, *16*(4), 641–655. doi: 10.1093/icc/dtm020

Link, A. N., & van Hasselt, M. (2020). Exploring the impact of R&D on patenting activity in small women-owned and minority-owned entrepreneurial firms. *Small Business Economics*, *54*(4), 1061–1066. doi: 10.1007/s11187-018-00130-9

Link, A. N., & Wright, M. (2015). On the Failure of R and D Projects. *IEEE Transactions on Engineering Management*, *62*(4), 442–448. doi: 10.1109/TEM.2015.2404873

Malmström, M., Johansson, J., & Wincent, J. (2017). Gender Stereotypes and Venture Support Decisions: How Governmental Venture Capitalists Socially Construct Entrepreneurs' Potential. *Entrepreneurship: Theory and Practice*, *41*(5), 833–860. doi: 10.1111/etap.12275

Mankins, J. C. (2009). Technology readiness assessments: A retrospective. *Acta Astronautica*, *65*(9-10), 1216–1223. Retrieved from `http://dx.doi.org/10.1016/j.actaastro.2009.03.058` doi: 10.1016/j.actaastro.2009.03.058

Marcus, J. (2013). The effect of unemployment on the mental health of spouses - Evidence from plant closures in Germany. *Journal of Health Economics*, *32*(3), 546–558. Retrieved from `http://dx.doi.org/10.1016/j.jhealeco.2013.02.004` doi: 10.1016/j.jhealeco.2013.02.004

Marlow, S., & Swail, J. (2014). *Gender, risk and finance: why can't a woman be more like a man?* (Vol. 26) (No. 1-2). Taylor & Francis. Retrieved from `http://dx.doi.org/10.1080/08985626.2013.860484` doi: 10.1080/08985626.2013.860484

Martinez, E. D., Botos, J., Dohoney, K. M., Geiman, T. M., Kolla, S. S., Olivera, A., ... Cohen-Fix, O. (2007). Falling off the academic bandwagon. Women are more likely to quit at the postdoc to principal investigator transition. *EMBO Reports*, *8*(11), 977–981. doi: 10.1038/sj.embor.7401110

Mullen, L. (2018). *gender: Predict Gender from Names Using Historical Data.* Retrieved from `https://github.com/ropensci/gender`

Murray, F., & Graham, L. (2007). Buying science and selling science: Gender differences in the market for commercial science. *Industrial and Corporate Change*, *16*(4), 657–689. doi: 10.1093/icc/dtm021

National Aeronautics and Space Administration. (2007). NASA Systems Engineering Handbook. *NASA Special Publication*, 360. Retrieved from `http://adsabs.harvard.edu/full/1995NASSP6105.....S`

National Research Council. (2005). *Analyzing Information on Women-Owned Small Businesses in Federal Contracting.* Washington, DC: The National Academies Press. doi: 10.17226/11245

National Research Council. (2009). *An Assessment of the Small Business Innovation Research Program at the National Aeronautics and Space Administration.* Washington, DC: The National Academies Press. Retrieved from `https://doi.org/10.17226/12441` doi: 10.17226/12441

National Research Council, Policy and Global Affairs, Board on Science Technology and Economic Policy, & Committee on Capitalizing on Science Technology and

Innovation. (2014). *SBIR at the Department of Defense.* National Academies Press. doi: 10.17226/18821

National Science Foundation. (2019). *Women, Minorities, and Persons with Disabilities in Science and Engineering Statistics: 2019* (Tech. Rep.). National Science Foundation. Retrieved from `www.nsf.gov/statistics/wmpd`

Nelson, J. A. (2016). Not-So-Strong Evidence for Gender Differences in Risk Taking. *Feminist Economics*, *22*(2), 114–142. doi: 10.1080/13545701.2015.1057609

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, *122*(3), 1067–1101. doi: 10.1162/qjec.122.3.1067

Robinson, D. H., McKay, D. W., Katayama, A. D., & Fan, A. C. (1998). Are Women Underrepresented as Authors and Editors of Educational Psychology Journals? *Contemporary Educational Psychology*, *23*(3), 331–343. doi: 10.1006/ceps.1997.0967

Rotefoss, B., & Kolvereid, L. (2005). Aspiring, nascent and fledging entrepreneurs: An investigation of the business start-up process. *Entrepreneurship and Regional Development*, *17*(2), 109–127. doi: 10.1080/08985620500074049

Samuelsson, M., & Davidsson, P. (2009). Does venture opportunity variation matter? investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics*, *33*(2), 229–255. doi: 10.1007/s11187-007-9093-7

Santamaría, L., & Mihaljević, H. (2018). Comparison and benchmark of name-to- gender inference services. *PeerJ Computer Science*, *2018*(7), 1–29. doi: 10.7717/peerj-cs.156

Servo, J. C., Verostek, J., Smith, V., Lidoro, K., Meade, D., Johnson, K., & Pipher, T. (2020). *Women's Inclusion in Small Business Innovation Research & Small Business Technology Transfer Programs* (Tech. Rep. No. August). Rochester, NY: Dawnbreaker, Inc.

Shane, S. (2016). Linked references are available on JSTOR for this article : Prior Knowledge and the Discovery of Entrepreneurial Opportunities. *Organization Science*, *11*(4), 448–469.

Shen, Y. A., Webster, J. M., Shoda, Y., & Fine, I. (2018). *Persistent Underrepresentation of Women's Science In High-Profile Journals.*

Smith, S., Hamilton, M., & Fabian, K. (2019). Entrepreneurial drivers, barriers and enablers of computing students: gendered perspectives from an Australian and UK university. *Studies in Higher Education*, *0*(0), 1–14. Retrieved from `https://doi.org/10.1080/03075079.2019.1637840` doi: 10.1080/03075079.2019.1637840

Stephan, P. E., & El-Ganainy, A. (2007). The entrepreneurial puzzle: Explaining the gender gap. *Journal of Technology Transfer*, *32*(5), 475–487. doi: 10.1007/s10961-007-9033-3

Sugimoto, C. R., Ni, C., West, J. D., & Larivière, V. (2015). The academic advantage: Gender disparities in patenting. *PLoS ONE*, *10*(5), 1–10. doi: 10.1371/journal.pone.0128000

Szymkowiak, M., & Rhodes-Reese, M. (2020). Addressing the Gender Gap: Using Quantitative and Qualitative Methods to Illuminate Women's Fisheries Participation.

Frontiers in Marine Science, *7*(May), 1–16. doi: 10.3389/fmars.2020.00299

Thébaud, S. (2010). Gender and entrepreneurship as a career choice: Do self-assessments of ability matter? *Social Psychology Quarterly*, *73*(3), 288–304. doi: 10.1177/0190272510377882

Thursby, J. G., & Thursby, M. C. (2005). Gender patterns of research and licensing activity of science and engineering faculty. *Journal of Technology Transfer*, *30*(4), 343–353. doi: 10.1007/s10961-005-2580-6

Uhlmann, E. L., & Cohen, G. L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*, *16*(6), 474–480. doi: 10.1111/j.0956-7976.2005.01559.x

United States Government Accountability Office. (2014). *WOMEN-OWNED SMALL BUSINESS PROGRAM: Certifier Oversight and Additional Eligibility Controls Are Needed* (Tech. Rep. No. GAO-15-54). Washington, DC: United States Government Accountability Office.

U.S. Small Business Administration. (2020). *About.* Retrieved from `https://www.sbir.gov/about`

Vilanova, L., & Vitanova, I. (2020). Unwrapping opportunity confidence : how do different types of feasibility beliefs affect venture emergence ? *Small Business Economics*, *55*, 215–236.

Wais, K. (2006). Gender Prediction Methods Based on First Names with genderizeR. *The R Journal*, *8*(1), 17–37. Retrieved from `https://doi.org/10.32614/RJ-2016-002` doi: 10.32614/RJ-2016-002

Wang, Y., Jones, B. F., & Wang, D. (2019). Early-career setback and future career impact. *Nature Communications*, *10*(1), 1–10. Retrieved from `http://dx.doi.org/10.1038/s41467-019-12189-3` doi: 10.1038/s41467-019-12189-3

West, J. D., Jacquet, J., King, M. M., Correll, S. J., & Bergstrom, C. T. (2013). The Role of Gender in Scholarly Authorship. *PLoS ONE*, *8*(7). doi: 10.1371/journal.pone.0066212

Wheadon, M., & Duval-Couetil, N. (2019). Token entrepreneurs: a review of gender, capital, and context in technology entrepreneurship. *Entrepreneurship and Regional Development*, *31*(3-4), 308–336. Retrieved from `https://doi.org/10.1080/08985626.2018.1551795` doi: 10.1080/08985626.2018.1551795

Whittington, K. B., & Smith-Doerr, L. (2005). Gender and commercial science: Women's patenting in the life sciences. *Journal of Technology Transfer*, *30*(4), 355–370. doi: 10.1007/s10961-005-2581-5

Zeffane, R. (2015). Gender, trust and risk-taking: A literature review and proposed research model. *Journal of Enterprising Communities*, *9*(3), 221–232. doi: 10.1108/JEC-03-2014-0004

Table 1
Definitions of Technology Readiness Levels (TRLs)

TRL	Definition
1	Basic principles observed and reported
2	Technology concept and/or application formulated
3	Analytical and experimental critical function and/or characteristic proof of concept
4	Component and/or breadboard validation in laboratory environment
5	Component and/or breadboard validation in relevant environment
6	System/sub-system model or prototype demonstration in a relevant environment
7	System prototype demonstration in an operational environment
8	Actual system completed and ’flight qualified’ through test and demonstration
9	Actual system flight proven through successful mission operations

Table 2

Sample selection

	Sample
Number of records	10502
Gender prediction	9970
Patent history confirmed*	8293
Phase 1	6903
Phase 2	1390

* Patent history confirmed = disambiguated or not present in patent database

Table 3

Descriptive statistics and p-values of associated T-tests comparing means

	Mean (FPI)	Mean (MPI)	p-value
Competitive History, P1	0.30	0.38	0.00
Competitive History, P2	0.16	0.25	0.00
Previous Patent(s), P1	0.10	0.35	0.00
Previous Patent(s), P2	0.12	0.39	0.00
Init. TRL, P1	2.36	2.41	0.16
Init. TRL, P2	3.35	3.42	0.41
Technical Score, P1	85.24	85.84	0.21
Technical Score, P2	91.48	92.70	0.09
Awardee-only sample			
Technical Score, P1	93.67	94.12	0.21
Technical Score, P2	95.87	96.06	0.59

Table 4

Impact of experience on selection, MPI sample only

	Selection, P1				Selection, P2			
	logistic				*logistic*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Competitive History	0.226^{***}	-0.093	0.229^{***}	-0.092	-0.251^{*}	-0.261	-0.230	-0.255
	(0.070)	(0.092)	(0.070)	(0.092)	(0.148)	(0.175)	(0.149)	(0.176)
Previous Patent(s)	0.085	0.011	0.080	0.008	0.288^{**}	0.277^{*}	0.284^{**}	0.275^{*}
	(0.064)	(0.084)	(0.064)	(0.085)	(0.124)	(0.147)	(0.124)	(0.147)
Technical Score		0.401^{***}		0.400^{***}		0.326^{***}		0.325^{***}
		(0.012)		(0.012)		(0.023)		(0.023)
Initial TRL			0.062^{**}	0.023			0.156^{**}	0.041
			(0.029)	(0.039)			(0.068)	(0.079)
Constant	-0.682^{***}	-37.224^{***}	-0.842^{***}	-37.267^{***}	0.556	-29.525^{***}	0.015	-29.605^{***}
	(0.193)	(1.160)	(0.207)	(1.163)	(0.342)	(2.150)	(0.414)	(2.156)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,245	6,245	6,245	6,245	1,262	1,262	1,262	1,262
Log Likelihood	−3,266.195	−1,936.846	−3,263.915	−1,936.677	−795.851	−598.564	−793.171	−598.428
Akaike Inf. Crit.	6,566.390	3,909.692	6,563.831	3,911.353	1,625.701	1,233.128	1,622.341	1,234.856

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5

Impact of experience on selection, FPI sample only

	Selection, P1				Selection, P2			
	logistic				*logistic*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Competitive History	0.188	-0.377	0.197	-0.375	0.160	-0.897	0.156	-0.894
	(0.227)	(0.296)	(0.228)	(0.296)	(0.629)	(0.831)	(0.630)	(0.831)
Previous Patent(s)	0.113	0.096	0.148	0.105	-0.835	-0.372	-0.852	-0.397
	(0.322)	(0.397)	(0.324)	(0.400)	(0.724)	(0.998)	(0.728)	(1.004)
Technical Score		0.361^{***}		0.360^{***}		0.466^{***}		0.465^{***}
		(0.035)		(0.035)		(0.114)		(0.113)
Initial TRL			-0.083	-0.025			0.061	0.128
			(0.098)	(0.135)			(0.257)	(0.332)
Constant	0.206	-29.878^{***}	0.363	-29.788^{***}	2.024^{*}	-41.734^{***}	1.859	-41.975^{***}
	(0.724)	(3.507)	(0.747)	(3.535)	(1.201)	(10.818)	(1.389)	(10.766)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	658	658	658	658	128	128	128	128
Log Likelihood	−326.879	−203.436	−326.517	−203.419	−66.627	−43.881	−66.598	−43.808
Akaike Inf. Crit.	687.759	442.872	689.035	444.838	167.253	123.763	169.197	125.615

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6

Covariate T-test results for entropy balancing

	FPI - mean	**MPI - mean**	**Statistic**	**p-value**
Phase 1, N=6903				
Competitive History	0.30	0.38	-4.05	0.00
Competitive History (wt)	0.30	0.30	0.07	0.95
Previous Patent(s)	0.10	0.35	-18.98	0.00
Previous Patent(s) (wt)	0.10	0.10	0.01	0.99
Phase 2, N=1390				
Competitive History	0.16	0.25	-2.85	0.00
Competitive History (wt)	0.16	0.15	0.06	0.95
Previous Patent(s)	0.12	0.39	-8.25	0.00
Previous Patent(s) (wt)	0.12	0.12	0.02	0.99

Table 7

Predictors of TRL using entropy-balanced weights for MPI observations

	Initial TRL, P2			
	OLS			
	Full Sample		Competitive	
	(1)	(2)	(3)	(4)
FPI	-0.098^{**}	0.007	-0.209^{***}	-0.089^{*}
	(0.049)	(0.041)	(0.060)	(0.048)
NASA final TRL, P1		0.574^{***}		0.593^{***}
		(0.023)		(0.027)
Constant	3.305^{***}	1.300^{***}	3.448^{***}	1.321^{***}
	(0.138)	(0.139)	(0.151)	(0.154)
Year fixed effects	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes
Observations	1,390	1,390	817	817
Adjusted R^2	0.042	0.344	0.050	0.406
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 8

Predictors of Technical Scores using entropy-balanced weights for MPI observations

	Technical Score							
	OLS							
	Full Sample, P1		Full Sample, P2		Competitive, P1		Competitive, P2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FPI	−0.268	−0.264	-1.259^{***}	-1.230^{***}	0.538^{***}	0.538^{***}	-0.374^{*}	−0.320
	(0.276)	(0.276)	(0.374)	(0.374)	(0.162)	(0.162)	(0.226)	(0.228)
Initial TRL		0.076		0.292		−0.014		0.254^{*}
		(0.132)		(0.205)		(0.086)		(0.132)
Constant	82.846^{***}	82.669^{***}	94.616^{***}	93.650^{***}	93.340^{***}	93.370^{***}	97.153^{***}	96.276^{***}
	(1.047)	(1.091)	(1.049)	(1.249)	(0.467)	(0.503)	(0.565)	(0.726)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,903	6,903	1,390	1,390	1,594	1,594	817	817
Adjusted R^2	0.016	0.016	0.068	0.069	0.105	0.104	0.056	0.059

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 9

Predictors of selection, weighted

	Selection, P1				Selection, P2			
	glm: quasibinomial *link = logit*				*glm: quasibinomial* *link = logit*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FPI	−0.065	−0.066	−0.031	−0.032	−0.167	−0.159	−0.017	−0.014
	(0.060)	(0.060)	(0.614)	(0.609)	(0.121)	(0.121)	(0.140)	(0.140)
Initial TRL		−0.025		−0.015		0.092		0.016
		(0.029)		(0.306)		(0.066)		(0.076)
Technical Score			0.367^{***}	0.367^{***}			0.376^{***}	0.376^{***}
			(0.087)	(0.087)			(0.024)	(0.024)
Constant	-0.362^{*}	−0.304	-33.457^{***}	-33.411^{***}	1.324^{***}	1.024^{***}	-33.952^{***}	-33.989^{***}
	(0.193)	(0.205)	(8.285)	(8.271)	(0.333)	(0.396)	(2.291)	(2.298)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,903	6,903	6,903	6,903	1,390	1,390	1,390	1,390

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

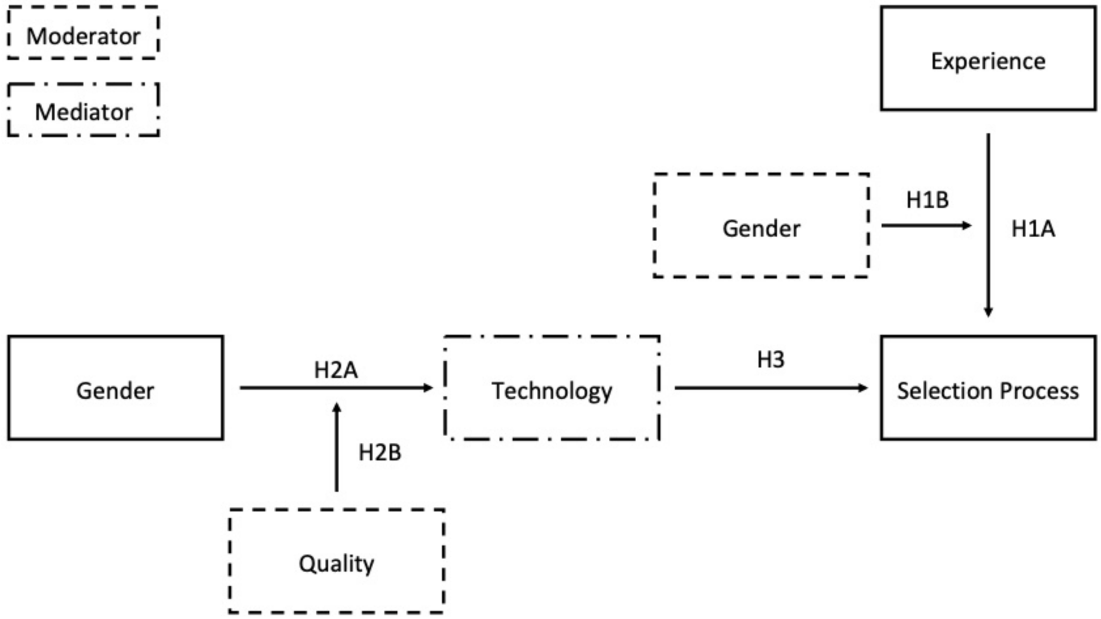

Figure 1. Proposed hypotheses for both Phases

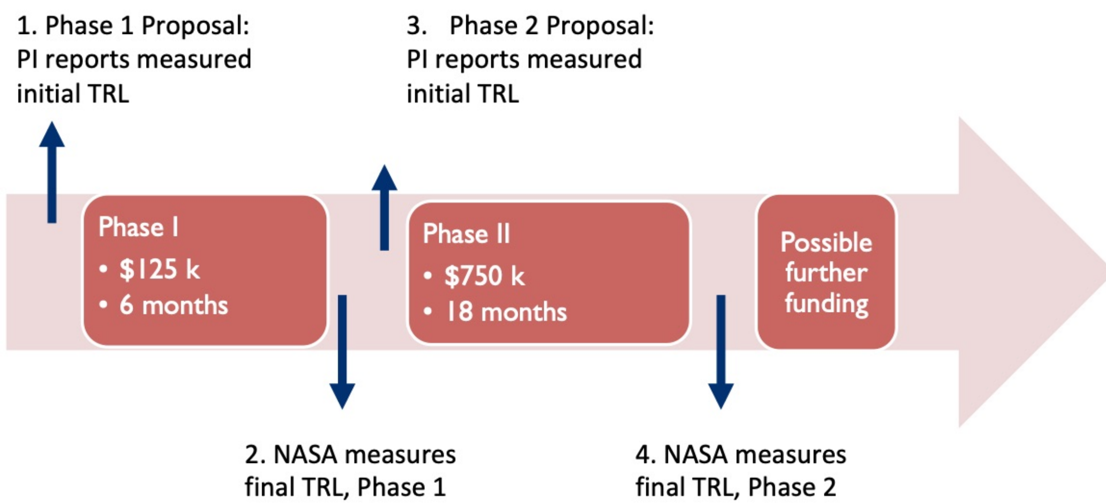

Figure 2. Sequence of measurements

Appendix

Table A1

Predictors of TRL using entropy-balanced weights for MPI observations with alternate GPT

	Init. TRL, P2			
	OLS			
	Full Sample		Competitive	
	(1)	(2)	(3)	(4)
FPI	-0.103^{**}	0.001	-0.270^{***}	-0.149^{***}
	(0.050)	(0.042)	(0.060)	(0.049)
NASA final TRL, P1		0.564^{***}		0.554^{***}
		(0.024)		(0.028)
Constant	3.314^{***}	1.330^{***}	3.497^{***}	1.494^{***}
	(0.136)	(0.141)	(0.143)	(0.153)
Year fixed effects	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes
Observations	1,346	1,346	787	787
R^2	0.061	0.343	0.084	0.399
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A2

Predictors of Technical Scores using entropy-balanced weights for MPI observations with alternate GPT

	Technical Score							
	OLS							
	Full Sample, P1		Full Sample, P2		Competitive, P1		Competitive, P2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FPI	−0.461	−0.458	−1.253***	−1.221***	0.813***	0.814***	−0.053	0.069
	(0.284)	(0.284)	(0.391)	(0.391)	(0.161)	(0.161)	(0.230)	(0.231)
Init. TRL		0.061		0.316		0.029		0.453***
		(0.135)		(0.213)		(0.086)		(0.137)
Constant	82.398***	82.255***	95.023***	93.977***	93.456***	93.391***	96.985***	95.400***
	(1.047)	(1.095)	(1.057)	(1.271)	(0.446)	(0.485)	(0.548)	(0.725)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,642	6,642	1,346	1,346	1,522	1,522	787	787
R^2	0.018	0.018	0.085	0.086	0.143	0.143	0.086	0.099

Note: *p<0.1; **p<0.05; ***p<0.01

Table A3

Predictors of TRL using entropy-balanced weights for MPI observations: Full sample (Multinomial logit)

	Initial TRL 3	Initial TRL 4 Model 1	Initial TRL 5+	Initial TRL 3	Initial TRL 4 Model 2	Initial TRL 5+
FPI	0.149	−0.252	−0.142	0.086	−0.160	0.301
	(0.421)	(0.447)	(0.566)	(0.441)	(0.504)	(0.678)
NASA final TRL, P1				1.013***	2.540***	3.486***
				(0.348)	(0.415)	(0.504)
Constant	1.690	0.810	−0.268	−1.478	−7.678***	−12.621***
	(1.332)	(1.401)	(1.687)	(1.745)	(2.025)	(2.624)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	647	458	128	647	458	128

Note: *p<0.1; **p<0.05; ***p<0.01

Table A4

Predictors of TRL using entropy-balanced weights for MPI observations: Competitive sample (Multinomial logit)

	Initial TRL 3	Initial TRL 4 Model 1	Initial TRL 5+	Initial TRL 3	Initial TRL 4 Model 2	Initial TRL 5+
FPI	−0.185	−0.421	−1.270	−0.188	−0.321	−1.235
	(0.605)	(0.632)	(0.911)	(0.645)	(0.742)	(1.185)
NASA final TRL, P1				0.898*	3.042***	4.322***
				(0.506)	(0.629)	(0.821)
Constant	2.782	1.890	1.021	−0.084	−8.365***	−14.102***
	(1.891)	(1.960)	(2.420)	(2.555)	(2.966)	(3.944)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	379	286	69	379	286	69

Note: *p<0.1; **p<0.05; ***p<0.01

Table A5

Predictors of Technical Scores using entropy-balanced weights for MPI observations (Binned TRL)

	Technical Score							
	OLS							
	Full Sample, P1		Full Sample, P2		Competitive, P1		Competitive, P2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FPI	−0.268	−0.223	−1.259***	−1.212***	0.538***	0.550***	−0.374*	−0.310
	(0.276)	(0.275)	(0.374)	(0.375)	(0.162)	(0.162)	(0.226)	(0.227)
Initial TRL 2, P1		0.909**				0.528**		
		(0.387)				(0.257)		
Initial TRL 3, P1		1.628***				0.412		
		(0.411)				(0.260)		
Initial TRL 4+, P1		−0.009				0.148		
		(0.510)				(0.330)		
Initial TRL 3, P2				0.841				1.322***
				(0.608)				(0.376)
Initial TRL 4, P2				1.532**				0.991**
				(0.646)				(0.394)
Initial TRL 5+, P2				0.690				1.789***
				(0.818)				(0.547)
Constant	82.846***	82.085***	94.616***	93.724***	93.340***	92.973***	97.153***	95.887***
	(1.047)	(1.088)	(1.049)	(1.167)	(0.467)	(0.505)	(0.565)	(0.663)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,903	6,903	1,390	1,390	1,594	1,594	817	817
Adjusted R^2	0.016	0.019	0.068	0.070	0.105	0.106	0.056	0.071

Note: *p<0.1; **p<0.05; ***p<0.01

Table A6

Predictors of Phase 2 Technical Scores with an alternate mediator, using entropy-balanced weights for MPI observations

	Technical Score, P2					
	OLS					
	Full Sample			Competitive		
	(1)	(2)	(3)	(4)	(5)	(6)
FPI	−1.259***		−0.916***	−0.374*		−0.376*
	(0.374)		(0.353)	(0.226)		(0.208)
Technical Score, P1		0.667***	0.658***		0.470***	0.470***
		(0.049)	(0.049)		(0.039)	(0.039)
Constant	94.616***	31.408***	32.749***	97.153***	52.506***	52.738***
	(1.049)	(4.714)	(4.733)	(0.565)	(3.702)	(3.699)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Center fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,390	1,390	1,390	817	817	817
Adjusted R^2	0.068	0.171	0.175	0.056	0.200	0.202
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A7

Heckman two-step model for robustness. Left: Selection equation for Phase 1 selection (probit); Right: Technical Score in Phase 2 (tobit)

	Dependent variable:	
	Selection, P1	Technical Score, P2
	(1)	(2)
Competitive History	-0.589^{***}	
	(0.055)	
Previous Patent(s)	0.084^{*}	
	(0.050)	
Technical Score, P1	0.219^{***}	
	(0.006)	
Initial TRL, P1	-0.016	
	(0.023)	
FPI	0.022	-0.990^{*}
	(0.080)	(0.578)
Initial TRL, P2		0.572^{***}
		(0.179)
Constant	-20.374^{***}	93.604^{***}
	(0.591)	(1.182)
Year fixed effects	Yes	Yes
Center fixed effects	Yes	Yes
Observations	6,735	6,735
Note:		*p<0.1; **p<0.05; ***p<0.01