SEX AND STARTUPS

Jens Frankenreiter, Talia Gillis, & Eric Talley*

Venture Capital (VC) is widely thought to have a gender problem. The sector has long been dominated by men on both the investor and founder sides, and the resulting skew can affect who gets to play, who gets funded, and who gets left behind. Nevertheless, reliable empirical research on early-stage startups is sparse, and we still know next to nothing about whether (and how) internal governance systems within startups manifest gender effects. This paper develops and studies a first-of-its-kind data set that tracks VC-backed startup governance as reflected in corporate charters. It affords a unique window to observe how cash flow and control rights emerge and evolve through successive rounds of startup funding, as well as the detailed structure of governance provisions that regulate internal corporate affairs. After unveiling this resource, we deploy it to assess whether the gender of startup founders predicts differential governance terms that regulate, enable, and impede founders' discretion in relation to their VC investors. Our ultimate findings are mixed. On the one hand, we uncover evidence that female founders face governance structurers that are distinguishable from a matched sample of male founders (as measured by the latent semantic content of charters), and that these differences have increased over time. On the other hand, when we analyze specific, hand-coded attributes of our data that correspond with canonical VC governance provisions, we find no strong patterns that work either to systematically advantage or impair women founders. Our results thus pose an intriguing mystery about what <u>is</u> driving the gendered semantic divergence in our corpus. Pending that mystery's resolution, however, our results are consistent with the conclusion that governance structure does not further aggravate other recognized hurdles that female founders face; but neither, for that matter, do governance regimes ameliorate such disparities.

* Many thanks to Kerry Baldwin, Molly Calkins, Katherine McCormick, Elizabeth Pollman, Darren Rosenblum, Marco Ventoruzzo, and participants in the 2022 BYU Deals Conference, the 2022 Penn/NYU Contracts Conference, the 2022 ECGI Conference on Private Equity and Venture Capital, ALEA 2022, CELS 2022, George Mason Law & Economics Workshop and the Conference on Data Science & Law for helpful comments and discussions. Jiyoung Kim provided outstanding research coordination and made substantive contributions to this project. Thanks as well to a large team of research assistants too numerous to mention here, but whose names are listed with tremendous gratitude in Appendix A. All errors are ours.

TABLE OF CONTENTS

INTRODUCTION

Much of the "dark matter" in the venture capital financing universe consists of star-crossed startups. It is now common wisdom that VC-backed startups overwhelmingly fail,[1] typically because their

[1] *See* Patrick Ward, *Is It True That 90% of Startups Fail?*, NANOGLOBALS (June 29, 2021), https://nanoglobals.com/startup-failure-rate-myths-origin/ (documenting failure rates of VC-backed startups in the range of 75% to 90%); Deborah Gage, *Venture Capital Secret: 3 Out of 4 Start-Ups Fail*, WALL STREET JOURNAL (Sept. 20, 2012), https://www.wsj.com/articles/

founders' visions prove untenable, unworkable, underwhelming, or underappreciated. But it is exceedingly rare for a startup to flounder *and then* for its floundering founder to founder on the shoals of a criminal conviction. Yet it was exactly that fate that awaited Theranos Inc.'s charismatic founder Elizabeth Holmes, who in January 2022 was convicted on multiple federal counts of wire fraud and conspiracy, allegations that stemmed from her fantastical claims about Theranos' revolutionary blood test technology—a spiel that ultimately proved spurious. Holmes was sentenced to over 11 years in federal prison, which she commenced serving in May 2023.[2]

By contrast, consider Adam Neumann, the comparably charismatic founder of WeWork Inc. While coaxing billions out of VC backers, Neumann made similarly lofty assurances about the bright prospects of his newly reimagined commercial lease business plan. And, like Holmes, his house of cards collapsed amid discoveries of corporate excess and woefully undisciplined governance.[3] Neumann was eventually (and unceremoniously) ousted from his executive perch, but he never faced criminal charges. Rather, he pocketed a $1.7 billion severance bounty as he decamped from his gilded WeWork cubicle. No longer affiliated with WeWork, Neumann now oversees several new real estate investment funds,[4] reportedly whiling away his time musing about

---

SB10000872396390443720204578004980476429190 (describing higher failure rates than indicated by the industry rule of thumb of that only three or four out of every ten startups fail completely); Elizabeth Pollman, *Startup Failure* (Working Paper, Mar. 20, 2022), https://law-economic-studies.law.columbia.edu/sites/default/files/content/Pollman_Startup%20Failure_rev%202022%2003%2020_for%2003%2028%20Law%20&%20Econ%20Workshop.pdf (describing law's role in both creating a system where startup failure is so common and shaping startups' ability to "fail with honor" through soft-landing acquisitions, acqui-hires, and other alternative mechanisms).

[2] *See* Bernd Debusmann Jr & James Clayton, Theranos CEO Elizabeth Holmes begins 11-year prison sentence, BBC News (May 30, 2023) https://www.bbc.com/news/world-us-canada-65756588.

[3] *See* Donald C. Langevoort & Hillary A. Sale, *Corporate Adolescence: Why Did 'We' Not Work?,* 99 TEXAS L. REV. 1347 (2021) (using the story of WeWork to illustrate the risks of "a build-up of bad choices and test behaviors commonly observed in human adolescents" occurring before a start-up reaches "public adulthood"); Dominic Rushe, *Troubled WeWork Scraps Share Sale After Ousting Founder Adam Neumann*, THE GUARDIAN (Sept. 30, 2019), https://www.theguardian.com/business/2019/sep/30/wework-scraps-share-sale-adam-neumann (describing how Neumann took $700 million out of the company before the IPO, initiated a questionable dual-class share sale that would have given him total control after the IPO, and engaged in generally "eccentric behavior").

[4] *See* Alexandra Tremayne-Pengelly, *WeWork Founder Adam Neumann Is Back With Another Real Estate Venture*, OBSERVER (Feb. 7, 2023), https://observer.com/2023/02/wework-founder-adam-neumann-is-back-with-

"becoming leader of the world, living forever, and amassing more than $1 trillion in wealth."[5] WeWork, meanwhile, is waging a losing battle for its very survival.[6]

Many factors no doubt distinguish the Holmes and Neumann narratives, ultimately rendering their side-by-side comparison to be questionable—both anecdotally and statistically. There is, for example, a compelling argument that Holmes's fraudulent behavior (in the medical field) had significantly greater social impact than Neumann's insouciant profligacy (in commercial real estate).[7] Nevertheless, it is hard *not* to notice the founders' divergent fates and wonder how and whether their paths may reflect—at least in some small way—a larger story about gender.[8] Did Holmes's and Neumann's distinct gender identities play any role in driving their differing receptions, treatments, trajectories, and outcomes? This question is relevant not only for the Holmes's and Neumann's specific fates but also for generations of startup founders who now follow in their collective wake.[9]

---

another-real-estate-venture/ (explaining how Neumann's new company, Flow, received a $350 million investment from a VC firm in 2022).

[5] *See* Mohammed Abrar Asif, *Adam Neumann: A Saga Of Lies And Fraud*, THE FINANCIAL PANDORA (Nov. 5, 2020), https://thefinancialpandora.com/adam-neumann-a-saga-of-lies-and-fraud/.

[6] *See* Lynn Doan et al., WeWork's shares are so low that the struggling office landlord is praying a 1-for-40 reverse stock split will salvage its listing, Fortune (Aug. 18, 2023) https://fortune.com/2023/08/18/wework-reverse-stock-split/.

[7] *See, e.g.*, Bobby Allyn, *The Elizabeth Holmes Trial Sparks A Silicon Valley Debate: Why Not Other Tech CEOs?*, NPR TECHNOLOGY (Sept. 25, 2021), https://www.npr.org/2021/09/25/1040442689/elizabeth-holmes-trial- why-her-not-other-ceos (listing potential explanations for the disparate treatment of Elizabeth Holmes, including sexism, the egregiousness of the allegedly fraudulent behavior, and evidence that Holmes acted with intent).

[8] We are hardly the first to posit this question. *See, e.g.*, Ellen Pao, *Sexism in Tech is Real and Alive. How Big a Role Is It Playing in Elizabeth Holmes's Trial?*, THE ECONOMIC TIMES (Sept. 17, 2021) ("Male chief executives and founders just aren't held accountable in ways that would lead to reform across the tech industry. And even when they are made to answer for their actions, they find their ways back into the fold very quickly."); Allyn, *supra* note 7. That said, at least one other (male) principal at Theranos was also convicted of criminal fraud. *See* Erin Griffith, *No. 2 Theranos Executive Found Guilty of 12 Counts of Fraud*, N.Y. TIMES (July 7, 2022), https://www.nytimes.com/2022/12/07/technology/sunny-balwani-theranos-sentenced.html#:~:text=No.,2%20Theranos%20Executive%20Is%20Sentenced%20to%20Nearly%2013%20Years%20for,counts%20of%20fraud%20in%20July ("Ramesh Balwani, the former chief operating officer of the failed blood testing start-up Theranos, was sentenced… to nearly 13 years in prison…").

[9] *See, e.g.*, Erin Griffith, *They Still Live in the Shadow of Theranos's Elizabeth Holmes*, N.Y. TIMES (Aug. 24, 2021), https://www.nytimes.com/2021/08/24/technology/theranos-elizabeth-holmes.html (reporting on female founders who must now distinguish themselves from Elizabeth

Concerns about gendered capitalism are hardly novel in the land of the unicorns (and unicorn wannabes). Indeed, the last half decade bears witness to significantly greater scrutiny of the gender dynamics within the entire venture capital industry, an area long dominated by males (on both founder and funder sides.)[10] Even as recently as 2022, for example, women-founded companies garnered somewhere between 2% and 16.5% of the total capital invested in venture-backed startups in the US (depending on how one adds up[11]). The gender challenges in VC echo—at least to some extent—similar inquiries within public companies, where a growing literature has documented a variety of ways that female management and director representation predicts different corporate attributes,[12] and yet also seems to increase the likelihood of being targeted by activists, critics, and cranks.[13] As one legal

Holmes in meetings with investors); Elaine Moore, *Silicon Valley Has Learnt Little From Elizabeth Holmes: Making Big Claims Remains the Starting Point for New Companies*, FINANCIAL TIMES MAGAZINE (Feb. 22, 2022), https://www.ft.com/content/66f11e1f-9ec9-406d-868f-d0f757a915d6 ("Instead of seeing the case as a spur to toughen up due diligence, the tech sector is choosing to dismiss it as an outlier.").

[10] *See, e.g.*, Benjamin P. Edwards & Ann C. McGinley, *Venture Bearding*, 52 U.C. DAVIS L. REV. 1873, 1884 (2019) (coining the term "venture bearding" to examine "how the current startup, technology, and venture capital landscape causes persons with stigmatized identities to strategically conceal facets of their female identities in favor of presenting masculinized identities to conduct business and raise capital"); Kellye Y. Testy, *From Governess to Governance: Advancing Gender Equity in Corporate Leadership*, 87 GEO. WASH. L. REV. 1095, 1096–98 (2019) (debunking the reasons justifying the slow progression toward more diverse board representation and outlining steps to improve gender equality in corporate governance). *See also US VC Female Founders Dashboard*, PITCHBOOK (Feb. 1, 2023), https://pitchbook.com/news/articles/the-vc-female-founders-dashboard [hereinafter *Female Founders Dashboard*].

[11] The number subdivides by whether one counts firms founded by woman-only founders (2%) or women co-founded teams (16.5%). *See Female Founders Dashboard*, *supra* note 10.

[12] Pressure to increase board gender diversity has led to different director pools. *See* Todd A. Gormley, Vishal K. Gupta, David A. Matsa, Sandra C. Mortal & Lukai Yang,*The Big Three and Board Gender Diversity: The Effectiveness of Shareholder Voice* (Nat'l Bureau of Econ. Rsch., Working Paper No. 30657, 2021), https://www.nber.org/papers/w30657 (showing that the pressure to increase diversity from "The Big Three" led to firms hiring female directors that were less connected to the existing CEO and board members, and had less executive experience). *See also* Miriam Schwartz-Ziv, *Gender and Board Activeness: The Role of a Critical Mass*, 52 J. FIN. & QUANTITATIVE ANALYSIS 751 (2017) (demonstrating that more gender diverse boards conduct more active board meetings).

[13] *See* Andrew Ross Sorkin, *Do Activist Investors Target Female C.E.O.s?*, N.Y. TIMES DEALBOOK (Feb. 9, 2015), https://archive.nytimes.com/dealbook.nytimes.com/2015/02/09/the-women-of-the-s-p-500-and-investor-activism/ (describing evidence of "a subconscious gender bias among activist investors" and the relationship between gender and power);

commentator has recently pointed out, most conventional company law statutes and doctrines have traditionally sidelined discrimination and harassment-related claims between shareholders and managers, granting them only incidental significance within other categories.[14]

Within VC-backed startups, there are some general signs that the gender gap is starker still. VC funds have tended (until recently) to be disproportionately male-dominated, and there is evidence that this composition affects their investments, particularly in their tepid investments in women-founded or women-led startups.[15] Other studies have examined differential treatment of women in managerial positions in other areas of finance, as female-managed funds have lower average inflows than male-managed funds despite showing no gender

---

Vishal K. Gupta, Seonghee Han, Sandra C. Mortal, Sabatino (Dino) Silveri & Daniel B. Turban, *Do Women CEOs Face Greater Threat of Shareholder Activism Compared to Male CEOs? A Role Congruity Perspective*, 103 J. APPLIED PSYCH. 228 (2018) "[O]ur results suggest that female (compared to male) CEOs have to deal with additional challenges imposed by activist investors and are more vulnerable to activists' efforts towards wielding power in the firm."); Bill B. Francis, Iftekhar Hasan, Yinjie (Victor) Shen & Qiang Wu, *Do Activist Hedge Funds Target Female CEOs? The Role of CEO Gender in Hedge Fund Activism*, 141 J. FIN. ECON. 372 (2021) ("Using a comprehensive US hedge fund activism dataset from 2003 to 2018, we find that activist hedge funds are about 52% more likely to target firms with female CEOs compared to firms with male CEOs."); Anna Domanska, *The Intense Scrutiny on Female CEOs by Activist Investors*, INDUS. LEADERS MAG. (Aug. 17, 2016), https://www.industryleadersmagazine.com/intense-scrutiny-female-ceos-shareholders-activist-investors/ (noting that using gendered language in PR materials "increases the likelihood of shareholder activism by 31%").

[14] *See* Ann M. Lipton, *Capital Discrimination*, 59 HOUS. L. REV. (May 10, 2022) ("[B]usiness law itself has no vocabulary to engage the influence of sex and gender, or to correct for unfairness traceable to discrimination.").

[15] *See Female Founders Dashboard*, *supra* note 10; Valentina Zarya, *Venture Capital's Funding Gender Gap Is Actually Getting Worse*, FORTUNE (Mar. 13, 2017), https://fortune.com/2017/03/13/female-founders-venture-capital/; Sophie Calder-Wang, Paul Gompers & Patrick Sweeney, *Venture Capital's "Me Too" Moment* (Nat'l Bureau of Econ. Rsch. Working Paper No. 28679, 2021), https://www.nber.org/papers/w28679 (analyzing how the increase in hiring of female venture capitalists following the Ellen Pao v. Kleiner Perkins gender discrimination trial led to an increase in female venture capitalists investing in female founders but no difference in male venture capitalists investing in female founders); Nitasha Tiku, *Gen Z Women Are Breaking into the Venture-Capital Boys Club*, WASH. POST (Apr. 23, 2021), https://www.washingtonpost.com/technology/2021/04/23/gen-z-venture-capital/; Paul A. Gompers & Sophie Calder-Wang, Diversity in Innovation 10–11 (Nat'l Bureau of Econ. Rsch. Working Paper No. 23082, 2017), https://www.nber.org/papers/w23082 ("From 1990-2016 women have been less than 10% of the entrepreneurial and venture capital labor pool.").

differences in performance.[16] And, as noted above (and documented below[17]), women founders have remained largely peripheral when it comes to both giving and receiving VC funding in the first instance.

Nevertheless, our ability to peer directly inside VC-backed companies has remained frustratingly opaque, courtesy of the non-public nature of both startups and their typical financiers, allowing them to collectively evade the radar of public securities databases. Most pointedly, we have virtually no visibility into the nature and characteristics of the internal governance of VC-backed startups. Although certain states (such as California) have promulgated general statutory protections prohibiting discriminatory treatment of female startup entrepreneurs,[18] understanding such mandates (or even enforcing them effectively) requires one to make comparisons *between* companies regarding how founders are treated in their respective corporate governance regimes. Do female-founded firms systematically face more onerous governance terms in relation to VC investors? Or, might their relative rarity on the VC landscape result in women founders garnering *more* generous terms after funding? The stakes in this context are particularly high, as an increasing number of women and other underrepresented groups pursue self-employment options as a response to perceptions of discrimination in conventional job markets.[19] And yet,

---

[16] *See* Alexandra Niessen-Ruenzi & Stefan Ruenzi, *Sex Matters: Gender Bias in the Mutual Fund Industry*, 65 MANAGEMENT SCIENCE 3001 (2019) ("document[ing] significantly lower inflows in female-managed funds than in male-managed funds."). Some studies have found the reverse—that women-founded startups outperform male-founded startups by a significant margin. *See* Katie Abouzahr, Matt Krentz, John Harthorne, & Frances Brooks Taplett, *Why Women-Owned Startups Are a Better Bet*, BOSTON CONSULTING GROUP (June 6, 2018), https://www.bcg.com/publications/2018/why-women-owned-startups-are-better-bet ("[B]usinesses founded by women ultimately deliver higher revenue—more than twice as much per dollar invested—than those founded by men.").

[17] *See infra* Part I.

[18] *See* Cal. Civ. Code § 51.9 (prohibiting sexual harassment pertaining to persons with a "business, service, or professional relationship"). This section was amended after several reports that female founders were being harassed by venture capital providers. *See* Lipton, *supra* note 14; Luke Stangel, *New State Bill Would Make Sexual Harassment in Venture Capital Illegal*, SILICON VALLEY BUS. J. (Aug. 22, 2017, 10:07 AM), https://www.bizjournals.com/sanjose/news/2017/08/22/vc-harassment-bill-sb-224-state-sen-jackson.html (explaining how this bill "would explicitly make sexual harassment between venture investors and startup founders illegal").

[19] *See, e.g.*, Madeline E. Heilman & Julie J. Chen, *Entrepreneurship as a Solution: The Allure of Self-Employment for Women and Minorities*, 13 HUMAN RESOURCES MANAGEMENT REV. 347 (2003) (discussing "the experiences that women and minorities encounter in organizational settings that result in frustration with corporate life and their opportunities for advancement").

our sightlines into these questions remain disconcertingly obscured.[20] Perhaps consequently, there has been no serious effort in the literature thus far to measure and quantify such considerations.[21]

Until now, that is. In this paper, we deploy a first-of-its-kind, hand-collected data set to peek inside the governance systems of VC-backed startups, asking whether women founders face materially different governance landscapes than those of comparable male counterparts. Our inquiry starts with a simple proposition: Corporate governance is foundational not just to value creation but also to the distribution of cash-flow and control rights between founders and funders.[22] The formal provisions of corporate governance thus constitute a critical, authoritative framework for allocating and distributing rights, duties, and privileges of founders, key employees, and VC investors in early stage companies.[23] Moreover, the multiple rounds of standard VC financing foreordain not only that these foundational documents may evolve as the VC-backed company matures but also that their initial structures may determine whether future evolution occurs at all.

Conventional accounts often posit that corporate governance regimes should evolve towards those that maximize the collective joint surplus of entrepreneurs and investors.[24] However, several real-world factors can conspire to frustrate that outcome, including bias, transaction costs, information disparities, liquidity constraints, market access and differential degrees of bargaining power[25]—many of which may be highly correlated to and/or causally driven by gender effects.

---

[20] Although the work here is still spotty, one notable study purports to demonstrate that female entrepreneurs often face different funding terms than male entrepreneurs. *See* Dana Kanze, Mark A. Conley, Tyler G. Okimoto, Damon J. Phillips & Jennifer Merluzzi, *Evidence That Investors Penalize Female Founders for Lack of Industry Fit*, 6 SCIENCE ADVANCES (2020) (documenting that female-led ventures catering to male-dominated industries receive less funding at lower valuations than female-led ventures catering to female-dominated industries).

[21] A few studies in the literature have come close on this score, but they score only glancing shots. *See infra* Part I.

[22] *See, e.g.*, Robert Bartlett & Eric Talley, *Law and Corporate Governance*, *in* THE HANDBOOK OF THE ECONOMICS OF CORPORATE GOVERNANCE (Oxford Press; Hermalin & Weisbach eds. 2017).

[23] *See* Jens Frankenreiter, Cathy Hwang, Yaron Nili & Eric Talley, *Cleaning Corporate Governance*, 170 U. PENN. L. REV. 1, 6 (2021) (describing the foundational nature of corporate governance structures).

[24] *See, e.g.*, Frank Easterbrook & Daniel Fischel, THE ECONOMIC STRUCTURE OF CORPORATE LAW 6 (1991).

[25] *See* Sarath Sanga & Eric Talley, *Don't Go Chasing Waterfalls: Fiduciary Duties in Venture Capital Backed Startups*, 52 J. OF LEGAL STUD. (forthcoming 2023) ("Venture-capital-backed startups are often crucibles of conflict between common and preferred shareholders, particularly around exit decisions.").

Thus, our project seeks to examine whether corporate governance provisions vary based on the gender characteristics of the founder team, and whether such observed variation appears to advantage or disadvantage diverse founders. As noted above, prior research has documented worse *funding* outcomes for women and racial minorities, and it has explored the possible mechanisms that lead to these outcomes. Our inquiry takes that program one step further, asking whether differences in gender predict not only funding differences but also differential allocations of formal governance rights for those who receive VC funding.

To conduct our inquiry, we focus on *the* foundational governance document of startups: Certificates of Incorporation (or "charters"). While public company charters have recently become more readily available for study by scholars,[26] private company charters—far more detailed than their public counterparts—remain far more elusive (notwithstanding the fact that they are, in theory, public documents[27]). With considerable effort, however, we were able to overcome these limitations with a sizeable sample, obtaining the full chartering history of hundreds of female-founded startups between 2003 and 2021. We further analyzed the content of the charters along several lines, including their latent semantic content, their core financial terms, and their non-financial control rights. We did the same for a large matched sample of "similar" male-founded startups, thereby facilitating an apples-to-apples comparison of governance regimes.

Our ultimate findings present something of a mixed bag. At a general level, a comparison of the overall semantic content of our charters indicates distinguishable governance differences between female-founded and male-founded firms. More specifically, we show that charters of female-founded startups resemble their male-founded counterparts *substantially less than* the male-founded counterparts resemble one another. This gender divergence, moreover, neither dissipates nor remains constant over time. Rather, after training an algorithmic classifier to predict the presence of a female founder solely from the charter's text, we find that prediction power modestly *grows* over the timespan of our data set, accelerating after the Elizabeth Holmes indictment dropped. Broadly, this finding suggests that female founders face a formal governance landscape that is not only predictably distinct from their male counterparts but remains persistently so.

That said, the overall semantic divergence of charters *in the aggregate* does not necessarily imply that such differences are manifested

[26] *See, e.g.*, Frankenreiter et al., *supra* note 23.
[27] *See, e.g.*, *id.* at 21-23 (documenting the surprising difficulty of sourcing public chartering histories from the State of Delaware).

for several key governance attributes that typically draw the attention of practitioners, judges and researchers. To explore this possible connection, we extracted labels on over four dozen specific procedures for our data, including a variety of financial terms (such as liquidation preferences, anti-dilution rights, and conversion rights) and non-financial terms (such as veto/approval rights,[28] fiduciary waivers,[29] and board representation). Here, we do not find a consistent pattern of gender differences across these various fields that appears systematically to advantage or disadvantage women founders. Although women founders appear to face disadvantages in some areas (such as the higher frequency of cumulative dividends and board appointment rights for VCs as well as the lower frequency of "pay-to-play" provisions), they tend to receive better treatment in others (such as participation rights of preferred stock, certain preferred veto rights, and fiduciary waivers for VC investors).[30] In most cases, moreover, the differences we do observe are economically modest and statistically negligible.[31]

Our findings—the first of their kind as far as we are aware—have several intriguing implications. On one level, they raise something of a mystery for future scholarship. Although female-founded firms manifest distinct governance documents from their male-founded counterparts in the semantic content of their governance documents, our more targeted inquiries expose largely where those differences *are not* located as opposed to where they *are*. Although the key driver of differences does not appear to reside in these targeted provisions, there may well be other substantive governance categories that we did not attempt to label for this project—and which future endeavors may tease out—that better manifest the same patterns as in the semantic analysis.

Even so, and pending the resolution of that mystery, our findings still carry material implications. To appreciate them, it is important to keep in mind that our data do not emerge from a vacuum: most critically, to be included in our sample, every startup must have been successful in procuring at least one round VC funding. Yet as noted above, prior work has documented that female founders appear to receive differential treatment *at the financing stage* by VC financiers—a finding that complicates the interpretation of gender differences that we can measure in governance structure. The fact that we uncover little

---

[28] *See, e.g.*, Sanga & Talley, *supra* note 25; PWP Xerion Holdings v. Red Leaf Resources, Inc., C.A. No. 2017-0235-JTL (Del. Ch. Oct. 23, 2019).
[29] *See, e.g.*, Gabriel Rauterberg & Eric Talley, *Contracting Out of the Fiduciary Duty of Loyalty: An Empirical Analysis of Corporate Opportunity Waivers*, 117 COLUM. L. REV. 1075 (2017).
[30] *See infra* Part IV.
[31] *See infra* Part IV and Appendix B.

systematic gender patterns key governance terms could be consistent with at least two different hypotheses. First, it could mean that VC actors' biases simply dissipate upon funding, possibly through a variety of market pressures.[32] Alternatively, our findings might constitute merely one stage of a complex sequence of gender interactions. For example: (i) women founders might face adverse treatment in attracting investment, leaving only "high quality" female-founded firms to receive funding in comparison to male counterparts of lower average quality; and (ii) instead of receiving *more lenient* governance terms befitting their higher average quality, women become saddled with identical regimes to those of their male comparators (who by hypothesis have lower average quality). Interpreted in this light, our findings are consistent with the conclusion that while startup governance may not *exacerbate* existing gender disadvantages, neither does it *ameliorate* them.

Our analysis proceeds as follows. Part I frames our project within the relevant literature on gender effects in startups. Part II describes our data sources and the architecture of our matched-sample data set. Part III provides a computational textual analysis of the semantic structure of startups' charters, revealing a gender effect that tends to amplify over time. Part IV then focuses in on a large set of canonical cash-flow and governance provisions that are frequently the focus of VC-founder negotiations. There we show that overall gender effects appear to be inconsistent and usually small statistically. Part V discusses implications of our results, both for legal and social policy and for future researchers.

I. THE BACKSTORY ON VENTURE CAPITAL AND GENDER

Although our project develops a novel data set, we hardly write from a blank slate—far from it. There are now developed academic and practitioner literatures on gender effects in corporate governance and entrepreneurship generally, and the insights (as well as many open questions) that flow therefrom substantially frame and motivate our inquiry here. This Part provides a brief overview of prior literature, emphasizing aspects that are the most germane to our central questions in this enterprise.

[32] *Compare* Gary S. Becker, THE ECONOMICS OF DISCRIMINATION (1957) (arguing discrimination would be competed away in thick markets) *with* Dan A. Black, *Discrimination in an Equilibrium Search Model*, 13 J. LABOR ECON. 309 (1995) (showing that search costs can perpetuate discrimination even with a large number of market participants).

Researchers and commentators have spilled substantial ink documenting the relatively anemic rate at which female-founded and female-controlled startups receive VC funding. Anecdotal accounts of this gap have been around for years,[33] but there is now a more rigorous set of qualitative, experimental and empirical accounts that lend support to those anecdotal accounts. Both sides of the market appear to have contributed to this shortfall. From the investor side, several recent studies have highlighted the male-gendered composition of individual venture funds and of the industry writ large.[34] Some researchers have suggested that gender effects manifest as early as a first pitch of an idea, where embedded forms of bias systematically tilt funding decisions towards male entrepreneurs.[35] Indeed, recent quasi-experimental

[33] *See, e.g.*, Dana Kanze, Laura Huang, Mark A. Conley and E. Tory Higgins, *Male and Female Entrepreneurs Get Asked Different Questions by VCs—And it Affects How Much Funding They Get*, HARV. BUS. REV. (June 27, 2017) (reporting on interactions between VCs and founders at an NYC tech crunch event, observing that male founders were asked "promotion questions" (e.g., aspirations/dreams) whereas female founders were asked prevention questions (e.g., safety, security, responsibility)); Helen Thomas, *Start-up Finance is a Closed Shop for Women*, FIN. TIMES (Sept. 29, 2021), https://www.ft.com/content/60caa57e-d40d-4d6f-974a-1d14a3798d27 (exploring how to address the problem of female founders generally raising less money); Josh Constine, *The Gap Table: Women Own Just 9% of Startup Equity*, TECHCRUNCH (Sept. 18, 2018, 11:08 AM), https://techcrunch.com/2018/09/18/the-gap-table/ (assessing that only 9 percent of founder and employee startup equity is owned by women, even though women constitute 35 percent of startup equity-holding employees); *Why VCs Aren't Funding Women-led Startups*, KNOWLEDGE AT WHARTON (May 24, 2016) (interviewing Wharton faculty who discuss several of the biases and challenges facing female founders); Helen Fitzwilliam, *Female-led Start-ups Embrace Plan B—then C, D, E . . .*, FIN. TIMES (Mar. 7, 2022), https://www.ft.com/content/b41e1ada-6b0a-4aac-9b52-ad811e759336 ("The upheavals of Covid-19 forced female founders to make the most of their skills at achieving more with fewer resources than many of their male counterparts.").

[34] Recent experimental work suggests that startups themselves have a preference for male investors, which might itself contribute to the low representative of women in VCs. *See* Ofir Gefen, David Reeb & Johan Sulaeman, *Choosing Startup Investors: Does Gender Matter?* (Oct. 28, 2022), available at https://www.ecgi.global/sites/default/files/Paper%3A%20Choosing%20Startup%20Investors%3A%20Does%20Gender%20Matter%3F.pdf

[35] *See* Kamal Hassan, Monisha Varadan, & Claudia Zeisberger, *How the VC Pitch Process is Failing Female Entrepreneurs*, HARV. BUS. REV. (Jan. 13, 2020), https://hbr.org/2020/01/how-the-vc-pitch-process-is-failing-female-entrepreneurs (describing how "relying upon data-driven processes in the initial vetting of candidates" can mitigate some of the gender-based biases associated with the pitch process); Laura Huang, Alison Wood Brooks, Sarah Wood Kearney, Fiona E. Murray, *Investors prefer entrepreneurial ventures pitched by attractive men*, 111 PNAS 12 (Mar. 10, 2014) (Investors prefer pitches presented by male entrepreneurs compared with pitches made by female entrepreneurs, even when the content of the pitch is the same).

findings suggest that even established funders appear systematically resistant to pitches that sound in a stereotypically female registrar, regardless of the gender identity projected by the person pitching.[36] Even the composition of VC funds has gender dynamics, and recent research shows that when VC partners have more daughters, their propensity to bring on additional female VC partners increases substantially.[37]

From the startup side, much attention has been devoted to asking whether there are founder-gender effects in the success rates at which startups successfully procure VC funding. Research on contemporary startup activity estimates (based on deal counts) that approximately one quarter of the founders of startups are women, up from nearly nothing two decades ago.[38] However, far less than that number receive VC backing (around 16.5% even by generous estimates[39]). There is evidence, moreover, that these differential rates of funding do not simply reflect "quality" differences among recipients. One empirical study, for example, finds that that notwithstanding the low rates of funding for women startup entrepreneurs, the return on invested capital (ROIC) of female-founded firms is higher than that of male-founded firms.[40] Several other recent contributions corroborate the notion that women founders typically confront largely unfounded negative gender stereotypes related to future success.[41] Moreover, there is at least some evidence that these tendencies have been surprisingly durable. For example, there is some recent evidence that female founders have begun

---

[36] *See* Lakshmi Balachandra, *Research: Investors Punish Entrepreneurs for Stereotypically Feminine Behaviors*, HARV. BUS. REV. (Oct. 19, 2018), https://hbr.org/2018/10/research-investors-punish-entrepreneurs-for-stereotypically-feminine-behaviors (explaining the influence of masculinity and femininity in how investors perceive entrepreneurs).

[37] *See* Sophie Calder-Wang & Paul A. Gompers, *And the children shall lead: Gender diversity and performance in venture capital*, 142 J. FIN. ECON. 1 (Oct. 2021).

[38] *See* PITCHBOOK RSCH. CTR., ALL IN: FEMALE FOUNDERS IN THE VC ECOSYSTEM (2022), https://pitchbook.com/news/reports/2022-all-in-female-founders-in-the-us-vc-ecosystem.

[39] *See Female Founders Dashboard*, *supra* note 10.

[40] *See* Abouzahr et al., *supra* note 16 ("[B]usinesses founded by women ultimately deliver higher revenue—more than twice as much per dollar invested—than those founded by men…").

[41] *See* Malin Malmstrom, Aija Voitkane, Jeaneth Johansson and Joakim Wincent, *VC Stereotypes About Men and Women Aren't Supported by Performance Data*, HARV. BUS. REV. (Mar. 15, 2018); Candida Brush, Patricia Green, Kashmi Balachnadra, Amy Davis, *The gender gap in venture capital- progress, problems, and perspectives*, 20 VENTURE CAPITAL 115 (2018); John Paul Titlow, *These Women Entrepreneurs Created A Fake Male Cofounder to Dodge Startup Sexism*, FAST COMPANY (Aug. 29, 2017).

to face an even *more* unfriendly funding environment since the initial story began to break on Theranos and Elizabeth Holmes.[42]

Outside of funding dynamics, the oeuvre of empirically grounded research focused on the corporate governance of startups is substantially thinner. Part of the challenge, as alluded to above, is that many of the core governance documents of startup companies—such as board resolutions/minutes, shareholder agreements, and general books and records—are not publicly available, a limitation that hampers one's ability to produce even modestly powered empirical studies. Although certificates of incorporation (or "charters") and bylaws for *public* corporations can be found in scattered locations on the SEC's EDGAR website, they are poorly organized and exclude not-yet-public startups by definition.[43] In contrast, the charters of all companies (public and private) are publicly available *in principle* from the Secretary of State's office in the state of incorporation; but here too gaining access to them is surprisingly difficult, cumbersome, and expensive in practice, possibly as a byproduct of deliberate throttling (and perhaps some technological limitations) by government actors.[44] That dearth of documentation has left researchers, by and large, to rely on conceptual, institutional, and theoretical analyses (of which there are now many examples[45]), as well as a hodgepodge of surveys and experimental settings for insights about startup governance. A recent paper by Jennifer Fan, for example, makes use of interviews and surveys to conclude that much of startup governance is ad hoc and informal, making it difficult in general to study how governance interacts with gender.[46] In a different industry survey of more than 1,200 entrepreneurs across eight markets, researchers

[42] *See* Erin Griffith, *supra* note 9; Gené Teare, *Global VC Funding To Female Founders Dropped Dramatically This Year*, CRUNCHBASE (Dec. 21, 2020), https://news.crunchbase.com/venture/global-vc-funding-to-female-founders/ (tracking a decrease in funding to female-led startups and a decline in the "proportion of dollars to female-only founders").

[43] *See* Frankenreiter et al., *supra* note 23, at 23-24.

[44] *Id.*

[45] *See, e.g.*, Dorothy S. Lund & Elizabeth Pollman, *The Corporate Governance Machine,* 121 COLUM. L. REV. 2563 (2021); Jesse M. Fried & Mira Ganor, *Agency Costs of Venture Capitalist Control in Startups*, 81 N.Y.U. L. REV. 967 (2006); Elizabeth Pollman, *Startup Governance*, 168 U. PENN. L. REV. 155 (2019); D. Gordon Smith, *The Exit Structure of Venture Capital*, 53 UCLA L. REV. 315, 347-48 (2005); Sanga & Talley, *supra* note 25; Michael Klausner & Stephen Venuto, *Liquidation Rights and Incentive Misalignment in Start-up Financing*, 98 CORNELL L. REV. 1399 (2013).

[46] *See* Jennifer S. Fan, *The Landscape of Startup Corporate Governance in the Founder-Friendly Era,* 18 N.Y.U. J. L. & BUS. 317 (2022).

found that 35% of women entrepreneurs reported experiencing gender bias and, on average, raised 5% less capital than men.[47]

There do exist a handful of important and thought-provoking empirical analyses of select corporate governance features around a limited set of cash flow rights and board structure in VC-backed startups,[48] which are more closely related to our enterprise here. However, they by and large make use of modestly sized data sets. Broughman and Fried, for example, study a limited sample of fifty startups that successfully negotiated an exit (by acquisition), studying the incidence and renegotiation of liquidation rights between shareholder constituencies.[49] Amornsiripanitch, Gompers, and Xuan study board representation for VC financiers, finding that a prior common relationship with the founder's network predicts a higher probability of taking a board seat.[50] The empirical governance literature appears *exceptionally* thin in engaging the question of how gender interacts with governance, a fact that is not surprising given the relative recent interest in VC gender effects, not to mention poor access to data in the field generally.

All told, the existing literature in VC finance and law documents what appears to be a discernible funding shortfall for women entrepreneurs, but we have little purchase thus far on how (or even whether) this shortfall interacts with the governance of startups once they obtain funding. This constitutes a significant gap in our current

---

[47] *See* Stefan Wagstyl, *Female Entrepreneurs Face Gender Bias When Raising Capital*, FIN. TIMES (Sept. 30, 2019), https://www.ft.com/content/24689200-e141-11e9-9743-db5a370481bc.

[48] *See, e.g.*, Steven N. Kaplan & Per Strömberg, *Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts*, 70 REV. OF ECON. STUD., 70, 281; Paul A. Gompers, *Optimal Investment, Monitoring and the Staging of Venture Capital*, 50 J. OF FIN. 1461 (1995); Natee Amornsiripanitch, Paul A. Gompers & Yuhai Xuan, *More than Money: Venture Capitalists on Boards*, 35 J. L. ECON. & ORG. 513 (2019); William A. Sahlman, *The Structure and Governance of Venture-Capital Organizations*, 27 J. FIN. ECON. 473 (1990); Andrei Shleifer & Robert W. Vishny, *A Survey of Corporate Governance,* J. OF FIN. 737 (1997); Brian Broughman & Jesse Fried, *Renegotiation of Cash Flow Rights in the Sale of VC-Backed Firms*, 95 J. FIN. ECON. 384, 389 (2010); Erik Berglof, *A Control Theory of Venture Capital Finance*, 10 J. L. ECON. & ORG. 247 (1994); Duncan Davidson, *Venture 101: Participating Preferred*, BULLPEN CAPITAL (Apr. 6, 2011), http://bullpencap.com/2011/04/06/venture-101-participating-preferred/.

[49] *See* Broughman & Fried, *supra* note 48, at 389 (2010) ("[R]enegotiation is more likely when governance arrangements, including the firm's choice of corporate law, give common shareholders more power to impede the sale.").

[50] *See* Amornsiripanitch et al., *supra* note 48 ("[L]ead investor status, prior investor-founder relationship, geographical proximity, the venture capital firm's track record, and the size of the venture capital firm's network of outsider board members and managers are all positively correlated with board membership.").

knowledge, and one that bears directly on the overall heft of the funding gap. For example, the difficulties women face in attracting VC investments might produce a pendulum effect for those who successfully attract investments, whereby female recipients are especially "high quality" entrepreneurs who can command more generous treatment in corporate governance. In such a case, corporate governance structures would work to *dampen* gendered funding biases at the initial stage. Alternatively, it might be the case that female founders are disadvantaged twice over, first at the funding stage, and *then again* through a governance structure that confronts them with relatively unattractive rights against their VC investors. Here, the levers of firm governance would work to *exacerbate* initial gendered funding imbalances. The interaction thus remains an important yet under-analyzed question for empirical analysis within both law and finance. In the balance of this paper, we will start filling that gap.

II. ASSEMBLING CORPORATE GOVERNANCE DATA

A significant contribution of this Article is an original data set of VC-backed startups that we constructed over the course of two years, so that we might analyze whether gender differences predict differential allocations of formal governance rights for those who receive VC funding. For the sake of transparency and replicability, we describe the process in some detail below. At the highest level, however, our overarching strategy was *first* to build an inventory of female entrepreneurs and their associated startups, and *second* to identify a set of comparable male-founded startups that share key characteristics with each of to the female-founded startups prior to the first round of investment. We base all statistical comparisons on the resulting "matched sample" of female- to male-founded firms, as detailed below.

A. Data Sources

As noted in the Introduction, early-stage startups are often difficult to study empirically because they make few if any publicly available disclosures, and what few public disclosures they make are often difficult to access from governmental repositories. That said, several non-governmental organizations have begun in recent years to provide collections of informal and official data related to privately held startups. We use two such databases here: Pitchbook[51] and VC Experts

[51] *See* https://pitchbook.com/solutions/venture-capital.

"Genesis"[52], using them in combination with one another as described below.

First, we employed the Pitchbook Venture Capital database to identify VC-backed startups that had at least one female founder. Specifically, we drew from a compilation of over 6,000 female startup founders and their companies over the period 2003-2021. (We largely relied on Pitchbook's designation to identify female founders rather than devising our own.[53]) This left us with a large compilation of "female-founded" companies as designated by Pitchbook.[54]

While Pitchbook is an informative resource to track several demographic and financing variables for startups, it provides surprisingly little in the way of actual, granular governance details. We therefore hand match our Pitchbook list of startups with a second database—Genesis—which has particularly granular and detailed information about governance. Genesis itself tracks several transactional variables pertaining to capital-structure (such as liquidation preferences, anti-dilution rights, redemption rights, etc.). But just as important, Genesis also provides access to the historical record of corporate governance documents for each company, as filed with state authorities in the state of incorporation (typically—though not always—Delaware). We extracted both types of data from Genesis, deploying optical character recognition (OCR) technology to extract the textual content of the official certificates of incorporation (a.k.a., "charters"), and amendments thereto.

Because Genesis typically tracks the longitudinal record of firm filings, we are able to observe multiple rounds of financing, including a host of labeled data for each of those rounds. Because each round of funding requires amending the charter,[55] we also observe the evolution of charters (and all their governance provisions) over time.[56] By assessing these charter terms and amendments, we can make inferences about the

---

[52] *See* https://lanyaplabs.com/.

[53] We nevertheless also double checked our designations once the fully matched sample was formulated (where we did uncover—and correct—misclassifications by Pitchbook). *See infra* note 63.

[54] In our baseline analysis, we emulate Pitchbook's methodology and deem a startup to be "female-founded" if it has at least one female member of its founding team. *See infra* note 63.

[55] *See, e.g.*, Del. Gen. Corp. Law §§ 102(a)(4); 151.

[56] We do not directly observe actual bargaining between investors and founders. These bargaining/pitching sessions are thought to be critical within the VC industry. *See* Hassan et al., *supra* note 35. Thus, we consider the round-by-round corporate charter provisions and contemporaneous transactional information recorded in Genesis to reflect the output of those negotiations.

state and evolution of corporate governance structures at various funding stages of private corporations.

In our baseline analysis, we deem a firm to be "female-founded" if at least one of the founders is female. This means that a company with several male founders and only one female founder would still be classified as "female-founded" in our baseline rubric. We follow Pitchbook's lead in defining female-founded firm this way for several reasons. First, prior research shows that founding teams with only one female member can receive different treatment.[57] Second, the small number of startups with only female founders poses challenges for meaningful analysis. Even with our broader definition of female-founded firms we have identified only 4.2% of the companies in the Genesis database as being female-founded. Data from recent years have shown that the percentage of startups with only female founders is about half of the percentage of startups with a mix of female and male founders.[58] Nevertheless, we later report on robustness checks that consider alternative definitions (with more modest sample sizes), including *majority* female founding teams and *exclusively* female-founded startups.[59]

B. Matching Formulation

It is impossible to analyze meaningfully the governance or finance traits of female-founded firms without a comparison set of male-founded firms. Conducting such a comparison can be tricky, however, because female founders are decidedly *not* a representative sample of startup entrepreneurs, and thus an unalloyed comparison of the two groups would be difficult to interpret. We therefore attempt to match our female-founded firms with other deemed male-founded firms that share similar observable characteristics. That said, one must further take care that the matching criteria are themselves independent from the outcomes of funding (such as financial and governance terms). We therefore focus our matching exercise on information that was available only up until the first recorded VC investment in the startup. We identify other firms similar to our female-founded firms (our deemed male-founded firms) using information on the first recorded funding date, the

[57] *See, e.g.*, Christopher Cassion, Yuhang Qian, Constant Bossou & Margareta Ackerman, *Investors Embrace Gender Diversity, Not Female CEOs: The Role of Gender in Startup Fundraising*, 377 Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (2021).
[58] *See, e.g.*, Gene Teare, *Global VC Funding To Female Founders Dropped Dramatically This Year*, CRUNCHBASE NEWS (Dec. 21, 2020).
[59] Our results do not substantially change along most dimensions. *See infra* Part IV.

round/series of that funding,[60] geographical region, and industry group. We build a statistical propensity score for matching our female-founded firms to other firms, which we refer to as "male-founded" startups for convenience.[61] We then match each female-founded startup with its three closest neighbors in propensity-score space.[62] Our 3-to-1 sampling and matching protocol resulted in 257 distinct female-founded startups and 771 matched male-founded firms, 620 of which are unique, per our strategy of sampling with replacement.[63] The results reported below focus largely on these two groups of female-founded firms and matched male-founded firms, and often only consider the governance terms identified in the first observed round of investment reported in Genesis and not later rounds of investment.

A standard way to assess the success of a matching protocol is to analyze covariate balance across the variables used for the match. Table 1 reports the percentage of companies in each region, industry sector, and round for our female-founded startups (first column), our *matched* male-founded firms (second column) and *all* male-founded firms (fourth column). The third column reports the balance of the covariates for female-founded firms and matched male-founded firms by reporting the standardized mean differences (SMD) between the two samples. Over 95% of the covariates have a SMD below the accepted threshold of 0.1, although a two have covariates slightly exceeding 0.1—the industry sector of "Computers and Peripherals" with only a few firms and

[60] For our matching protocol we use the first recorded funding deal in Genesis, which may not be the true first funding deal for the firm. Table 2 provides details on the breakdown of the funding series of the first recorded deal in Genesis.

[61] To construct this score, we employ a LASSO-based scoring methodology to estimate the probability of a startup being "female-founded" based on pre-investment observable characteristics. While traditional propensity score approaches often rely on logistic regression, we opted for a LASSO prediction due to the presence of numerous pre-treatment categorical covariates, making LASSO a more suitable prediction technique in this context.

[62] For each new firm, we replace all prior matches into the pool, so that in some cases two (or more) female-founded startups may be matched with common male-founded counterparts.

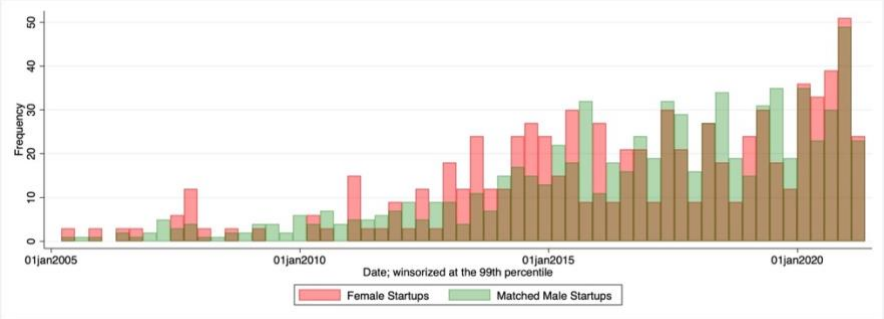
[63] As flagged above in note 53, we manually verified the gender classification for each of our female-founded firms and deemed male-founded firms, and this process identified misclassifications of several firms. Specifically, we found that Pitchbook had wrongly classified several firms as female-founded when hand inspection revealed no female founders (20 firms). These firms were dropped from the set of female-founded companies. We also hand inspected our deemed male-founded match set and found some that in fact had female founders on the team (9 firms). For these misclassified matches, we provided new male-founded matches for female-founded firms that lost a male counterpart by choosing their next nearest coded match among male-founded companies.

"Healthcare Services." We assess that the matching protocol on region, industry and round was implemented successfully.

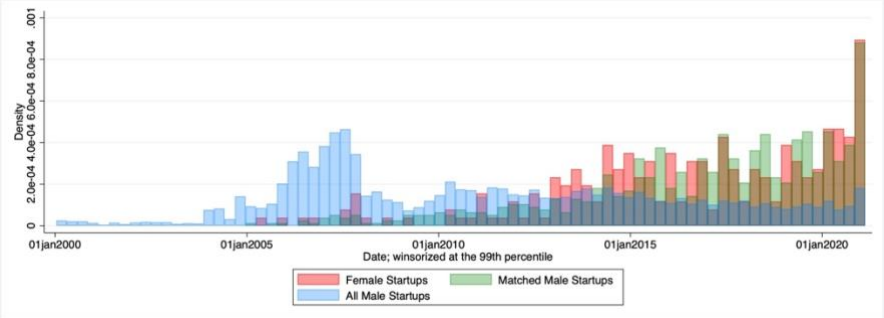
To compare the first funding date, an additional variable used for matching, we plot the histogram of the dates for our female-founded firms and matched male-founded firms in Panel (a) of Figure 1. This plot weights male-founded firms according to our 3-to-1 matching with female-founded firms. The histogram demonstrates a high overlap in the timing of the first funding date for our matched firms, with most companies receiving their first round of funding between 2012 and 2021. Both female-founded and matched male-founded firms generally increase in frequency as time passes, with the sole exception of 2021, where we ceased sampling mid-year. The overlap in the distribution for female-founded and matched male-founded firms further supports the successful implementation of the matching protocol.

Figure 1: First Observed Deal Date

(a) Female-Founded Startups as compared to Matched Male-Founded Startups (3:1)



(b) Female-Founded Startups, Matched Male-Founded Startups, and All Male-Founded Startups



Panel (b) of Figure 1 adds the distribution of all male-founded firms in the Genesis database. The distribution of all male-founded firms (in blue) is skewed to the left meaning that the full sample of male-founded firms is more likely to have received their first round of funding in earlier years relative to the matched male-founded firms (in green) and female-founded firms (in red). This distributional skew further supports our decision to perform most of our analysis comparing female-founded startups to the matched male-founded sample (see Table 1).

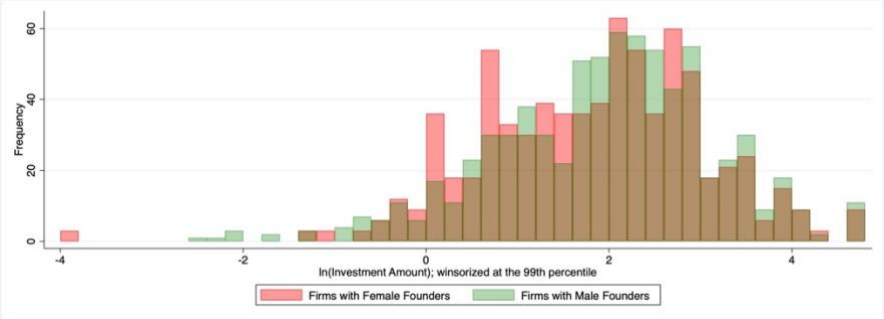
Table 1: Covariate Comparisons

	(1) Female Startups	(2) Matched Male Startups	(3) Standardized Mean Difference (SMD)	(4) All Male Startups
Region				
Alaska/Hawaii/Puerto Rico	0.00%	0.00%		0.05%
Canada	0.00%	0.00%		0.05%
Colorado	1.00%	0.80%	0.0212	1.11%
DC/Metroplex	0.90%	1.60%	0.0630	2.33%
India	0.00%	0.00%		0.01%
Midwest	1.20%	0.80%	0.0402	5.24%
New England	6.00%	6.60%	0.0247	10.13%
New York State	6.80%	7.00%	0.0079	7.12%
Northwest	1.90%	1.20%	0.0567	2.53%
Other – World	0.10%	0.00%	0.0447	0.03%
Philadelphia Metro	0.10%	0.00%	0.0447	0.92%
Sacramento/Northern California	1.60%	0.80%	0.0735	0.63%
Silicon Valley	72.80%	73.90%	0.0249	54.10%
Southeast	0.80%	0.40%	0.0518	3.69%
Southern California	3.60%	4.70%	0.0552	7.23%
Southwest	1.80%	2.30%	0.0353	4.40%
United Kingdom	0.20%	0.00%	0.0633	0.03%
Sector				
Advanced_Special_Materials_Chemi	0.00%	0.00%		0.40%
Biotechnology	15.90%	17.50%	0.0429	13.78%
Business Products and Services	6.70%	6.60%	0.0040	7.97%
Computers and Peripherals	0.50%	0.00%	0.1003	1.60%
Consumer Products and Services	15.60%	14.80%	0.0223	6.64%
Electronics Instrumentation	4.70%	5.10%	0.0185	3.94%
Financial Services	5.00%	5.10%	0.0046	3.97%
Healthcare Services	5.10%	8.20%	0.1247	3.02%
IT Services	10.80%	9.70%	0.0363	16.50%
Industrial Energy	2.30%	3.50%	0.0716	4.16%
Media and Entertainment	2.40%	3.90%	0.0860	7.29%
Medical Devices and Equipment	6.40%	4.70%	0.0743	8.07%
Networking and Equipment	0.40%	0.40%	0.0000	1.62%
Other	3.80%	2.30%	0.0873	0.68%
Retailing ~n	5.10%	6.20%	0.0477	1.82%
Semiconductors	0.20%	0.00%	0.0633	3.68%
Software	49.00%	48.20%	0.0160	42.87%
Telecommunications	1.20%	2.30%	0.0840	3.97%
Round				
Seed	27.40%	27.20%	0.0045	11.36%
Series A	46.90%	47.90%	0.0200	49.45%
Series B	18.60%	18.30%	0.0077	20.64%
Series C	3.60%	4.70%	0.0552	10.87%
Series D	2.00%	1.20%	0.0638	4.16%
Series E or Greater	1.40%	0.80%	0.0575	3.53%

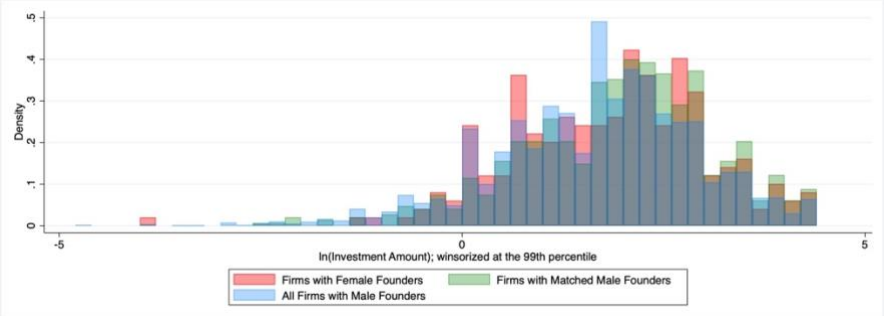
While generally satisfactory, our matching approach has certain limitations. Several of our matching variables are fairly coarse and it is possible that they mask material forms of unobserved heterogeneity. For example, the firm sector variable includes broad categories so that sectors, like "Software," may cover many very different types of products mean for different markets. It is also important to recognize that we have very few variables that reflect information up until the first recorded investment. Because gender is not randomly assigned, there may be many further unobservable variables that reflect real differences between our female-founded firms and matched male-founded firms. Lastly, even the few variables we use in our matching protocol, such as region and sector, may be endogenously selected by founders and so these variables could themselves be impacted by gender and correlate with other unobserved differences.

Figure 2: Amount Invested

(a) Female-Founded Startups as compared to Matched Male-Founded Startups (Natural Logarithm)



(b) Female-Founded Startups, Matched Male-Founded Startups, and All Male-Founded Startups (Natural Logarithm)



C. Summary Characteristics

By limiting our comparison of female-founded firms to matched male-founded firms, rather than all male-founded firms, we are able to compare firms that are more similar on dimensions other than the variables that are the basis of our propensity scores. We consider, for example, the investment amount in the first observed investment round—an outcome variable that was not used for matching firms—in Figure 2. Panel (a) of Figure 2 compares the amount invested, expressed in natural logs to remove skew, in female-founded firms to the matched male-founded sample. Panel (b) of Figure 2 includes the distribution of investment amount for all male-founded firms in blue. Here too we observe that the female-founded sample is more similar to the matched male-founded sample along the dimension of investment amount than to the full sample of male-founded firms.

Table 2: Summary Statistics

	(1) Female Startups	(2) Matched Male Startups	(3) All Male Startups
Year of the First Recorded Deal			
Median	2016	2017	2010
Mean	2016.16	2016.39	2010.80
Std. Dev.	3.62	3.47	4.83
Min.	2005	2005	2000
Max.	2021	2021	2021
Year of the Last Recorded Deal			
Median	2019	2018	2013
Mean	2017.86	2017.61	2012.99
Std. Dev.	3.00	2.98	4.90
Min.	2005	2005	2000
Max.	2021	2021	2021
First Observed Round Valuation (mil.)			
Median	25.42	27.73	21.76
Mean	49.93	81.77	51.98
Std. Dev.	88.75	533.45	201.01
Min.	0.00	0.00	0.00
Max.	938.80	14360.99	14360.99
Final Observed Round Valuation (mil.)			
Median	55.94	50.21	48.95
Mean	190.08	429.32	272.50
Std. Dev.	462.35	2700.08	1648.25
Min.	0.00	0.00	0.00
Max.	4362.38	33916.59	75188.23
First Observed Round Valuation (ln)			
Median	3.24	3.33	3.11
Mean	3.22	3.40	3.14
Std. Dev.	1.15	1.18	1.23
Min.	-1.36	0.54	-10.08
Max.	6.84	9.57	9.57
Final Observed Round Valuation (ln)			
Median	4.03	3.94	3.93
Mean	4.13	4.09	4.04
Std. Dev.	1.39	1.53	1.53
Min.	1.10	0.54	-3.02
Max.	8.38	10.43	11.23

Table 2 summarizes descriptive statistics comparing our female-founded firms, the matched male-founded firms, and the full sample of male-founded firms. Table 2 shows that the year of the first observed round of investment is similar for female-founded firms and matched male-founded firms, as reflected in Figure 1. When comparing the investment amount for the first observed round, the median investment amount is similar for female-founded firms and matched male-founded firms, but the distribution is skewed to the right for the matched male-founded firms who have a higher mean investment amount. These differences are attenuated when consider the natural log of investment amount. Table 2 also compares the investment amount of the last recorded investment round that we have in our data set. Again, the

median investment amount in the last recorded investment round is similar for female-founded firms and matched male-founded firms, but the distribution is skewed to the right for those matched male-founded firms who have a higher mean investment amount.

III. CHARTER ARCHITECTURE

We begin our analysis from a relatively high altitude, investigating differences in the broad textual and semantic content of the charters in our sample. Like the later parts of our analysis, this investigation exploits the fact that the startup charters embody the most central and durable governance rights that investors bargain for in their negotiations with the founders and previous investors. They additionally ordain the rights enjoyed by the holders of the various series of preferred shares issued in subsequent financing rounds. The charter of a VC-backed startup also usually contains the provisions that govern the payment waterfall in the event of an exit. If gender effects manifest in the internal governance systems of startups, one would expect differences in the content and style of charters, too.

To conduct this inquiry, we deploy several tools from machine learning and computational text analysis. These techniques have recently seen an upswing in legal research,[64] and a burgeoning literature in

[64] *See, e.g.*, Jens Frankenreiter & Michael A. Livermore, *Computational Methods in Legal Analysis*, 16 ANN. REV. L. & SOC. SCI. 39, 40 (2020) ("Techniques from the fields of artificial intelligence, natural language processing, text mining, network analysis, and machine learning are now routinely taken up by legal practitioners and law scholars."); Kellen Funk & Lincoln A. Mullen, *The Spine of American Law: Digital Text Analysis and U.S. Legal Practice*, 123 AM. HIST. REV. 132, 136 (2018) (arguing in favor of combining computational text analysis with traditional historical research techniques); Michael A. Livermore, Allen B. Riddell & Daniel N. Rockmore, *The Supreme Court and the Judicial Genre*, 59 ARIZ. L. REV. 837, 840 (2017) (exploring differences in the writing styles of U.S. Supreme Court Justices through the lens of computational techniques); Jonathan Macey & Joshua Mitts, *Finding Order in the Morass: The Three Real Justifications for Piercing the Corporate Veil*, 100 CORNELL L. REV. 99, 103 (2014) (using computational techniques to develop a classification of piercing the corporate veil cases); Marian Moszoro, Pablo T. Spiller & Sebastian Stolorz, *Rigidity of Public Contracts*, 13 J. EMPIRICAL LEGAL STUD. 396, 396 (2016) (applying quantitative techniques to investigate contracts in regulated industries); Julian Nyarko, *Stickiness and Incomplete Contracts*, 88 U. CHI. L. REV. 1, 30 (2021) (using supervised machine learning techniques to determine the existence of dispute resolution clauses in contracts); David E. Pozen, Eric L. Talley & Julian Nyarko, *A Computational Analysis of Constitutional Polarization*, 105 CORNELL L. REV. 1, 3-4 (2019) (using supervised machine to investigate the growth of polarization in constitutional debate); Eric L. Talley, *Is the Future of Law a Driverless Car?: Assessing How the Data-Analytics Revolution Will Transform Legal Practice*, 174 J. INSTITUTIONAL & THEORETICAL ECON. 183, 184 (2018) ("Although quantitative analysis of law (also

business law has begun to use them to great effect.[65] At their broadest level, machine learning approaches treat written texts as data,[66] converting documents into numerical representations, and thereby making them amenable to various types of statistical analysis.

We apply two different approaches from the field of computational text analysis to study our corpus. First, we obtain a range of easily obtainable metrics that capture basic characteristics of startup charters, including their length and their lexical variability.[67] We then use these and other metrics to compare across charters, measuring whether there are substantial differences between the charters of female- and male-founded firms.

Second, we employ a so-called *bag-of-words* approach that captures aspects of the semantic content of a document. Bag-of-word techniques provide numerical representations of the vocabulary used in a corpus of documents. More specifically, they convert documents—in our case, corporate charters—into high-dimensional vectors whose individual elements depict whether a particular word is featured in a document or

called empirical legal studies) is nothing new, textual analysis methods have become significantly more powerful over the last half decade."); Eric Talley & Drew O'Kane, *The Measure of a MAC: A Machine-Learning Protocol for Analyzing Force Majeure Clauses in M&A Agreements*, 168 J. INSTITUTIONAL & THEORETICAL ECON. 181, 183 (2012) (discussing ways to use quantitative techniques to improve our understanding of MAC provisions in M&A agreements).

[65] Probably the most similar analysis to the one presented in this paper can be found in Frankenreiter et al., *supra* note 23, who investigate the semantic content of the charters of publicly traded firms using similar techniques. Aside from this paper, there is little work that would use these tools to investigate corporate governance documents. However, various authors have used these tools with other types of documents, including financial disclosures and credit agreements. *See* Adam B. Badawi, Scott D. Dyreng, Elisabeth de Fontenay & Robert W. Hills, *Contractual Complexity in Debt Agreements: The Case of EBITDA* (Duke L. Sch. Pub. L. & Legal Theory Series, Paper No. 2019-67), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3455497# (using machine learning techniques to analyze EBITDA definitions in credit agreements); Rauterberg & Talley, *supra* note 29, at 1078 (2017) (building a targeted corpus of corporate opportunity waivers from public filings); Elvis Hernandez-Perdomo, Yilmaz Guney & Claudio M. Rocco, *A Reliability Model for Assessing Corporate Governance Using Machine Learning Techniques*, 185 Reliability Eng'g & Sys. Safety 220, 222 (2019) (marshaling select financial disclosure items related to corporate governance to assess "systems failure" in firms); Ryan Bubb & Emiliano Catan, *The Party Structure of Mutual Funds* 1-2 (Eur. Corp. Governance Inst., Law Working Paper No. 560, 2020), https://ssrn.com/abstract=3124039 (using machine learning techniques to study mutual fund voting patterns).

[66] *See also* Frankenreiter et al., *supra* note 23.

[67] *See also* Frankenreiter & Livermore, *supra* note 64, at 43 (giving examples of research using similar metrics).

not.[68] Because of the high-dimensional nature of data generated by means of bag-of-word approaches, it is impossible to just compare them like one can compare, say, the length of different documents. Therefore, it has become common to further analyze such data using machine learning techniques.[69]

From the myriad machine learning tools that can be applied to such data, we concentrate on three: We apply *unsupervised machine learning* techniques to obtain a two-dimensional representation of the charters in our corpus, which we use to determine whether there are clusters of charters that consist primarily of either male-founded or female-founded firms. We also calculate *similarity scores* between the charters in our data set to determine whether the charters of male-founded firms are, on average, more similar to each other than they are to female-founded firms. Finally, we use *supervised machine learning* to ascertain whether it is possible to successfully predict whether a startup was founded by a female founder from the vocabulary used in its charter.

While our corpus contains the full chartering histories of all the startups we include in our data set, the analysis below (as well as the following analyses) focuses exclusively on the first charter that a startup adopts after receiving its first round of VC funding. There are several reasons for this restriction. First, including more than one charter per company would render the analysis substantially more complex than it already is. Second, because our data set includes companies founded over nearly a twenty-year span, the length of the chartering histories available to us varies greatly between different companies, raising tricky questions about how to compare an older company after various rounds of investments with a younger company that has gone through one financing round only. Finally, to the extent that the evolution of a startup is in part the path-dependent byproduct of its initial governance regime, the content of later charters (when they exist) is even more challenging to interpret.

A. Simple Document Metrics

We begin with a set of high-level comparisons of charter contents, distinguishing our female-founded firms from the male-founded

[68] More precisely, we obtain "binary term-frequency/inverse-document-frequency" representations for the documents in our corpus. Before obtaining these representations, we apply familiar pre-processing steps including stopword removal and stemming (using the so-called Snowball Stemmer). For more details on these techniques, *see* Frankenreiter & Livermore, *supra* note 64.
[69] *Id.*

matches. The three panels of Figure 3 display common document-level metrics for our chartering corpus on an annual basis, depending on the year in which the company first received VC funding.

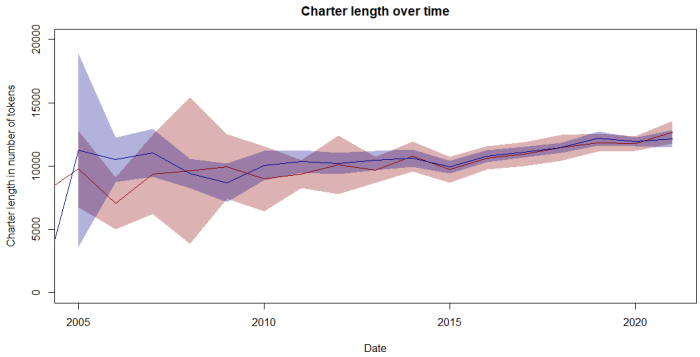
Panel (a) illustrates the mean length of charters (in words). Panel (b) illustrates the "readability" of charters in each year, as measured by the well-known Flesch-Kincaid (F-K) scale. Originally developed to gauge the content of mechanical instructional manuals, F-K scores are calculated on the basis of the average length of words and sentences in a document. The score proxies proportionally to readability, so that higher scores denote greater readability. An F-K score below 10.0 is considered to be the most challenging, appropriate to a professional trained in the field.[70] Panel (c) tracks lexical variation in the form of "Type-Token Ratios" (TTRs) of startups' charters by year. The TTR is a common metric that represents the ratio of unique terms divided by the total number of words in the document. This metric helps researchers understand a document's repetitiveness and redundancy: As the TTR shrinks, the document becomes more repetitive (i.e., has less lexical variation).[71]

[70] For more details, *see* Frankenreiter et al., *supra* note 23.
[71] *Id.*

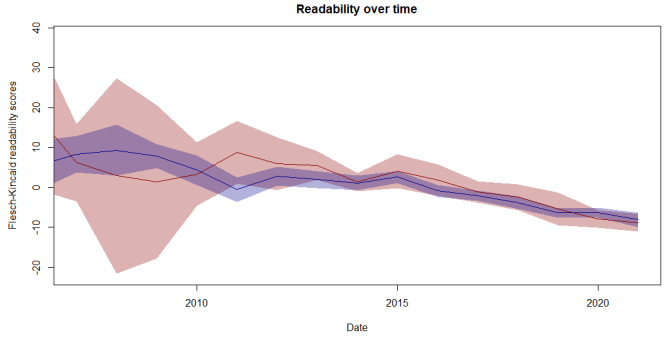
Figure 3: Charter Contents

Red = female-founded mean; Blue=matched male-founded mean; Shaded Regions = 95% confidence intervals.

(a) Length (in words)



(b) Readability (F-K Readability Score)



(c) Lexical variation (Type-Token Ratio)

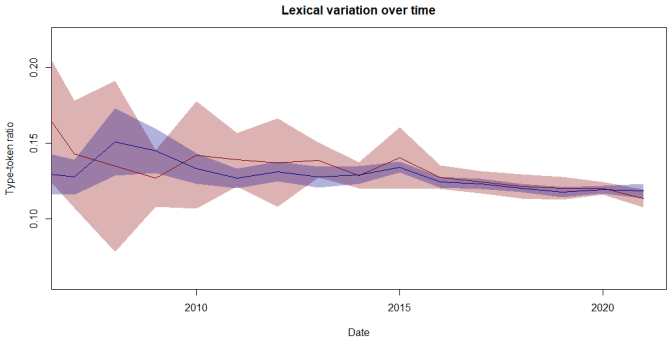


Figure 3 illustrates several notable trends. Charters tend to become longer over time (Panel (a)). However, there seems to be little difference between female- and male-founded firms: Although male-founded charters appear lengthier than female-founded charters initially, the difference remains noisy and converges over time. Panel (b) reflects a

similar trend for readability. Charters' readability score decreases over time, with more recent first charters being less readable. Here too, point estimates suggest that male-founded firm charters start off as less readable than female-founded firm charters and converge over time, although estimates are somewhat noisy. Lexical variation, in Panel (c), moderately decreases over time, with overall variation between female-founded firms and male-founded firms, also decreasing over time.

B. Low Dimensional Representation of Charters

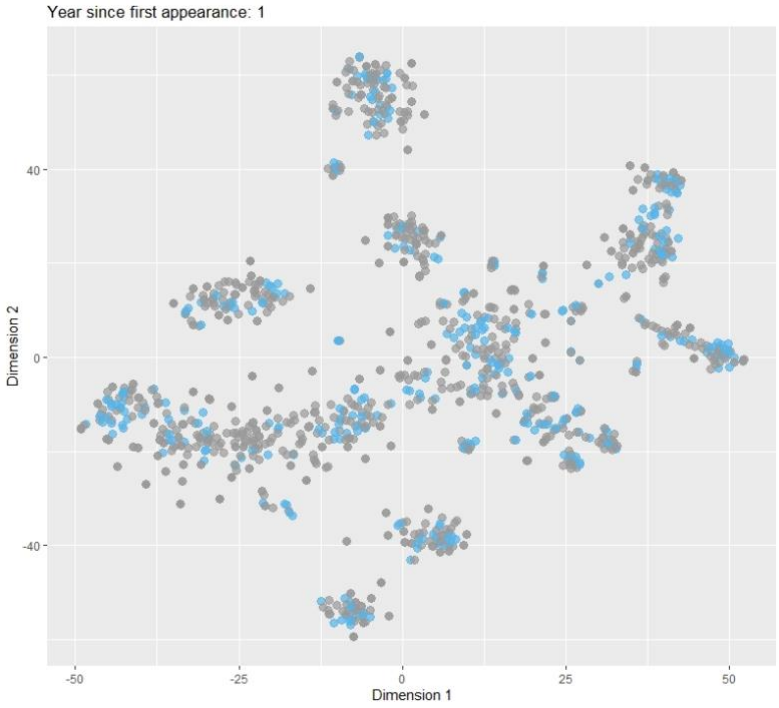
As described above, the bag-of-words approach we deploy to represent charters converts them into high-dimensional numerical vectors. Due to their high-dimensional nature, it is impossible to display these vector representations in a single two-dimensional graph. However, various computational techniques allow for the "mapping" of such vectors into reduced dimensional space. Applying those techniques here allows us to represent each document as a point in two-dimensional component space, with their spatial proximity providing a visual proxy for similarity of the documents: Those that use similar vocabulary tend to cluster closely together, while those that use differentiated words tend to be displayed in different parts of the plot.

Figure 4 depicts a syntactical representation of all initial charters in our matched data set, color-coded by the existence of a female founder at the startup.[72] Light blue dots represent female-founded firms and gray dots represent male-founded firms. The wide dispersion of the scatter field is indicative of substantial variation in the contents of the charters within our data set. At the same time, this graph does not suggest that there are systematic differences between both types of firms. Rather, female-founded firms seem to be represented in all major clusters that appear in the graph, and there also appears to be no cluster in which female-founded firms are overly represented.[73]

[72] In order to obtain two-dimensional representations, we proceed in two steps. First, we reduce the dimensionality of our dummy TF-IDF vectors to 50 using the SVD algorithm. Second, we use the T-SNE algorithm to generate two-dimensional representations.

[73] To validate these results, we divide firms into clusters using the k-means clustering algorithm and test whether female-founded firms are unequally distributed across clusters. These tests do not suggest that there are systematic differences between female-founded and male-founded firms.

Figure 4: Two dimensional charter representations



C. Similarity Comparisons

In a next step, we address the question whether the semantic content of female-founded firms differs in measurable ways from those of male-founded firms more systematically. For this, we compute the cosine similarity[74] between all initial charters in our data set and determine whether the differences between female-founded firms and male-founded firms are more pronounced than differences between male-founded firms.

To survey all relevant comparisons, we assess each inter-firm permutation afforded by our three-to-one matching protocol. Figure 5 offers a conceptual illustration. For each female-founded firm (denoted "F"), our protocol generates three matched male-founded firms ("M1," "M2," and "M3"). Within each matched 4-tuple, we first generate cosine similarity scores between the female-founded firm and each of the three male-founded matches (F v. M1, F v. M2, and F v. M3), represented by the black dashed arrows in the Figure. We then generate analogous

[74] The cosine similarity score is bounded between a minimum of 0.0 and maximum of 1.0, and it measures the cosine of the angle between the vector representations of two documents. *See* Pozen et al., *supra* note 64, at 34 (2019).

measures for the three remaining permutations of the male-founded match firms (M1 v. M2, M1 v. M3, and M2 v. M3), represented by gray dashed arrows. Thus, each matched 4-tuple allows us to extract three male-female "treatment" comparisons and three male-male "control" comparisons. We aggregate all six comparisons across every 4-tuple in our data set, resulting in a population distribution for matched-firm similarities.

Figure 5: Permutations of Female-Founded Startups & Male-Founded Matches (1x3 match)

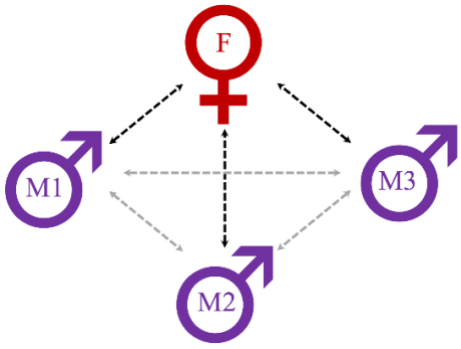
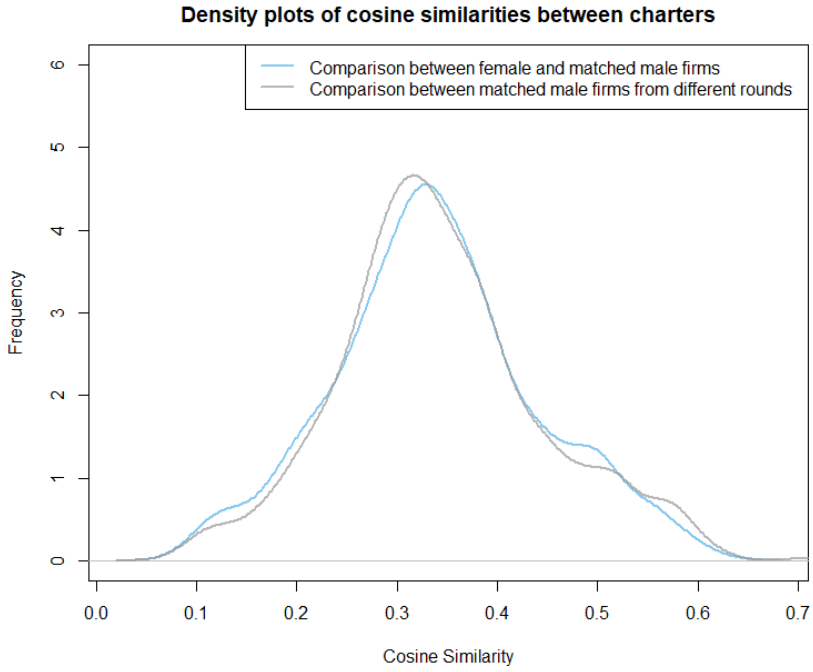


Figure 6 plots the smoothed population histogram of cosine similarities by comparison group, based on the content of the firm's first observable charter. The blue curve represents the density of similarities between female-founded firms and matched male-founded firms, while the gray curve represents the density of the analogous scores only among male-founded firms in our sample. Both curves manifest significant heterogeneity in syntactic variation between charters. In other words, while startup charters may all be long, complicated, and lexically boring, they tend to do those things in different ways. In addition, however, note that for the female-male distribution, there appears to be slightly more weight on the lower end of the distribution, suggesting that as a whole, the charters of female-founded startups diverge from their male-founded counterparts more than those counterparts differ from one another. This eyeballing impression is also borne out in numbers: The average cosine similarity between female-founded and male-founded companies is slightly higher than the average cosine similarity between different male-founded companies. However, this difference does not appear as statistically significant in standard statistical tests: We can neither reject the null hypothesis of no difference in means under a two-

sample t-test (t = -0.7176; p-value = 0.4731) nor under a Wilcoxon rank sum test with continuity correction (W = 294389, p-value = 0. 7461).[75]

Figure 6: Kernel Density of Cosine Similarity

(Female-Male versus Male-Male subpopulations)



D. Predicting Founder Type from Textual Content of Charter

Finally, we make use of the vector representations obtained above in an indirect test for whether the contents of charters of firms led by female founders are "predictably" different from the contents of charters of matched male-founded startups. Similar to the content analysis literature on polarized political partisanship,[76] we train several machine learning classifiers to determine whether (and how well) a calibrated algorithm is able to predict the founder type solely on the basis of the vocabulary used in a charter. The ease with which an algorithm can make this prediction can be thought of as a proxy for how "gender-specific" the semantic structure of the charter is. Using 10-fold cross validation to evaluate the relative performance of different algorithms in

[75] Similarly, a two-sample Kolmogorov-Smirnov test does not reject the hypothesis that the similarity scores for female-male comparisons and male-male comparisons are drawn from the same distribution (p-value: 0.8124).
[76] *See* Pozen et al., *supra* note 64, at 34 (2019).

various configurations, we settle on random forests using the first 880 principal components of the document-level feature vectors. Figure 7 depicts a receiver operating characteristic (or "ROC") curve, which embodies information on the overall performance of this algorithm. An ROC curve for a *random* classifier (i.e., one that is unable to obtain any meaningful information on the founder type) would be expected to lie near the black diagonal line from the lower left to the upper right. A highly predictive algorithm would be highly concave, bending far to the upper right, indicating that it was able to correctly predict the female-founded firms in the data set without also (incorrectly) flagging many male-founded firms as female-founded. Standard machine learning diagnostics often compute the area under the curve of the ROC (or "ROC-AUC") as a prediction diagnostic, which will range between 0.5 (essentially random) to 1.0 (maximally predictive).

Figure 7: ROC Curve (Full Matched Sample)

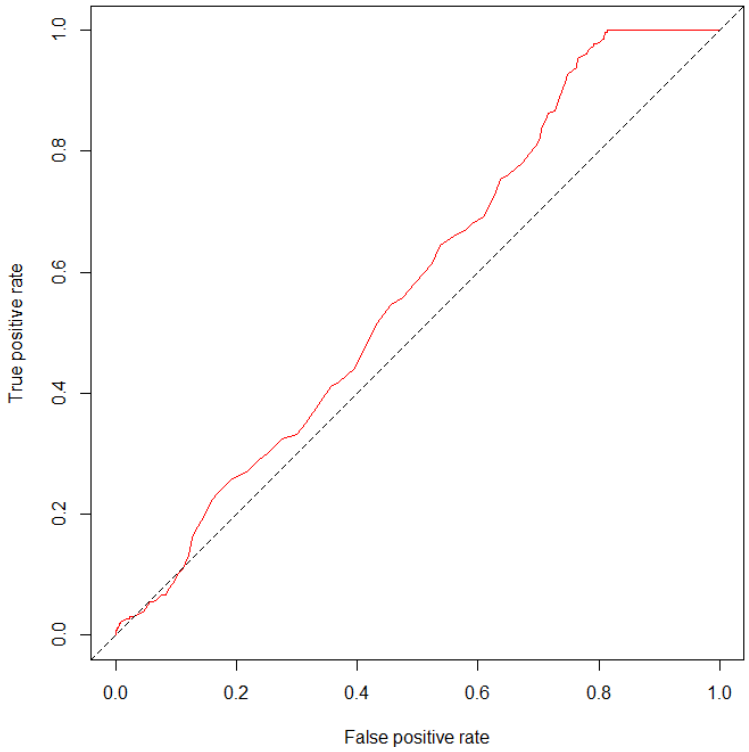
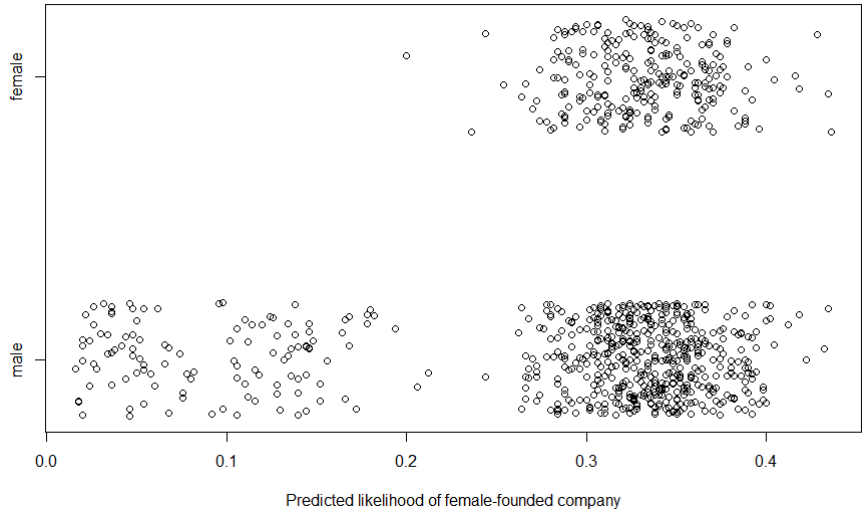


Figure 7 suggests that our algorithm performs acceptably well overall. A particularly noteworthy observation is that the algorithm reaches an almost perfect true positive rate at a false positive rate of around 0.8. This suggests that there are at least *some* male-founded firms

in our sample that are sufficiently different from female-founded firms for the algorithm to clearly distinguish between the two groups. But not vice versa: The ROC line's moderate slope in the lower parts of the graph indicates that there is no group of female-founded firms whose charters make it easy for the algorithm to clearly distinguish them from *all* male-founded firms. The finding that our algorithm predicts passably well is also borne out by standard numerical measures of predictive accuracy: We obtain an overall predictive accuracy of .709, an F-1 value of .504, and an AUC (area under the curve) value of .578.

Figure 8 provides another perspective on the main finding described above. This figure depicts, on the x-axis, the predicted likelihood of a company being female-founded, based on the vocabulary used in its charter. On the y-axis, the figure depicts the true "label" associated with this company, i.e., whether the company is female-founded or not. It can be seen again that the algorithm is mostly unable to correctly identify female-founded companies without also falsely labeling at least some male-founded companies as female-founded. However, the algorithm is able to correctly identify a set of male-founded companies without ever confusing them with female-founded firms. This finding suggests that there is a subset of male-founded firms in our data set whose charters differ substantially from those of female-founded firms.

Figure 8: Predictions vs. Labels



Of course, this finding on its own does not shed significant light on the *precise sources* of the governance differences between female- and male-founded firms. Are these differences in the wording only (and thus epiphenomenal to governance), or are these differences reflective of a

differential treatment of female founders at the level of governance terms or financial provisions that investors reserve for themselves when deciding whether to invest in the company? Without a closer understanding of the substantive content of the charters, this question cannot be answered. It is to this set of questions we next turn, comparing charters along an enumerated set of financial and nonfinancial provisions that frequently garner attention from practitioners, commentators, judges, and other relevant actors.

IV. INDIVIDUAL LEVERS OF VC GOVERNANCE

Our analysis in Part III above is at least suggestive that there are measurable differences between female- and male-founded firms, which merit further exploration by considering more granular information about specific cash-flow and control rights. In this Part, we attend to that task, offering a series of comparisons between female-founded and male-founded firms across a variety of individual financial and governance provisions as of the first funding deal, which typically reflects VC funding. We start by describing the sample of firms for this analysis, and we then discuss how we coded (or "labeled") each provision for analysis. As described below, some of the provision labels come from the meta data on the Genesis database. Other labels were created by a team of research assistants that manually coded the provisions from the text of the corporate charters. We then use the meta data and hand-coded labels to compare female-founded and male-founded firms, starting with the financial provisions and then discussing non-financial, governance provisions.

A. Subsample Description

In the analysis below we focus on the financial and control rights provisions of the first round of funding. New rounds of funding typically create new classes of preferred shareholders with their own set of financial and governance rights, possibly leading to a change in the rights of founders and legacy shareholders. The rights of preferred shareholders can constrain the startup's ability to raise capital in subsequent rounds and can affect the financial and governance rights determined in later investment rounds. For this reason, we choose to focus on the rights of preferred shareholders determined in the first round of investment, thereby attenuating the path dependency impact of earlier investment round terms on the investment terms of multiple successive rounds. We treat the first *observed* round of funding reported in Genesis as the first round of funding for our analysis, although it is

possible that for some companies with partial records in Genesis there may be an earlier investment round that we are unable to observe.[77]

B. Coding Description and Rubric

To examine funding differences and differential allocation of formal governance rights for female-founded startups and male-founded startups, we use two sources of labeled information. Our first source is investment deal information labeled by the Genesis database. The Genesis meta data tends to concentrate on financial (as opposed to governance) provisions. Our second source is the hand-coded labels created from the text of the corporate charter filed after an investment round, which concentrates greater weight on governance. We describe both sources in detail below.

Genesis: The Genesis database contains a record of investment deals for each firm and includes metadata on several provisions of those deals. The labeled financial provisions tracked by Genesis include the firm's valuation and cash-flow rights of the preferred shareholders such as the liquidation preference, liquidation multiple, cumulative dividends, pay-to-play provision incidence, and option pool. This metadata provides some important information about the cash-flow rights of investors; however, some of the information is coarse and it does not contain any information about governance and control rights.

Hand-Coded Fields: To supplement the Genesis meta data with governance rights and more granular financial information, we hand-coded the provisions found in a firm's charters. We do not directly observe side agreements between investors and founders. That said, the corporate charter, which we do observe, is required by law to reflect most important financial and governance rights, so we consider the corporate charter provisions as a trustworthy reflection the allocation of governance rights and financial rights that investors and founders have agreed upon. Charter amendments frequently coincide with new rounds of financing, as state law typically requires a charter amendment when new stock is created for a round of financing. By selecting the corporate charter that is filed with Secretary of State's office in the state of incorporation at a date closest to the first deal date (documented in Genesis), we expect the provisions in that charter to reflect the most recent investment deal.[78]

[77] As discussed in Part III.C, over 70% of companies in our sample have either "Seed" or "Series A" round recorded as their first investment deal.

[78] For most companies we are able to identify a new corporate charter filed within days of the investment deal date. In some cases, we use the corporate charter filed just before the documented investment date. When there is no corporate charter filed

The labels for the governance rights and financial provisions were created by meticulously hand-coding the contents of corporate charters. A team of research assistants ("coders") were supplied with a detailed rubric that asked them to read the text of each version of a corporate charter, and then elicited information on 80 detailed attributes from the text. The complexity of language and use of legal syntax in corporate charters meant that coding the charters required legal training and familiarity with corporate governance and finance terms so that we required our coders to go through a lengthy period upfront training before they were randomly assigned corporate charters from our sample for the labeling tasks.

There are three categories of questions we tracked based on the text of the charters. The first category relates to rights of preferred shareholders, often focusing on the most recent class of preferred shareholders in the charter. Coders were asked to label the rights of preferred shareholders along several dimensions related to cash-flow rights (e.g., detailed information on liquidation preferences and contractual dividends) and control rights (e.g., veto and voting rights).

The second category of questions relates to corporate opportunity waivers (i.e., the ability to pursue business opportunities without offering them first to the corporation) for directors, officers, and shareholders. The third category we required coders to answer questions about waivers and indemnification for liability for directors, officers, and shareholders in the charter. In total, coders were asked just over 80 questions regarding each charter, creating labels for further analysis of charter content.

Members of our research team were aware of the purpose of this study and how their coding of the corporate charters would eventually be used for our analysis. Charters contain identifying information about the corporation, such as the corporation's name and purpose, but charters typically *did not* include information that identifies founders. This means that based on the charter alone, student coders could not infer whether the charter they are coding was of a female-founded startup or male-founded startup. We therefore do not expect that knowledge of the research question biased the coding enterprise.

We employed several mechanisms to address challenges related to the complexity of legal texts and labelling multi-dimensional provisions into simple and pre-determined rubrics for coding purposes. In addition

just after or just before the investment date, we use the first corporate charter filed following the date of the investment deal. If the charter filed closest to the documented deal date does not have a class of preferred shareholders, we assume that the charter does not reflect an investment round and therefore look for the first subsequent charter that includes a class of preferred shareholders.

to providing detailed guidance on the coding questions with examples of borderline cases, as issues arose and rarer provisions were discovered, these guidelines were updated. Lastly, randomly selected charters were assigned to more than one coder, allowing us to detect and address labeling inconsistencies.

C. Financial Provisions

From our labeled data fields, we have information on over fifteen financial provisions. We begin by describing and comparing several common financial provisions in startup corporate charters. The provisions highlighted below[79] relate most immediately to cash-flow rights,[80] even if they indirectly affect control rights as well, since they can skew the parties' incentives relating to fundamental questions about whether and when to exit.[81] We analyze provisions that relate more directly to governance and control rights in the next Part.

Table 3: Summary Statistics of Financial Provisions

Variable	Female-Founded Startups	Male-Founded Matches	F-M Diff	Fisher's Exact P-Value
Financial Provisions				
(1) Up Round	90.7	87.4	3.3	0.165
(2) Liquidation Preference	98.8	98.7	0.1	1.000
(3) Participating Preferred	12.1	19.5	-7.4	0.000
(4) Redemption Rights	12.6	12.6	0	1.000
Dividends				
(5) Contractual Dividend	5.6	6.4	-0.8	0.518
(6) Cumulative Dividend (If Dividend=1)	47.1	37.9	9.2	0.437
Pro-founder Provisions				
(7) Pay to Play	3.5	5.4	-1.9	0.083
(8) Option Pool	57.6	57.1	0.5	0.877

We start by examining the direction of the first observed financing round, reported in Line 1 of Table 3. A financing round is said to be an

[79] Several other financial provisions were coded up by our research assistant team but are not detailed in the main text. Our emphasis in the main text is on the most important and interesting financial provisions. The additional financial provisions documented include: the original issue price of preferred stock; the entitlement of preferred stock to a share of discretionary dividends; the liquidation preference's multiple; and the convertibility of preferred stock to common stock. Data on these additional provisions can be provided upon request to the authors.

[80] These cash-flow can significantly alter the true valuation of VC-backed companies. *See, e.g.*, Will Gornall and Ilya Strebulaev, *Squaring Venture Capital Valuations with Reality*, 135 J. FIN. ECON. 120, 135 (2020) ("Overvaluation arises because the most recently issued preferred shares have strong cash flow rights.").

[81] *See, e.g.*, Ola Bengtsson & Berk Sensoy, Changing the Nexus: The Evolution and Renegotiation of Venture Capital Contracts, 50 J. FIN AND QUANT. ANALYSIS 349 (2015).

"up" ("down") round if the per-share price that the round attracts is above (below) that of the prior round. If the price remains unchanged it is labeled a "flat" round. Based on the Genesis database meta data, we observe that a vast majority of all subpopulations of firms experience an up round in their initial observed round. Our matched firms track particularly closely, both manifesting somewhere between 87-91% up rounds. Male-founded firms demonstrate a marginally higher tendency for flat rounds than female-founded firms. A Pearson Chi-squared test fails to reject the null hypothesis that female-firm distribution is distinct from matched male-founded firms (p=0.091).

A highly relevant capital-structure consideration is the VCs' liquidation preference. Typically, VC investors assume a senior position during a company's liquidation or exit event. In other words, if the startup is acquired or liquidates, the VC investor is entitled to receive a specified multiple of their original investment, prior to any payout to shareholders. This liquidation preference plays a crucial role, as it can constrain the startup's ability to raise capital in subsequent investment rounds.[82] Line 2 of Table 3 indicates whether the latest round of preferred shareholder class has a liquidation preference in the event of a liquidation or exit event, using our hand-coded labels. We find that nearly all companies—both female-founded and male-founded—include a liquidation preference for the most recent preferred shareholders. A Fisher's Exact Test fails to reject the null hypothesis that the female-founded firm distribution is distinct from that of matched male-founded firms (p=1.000).

Though not technically the same as liquidation multiple, preferred shareholders sometimes have an additional right to periodic dividends.[83] In some cases, the right to receive a dividend accumulates over time for periods where the company does not pay it (which is commonplace). Any accumulated arrears must then be settled (in addition to the liquidation right) during a startup's liquidation. The implication of a cumulative dividend is to amplify the magnitude of the liquidation right, but in a subtle way that is not always immediately apparent.[84] Line 5 of Table 3 shows the frequency of a contractual dividend right in both our

[82] *See, e.g.*, Brian J. Broughman & Jesse M. Fried, *Carrots and Sticks: How VCs Induce Entrepreneurial Teams to Sell Startups*, 98 CORNELL. L. REV. 1319, 1343 (2013).

[83] In addition to contractual dividends, preferred shareholders may also be entitled to dividends if dividends are paid to common stock. 95.2% of female-founded firms give preferred shareholders the right to a discretionary dividend, while 93.7% of male-founded gives provide this right (p=0.560)

[84] *See, e.g.*, Broughman & Fried, at 1327 ("When the preferred shareholders are entitled to cumulative dividends, the liquidation preferences are even larger because the preferences include, in addition to the multiple, any unpaid dividends (even if not declared)"); Klausner & Venuto, *supra* note 45.

samples, using our hand-coded labels. Here, female-founded firms are subject to contractual dividends slightly less frequently (5.6%) than the matched male-founded group (6.4%). However, a Fisher's Exact Test fails to reject the null hypothesis (p=0.518). Further, Line 6 of Table 3 highlights the proportion of contractual dividends that are cumulative. Despite a lower number of female-founded firms facing contractual dividends, those that do more face cumulative dividends more frequently (47.1% and 37.9% respectively, with associated p=0.437).

Preferred stockholders may also enjoy cash flow rights in the form of upside "participation." This allows them to partake in gains of the common shareholders, such as during a liquidation event, without having to convert their shares to common stock.[85] Such participation rights enhance the VC investor's financial position and simultaneously diminish the cash flow rights of the founders and other common stockholders. Line 3 of Table 3 compares the prevalence of participation rights in preferred stock grants across our two groups, using our hand-coded labels. Here, female-founded firms in our sample face a significantly *lower* incidence of participation rights (12.1%) than matched male-founded firms (19.5%, with associated p=0.000).

Venture capital investors might also be allocated a redemption option incorporated into their stock grant. This essentially offers them a "put" option to force the startup to back their shares at a specified price (or formula). In practice, such redemption provisions can create significant liquidity crises inside illiquid startups. Given the obligation of the startup to secure cash to satisfy redemptions, it may find itself in the unenviable position of either breaching the redemption demand or acceding to a sale of the company at a price less favorable to the common shareholders.[86] Line 4 of Table 3 illustrates redemption rights across our two relevant groups, using our hand-coded labels. Here we see that both our female-founded firms and the matched sample have identical and rather modest exposure to redemption rights with identical incidence rates of 12.6% each (and associated p=1.000).

In anticipation of multiple stages of investment, VC-backed firms may create incentives to encourage early investors to continue to participate in subsequent financing rounds. One contractual device that aligns with this staged financing model is known as a "pay-to-play" provision. While not mandating further investments, such provisions

[85] Klausner & Venuto, *supra* note 45, at 1405 ("Once all preferred stockholders' initial liquidation preferences are fulfilled, the preferred shareholders' participation and conversion rights determine the allocation of the remaining proceeds of a sale")
[86] Charters were also coded for whether preferred stock was convertible to common share. 100% of female-founded firms and 99.60% of male-founded firms (p=1.000) provided preferred shareholders this possibility.

tend to impose significant costs on investors who choose not to partake in later rounds—commonly through the forfeiture antidilution rights.[87] This helps ease constraints on the startup's ability to raise capital in subsequent rounds. Line 7 of Table 3 tracks pay-to-play terms across our cohorts, illustrating that women-founded firms are slightly *less* likely to have the provisions (3.5%) relative to the matched male-founded startups (5.4%, with a borderline significant p=0.083).

Finally, another consideration is the allocation of a reserved stock option pool for future employee compensation. Option pool allocations have complicated strategic implications. While an ample option pool equips founders with flexibility for compensating both existing and future employees, expanding this pool as part of the investment round essentially incorporates the dilutive impact of the new options into the deal. This can reduce the per-share price paid by the investor. If an option pool is introduced or expanded after the investment, both the founder and investor would effectively have to share the dilution burden. Typically, therefore, the inclusion of an option pool in the funding round is less favorable to the founder and more advantageous for the investor. Line 8 of Table 3 illustrates the fraction of deals in each cohort that provide for an option pool using Genesis meta data. Here, female-founded firms and male-founded firms face a nearly identical prevalence of option pool allocations of 57.6% and 57.1% (with associated p=0.877).

In summary, the analysis of financial provisions in Table 3 reveals more similarities than disparities between female-founded and male-founded counterparts. In some cases, women face slightly *less favorable* provisions. The lower likelihood of female-founded startups having pay-to-play and higher incidences of a preferred shareholder liquidation preference put female founders in a less advantageous position. In other instances, however, female founders seem to be treated *more favorably* than their matched counterparts, such as with participating preferred rights. And, in most other cases, we do not find consistent and/or significant statistical trends in either direction. Based on our sample, then, cash-flow provisions do not appear to provide considerable traction in differentiating between female and male founder teams.

D. Non-Financial Provisions

Beyond financial provisions analyzed above, our hand-labeled data also support comparisons of several non-financial provisions related to governance and control rights. We highlight many of these dimensions

[87] *See, e.g.*, Robert P. Bartlett, *Venture Capital, Agency Costs, and the False Dichotomy of the Corporation*, 54 UCLA L. REV. 37, 57 (2006-2007).

in Table 4.[88] Board membership is typically viewed as a significant leverage point for influence, making it especially crucial in firm governance.[89] Line 1 of Table 4 reports percentage of startups where preferred shareholders have rights to appoint at least one board member for female-founded firms with male-founded matches. Here, female-founded firms are more frequently encumbered with VC board representation rights (at 56.1%), mildly exceeding the same entitlement within matched male-founded companies (54%). Nevertheless, the difference is not statistically significant (p=0.439).

Table 4: Summary Statistics of Non-Financial Provisions

Variable	Female-Founded Startups	Male-Founded Matches	F-M Diff	Fisher's Exact P-Value
Board Appointments				
(1) Board Appointment Rights	56.1	54	2.1	0.439
(2) Mean Ratio of Appointments	1.00346	1.151	-0.147476	0.0026 (t-test)
Voting Rights				
(3) General Voting Rights	98.8	99.4	-0.6	0.299
(4) Changes to Business Plan	10.2	14.6	-4.4	0.010
(5) Changes to Articles of Incorporation	83.0	85.2	-2.2	0.262
(6) Changes to common / preferred stock (including number of shares)	96.1	96.4	-0.3	0.788
(7) Issuing new series of stock and/or new capital raising round	93.7	94.7	-1	0.442
(8) Incurring additional capital debt that has a security interest	34.5	33.6	0.9	0.705
(9) Incurring additional capital debt that does not have a security interest	41.6	44.9	-3.3	0.197
(10) Changes to board size or membership rules	81.5	86.2	-4.7	0.015
(11) Payment / Declaration of dividends or distributions	90.2	93.3	-3.1	0.032
(12) Liquidation events	90.6	94.4	-3.8	0.006
(13) Entry into agreement as to merger, acquisition, asset sale, IPO, SPAC or other "exit" event	73.3	78	-4.7	0.037
(14) Mean Special Veto Rights Score	6.875	7.118	-0.243	0.0111 (t-test)
Waivers				
(15) Waiver for Directors	82.8	85.8	-3	0.123
(16) Waiver for Shareholders	72.7	74.9	-2.2	0.324
(17) Waiver for Officers	12.2	11.7	0.5	0.814

However, the mere capability to nominate board members may not be meaningful if another shareholder constituency has even more generous board appointment rights. A more nuanced analysis would examine the ratio of board members that are "owned" by preferred VC

[88] Our hand-labeled data encompasses various other non-financial provisions that are not reported in Table 4. These additional provisions include the pooling of preferred stock voting with other votes and the specific rights of preferred and common stockholders in appointing board members (including the number of board members chosen through the combined voting of both shareholder groups). We also document whether the charter specifies the total board size. Data on these additional variables is available upon request to the authors.

[89] The literature on startup governance has highlighted the importance of board appointments for allocation of control within startup firms and its impact of cash flow rights. *See* Broughman & Fried, *supra* note 48, at 385 (2010) (showing that when VCs are control of the board, there is less likely to be a deviation from previously agreed upon cash flow rights).

shareholders versus those controlled by common.[90] For this analysis, we considered the board appointments made by *any* preferred shareholders and not just board appointments made by the most recent class of preferred shareholders (as reported in Line 1 of Table 4). Line 2 of Table 4 reports the mean ratio of preferred shareholder appointments to common shareholder appointments for all cases in our data set where the ratio is defined. For female-founded firms, the mean ratio is 1.003, suggesting that when the number of the appointments is defined in the charter, the preferred and common shareholders have the exclusive right to appoint a roughly equal number of directors. For male-founded firms, in contrast, the mean ratio is higher (1.151), reflecting greater power of preferred shareholders over common shareholders in board appointment rights. We find this difference to be statistically significant ($p=0.003$).

Another important governance right is access to the corporate ballot box. Although preferred shareholders traditionally have limited or no voting rights, VC investors often receive the immediate right to vote alongside common shareholders in typical corporate governance matters (e.g., amending bylaws, cleansing, etc.) Table 4, Line 3 reveals a negligible difference between female- and male-founded companies, with both overwhelmingly allowing preferred shareholders voting rights on such matters.

Going beyond affirmative governance rights, investors may also demand a slew of "negative" control rights—typically manifested in a right to veto certain types of decisions made by the company.[91] The practice of giving VC investors certain veto rights is commonplace in startup governance, but there is some variation in the breadth and type of veto rights employed.[92] Table 4, Lines 4 and 5 illustrate two examples of veto rights, relating to (i) changes made to the business plan of the company; and (ii) changes to the articles of incorporation. Male-founded

[90] Given this is an early stage of funding, there is likely to be more common shareholder appointed board members relative to preferred shareholder appointed board members than in later rounds. *See* Pollman, *supra* note 45, at 181 (2019) ("Researchers have found a general trend in the evolution of a typical startup board over its life cycle-frequently starting out dominated by founders and transforming to shared or investor control at some time within the first few rounds of venture financing. This pattern occurs because investors typically build their voting power and seek additional board seats with each round of financing.").

[91] On the distinction protective provisions provided by special voting rights and the potentially more robust power of board appointments and board control, *see* Broughman & Fried, *supra* note 48, at 386 (2010); Fried & Ganor, *supra* note 45, at 987 (2006).

[92] For this analysis, we considered negative control rights given to *any* preferred shareholders and not just special voting rights given to the most recent class of preferred shareholders.

firms are discernibly more likely than female-founded counterparts to grant veto rights to preferred shareholders when it comes to changes in the business plan (14.6% to 10.2%)—a difference that is also statistically significant.[93] In contrast, veto rights are common (and highly comparable) for other matters. For example, as to changes to the charter, both female- and male-founded startups allocate such veto rights over four-fifths of the time.

All told, our data track ten unique veto right categories (including the illustrations above). The different veto right dimensions are reported in Table 4, Lines 4-13. To offer a composite view, we also formulated a "veto score" consisting of the additive sum of each of the 10 indicator variables we tracked. Table 4, Line 14 illustrates a marginally lower mean veto score for female-founded firms (6.875 versus 7.118) and an identical median score (of 7). The mean different is statistically significant (p=0.011).

Finally, startups may grant VC investors the right to compete with or remove business from the firm. Starting in 2000, several states began to empower corporate entities to waive the (so-called) corporate opportunity doctrine—a subspecies of the duty of loyalty—for their directors, officers, and investors.[94] Most commentators have observed that the statutory change permitting such waivers was driven substantially by private equity and venture capital investors, who feared liability when they placed nominees on two portfolio companies in the same industry.[95]

Table 4, Lines 15-17 depict the incidence of corporate opportunity waivers in female- and male-founded firms. Regarding ordinary directors, female-founded startups are slightly less likely to have waivers protection directors (82.8% versus 85.8%), but the differences are not statistically meaningful. Similarly, waivers for large shareholders are also relatively frequent, though slightly less prevalent across all companies (72.7% versus 74.9%), a difference that remains statistically indistinguishable. Finally, waivers are much less prevalent for officers (12.2% and 11.7%), a position that is not uncommonly occupied by founders—thus indicating that founders are typically prohibited from pursuing competing business opportunities in their individual capacities. Once again, however, the distinction between female- and male-founded firms is statistically insignificant.[96]

---

[93] Fisher's exact p-value = 0.010.

[94] *See, e.g.*, DGCL § 122(17).

[95] *See* Rauterberg & Talley, *supra* note 29, at 1089.

[96] There are several additional dimensions related to the waivers and protections afforded to directors, shareholders and officers that were coded up. These include: Exculpation of directors from monetary liability for breach of any fiduciary duties;

E. Summing Up

Overall, while our inquiry into individual financial and governance measures uncovers some interesting distinctions, a side-by-side comparison of non-financial and governance terms between female-founded and male-founded reveals far more similarities than differences. In some cases, women face slightly more onerous provisions, while in some they face a less constraining road. Notably, more disparities lack robust statistical backing. Based on more targeted inquiries, then, the collection of standard financial and governance terms fail to distinctly differentiate the formal governance rights of female and male founder teams.

F. Robustness of "Female-Founded" definition.

The above analysis presents a comprehensive comparison of startup governance as a function of founder gender. At the same time, however, our analysis utilizes a particular definition of the founding team's gender characteristic: any firm with at least one female founder was classified under our rubric as a "female-founded" startup. This approach afforded us with a large sample of data for comparisons. Nevertheless, it may be overinclusive, too, sweeping in companies that (for example) add a "token" female founder for optics rather than substantive contributions. Such firms might better resemble (and should be treated as) male-founded firms, and thus their inclusion here would tend to dilute genuine gender effects in the aggregate. Given that the findings presented above also reveal few systematic significant differences, our definition of a female-founded firm warrants further scrutiny.

To test the robustness of our definition on our results, we replicated the above analysis with two alternative and more stringent definitions of female founder teams: (A) Startups with majority female founder teams ("MFF"); and (B) Startups with fully female founder teams ("FFF"). Adopting these stricter definitions can lead to two conflicting effects. On the positive side, they could reduce the noise arising from a broader "female-founded" definition that encompasses companies with a majority of male founders, where the influence of a female founder might be diminished. Conversely, each successive robustness measure

---

exculpation of officers from monetary liability for breach of any fiduciary duties; allowing/requiring indemnification of directors for liability; allowing/requiring indemnification of officers for liability; allowing/requiring advancement of fees of directors for liability; allowing/requiring advancement of fees of officers for liability; allowing/requiring the purchase of directors insurance against liability; allowing/requiring the purchase of officers insurance against liability. Data on these additional provisions is available upon request to the authors.

implies an incrementally stricter standard for female founding teams which considerably shrinks our sample size. Specifically, our data set of female-founded companies (and the corresponding 3x1 male-founded matches) dropped with each stricter criterion, decreasing from 257 in our primary analysis to 141 in the MFF category[97] and down to 50 for the FFF category.[98] This reduction in sample size can amplify the noise in our estimates.

Table 5 presents the results using these two alternative definitions in comparison with our baseline more expansive definition of female-founded firms.

Table 5: Robustness Checks

Variable	Original Ssample F-M Diff	Fisher's Exact P-Value	MFF F-M Diff	Fisher's Exact P-Value	FFF F-M Diff	Fisher's Exact P-Value
Financial Provisions						
(1) Up Round	3.3	0.165	7.3	0.006	6.1	0.237
(2) Liquidation Preference	0.1	1.000	-0.7	0.340	-1.3	0.622
(3) Participating Preferred	-7.4	0.000	-6.5	0.008	-10.8	0.007
(4) Redemption Rights	0	1.000	0.7	0.755	-1.2	0.852
Dividends						
(5) Contractual Dividend	-0.8	0.518	-4.9	0.002	-3.4	0.138
(6) Cumulative Dividend (If Dividend=1)	9.2	0.437	17.1	0.358	13.6	0.644
Pro-founder Provisions						
(7) Pay to Play	-1.9	0.083	-2.2	0.143	-3.3	0.318
(8) Option Pool	0.5	0.877	1.5	0.678	-2.0	0.817
Board Appointments						
(1) Board Appointment Rights	2.1	0.439	3.7	0.298	-1.1	0.907
(2) Mean Ratio of Appointments	-0.147476	0.0026 (t-test)	-0.2	0.0002 (t-test)	-0.2	0.1398 (t-test)
Voting Rights						
(3) General Voting Rights	-0.6	0.299	-0.5	0.373	0.0	.
(4) Changes to Business Plan	-4.4	0.010	-2.9	0.218	-8.1	0.034
(5) Changes to Articles of Incorporation	-2.2	0.262	4.8	0.057	-1.2	0.865
(6) Changes to common / preferred stock (including number of shares)	-0.3	0.788	-1.2	0.465	0.0	1
(7) Issuing new series of stock and/or new capital raising round	-1	0.442	-0.7	0.773	-2.6	0.488
(8) Incurring additional capital debt that has a security interest	0.9	0.705	-1.1	0.767	-2.9	0.625
(9) Incurring additional capital debt that does not have a security interest	-3.3	0.197	-2.9	0.405	-19.0	0.001
(10) Changes to board size or membership rules	-4.7	0.015	-6.6	0.014	-5.9	0.000
(11) Payment / Declaration of dividends or distributions	-3.1	0.032	-3.8	0.067	-3.8	0.139
(12) Liquidation events	-3.8	0.006	-6.4	0.002	-10.6	0.004
(13) Entry into agreement as to merger, acquisition, asset sale, IPO, SPAC or other "exit" event	-4.7	0.037	-6.0	0.06	-5.2	0.339
(14) Mean Special Veto Rights Score	-0.243	0.0111 (t-test)	-0.3	0.050 (t-test)	-0.7	0.0038 (t-test)
Waivers						
(15) Waiver for Directors	-3	0.123	-6.5	0.011	-14.5	0.002
(16) Waiver for Shareholders	-2.2	0.324	-8.8	0.006	-17.7	0.001
(17) Waiver for Officers	0.5	0.814	1.0	0.676	0.6	1.000

Table 5 indicates that the general patterns observed in our baseline analysis are consistent with the results from our more stringent definitions of female-founded teams. In these narrowed samples, many similarities between female- and male-founded startups continue to

[97] With 423 male-founded matches and 369 unique male matches.
[98] With 150 male-founded matches and 140 unique male matches.

emerge. Where prior sections highlighted disparities between the two groups, these differences often become more pronounced under the stricter definitions of female founder teams. For instance, the tendency of male-founded firms to grant special veto rights to preferred shareholders on certain matters is even more evident and is often more statistically significant in the narrower definitions of female-founded firms. The tighter confidence intervals under the stricter definition, despite the reduced sample size, suggests there might be genuine distinctions between the treatment of male- and female-founded enterprises by investors on some dimensions. Nonetheless, the multifaceted nature of these differences—with female founder firms including more founder-friendly terms in some areas and less founder-friendly terms in others—defy a simple overarching narrative.

V. INTERPRETATIONS, IMPLICATIONS, AND EXTENSIONS

The previous Parts collectively present an intriguing puzzle for our research enterprise—one that carries material policy implications. On the one hand, in Part III we employed a variety of machine learning tools to demonstrate that the *overall contents* of startups charters are sufficiently distinct to predict the existence of a woman founder based solely on the un-interpreted text alone. Indeed, even within an "apples to apples" matched sample, the control group of all-male founder teams appear (on average) to attract overall governance structures that are predictably distinct from the treatment group of women-founded companies. On the other hand, Part IV demonstrates that whatever the source of this distinct semantic structure, it does not seem to manifest itself (at least very clearly) within a collection of canonical governance provisions that we meticulously hand-coded in our data set. In other words, the consistent *overall* differentiation in governance regimes does not appear to generate clear patterns when we focus on specific provisions over which founders, funders, and lawyers of all stripes commonly obsess. Jointly, these conclusions pose at least two interrelated questions. Foremost, if the semantic differences in the control and treatment groups do not manifest in canonical governance measures, what might be driving the overall difference? And secondarily, how do these results bear on more general policy questions related to gender and VC financing?

As to the first question, we can think of three plausible explanations that might explain why our hand-coded data labels do not appear to echo

the semantic distinctions we measure in the aggregate.[99] First, it may be that our hand-labeling protocols simply omit certain critical items that are central to governance. Although we interviewed dozens of experts in order to isolate key variables that warranted labeling, there is no guarantee that we netted all of them in our hand-coding enterprise. That said, we *are* pretty confident that our research design captured many (if not most) of the terms that are thought generally to be important for corporate governance.[100] It thus would be genuinely surprising if (a) we omitted a subset of terms that were—unbeknownst to us—the most critical, *and* (b) these omitted provisions systematically differentiated our treatment and control companies. Though the confluence of these scenarios is not impossible *per se,* we think it improbable. Future research endeavors might nevertheless attempt to revisit the protocols we have employed, expanding them as appropriate to sweep in other specific governance levers that our research design might have neglected.

A second possibility is that there *are* patterns in the specific hand-coded provisions from Part IV, but that their interactions are too complex to reveal themselves with univariate comparisons. In other words, there may be a series of latent distinctions inherently present within *combinations* of labeled provisions that can only shine through in the aggregate. This possibility seems worthy of further investigation, since governance itself is complex, and its levers are highly interdependent. To get some purchase on this issue we revisited the semantic predictive exercises from Part III, this time appending the unigram texts with the labels available in the subsample studied Part IV. From here one can ask whether including the added hand-coded labels enhances predictive power of a machine learning classifier over using the text alone. To the extent that the labels do augment predictive power, it would be fair to conclude that it suggests that the data labels do in fact help predict founder genders through more complex linear combinations of labels.

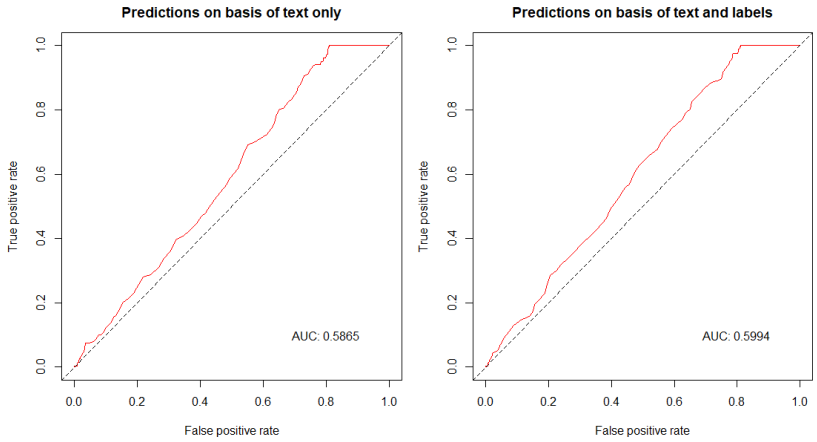
The results of this inquiry are illustrated in Figure 9. The left panel of the figure depicts the ROC curve when one uses only the textual content of the charters to predict founder gender. The right panel of the Figure does the same but for a combined raw data set that merges both the textual content and the hand-coded labels. As one can see from the figure, the inclusion of labels in addition to texts has virtually no effect

[99] This assumes the integrity of our labeling enterprise. Here we are pretty confident that we trained our research team to produce reliable and replicable labels. *See supra* Part IV.

[100] *See supra* Part IV.

on the classifier's ability to predict founder gender.[101] To the extent that the hand-coded labels are capable of generating additional predictive power, then, it does not appear to come (if at all) through various linear combinations of them.[102]

Figure 9: Machine Learning Classifier of Founder Gender



A final possibility is that the semantic differences we detected in Part III were simply epiphenomena that did not carry over to the targeted governance provisions in Part IV. In other words, the argument goes, our machine learning classifier may simply have seized on technical linguistic differences between charters that track gender (such as proper names, pronouns, or gendered sub-industries)—differences that ultimately prove orthogonal to corporate governance *per se*, even while such linguistic tokens can mechanically help predict founders' gender identity. To the extent that this final hypothesis holds, then, one could reasonably conclude that women founders (as we have defined them) do not actually face different types of firm governance, notwithstanding the measurable semantic differences in charters.

There are some reasons to be skeptical about this possibility, too. For example, our matched sample was constructed by finding industry

[101] There is similarly little added predictive power when one moves from text only to text plus labels when measured by predictive accuracy (0.7080958 versus 0.7069555) or F-1 score (0.5059055 for both predictions).

[102] It is possible that a variety of non-linear transformations of the labeled data (a typical approach of neural net / word embedding based classifiers) could generate greater predictive power, but it would do so for the textual data as well. *See* Justin Grimmer, Margaret E. Roberts & Brandon M. Stewart, TEXT AS DATA: A NEW FRAMEWORK FOR MACHINE LEARNING AND THE SOCIAL SCIENCES 82-84 (Princeton Press 2022).

matches, and thus it should have controlled for certain types of epiphenomenal gender segregation (in particular, the problem of possibly gendered sub-industries and their impact on charter content) by construction. Furthermore, governance documents generally tend to be neutral in their textual content, and it seems unlikely that such distinctions are driving our results. Moreover, if the use of gendered names and pronouns is driving our results, it is hard to understand why our ability to differentiate between male- and female-founded firms tends to get *better* (and certainly grows no worse) in the latter part of our data set, just as overall parlance has become more conscientious about avoiding the use of strongly gendered nouns and pronouns.

Nevertheless, even if the "epiphenomenon" hypothesis holds validity—and female founders face indistinguishable governance regimes from their male counterparts—there remains an important policy take-away for gender dynamics over the life cycle of a startup. To the extent that women startup entrepreneurs are disadvantaged at the funding stage, this initial imbalance does not get "fixed" or otherwise dampened subsequently through more attractive governance. By the same token, and in a glass-half-full sense, the initial funding imbalance is also not exacerbated through subsequent governance choices either. Nevertheless, a bottom line for policy makers remains that initial funding differences can and do matter not just at the funding stage, but they can propagate forward in time, resisting repair or compensation later through formal institutions of governance.

Moreover, even to the extent that our analysis above holds, our results help set a statistical governance baseline going forward. As discussed above, several states have promulgated legislation that formally scrutinizes acts of alleged gender discrimination in VC markets, including firm governance.[103] Without having a baseline for what types of governance provisions are consistent with market practices, it would be difficult (if not impossible) to prove disparate treatment of individual claimants. Our results can help clarify a sense of market comparisons for future claimants on an individual basis, regardless of whether there are aggregate governance differences by gender in the VC startup space.

Notwithstanding the novel contribution that we make at the *company* level, it is important to note that our collective efforts to understand sex and startups is still at a stage of relative infancy. Most centrally, our analysis in this paper has not ventured (so to speak) into the structure of the *funding* side of the market, to determine how and whether fund governance and startup governance interact. Controlling for the sources of funding (and the patterns by which funders and founders match with

[103] *See supra* Part I.

one another) can no doubt reduce significant statistical noise in efforts such as ours. But moreover, doing so may help uncover patterns and trends that a more complete accounting of the startup financing and governance market otherwise misses. Although we leave such efforts for future researchers (including ourselves), the contributions we have made here push constructively towards that goal.

Those caveats aside, our analysis suggests that parties hoping to address gender inequities in the VC space may have under-utilized an important weapon for the battle: corporate governance. Indeed, attorneys, entrepreneurs, and transactional designers may fruitfully be able to deploy governance tools more effectively than current practice seems to suggest. To the extent that women experience disadvantages at the funding stage, those who succeed at attracting investments will tend to have higher average quality than the "market" of male founder counterparts. As such women founders may be able to tap into underutilized benefits by insisting that financiers to give them broader rights, more enhanced cash flows, and fewer constraints in their formal governance arrangements, rather than merely emulating market norms.

CONCLUSION

This paper has presented and analyzed a first-of-its-kind data set of corporate governance documents in order to gain purchase on the question of whether the governance structures of VC-backed companies tend to track, dampen, or exacerbate widely documented gender imbalances for startup funding. Our results are simultaneously intriguing and paradoxical. From our collection of corporate charters, it is clear that the general semantic content of female-founded and male-founded startups differ discernibly from one another, and that these differences have not dissipated over time. In particular, there appear to be governance structures that are unique to male-founded firms and largely unavailable to female-founded counterparts, even when the comparison is constrained to a matched sample that makes comparisons on an "apples to apples" basis. On the other hand, the overall syntactical differences between and across charters does not appear to be borne out within the patterns of key, focal provisions that typically draw substantial attention to how corporate governance allocates cash flow and control rights between founders and funders. In essence, corporate governance appears to have remained on the sidelines in the gendered world of venture finance, neither exacerbating nor ameliorating documented gender imbalances within the sector. Those interested in effectuating change within the sector might thus consider innovative ways to

counteract funding imbalances through more and different governance concessions.

APPENDIX A: RESEARCH ASSISTANTS

The completion of this project was enormously labor intensive, and it required a rolling team of researchers from across three different universities. We list them below with our enormous gratitude:

Lead Research Assistant

Jiyoung Kim

Research Assistants

Kathryn T. Benvenuti
Jack Bradley
Elena Carroll-Maestripieri
Elana Michelle Cates
Joy Chow
Selin Deldag
William S. Donaldson
Banu Dzhafarova
Emily M. Erickson
Benjamin P. Faber
Moshe Gershenfeld
Ziyu Guo
Connor J. Hudson
Cailin Liu
Caleb Logan
DanLan Luo
Daniel Marzagalli
Kyle Mary Oefelein
Sonila Helen Pascale
Connor Rust
Deborah A. Sparks
Alexander H. Sugerman
Madison White
Ridglea K. Willard
Dajun Xu

APPENDIX B: REGRESSION ANALYSIS

The charts in Part III.C of the main text help us to assess each of the financial factors sequentially. These figures do not attempt to control for any other observables and are a raw average for female-founded and male-founded firms. In this Appendix we present the results in the form of regression analysis for a selected number of financial variables. Below we briefly present linear probability regressions with the following form:

$$Y_i = \beta_0 \cdot X_i + \beta_1 Female_i + \varepsilon_i ,$$

where Y_i denotes a binary variable of interest (see table), X_i denotes a vector of control variables (related to round, logged post-money valuation, logged investment, year, and region) and $Female_i$ denotes a female-founded firm. Because we focus only on the first observed charter, the regressions in (1) are effectively cross-sectional in nature (although we still control for region and year fixed effects).

The results are given in Table B1. As can be seen from the table, the "Female Startup" variable is statistically significant in only a few situations, predicting a smaller likelihood of participation preferred rights and a smaller likelihood of the charter containing a pay-to-play provision. The remaining Female Startup coefficients are statistically no different from zero under standard criteria. Overall, these results suggest that female founder status plays a surprisingly minor role in predicting several key cash-flow and capital-structure variables that are often treated as focal in the startup-VC relationship.

Table B1: Regression Analysis

	(1) Liquidation Preference	(2) Liquidation Preference	(3) Participation Preferred	(4) Participation Preferred	(5) Redemption	(6) Redemption	(7) Contractual Dividends	(8) Contractual Dividends	(9) Pay-to-Play Provision	(10) Pay-to-Play Provision	(11) Option Pool	(12) Option Pool
ln(Valuation)	-0.00879	-0.00863	-0.0429***	-0.0445***	-0.0320***	-0.0336***	0.00339	0.00362	-0.00722	-0.00701	0.0652***	0.0639***
	(0.00642)	(0.00632)	(0.0116)	(0.0118)	(0.00848)	(0.00849)	(0.00637)	(0.00636)	(0.00514)	(0.00518)	(0.0120)	(0.0118)
ln(Investment)	0.00322	0.00467	0.0106	-0.0251	0.0390***	0.0233	-0.00674	0.000883	0.0330***	0.0369***	0.0708***	0.0472**
	(0.00637)	(0.00707)	(0.0133)	(0.0175)	(0.0114)	(0.0145)	(0.00894)	(0.0120)	(0.00744)	(0.0102)	(0.0171)	(0.0209)
Female Startup	0.00104	0.000650	-0.0784***	-0.0780***	0.000936	0.00183	-0.0114	-0.0119	-0.0211	-0.0211	0.0167	0.0163
	(0.00980)	(0.00964)	(0.0258)	(0.0257)	(0.0244)	(0.0245)	(0.0188)	(0.0188)	(0.0138)	(0.0139)	(0.0339)	(0.0336)
Series_A		-0.0152		0.109***		0.0226		-0.0416		-0.00692		0.0245
		(0.0110)		(0.0357)		(0.0323)		(0.0299)		(0.0180)		(0.0500)
Series_B		0.000938		0.159***		0.0724		-0.0392		-0.0196		0.118**
		(0.0112)		(0.0485)		(0.0451)		(0.0365)		(0.0271)		(0.0589)
Series_C		-4.23e-05		0.230***		0.106		-0.00771		-0.0274		0.206**
		(0.0127)		(0.0734)		(0.0673)		(0.0550)		(0.0486)		(0.0872)
Series_D_or_greater		-0.0485		0.232**		0.172*		-0.0638		-0.0207		0.0855
		(0.0581)		(0.109)		(0.102)		(0.0429)		(0.0486)		(0.102)
Observations	1,462	1,462	1,464	1,464	1,446	1,446	1,441	1,441	1,466	1,466	1,466	1,466
R-squared	0.063	0.070	0.148	0.167	0.175	0.184	0.062	0.066	0.147	0.148	0.263	0.271
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1