european corporate governance institute

Sex & Startups

Law Working Paper N° 759/2024
June 2025

Jens Frankenreiter
Washington University in St. Louis

Talia Gillis
Columbia Law School

Eric Talley
Columbia Law School

ECGI Working Paper Series in Law

Sex & Startups

Working Paper N° 759/2024
June 2025

Jens Frankenreiter
Talia Gillis
Eric Talley

Abstract

Venture capital is widely perceived to have a gender problem. Both founders seeking capital and the investors themselves are overwhelmingly male, fomenting concerns about how - and how fairly - the VC sector distributes its economic gains. Although gender disparities in funding are well documented, we still know little about whether the governance of VC-backed startups similarly manifests gender imbalances. This knowledge gap is critical, since VC investments influence cash flow and control rights, which can vary substantially from deal to deal. This study unveils a first-of-its-kind dataset that offers detailed insights into the governance of VC-backed startups. Our data contain granular information on both cash flow and control rights within startups over a span of nearly two decades. Most significantly, our data permit one to interrogate whether a founder's gender predicts distinct governance patterns in relation to their VC investors. Our analysis yields several intriguing findings. First, with the help of machine learning techniques, we uncover persistent, gender-specific patterns in the linguistic architecture of startup governance documents. Digging deeper into the substantive governance provisions themselves, we show that female founders encounter a mix of advantages and disadvantages. For instance, they tend to face relatively unfavorable conditions in dividend policies and board representation, but more favorable treatment in profit participation and veto rights. ...

Jens Frankenreiter
Washington University in St. Louis

Talia Gillis
Columbia Law School

Eric Talley
Columbia Law School

Sex & Startups

Jens Frankenreiter,[†] Talia B. Gillis[††] & Eric L. Talley[†††]

Venture capital is widely perceived to have a gender problem. Both founders seeking capital and the investors themselves are overwhelmingly male, fomenting concerns about how—and how fairly—the VC sector distributes its economic gains. Although gender disparities in <u>funding</u> are well documented, we still know little about whether the <u>governance</u> of VC-backed startups similarly manifests gender imbalances. This knowledge gap is critical, since VC investments influence cash flow and control rights, which can vary substantially from deal to deal.

This study unveils a first-of-its-kind dataset that offers detailed insights into the governance of VC-backed startups. Our data contain granular information on both cash flow and control rights within startups over a span of nearly two decades. Most significantly, our data permit one to interrogate whether a founder's gender predicts distinct governance patterns in relation to their VC investors.

Our analysis yields several intriguing findings. First, with the help of machine learning techniques, we uncover persistent, gender-specific patterns in the linguistic architecture of startup governance documents. Digging deeper into the substantive governance provisions themselves, we show that female founders encounter a mix of advantages and disadvantages. For instance, they tend to face relatively unfavorable conditions in dividend policies and board representation, but more favorable treatment in profit participation and veto rights. On balance, however, our analysis reveals no systematic pattern of gender disparities in either direction, a surprising result in the light of our linguistic findings, and one that confounds theoretical predictions about how gender bias plausibly affects firm governance. Ultimately, our analysis reveals a complicated landscape: While female founders encounter linguistically distinct governance structures, those differences do not appear systematically to exacerbate the challenges women founders face at the initial funding stage. By the same token, startup governance also does not appear to counteract gendered funding disparities, either.

---

† Associate Professor of Law, Washington University in St. Louis.

†† Associate Professor of Law, Columbia Law School.

††† Marc and Eva Stern Professor of Law and Business, Columbia Law School. Many thanks to Ian Ayres, Kerry Baldwin, Robert Bartlett, Molly Calkins, Katherine McCormick, Dorothy Lund, John Morley, Chris Nicholls, Elizabeth Pollman, Mitch Polinsky, Roberta Romano, Darren Rosenblum, Sarath Sanga, Natasha Sarin, Megan Stevenson, Marco Ventoruzzo, Rebecca Wexler, 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, the Conference on Data Science & Law, and workshop participants at American University, Columbia, George Mason, Stanford, University of Amsterdam, Virginia, Western Ontario, and Yale for helpful comments and discussions. Jiyoung Kim provided outstanding research support and made substantive contributions to this project. Thanks as well to a veritable phalanx of research assistants too numerous to mention here, but whose names are listed with tremendous gratitude in Appendix A. All errors are ours.

Introduction

Much of the "dark matter" in the venture capital (VC) financing universe consists of star-crossed startups. It is now common wisdom that VC-backed startups overwhelmingly fail,[1] typically because their founders' visions prove untenable, unworkable, underwhelming, or underappreciated. But it is exceedingly rare for a startup to fail *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, charges 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 supercilious assurances about the bright prospects for 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 $445 million severance bonanza as he decamped from his gilded WeWork cubicle.[4] No longer affiliated with WeWork, Neumann now oversees several new real estate

1. *See* Patrick Ward, *Is It True That 90% of Startups Fail?*, NANOGLOBALS (June 29, 2021), https://nanoglobals.com/startup-failure-rate-myths-origin [https://perma.cc/SJ8G-L9XP] (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 ST. J. (Sept. 20, 2012, 12:01 AM ET), http://online.wsj.com/article/SB10000872396390443720204578004980476429190 [https://perma.cc/N8TJ-A33F] (describing higher failure rates than indicated by the industry rule of thumb that only three or four out of every ten startups fail completely); Elizabeth Pollman, *Startup Failure,* 73 DUKE L.J. 327 (2023) (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 [https://perma.cc/7X5L-9U24].

3. *See* Donald C. Langevoort & Hillary A. Sale, *Corporate Adolescence: Why Did "We" Not Work?,* 99 TEX. L. REV. 1347, 1349 (2021) (using the story of WeWork to illustrate the risks of "a build-up of bad choices and testy 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 [https://perma.cc/2PDK-RSXT] (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. Edward Helmore, *WeWork Founder Adam Neumann Receives $445m Payout in Exit Package*, THE GUARDIAN (May 27, 2021), https://www.theguardian.com/business/2021/may/27/wework-founder-adam-neumann-enormous-exit-package [https://perma.cc/K7MY-KTCF].

investment funds,[5] reportedly whiling away his time musing about "becoming leader of the world, living forever, and amassing more than $1 trillion in wealth."[6] WeWork, meanwhile, is waging a losing battle for its very survival in U.S. bankruptcy court.[7]

To be sure, myriad factors distinguish the Holmes and Neumann narratives, rendering an extended side-by-side comparison somewhat tenuous. There is, for example, a compelling argument that Holmes's fraudulent behavior (in the medical field) imposed significantly more social harm than Neumann's insouciant profligacy (in commercial real estate).[8] Nevertheless, it is hard *not* to notice Holmes's and Neumann's divergent fates and wonder whether their paths may reflect—at least in some small way—a larger story about gender.[9] Did their distinct gender identities play any role in driving their differing receptions, treatments, trajectories, and outcomes? This question is relevant not only for Holmes's and Neumann's specific stories, but also for generations of startup founders who now will follow in their respective wakes.[10]

5. *See* Alexandra Tremayne-Pengelly, *WeWork Founder Adam Neumann Is Back With Another Real Estate Venture*, OBSERVER (Feb. 7, 2023, 5:21 PM), https://observer.com/2023/02/wework-founder-adam-neumann-is-back-with-another-real-estate-venture [https://perma.cc/VYA4-L3M6] (explaining how Neumann's new company, Flow, received a $350 million investment from a VC firm in 2022).

6. *See* Eliot Brown, *How Adam Neumann's Over-the-Top Style Built WeWork. 'This Is Not the Way Everybody Behaves.'*, WALL ST. J. (Sept. 18, 2019), https://www.wsj.com/articles/this-is-not-the-way-everybody-behaves-how-adam-neumanns-over-the-top-style-built-wework-11568823827 [https://perma.cc/JA8J-V3JU].

7. *See* Sujeet Indap & Joshua Oliver, *WeWork Files for Bankruptcy Amid Office Market Downturn*, FIN. TIMES (Nov. 7, 2023, 4:11 AM), https://www.ft.com/content/775ec01c-1f57-4105-bbc2-b048865056f7 [https://perma.cc/3NZA-DGQJ].

8. *See, e.g.*, Bobby Allyn, *The Elizabeth Holmes Trial Sparks A Silicon Valley Debate: Why Not Other Tech CEOs?*, NPR (Sept. 25, 2021, 7:07 AM ET), https://www.npr.org/2021/09/25/1040442689/elizabeth-holmes-trial-why-her-not-other-ceos [https://perma.cc/RC8N-U6B7]. Moreover, in recent months we have also begun to see several criminal convictions of male founders whose fates went horribly awry (such as FTX's Sam Bankman-Fried and Nikola's Trevor Milton). *See, e.g.*, Britney Nguyen, *Sam Bankman-Fried Faces 110-Year Max Sentence After FTX Trial—Here's How Long Experts Think He'll Be Behind Bars*, FORBES (Nov. 3, 2023, 1:40 PM EDT), https://www.forbes.com/sites/britneynguyen/2023/11/03/sam-bankman-fried-faces-110-year-max-sentence-after-ftx-fraud-trial-heres-how-long-experts-think-hell-be-behind-bars [https://perma.cc/VUJ7-T9YT]; Joe Miller & Claire Bushey, *Nikola founder Trevor Milton Sentenced to 4 Years for Fraud,* FIN. TIMES (Dec. 18, 2023, 2:28 PM), https://www.ft.com/content/47341212-5a66-4e01-85ef-3a0ba8c1fea5 [https://perma.cc/6KF8-NJSZ].

9. We are hardly the first to posit this question. *See, e.g.*, Ellen Pao, *The Elizabeth Holmes Trial Is a Wake-Up Call for Sexism in Tech*, N.Y. TIMES (Sept. 15, 2021), https://www.nytimes.com/2021/09/15/opinion/elizabeth-holmes-trial-sexism.html [https://perma.cc/5LK2-SG5Q] ("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 8. 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 (Dec. 7, 2022), https://www.nytimes.com/2022/12/07/technology/sunny-balwani-theranos-sentenced [https://perma.cc/5ETF-JNNB] ("Ramesh Balwani, the former chief operating officer of the failed blood testing start-up Theranos, was sentenced . . . to nearly 13 years in prison . . . .").

10. *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 [https://perma.cc/R82R-HFCX] (reporting on female founders who must now distinguish themselves from Elizabeth Holmes in meetings with investors); Elaine Moore, *Silicon Valley Has Learnt Little From*

Concerns about gendered capitalism are hardly new to corporate governance, and a growing literature documents how female representation among officers and directors interacts meaningfully with corporate conduct,[11] while simultaneously increasing the likelihood of being targeted by activists, critics, and cranks.[12] Within the stubbornly male-dominated VC ecosystem, however, these concerns grow even starker.[13] Indeed, the entire industry has attracted critical scrutiny over the last decade for its anemic track record on gender. As recently as 2024, for example, women-founded companies garnered somewhere between 2.0% and 20.6% of the total capital invested in venture-backed startups in the US (depending on how one adds up).[14] VC funds also have tended (until recently) to be disproportionately male-dominated, and there is evidence that this composition affects their investments, particularly in their

Elizabeth Holmes: Making Big Claims Remains the Starting Point for New Companies, FIN. TIMES (Feb. 22, 2022), https://www.ft.com/content/66f11e1f-9ec9-406d-868f-d0f757a915d6 [https://perma.cc/53NL-UKRV] ("Instead of seeing the case as a spur to toughen up due diligence, the tech sector is choosing to dismiss it as an outlier.").

11. 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* 5 (Nat'l Bureau of Econ. Rsch., Working Paper No. 30657, 2021), https://www.nber.org/system/files/working_papers/w30657/w30657.pdf [https://perma.cc/JPT3-LENZ] (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, 751 (2017) (demonstrating that more gender diverse boards conduct more active board meetings).

12. *See* Andrew Ross Sorkin, *Do Activist Investors Target Female C.E.O.s?*, N.Y. TIMES DEALBOOK (Feb. 9, 2015, 9:09 PM), https://archive.nytimes.com/dealbook.nytimes.com/2015/02/09/the-women-of-the-s-p-500-and-investor-activism [https://perma.cc/QH56-WKQR] (describing evidence of "a subconscious gender bias among activist investors" and the relationship between gender and power); 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, 233 (2018) ("[O]ur results suggest that female CEOs have to deal with additional challenges imposed by . . . activist investors and are more vulnerable than male CEOs 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, 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 [https://perma.cc/4BSC-2JN2] (noting that using gendered language in PR materials "increases the likelihood of shareholder activism by 31 percent").

13. *See, e.g.*, Benjamin P. Edwards & Ann C. McGinley, *Venture Bearding*, 52 U.C. DAVIS L. REV. 1873, 1873 (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 (Mar. 4, 2025), https://pitchbook.com/news/articles/the-vc-female-founders-dashboard [https://perma.cc/3P9P-VT24] [hereinafter *Female Founders Dashboard*] (revealing sharp disparities in venture-funding rates for only female founded and female co-founded firms).

14. The number subdivides by whether one counts firms founded by woman-only founders (2.0%) or women co-founded teams (20.6%). *See Female Founders Dashboard*, *supra* note 13.

historically tepid engagement with women-founded or women-led startups.[15] Other studies have documented that female-managed funds have lower average inflows than male-managed funds despite showing no gender differences in performance.[16] And, as noted above (and documented below),[17] women have remained largely peripheral when it comes to both bestowing and receiving VC funding in the first instance.

Yet in important ways, VC funding disparities are but the tip of a much larger iceberg for women founders. Venture funds are hardly selfless altruists, and the capital they bestow is seldom gifted: their investments frequently have significant strings attached, reflected in dozens of cash flow and control rights that can collectively work to raise expectations, appropriate power, and dilute founders' economic stakes. The aggregation of such governance provisions can radically alter the balance of power between founders and funders, rendering the "lucky" recipients of VC money far less fortunate than it might at first appear.[18] If women founders must scale metaphorical mountains simply to get funded, it is hard not to wonder about the governance landscape that awaits them on the other side.

Unfortunately, our ability to analyze governance in VC-backed startups has remained frustratingly limited, courtesy of the non-public nature of both startups and their financiers, causing both groups to fall under the radar of most public disclosure databases. Beyond collected anecdotes, researchers have had virtually no visibility on the broad nature and characteristics of internal startup governance. Do female-founded firms systematically face more onerous governance terms than their male-founded counterparts? Or, might their relative

15. *See Female Founders Dashboard*, *supra* note 13; Valentina Zarya, *Venture Capital's Funding Gender Gap Is Actually Getting Worse*, FORTUNE (Mar. 13, 2017, 11:24 AM ET), https://fortune.com/2017/03/13/female-founders-venture-capital [https://perma.cc/44F8-SNDG]; Sophie Calder-Wang, Paul Gompers & Patrick Sweeney, *Venture Capital's "Me Too" Moment* 2 (Nat'l Bureau of Econ. Rsch. Working Paper No. 28679, 2021) (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 [https://perma.cc/D8YH-S7WK]; Paul A. Gompers & Sophie Q. Wang, *Diversity in Innovation* 10-11 (Nat'l Bureau of Econ. Rsch. Working Paper No. 23082, 2017) ("From 1990-2016 women have been less than 10% of the entrepreneurial and venture capital labor pool . . . .").

16. *See* Alexandra Niessen-Ruenzi & Stefan Ruenzi, *Sex Matters: Gender Bias in the Mutual Fund Industry*, 65 MGMT. SCI. 3001, 3001 (2019) (documenting "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*, BOS. CONSULTING GRP. 1 (November 2018), https://web-assets.bcg.com/img-src/BCG-Why-Women-Owned-Startups-Are-a-Better-Bet-May-2018-NL_tcm9-193585.pdf [https://perma.cc/UBY8-RVGW] ("[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, e.g.*, John Mullins, *VC Funding Can Be Bad for Your Start-Up,* HARV. BUS. REV. (August 4, 2014), https://hbr.org/2014/08/vc-funding-can-be-bad-for-your-start-up [https://perma.cc/VN2E-WLPF] (cataloging a variety of costs to taking VC money).

rarity on the VC landscape give them sufficient bargaining power to demand *more* generous provisions?

The stakes in answering such questions are high and growing. Women and other underrepresented groups increasingly pursue entrepreneurial ventures as alternatives to conventional job ladders,[19] but they have limited information about what to expect from their financial backers. Similarly, prominent states (such as California) have begun to promulgate statutory protections prohibiting discriminatory treatment of female startup entrepreneurs,[20] but monitoring compliance with such mandates requires one to make comparisons *between* companies about how founders are treated within the governance environments they occupy. And put simply, our sightlines into these questions remain stubbornly obscured.[21]

Until now, that is. This paper deploys 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 typical VC investments foreordain not only that these foundational documents may evolve as the VC-backed company matures, but also that their initial structures can affect whether evolution occurs at all.

---

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

20. *See* CAL. CIV. CODE § 51.9 (West 2024) (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* Ann M. Lipton, *Capital Discrimination*, 59 HOUS. L. REV. 843, 843 (2022) ("[B]usiness law itself has no vocabulary to engage the influence of sex and gender, or to correct for unfairness traceable to discrimination."); 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 [https://perma.cc/8VN7-AFK2] (explaining how this bill "would explicitly make sexual harassment between venture investors and startup founders illegal").

21. 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 SCI. ADVANCES 1 (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).

22. *See, e.g.*, Robert Bartlett & Eric Talley, *Law and Corporate Governance*, *in* THE HANDBOOK OF THE ECONOMICS OF CORPORATE GOVERNANCE (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).

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. This study 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 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 "lucky" enough to receive VC funding.

To focus our inquiry, we draw on *the* constitutional governance document for startups: Certificates of Incorporation (or corporate "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 hard to collect (notwithstanding the fact that they are, in theory, public documents).[27] With considerable effort, however, we were able to obtain 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 sizeable matched sample of "similar" male-founded startups, enabling an apples-to-apples comparison of governance regimes.[28]

Our ultimate findings are simultaneously interesting and perplexing. At the most general level, we find a measurable gender difference in the linguistic content of our charters between female-founded and male-founded firms. More specifically, we employ computational machine learning methodologies to show that *charters of female-founded startups resemble their male-founded counterparts substantially less than male-founded counterparts resemble one another*. Broadly,

---

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*, 53 J. LEGAL STUD. 21, 23 (2024) ("The class conflict between common and preferred shareholders is particularly acute when the [VC-backed] company is deciding whether to continue operations or exit.").

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).

28. While the gender categories we use are binary in nature, we acknowledge that such categorizations do not encompass all gender identities. Our choice of these categories is due to constraints in our primary data sources (described below). Reliable data classifications that include other gender identities are, unfortunately, not available at present.

this finding suggests that female founders face a formal governance landscape that is predictably distinct from their male counterparts.

That said, the gender differences we uncover in the *semantic content* in charters do not alone reveal whether such distinctions are also embodied in governance terms that typically draw lawyerly attention. To explore this possible connection, we meticulously hand-labeled over six-dozen specific features in our data, including a variety of cash flow rights (such as liquidation preferences, anti-dilution rights, and conversion rights) and control rights (such as veto/approval rights,[29] fiduciary waivers,[30] and board representation). Our analysis of these dimensions reveals a surprisingly complex topography. While women founders appear to get the short end of the stick in some governance 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 receive more favorable treatment in others (such as participation rights of preferred stock, certain preferred veto rights, and fiduciary waivers for VC investors).[31] Many of the differences we do observe, in fact, are numerically modest and not statistically significant.[32] In the aggregate, we do not uncover systematic gender patterns in our analysis of substantive governance provisions, a finding that we find surprising, and one that stands in contrast to the predictions that typically emerge from economic theories of discrimination in market settings.[33]

Our findings—the first of their kind as far as we are aware[34]—have several intriguing implications. On one level, they raise something of a mystery for future scholars. Although we show female-founded firms "look different" from their male-founded counterparts in the linguistic structure of their governance documents, they do not appear to "act different" when measured by the aggregation of familiar governance terms. In other words, our comparison of substantive provisions ultimately reveals where gender differences *are not* located as opposed to where they *are*.[35]

29. *See, e.g.*, Sanga & Talley, *supra* note 25 (discussing the role and importance of veto/approval rights in corporate governance); PWP Xerion Holdings III LLC v. Red Leaf Res., Inc., C.A. No. 2017-0235-JTL, 2019 WL 5424778 (Del. Ch. Oct. 23, 2019) (same).

30. *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) (discussing the role of fiduciary waivers in corporate governance).

31. *See infra* Part IV.

32. *See infra* Part IV and Appendix B.

33. *See infra* Appendix C at 641-54 (developing a theoretical framework for VC-founder contracting in the presence of bias and/or gender-based preferences and generating predictions therefrom).

34. In a roughly contemporaneous (albeit slightly later) research effort to ours, Robert Bartlett extracts chartering histories of startups using a similar approach, but deploys it for a different purpose of assessing text-analytic trends in startup governance relative to canonical templates (such as the National Venture Capital Association model documents). Bartlett does not attempt to hand-label the chartering data, however, as we present here. *See* Robert Bartlett, *Standardization and Innovation in Venture Capital Contracting: Evidence from Startup Company Charters* (Sept. 2023), https://download.ssrn.com/24/02/09/ssrn_id4721999_code363066.pdf [https://perma.cc/YXW4-CPHQ].

35. There may, of course, be other substantive governance categories that we did not attempt to label that better explain the distinctive gender patterns we uncover in semantic content. We return to this question in Part V, *infra*.

Even so, and pending the resolution of the mystery above, our findings carry material implications. To appreciate them, it is important to keep in mind that our data do not emerge from a vacuum: most critically, every startup in our data set must have been previously successful in procuring at least one round of VC funding. Yet as noted above, prior work has documented that female founders appear to receive differential treatment *at the financing stage*—a finding that complicates the interpretation of gender differences that we can measure in governance structure. The fact that we uncover little systematic gender patterns in key governance terms could be consistent with multiple different hypotheses. First, it might imply that observed gender differences in VC funding stem from factors far more heterogenous and complicated than VCs' beliefs or preferences. Alternatively, it could mean that VC biases may exist only at the funding stage but dissipate afterwards (possibly through a variety of market pressures[36]). Or our findings might reflect a single stage in a complex series 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 are saddled with regimes substantively similar 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 VC-backed 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 persists over time. Part IV then focuses in on a large set of canonical cash-flow and control provisions that are frequently the focus of VC-founder negotiations. There we show that overall gender effects appear to be cross-cutting in direction and statistically modest on the whole. Part V discusses implications of our results, both for legal and social policy and for future researchers. Appendix C contains additional statistical robustness results, and it develops several game-theoretic economic models of discrimination that generate testable predictions to help frame our empirical findings.[37]

I. The Backstory on Venture Capital and Gender

Although our project develops a novel data resource, we hardly write from a blank slate—far from it. There are now developed academic and practitioner

36. *Compare* GARY S. BECKER, THE ECONOMICS OF DISCRIMINATION 128 (1957) (arguing discrimination "should, on average, be less in competitive industries than in monopolistic ones") *with* Dan A. Black, *Discrimination in an Equilibrium Search Model*, 13 J. LAB. ECON. 309, 309 (1995) (showing that search costs can perpetuate discrimination even with a large number of market participants).

37. *See infra* Appendix C.

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.

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,[38] 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.[39] 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.[40] Indeed, recent quasi-experimental 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

38. *See, e.g.*, Dana Kanze, Laura Huang, Mark A. Conley & 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), https://hbr.org/2017/06/male-and-female-entrepreneurs-get-asked-different-questions-by-vcs-and-it-affects-how-much-funding-they-get [https://perma.cc/LW9Q-2JF2] (reporting on interactions between VCs and founders at an NYC TechCrunch 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 [https://perma.cc/JU25-2KSU] (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), https://techcrunch.com/2018/09/18/the-gap-table [https://perma.cc/JA47-82HL] (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), https://knowledge.wharton.upenn.edu/article/vcs-arent-funding-women-led-startups [https://perma.cc/S93Z-DWKH] (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 [https://perma.cc/4UU3-897C] ("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.").

39. Recent experimental work suggests that startups themselves have a preference for male investors, which might itself contribute to the low representation of women in VCs. *See* Ofir Gefen, David Reeb & Johan Sulaeman, *Choosing Startup Investors: Does Gender Matter?*, EUR. CORP. GOVERNANCE INST. (Oct. 28, 2022), https://www.ecgi.global/sites/default/files/Paper%3A%20Choosing%20Startup%20Investors%3A%20Does%20Gender%20Matter%3F.pdf [https://perma.cc/8H7D-YTY5].

40. *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 [https://perma.cc/5W6W-CQ36] (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); Alison Wood Brooks, Laura Huang, Sarah Wood Kearney & Fiona E. Murray, *Investors Prefer Entrepreneurial Ventures Pitched by Attractive Men*, 111 PNAS 4427, 4427 (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).

the person pitching.[41] 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.[42]

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.[43] However, far less than that number receive VC backing (around 20.6% even by generous estimates).[44] 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.[45] Several other recent contributions corroborate the notion that women founders typically confront largely unfounded negative gender stereotypes related to future success.[46] 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 to face an even *more* unfriendly funding environment since the initial story began to break on Theranos and Elizabeth Holmes.[47]

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

41. *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 [https://perma.cc/96WN-NTHG] (explaining the influence of masculinity and femininity in how investors perceive entrepreneurs).

42. *See* Sophie Calder-Wang & Paul A. Gompers, *And the Children Shall Lead: Gender Diversity and Performance in Venture Capital*, 142 J. FIN. ECON. 1 (2021).

43. *See* PITCHBOOK, ALL IN: FEMALE FOUNDERS IN THE VC ECOSYSTEM (2022), https://files.pitchbook.com/website/files/pdf/2022_All_In_Female_Founders_in_the_US_VC_Ecosystem.pdf [https://perma.cc/2F3U-S2PP].

44. *See Female Founders Dashboard*, *supra* note 13.

45. *See* Abouzahr et al., *supra* note 16, at 1.

46. *See, e.g.*, Malin Malmström, Aija Voitkane, Jeaneth Johansson & Joakim Wincent, *VC Stereotypes About Men and Women Aren't Supported by Performance Data*, HARV. BUS. REV. (Mar. 15, 2018), https://hbr.org/2018/03/vc-stereotypes-about-men-and-women-arent-supported-by-performance-data [https://perma.cc/ZM9T-5GRM]; Candida Brush, Patricia Green, Kashmi Balachandra & Amy Davis, *The Gender Gap in Venture Capital: Progress, Problems, and Perspectives*, 20 VENTURE CAP. 115 (2018); John Paul Titlow, *These Women Entrepreneurs Created a Fake Male Cofounder to Dodge Startup Sexism*, FAST CO. (Aug. 29, 2017), https://www.fastcompany.com/40456604/these-women-entrepreneurs-created-a-fake-male-cofounder-to-dodge-startup-sexism [https://perma.cc/G6Y7-GX46].

47. *See* Griffith, *supra* note 10; 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 [https://perma.cc/S9Z2-7ESF] (tracking a decrease in funding to female-led startups and a decline in the "proportion of dollars to female-only founders").

powered empirical studies. Although certificates of incorporation (or "charters") and bylaws for *public* corporations can be found in scattered locations on the Securities and Exchange Commission's EDGAR website, they are poorly organized and exclude not-yet-public startups by definition.[48] 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.[49]

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),[50] as well as a hodgepodge of surveys and experimental settings for insights about startup governance. A recent paper by Jennifer S. 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.[51] In a different industry survey of more than 1,200 entrepreneurs across eight markets, researchers found that 35% of women entrepreneurs reported experiencing gender bias and, on average, raised 5% less capital than men.[52] There do exist a handful of thought-provoking empirical analyses of select corporate governance features around a limited set of cash flow rights and board structure in VC-backed startups,[53] 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

48. *See* Frankenreiter et al., *supra* note 23, at 23-24.

49. *Id.*

50. *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).

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

52. *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 [https://perma.cc/2RYR-5FHP].

53. *See, e.g.*, Steven N. Kaplan & Per Strömberg, *Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts*, 70 REV. ECON. STUD. 281 (2003); Paul A. Gompers, *Optimal Investment, Monitoring and the Staging of Venture Capital*, 50 J. 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,* 52 J. 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 Berglöf, *A Control Theory of Venture Capital Finance*, 10 J. L. ECON. & ORG. 247 (1994).

constituencies.[54] 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.[55] 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 relatively 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 an evident 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. There are many reasons to think that such patterns should be observable, however: from a theoretical perspective, there are sound reasons to predict that both taste-based and bias-based discrimination would likely manifest not only at the "extensive margin" of making an investment, but also on several "intensive margins" related to how founders and funders split cash flow and control rights.[56] And indeed, Appendix C of this Article develops several alternative theoretical models of gender discrimination by startup investors, each which delivers distinct testable predictions about how such phenomena (if present) should manifest through different patterns of cash flow rights, control rights, or both.

Given the sharpness of such predictions, the lack of empirical visibility around VC governance constitutes a significant constraint to our current knowledge—one that bears directly on the important policy implications. 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.

54. *See* Broughman & Fried, *supra* note 53, at 384 ("[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.").

55. *See* Amornsiripanitch et al., *supra* note 53, at 517 ("[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.").

56. *See infra* Appendix C at 641-54 (dynamic appendix available at https://drive.google.com/file/d/1nWO6BugjgJs3ujvKpm8sNBu3d4tkpeye/view).

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 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[57] and VC Experts "Genesis,"[58] 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 founders and their VC-backed companies over the period 2003-2021. We largely relied on PitchBook's designation to identify the gender of founders rather than devising our own, as PitchBook is a widely used industry database that systematically tracks founder characteristics.[59] Although the companies included in PitchBook and the founders for whom it provides gender information may not be fully representative—introducing potential biases in our analysis—in the absence of a more accurate and less biased alternative, we rely on PitchBook's data. 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-

57. *See Inform and Validate Your Intuition: PitchBook for Venture Capital*, PITCHBOOK, https://pitchbook.com/solutions/venture-capital [https://perma.cc/6T86-BZL8].

58. *See Genesis*, LAGNIAPPE LABS, LLC (last visited Mar. 9, 2025), https://lanyaplabs.com [https://perma.cc/VS79-TREQ].

59. 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 68.

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 technology to extract the textual content of the official certificates of incorporation (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,[60] we also observe the evolution of charters (and all their governance provisions) over time.[61] By assessing these charter terms and amendments, we can make inferences about the state and evolution of corporate governance structures at various funding stages of private corporations.

For our baseline analysis, we define a "female-founded" startup as any firm with at least one female founder. 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. This definition aligns with the PitchBook protocols in reporting on female-founded companies and reflects prior research suggesting that the presence of even a single female founder can meaningfully shape firm outcomes.[62] While this definition is broad, it captures the reality that female founders may face distinct governance challenges and funding dynamics even when they are not the majority on a founding team.

Moreover, 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.[63] Nevertheless, we recognize that this definition may be overinclusive, potentially incorporating firms where female founders play a marginal role in governance or strategic decision-making. This could influence governance outcomes and may dilute observed gender effects. However, defining "female-founded" is inherently challenging, as there is no single, principled way to determine the threshold at which a female founder's presence should be considered meaningful. To ensure our findings are not driven by this definitional choice, however, we also consider several

60. *See, e.g.*, DEL. CODE ANN. tit. 8, §§ 102(a)(4), 151 (West 2024).

61. 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 40. Thus, we consider the round-by-round corporate charter provisions and contemporaneous transactional information recorded in Genesis to reflect the output of those negotiations.

62. *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*, *in* PROCEEDINGS FOR THE INTELLIGENT TECHNOLOGIES FOR INTERACTIVE ENTERTAINMENT: 12TH EAI INTERNATIONAL CONFERENCE 145 (2020).

63. *See, e.g.*, Teare, *supra* note 47.

alternative definitions, such as firms where all founders are female or where a majority of the founding team is female. While these narrower classifications provide additional insight, they are broadly consistent with our baseline approach yet significantly reduce the sample size, reinforcing our decision to rely on the broader PitchBook definition in our main analysis.[64]

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). This is particularly challenging when investment outcomes, such as investment amount, can reflect both relevant outcome variables potentially influenced by founder gender—on which we would not want to match—and pre-investment differences between companies, such as the stage at which they seek funding, which may affect the amount raised and theoretically should be controlled for when matching firms. We are therefore careful to 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 round/series of that funding,[65] geographical region, and industry group. Ideally, we would match companies on a more detailed set of pre-investment characteristics, such as the stage of development at which they seek funding, but our matching is constrained by the available data. 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.[66] We then match each female-founded startup with its <u>three</u> closest neighbors in propensity-score space.[67] 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

64. Our results do not substantially change along most dimensions. *See infra* Part IV.

65. 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.

66. 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.

67. 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.

unique, per our strategy of sampling with replacement.[68] 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.

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%

68. As flagged above in note 59, 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.

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 an SMD below the widely accepted threshold of 0.1,[69] with only modest exceptions in sporadic industries.[70] 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.

69. *See* JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES (2d ed. 2013). Based on this criterion, Table 1 gives us appreciable confidence that our matching approach functioned well. To be sure, some of our matching variables are fairly coarse, and it is possible that they mask material forms of unobserved heterogeneity. Moreover, even though have made use of only a small number of matching variables prior to the first recorded investment (a deliberate choice), 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 already distorted by gender and thus correlated with other unobserved differences. That said, the covariate balance measures in Table 1 remain highly encouraging.

70. Even for the two industries where the SMDs exceed 0.1—Computers and Peripherals and Healthcare Services—the imbalance is minimal: in both cases, the SMD is only slightly above the conventional 0.1 threshold.

Figure 1a: Female-Founded Startups as Compared to Matched Male-Founded Startups (3:1), By First Deal Date

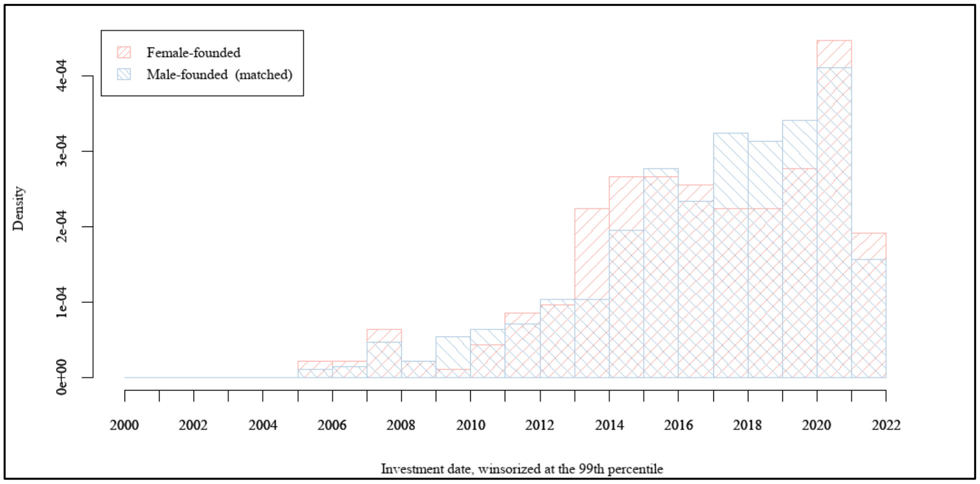
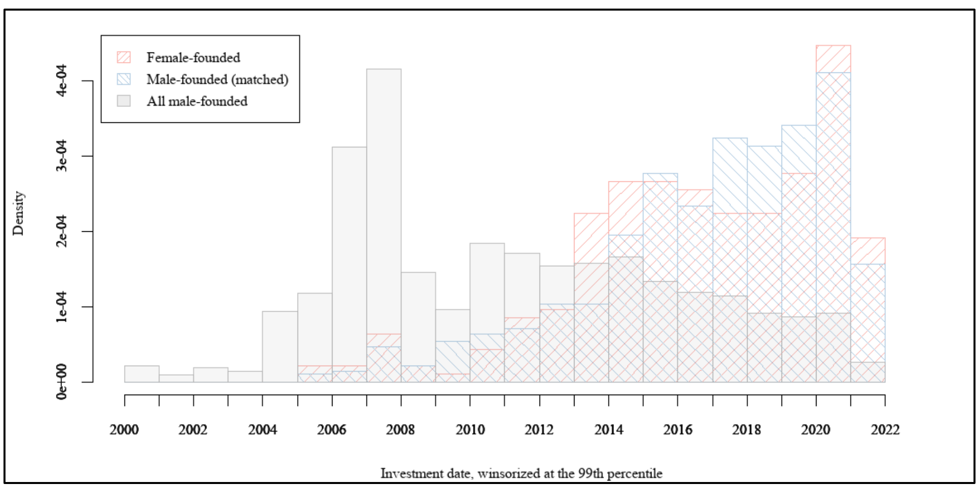


Figure 1b: Female-Founded Startups, Matched Male-Founded Startups, and All Male-Founded Startups, By First Deal Date



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 solid grey) 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 and female-founded firms. 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): there are clear differences in the timing of investment between male-founded firms and female-founded firms. A raw comparison across all firms could therefore conflate differences in governance terms with differences in investment timing, rather than isolating effects related to founder gender.

Figure 2a: Amount Invested in Female-Founded Startups as Compared to Matched Male-Founded Startups

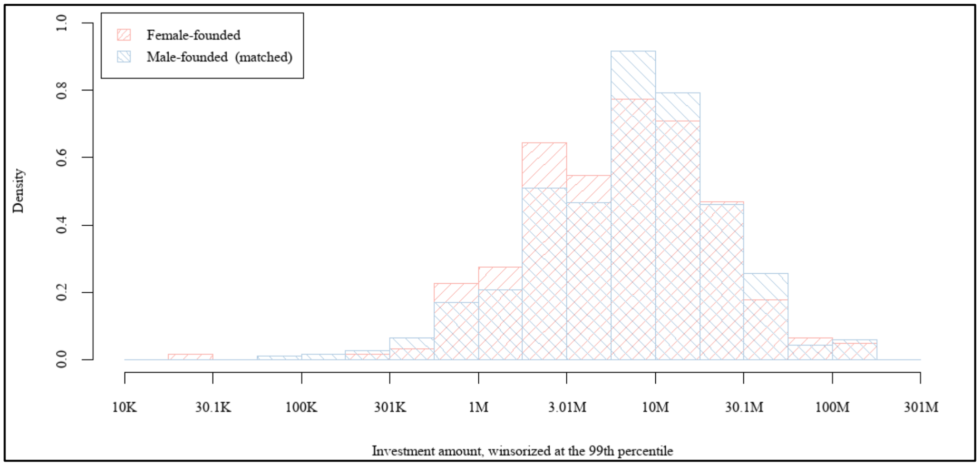
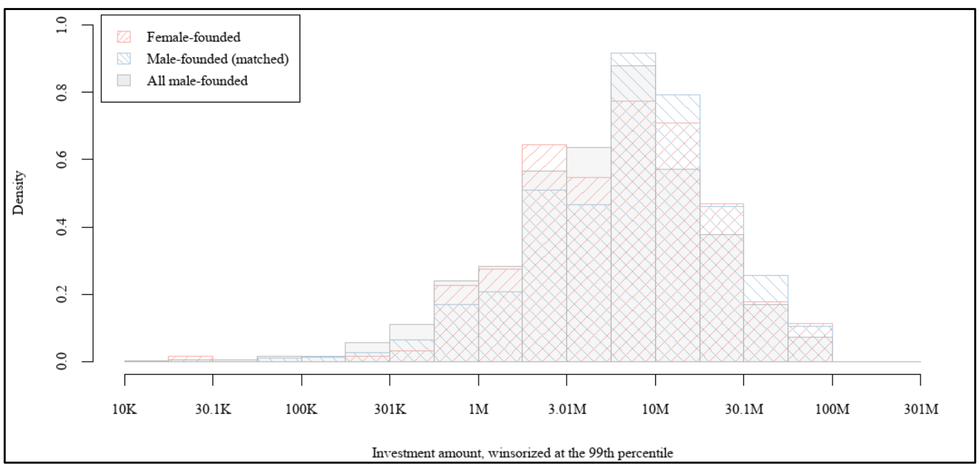


Figure 2b: Amount Invested in Female-Founded Startups, Matched Male-Founded Startups, and All Male-Founded Startups



C. Summary Characteristics

By limiting our comparison of female-founded firms to matched male-founded firms, rather than all male-founded firms, we can more confidently compare gender differences along a host of interesting dimensions. 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 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 several 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, which have a higher mean investment amount. These differences are substantially attenuated, however, when valuation is measured using 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 which 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 certificates of incorporation 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 become popular in legal scholarship,[71] and a burgeoning literature in business law has begun to use

71. *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 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, Christian Kellner & Alexander Stremitzer, *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).

them to great effect.[72] At their broadest level, machine learning approaches treat written texts as data,[73] 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.[74] 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 not.[75] 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.[76]

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

72. 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, 2022), https://ssrn.com/abstract=3455497 [https://perma.cc/F42B-TLEC] (using machine learning techniques to analyze EBITDA definitions in credit agreements); Rauterberg & Talley, *supra* note 30, at 1078 (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., Working Paper No. 560, 2020), https://ssrn.com/abstract=3124039 [https://perma.cc/BTW8-4SFE] (using machine learning techniques to study mutual fund voting patterns).

73. *See also* Frankenreiter et al., *supra* note 23, at 6-8 ("Lawyers can and must play a more central role in empirical corporate governance research, reclaiming the function for which they are professionally trained.").

74. *See also* Frankenreiter & Livermore, *supra* note 71, at 43 (giving examples of research using similar metrics).

75. 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 71, at 43-45.

76. *Id.* at 44 ("Depending on the techniques employed, the ensuing textual representation can be rather high dimensional, potentially necessitating an additional step of dimensionality reduction. Two main families of machine-learning approaches are commonly used to achieve this goal: supervised and unsupervised learning.").

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 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.[77] 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).[78]

77. For more details, *see* Frankenreiter et al., *supra* note 23.
78. *Id.* at 28.

There are various ways in which the differential treatment of female founders might manifest in high-level document metrics. For example, if male founders can extract more concessions from investors—particularly in the form of tailored charter provisions—one might expect the charters of female-founded companies to be shorter, have higher TTR, and be more readable. Conversely, if investors insist on non-standard clauses in negotiations with female-founded companies, the opposite may be true.

Figure 3a: Charter Contents, Length (in Words)

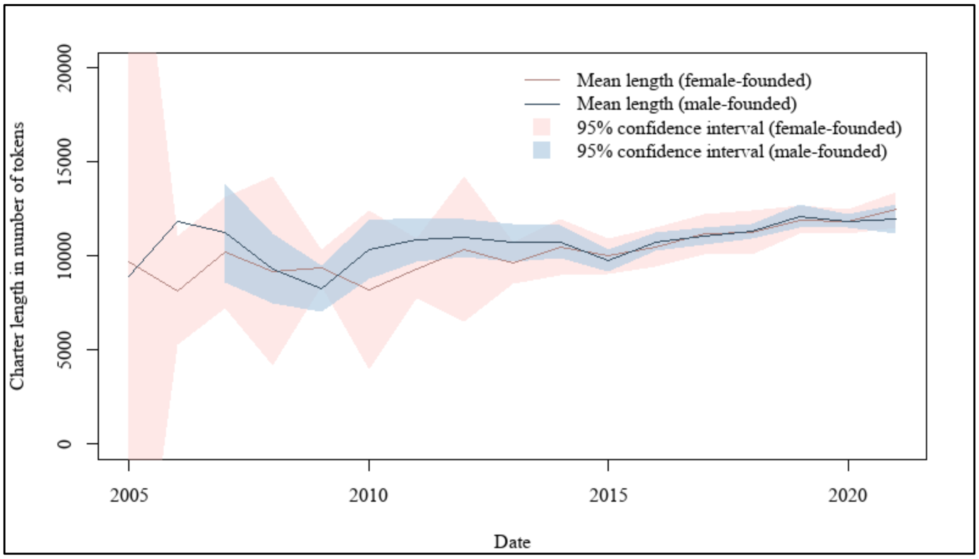


Figure 3b: Charter Contents, Readability (F-K Readability Score)

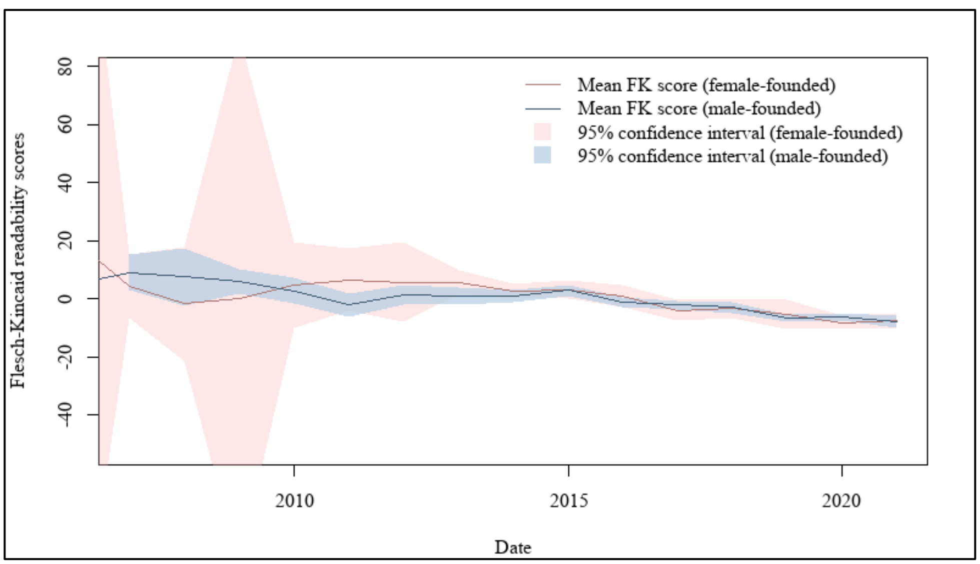


Figure 3c: Lexical Variation (Type-Token Ratio)

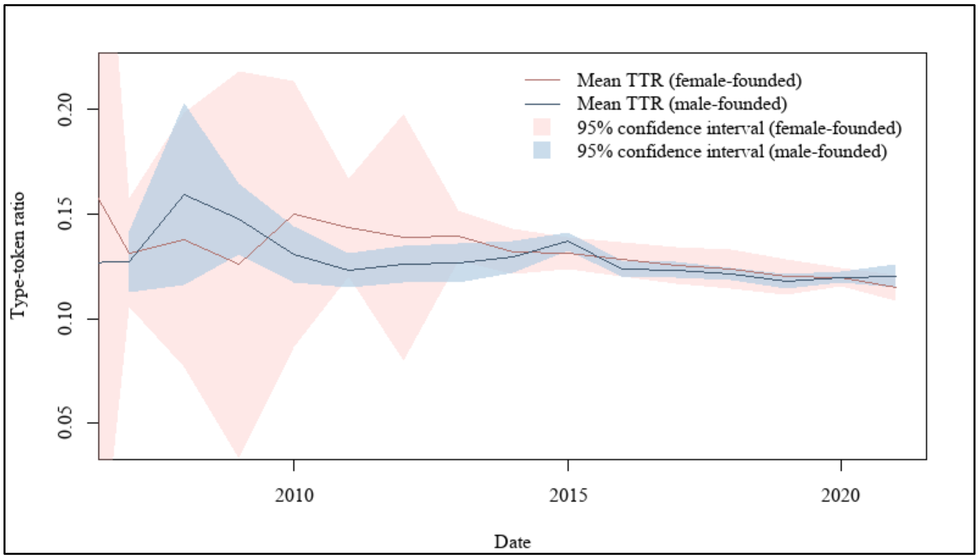


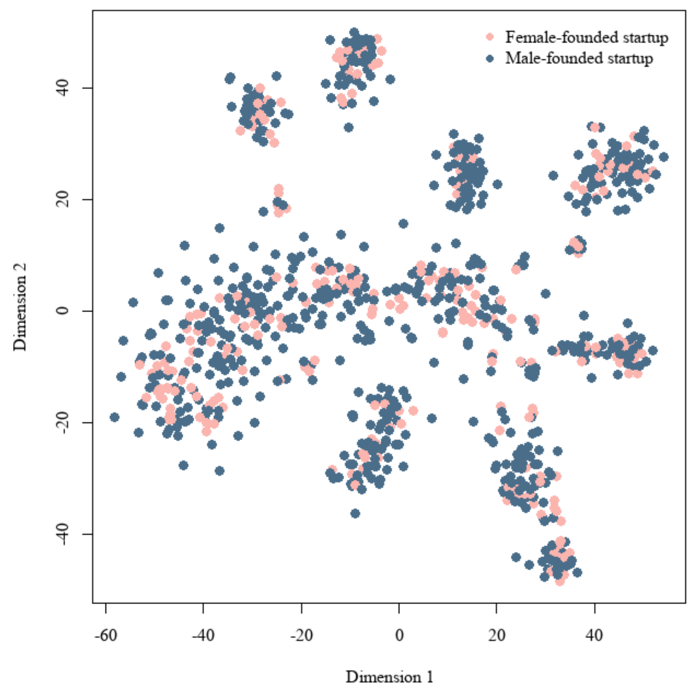
Figure 3 illustrates several notable trends. Charters tend to become longer over time (Panel (a)). However, there seems to be little difference between female- and male-founded firms: although male-founded charters appear lengthier than female-founded charters initially, the difference remains noisy 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.[79] Pink dots represent female-founded firms and blue 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.[80]

Figure 4: Two-Dimensional Charter Representations



79. 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.

80. 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.

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[81] 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 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)

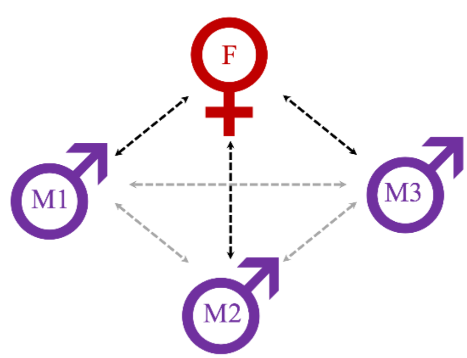
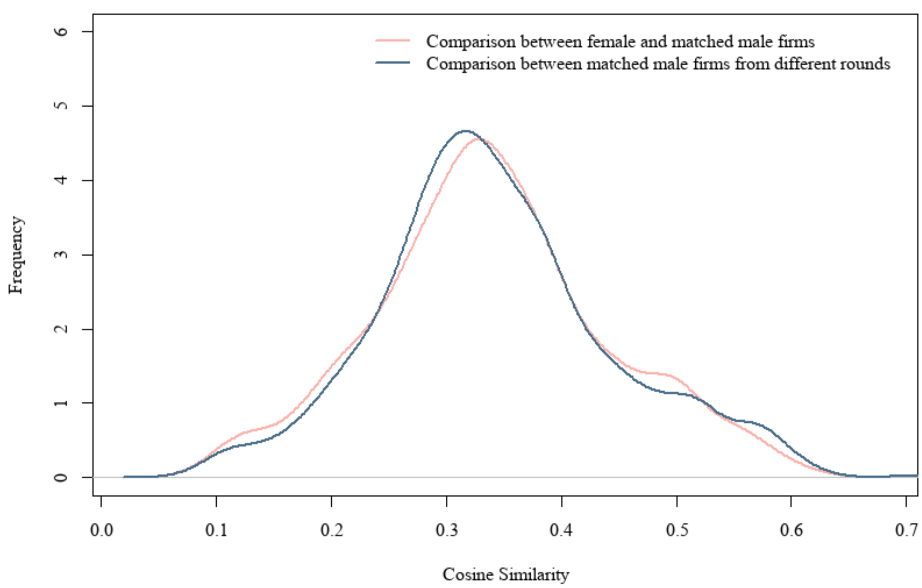


Figure 6 plots the smoothed population histogram of cosine similarities by comparison group, based on the content of the firm's first observable charter. The pink curve represents the density of similarities between female-founded firms and matched male-founded firms, while the blue curve represents the

81. The cosine similarity score measures the cosine of the angle between the vector representations of two documents. For documents represented as TF-IDF vectors, the score is bounded between a minimum of 0.0 (indicating no similarity) and maximum of 1.0 (indicating an identical distribution of words), and it measures the cosine of the angle between the vector representations of two documents. *See* Pozen et al., *supra* note 71, at 47 n.122.

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 be 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.[82] This eyeballing impression is also borne out in numbers: the average cosine similarity between female-founded and male-founded companies is slightly lower 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).[83]

Figure 6: Kernel Density of Cosine Similarity, Female-Male versus Male-Male Subpopulations



82. This divergence is especially pronounced for cosine similarity scores between 0.1 and 0.2.

83. 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).

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,[84] 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 various configurations, we settle on random forests using the first 1,000 principal components of the document-level feature vectors. Figure 7 depicts a receiver operating characteristic (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 (AUC) of the ROC as a prediction diagnostic, which will range between 0.5 (essentially random) to 1.0 (maximally predictive). While no universal standard exists for interpreting intermediate AUC scores, values between 0.6 and 0.8 are often considered fair to acceptable, while scores above 0.8 are usually regarded as good to very good.[85]

84. *See* Pozen et al., *supra* note 71, at 34.

85. Anne A.H. de Hond, Ewout W. Steyerberg & Ben van Calster, Comment, *Interpreting Area Under the Receiver Operating Characteristic Curve*, 4 LANCET E853 (2022).

Figure 7: ROC Curve (Full Matched Sample)

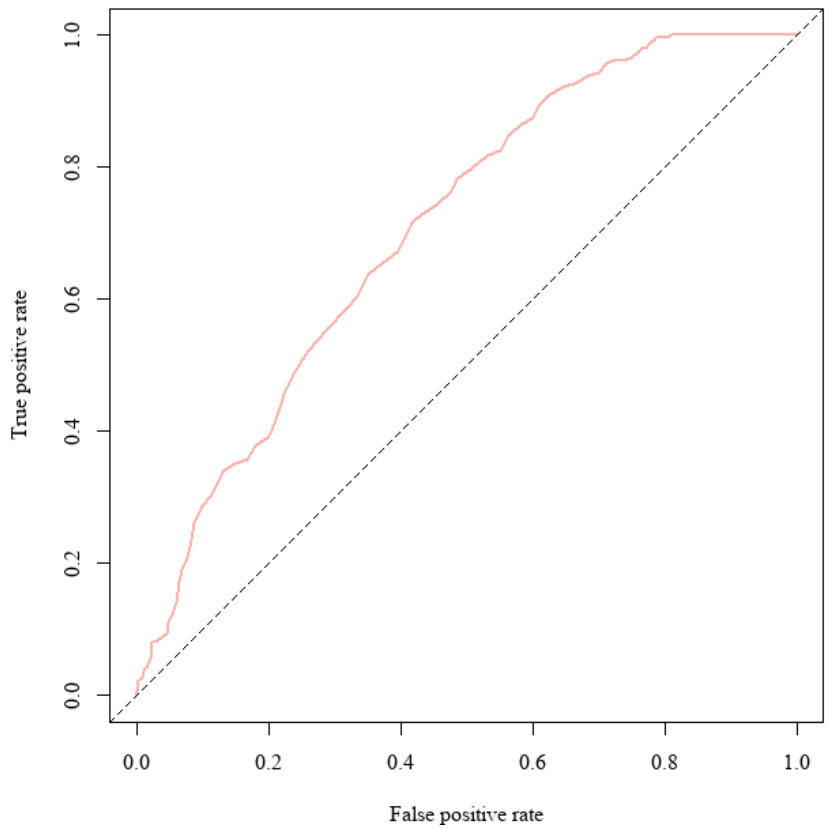
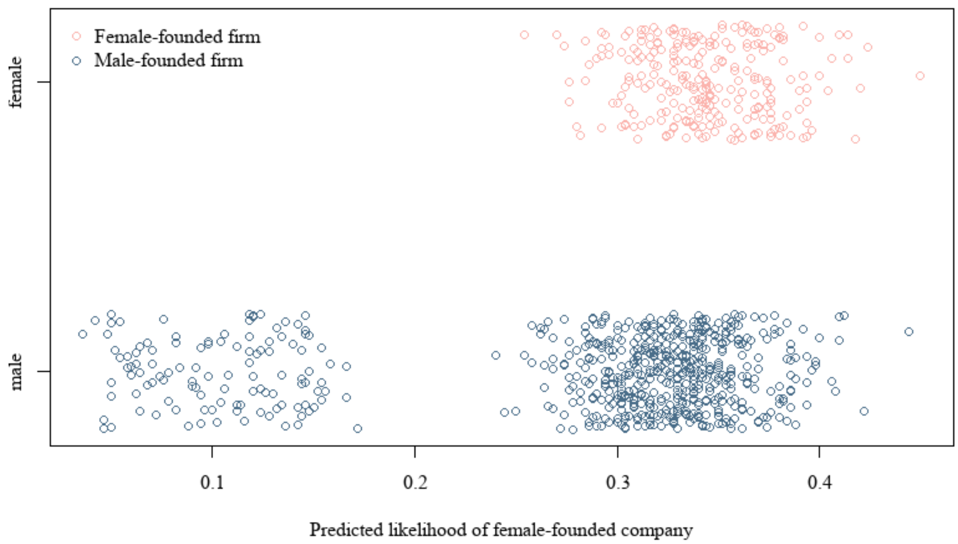


Figure 7 suggests that our algorithm performs well overall, substantially better than random guessing. A particularly noteworthy observation is that the algorithm reaches an almost perfect true positive rate at a false positive rate of around 0.8. This suggests that there are at least *some* male-founded 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 0.709, an F-1 value of 0.533, and an AUC (area under the curve) value of 0.699.

Figure 8 provides another perspective on the main finding described above. This figure depicts, on the horizontal axis, the predicted likelihood of a company being female-founded, based on the vocabulary used in its charter. On the vertical axis, the figure depicts the true outcome "label" associated with this company, i.e., whether the company is female-founded or not. Here it is also evident that the machine algorithm is mostly unable to correctly identify female-founded companies without also falsely labeling at least some male-founded companies as female-founded. In contrast, 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



We view these results to be consistent with the hypothesis that female-founded and male-founded firms receive governance regimes that manifest clear differences based on semantic content alone. This finding on its own, however, does not itself reveal the *precise nature* or significance of governance differences along gender lines. Do our results capture differences in the wording only (and are thereby epiphenomenal to governance)? Or, do these semantic differences betray substantive governance gender discrepancies within startups? Without a closer understanding of the substantive content of the charters, this question is hard to answer.[86] It is to that topic we next turn, examining startup governance along set of substantive cashflow and control provisions that frequently garner attention from practitioners, commentators, judges, and other relevant actors.

86. An alert reader might conjecture the semantic differences illustrated above are an artifact of law firm sorting—whereby female founders direct their legal work to a distinct set of firms than male founders. In Appendix C, we show that law firm sorting does not appear to be the culprit. While prominent VC law firms tend to produce semantically unique "clusters" of governance documents, we find no evidence that women founders direct business to them in proportions different from male-founded matches. *See infra* Appendix C at 654-57.

IV. Individual Levers of VC Governance

Our analysis in Part III suggested that there are measurable philological governance differences between female- and male-founded startups—differences that merit further analysis by exploring more granular inquiry into 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 in the Genesis database. Other labels were created by a team of research assistants that manually (and laboriously) 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, beginning first with the financial "cash flow" provisions and then considering non-financial "control" provisions.

A. Subsample Description

In the analysis below we focus on the cash flow and control rights provisions manifest in the *first observed 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 focus on the rights of preferred shareholders determined in the first observed round of investment, thereby attenuating the path dependency impact of earlier investment round terms on the investment terms of multiple successive rounds.[87]

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.

87. 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. 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.

Genesis: The Genesis database contains a record of investment deals for each firm and includes meta data 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 meta data 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,[88] 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 overwhelmingly 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.[89]

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).

88. *See, e.g.,* DEL. CODE ANN. tit. 8 § 102 (2024) (detailing required contents of a certificate of incorporation).

89. 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 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.

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. For 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 *do 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 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[90] relate most immediately to cash-flow rights,[91] 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.[92] We analyze provisions that relate more directly to governance and control rights in the next Part.

90. 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.

91. These cash-flow rights can significantly alter the true valuation of VC-backed companies. *See, e.g.*, Will Gornall & Ilya A. 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.").

92. *See, e.g.*, Ola Bengtsson & Berk A. Sensoy, *Changing the Nexus: The Evolution and Renegotiation of Venture Capital Contracts*, 50 J. FIN & QUANT. ANALYSIS 349 (2015) (exploding the relationship between cash flow rights, firm performance, and exit behavior).

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	0.338
(2) Liquidation Preference	98.8	98.7	0.1	1.000
(3) Participating Preferred	12.1	19.5	-7.4	0.006
(4) Redemption Rights	12.6	12.6	0	1.000
Dividends				
(5) Contractual Dividend	5.6	6.4	-0.8	0.763
(6) Cumulative Dividend (If Dividend=1)	47.1	37.9	9.2	0.578
Pro-founder Provisions				
(7) Pay to Play	3.5	5.4	-1.9	0.315
(8) Option Pool	57.6	57.1	0.5	0.942

We first consider the direction of the observed financing round, reported in Line 1 of Table 3. A financing round is said to be an “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 indistinct from matched male-founded firms (p=0.338).

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.[93] 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).

93. *See, e.g.*, Gornall & Strebulaev (2020), *supra note* 91 at 135; Robertt Bartlett III, *Understanding Price-Based Antidilution Principle: Five Principles to Apply When Negotiating a Down-Round Financing*, 59 BUS. LAW. 23, 29 (2003).

Though not technically the same as liquidation multiple, preferred shareholders sometimes have an additional right to periodic dividends.[94] 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.[95] Line 5 of Table 3 shows the frequency of a contractual dividend right in both our 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=0.763). 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=0.578).

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.[96] 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=0.006).

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 buy 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.[97] Line 4 of Table 3 illustrates redemption rights

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

95. *See, e.g.*, Broughman & Fried, *supra* note 93, 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).").

96. Klausner & Venuto, *supra* note 50, 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.")

97. 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.

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, forward-looking founders may seek to build in incentives encouraging early investors to renew their participation 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 tend to impose significant costs on investors who choose not to partake in later rounds—commonly through the forfeiture of antidilution rights.[98] 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%, p=0.315).

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=0.942).

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.

98. *See, e.g.*, Robert P. Bartlett, III, *Venture Capital, Agency Costs, and the False Dichotomy of the Corporation*, 54 UCLA L. REV. 37, 57 (2006) (explaining the penalties associated with a pay-to-play model).

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 in Table 4.[99] Board membership is typically viewed as a significant leverage point for influence, making it especially crucial in firm governance.[100] 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 (54.2%). Nevertheless, the difference is not statistically significant (p=0.610).

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	1.9	0.610
(2) Mean Ratio of Appointments	1.003	1.151	-0.147	0.040 (t-test)
Voting Rights				
(3) General Voting Rights	98.8	99.4	-0.6	0.420
(4) Changes to Business Plan	10.2	14.6	-4.4	0.090
(5) Changes to Articles of Incorporation	83.0	85.2	-2.2	0.421
(6) Changes to common / preferred stock (including number of shares)	96.1	96.4	-0.3	0.847
(7) Issuing new series of stock and/or new capital raising round	93.7	94.7	-1	0.527
(8) Incurring additional capital debt that has a security interest	34.5	33.6	0.9	0.819
(9) Incurring additional capital debt that does not have a security interest	41.6	44.9	-3.3	0.382
(10) Changes to board size or membership rules	81.5	86.2	-4.7	0.084
(11) Payment / Declaration of dividends or distributions	90.2	93.3	-3.1	0.100
(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	78.0	-4.7	0.145
(14) Mean Special Veto Rights Score	6.875	7.118	-0.24	0.066 (t-test)
Waivers				
(15) Waiver for Directors	82.8	85.8	-3	0.266
(16) Waiver for Shareholders	72.7	74.9	-2.2	0.508
(17) Waiver for Officers	12.2	11.7	0.5	0.911

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 shareholders versus those

99. 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.

100. 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 53, at 385 (showing that when VCs are in control of the board, there is less likely to be a deviation from previously agreed-upon cash flow rights).

controlled by common.[101] 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=0.040).

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.[102] 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.[103] 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 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 borderline statistically significant (p=0.090). In contrast, veto rights are common (and highly comparable) for other matters. For example, as to changes to the

101. 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 50, at 181 ("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." (footnotes omitted)).

102. On the distinction between protective provisions provided by special voting rights and the potentially more robust power of board appointments and board control, *see* Broughman & Fried, *supra* note 53, at 386 ("Unlike protective provisions, which give VCs the power only to block unfavorable transactions, board control enables VCs to replace managers as well as initiate fundamental transactions such as sales, IPOs, and dissolutions."); and Fried & Ganor, *supra* note 50, at 987 ("In short, board control gives VCs substantial power over and above whatever contractual provisions they have negotiated.").

103. 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.

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 difference is borderline statistically significant (p=0.066).

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.[104] 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.[105]

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 waiver protections for 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.[106]

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

104. *See, e.g.*, DEL. CODE ANN. tit. 8 § 122(17) (2024).

105. *See* Rauterberg & Talley, *supra* note 30, at 1089.

106. 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; 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.

more onerous provisions, while in some they face a less constraining road. Notably, most of the observed numerical disparities lack robust statistical significance. 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.[107]

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, as discussed earlier, 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.

Defining "female-founded" is inherently challenging, as there is no clear dividing line for when a female founder's presence should be considered meaningful. Various definitions—such as requiring a majority of female founders, equal gender representation, or female leadership—capture different aspects of gender dynamics but introduce their own trade-offs. In the absence of a single principled standard, for our main analysis, we follow PitchBook's primary classification.

To test the sensitivity of our results to our definition, 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 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

107. In Appendix B, *infra,* as well as Appendix C, *infra*, we deploy regression analysis techniques to reach similar conclusions as those discussed above.

the MFF category[108] and down to 50 for the FFF category.[109] 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	0.338	7.3	0.103	6.1	0.682
(2) Liquidation Preference	0.1	1.000	-0.7	0.602	-1.3	0.438
(3) Participating Preferred	-7.4	0.006	-6.5	0.086	-10.8	0.079
(4) Redemption Rights	0	1.000	0.7	0.882	-1.2	1.000
Dividends						
(5) Contractual Dividend	-0.8	0.763	-4.9	0.047	-3.4	0.454
(6) Cumulative Dividend (If Dividend=1)	9.2	0.578	17.1	0.642	13.6	1.000
Pro-founder Provisions						
(7) Pay to Play	-1.9	0.315	-2.2	0.385	-3.3	0.524
(8) Option Pool	0.5	0.942	1.5	0.770	-2.0	0.870
Board Appointments						
(1) Board Appointment Rights	2.1	0.61	3.7	0.493	-1.1	1
(2) Mean Ratio of Appointments	-0.147476	0.0403 (t-test)	-0.2	0.0128 (t-test)	-0.2	0.3002 (t-test)
Voting Rights						
(3) General Voting Rights	-0.6	0.42	-0.5	0.436	0.0	.
(4) Changes to Business Plan	-4.4	0.090	-2.9	0.476	-8.1	0.238
(5) Changes to Articles of Incorporation	-2.2	0.421	4.8	0.218	-1.2	0.811
(6) Changes to common / preferred stock (including number of shares)	-0.3	0.847	-1.2	0.59	0.0	1
(7) Issuing new series of stock and/or new capital	-1	0.527	-0.7	0.836	-2.6	0.501
(8) Incurring additional capital debt that has a security	0.9	0.819	-1.1	0.835	-2.9	0.734
(9) Incurring additional capital debt that does not have a security interest	-3.3	0.382	-2.9	0.558	-19.0	0.022
(10) Changes to board size or membership rules	-4.7	0.084	-6.6	0.081	-16.6	0.012
(11) Payment / Declaration of dividends or	-3.1	0.1	-3.8	0.15	-5.9	0.265
(12) Liquidation events	-3.8	0.006	-6.4	0.015	-10.6	0.031
(13) Entry into agreement as to merger, acquisition, asset sale, IPO, SPAC or other "exit" event	-4.7	0.145	-6.0	0.175	-5.2	0.438
(14) Mean Special Veto Rights Score	-0.243	0.0664 (t-test)	-0.3	0.154 (t-test)	-0.7	0.0297 (t-test)
Waivers						
(15) Waiver for Directors	-3	0.266	-6.5	0.071	-14.5	0.021
(16) Waiver for Shareholders	-2.2	0.508	-8.8	0.048	-17.7	0.02
(17) Waiver for Officers	0.5	0.911	1.0	0.765	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 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.

108. With 423 male-founded matches and 369 unique male matches.
109. With 150 male-founded matches and 140 unique male matches.

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 we detected along gender lines do not manifest in canonical governance measures, what is driving those differences? And secondarily, how do these results bear on more general policy questions related to gendered venture capitalism? We address each in turn below.

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.[110] 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, judges, and practitioners 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.[111] 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.

110. 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.

111. *See supra* Part IV.

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 term counts with the with the hand-coded 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 on the classifier's ability to predict founder gender.[112] 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.[113]

Figure 9: Machine Learning Classifier of Founder Gender

112. There is similarly little added predictive power when one moves from text only to text plus labels when measured by predictive accuracy (0.7218 versus 0.7195) or F-1 score (0.5324 vs. 0.5342).

113. 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 146-48 (2022).

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 non-substantive grammatical distinctions between charters that track gender (such as proper names, pronouns, gendered sub-industries, or law firm document templates)—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.

While there may be reasons to be skeptical about this final possibility, too, our study design worked hard to minimize such concerns. For example, our matched sample was constructed through industry matches, and thus it should have substantially controlled for hidden gender segregation (especially the problem of gendered sub-industries and their impact on charter content) by construction. Furthermore, governance documents generally tend to be gender neutral in their grammatical content, and it seems unlikely that such distinctions are driving our results. Finally, the entrepreneurs' choice of law firm does not appear to be driving the philological distinctions we detect, either. As shown in Appendix C, we cannot discern any strong gender patterns of startup sorting into top law firms that might be driving our results.[114]

To the extent that the "epiphenomenon" hypothesis has legs—and female founders face immaterially different governance regimes from their male counterparts—what are the policy implications of such a finding? Here we can offer two insights. The first relates our finding to underlying theories of how gender effects might manifest in contracting. As we demonstrate in Appendix C,[115] theoretical models of gender discrimination (based on either bias or gender preferences) tend to predict that gender effects *should* manifest *somewhere* in governance, either through cash flow rights or control rights (or possibly both). To the extent that our results fail to bear out those predictions, they are consistent with the hypothesis that gender effects become muted (or otherwise masked) once corporate governance design strategies are in play. The available structures may simply be too rigid or standardized to manifest meaningful gender effects.

Even if this is so, another important policy takeaway remains for those interested in gender effects over the "lifecycle" of a startup. As documented earlier,[116] a now sizeable literature documents areas where women startup entrepreneurs are disadvantaged in procuring VC funding. In such instances, one might hope that governance might "remedy" this funding imbalance, by allocating more attractive governance rights to previously disadvantaged women founders. Our findings

114. Appendix C contains a detailed investigation of how startups select among top VC-oriented law firms, and whether such sorting produces evident linguistic "clustering" in charters along gender lines (*e.g.*, as each law firm uses its preferred form charter as a starting point). While we indeed find evidence of semantic clustering of charters by law firm, female-founded startups appear to sort themselves amongst law firms in a manner that is statistically indistinguishable from their male-founded matches. *See infra* Appendix C at 654-57.

115. *Id.*

116. *See supra* Part I.

suggest that such "settling up" dynamics are not born out in the data. By the same token, our findings also suggest that initial funding imbalances are not *exacerbated* through subsequent governance choices. Nevertheless, a bottom line for policy makers remains that initial funding differences may well matter not just at the funding stage, as they can propagate forward in time, resisting later course corrections through the institutions of governance.

Above all, our results help additionally set a baseline measure of VC governance and gender for future researchers. As discussed above, several states have promulgated legislation that formally scrutinizes acts of alleged gender discrimination in VC markets, including firm governance.[117] 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 gender implications for startup governance 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 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 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

117. *Id.*

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
Joel Sontag
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, specifically predicting a smaller likelihood of participation preferred rights. 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.

In an Appendix C,[118] we investigate alternative specifications of these regressions, utilizing governance provisions and law firm identities to predict the existence of a female founder team. Our results are consistent with those in Table B1.

118. *See infra* Appendix C at 657-59.

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)	-0.0136	-0.0132	-0.0462***	-0.0471***	-0.0325***	-0.0342***	0.00705	0.00719	-0.00625	-0.00627	0.0686***	0.0681***
	(0.00846)	(0.00825)	(0.0131)	(0.0134)	(0.00875)	(0.00877)	(0.00754)	(0.00756)	(0.00491)	(0.00493)	(0.0128)	(0.0125)
ln(Investment)	0.00833	0.00919	0.018	-0.0187	0.0408***	0.0234*	-0.0145	-0.00807	0.0398***	0.0409***	0.0521***	0.0301
	(0.00808)	(0.00901)	(0.0151)	(0.0196)	(0.0111)	(0.0136)	(0.0113)	(0.0150)	(0.00859)	(0.0108)	(0.0180)	(0.0220)
Female Startup	0.00329	0.00277	-0.0779***	-0.0790***	0.00123	0.0018	-0.0124	-0.0129	-0.0174	-0.0176	0.0118	0.0103
	(0.00984)	(0.00970)	(0.0265)	(0.0265)	(0.0247)	(0.0247)	(0.0191)	(0.0191)	(0.0142)	(0.0143)	(0.0343)	(0.0342)
Series_A		-0.0082		0.0913**		0.0304		-0.0346		-0.00425		0.025
		(0.0114)		(0.0382)		(0.0279)		(0.0337)		(0.0129)		(0.0504)
Series_B		0.00202		0.176***		0.0864**		-0.0355		-0.014		0.112*
		(0.0146)		(0.0561)		(0.0405)		(0.0377)		(0.0253)		(0.0633)
Series_C		0.00387		0.306***		0.125*		0.00491		0.0202		0.243***
		(0.0167)		(0.0865)		(0.0663)		(0.0669)		(0.0581)		(0.0772)
Series_D_or_greater		-0.0577		0.162*		0.183**		-0.0576		-0.00028		0.0329
		(0.0721)		(0.0942)		(0.0916)		(0.0490)		(0.0586)		(0.119)
Observations	968	968	968	968	956	956	957	957	970	970	970	970
R-squared	0.049	0.056	0.142	0.165	0.186	0.196	0.054	0.058	0.149	0.15	0.233	0.243
Adjusted R-squared	0.0181	0.0213	0.1145	0.1349	0.1592	0.1664	0.0233	0.0228	0.1221	0.1193	0.2087	0.2153
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 clustered at the company level in parentheses. *** $p<.01$, ** $p<.05$, * $p<.1$

Appendix C

This Appendix supplements the analysis from the main text in two respects. First, we develop a theoretical framework to analyze strategic interactions between founders and venture capital investors (who potentially harbor conscious or unconscious gender biases). The framework generates predictions about whether/how VC biases plausibly manifest through cash flow rights or control rights (or both), thereby providing a basis for interpreting the empirical results shown in the main text. Second, we offer several additional empirical robustness results that were not included in the paper, including an analysis of whether law-firm effects might be partially responsible for the semantic patterns we uncover that differentiate charter texts along founder-gender lines.

Theoretical Framework of VC Gender Bias

We seek to develop a framework that is sufficiently flexible to analyze many species of gender biases, situating them within a transactional negotiation between (1) an entrepreneur-founder ("E") whose gender is observable (as well as other characteristics) and (2) a venture capitalist ("VC"), who may take founder gender into account when making decisions and forming beliefs and strategies.

The three versions of the model we investigate consider alternative types of possible VC discrimination against women founders (both belief-based and taste-based), and each has testable predictions about how such biases would manifest empirically. In each case, we assume that beyond the specific instantiation of bias being studied, the founder's gender is otherwise irrelevant to value-creation potential. After presenting each version of the model, we highlight (and prove) propositions that deliver empirical predictions, and we then compare those predictions to our empirical results from the main text.

Overall, the predictions generated by the modeling approaches presented below are difficult to square with our empirical findings, and in some cases are directly contradicted by them. For certain variations (*e.g.,* Model B below, involving gendered bias about a founder's managerial skills), our results could arguably be reconciled with a hypothesis of VC bias, but only in the special case where the venture capitalists believe that *both* male and female founders alike are highly susceptible to managerial errors. And even there, the reconciliation is a tenuous one, because it forces other model predictions that themselves conflict with well-known empirical results in prior literature.

Basic Setup

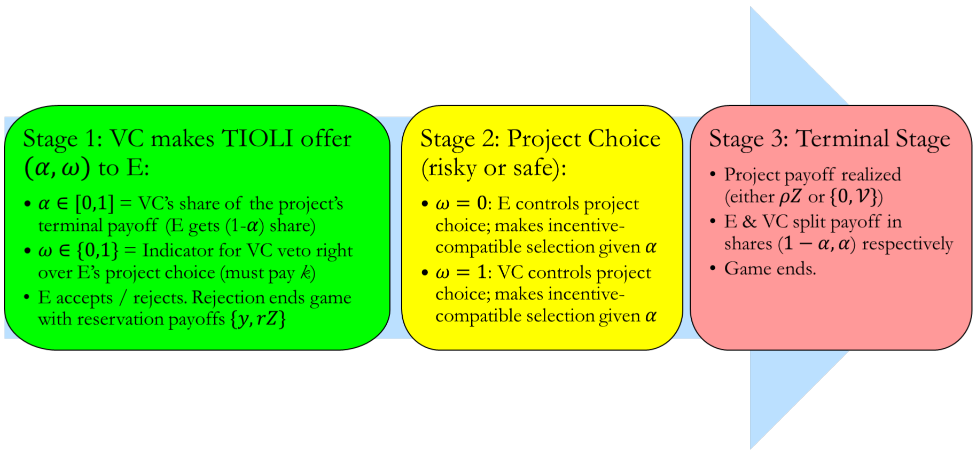
Before introducing the possibility of gender effects, we begin with a baseline version of the model to illustrate the key moving parts. Consider a contracting problem involving two players: an Entrepreneur ("E") and a Venture Capitalist ("VC"). To focus on core intuitions, suppose that both parties are risk-neutral and that there is no time discounting. The VC is able to invest in an outside project with known gross return r, but is considering investing instead in E's project should they execute a mutually beneficial contract. The entrepreneur E has outside option $y \in (0, \infty)$, which is privately known to E and is distributed according to a known cumulative distribution function $G(y)$; we assume that $g(y) \equiv G'(y) > 0$ for all y, and that the (inverse) hazard rate $\frac{G(y)}{g(y)}$ is everywhere increasing in y.

The startup venture needs a capital infusion of $Z > 0$ in order to operate, and E's illiquidity makes outside financing from the VC necessary if the project is to go forward. If the VC does not invest, the game ends, with E receiving the outside option y and the VC investing in the outside project with return r. If the VC and E do reach a deal, in contrast, E can select one of two projects, as follows:

- First, E can choose a "safe" project, where the gross return of ρ is distributed ex ante such that $\rho \sim H(\rho)$. The value of the safe project return is realized and becomes common knowledge prior to contracting.
- Alternatively, E can select a risky project, which has binary outcomes $\{0, \mathcal{V}\}$ corresponding with "failure" and "success," occurring with probabilities $(1 - q)$ and q, respectively. If the project succeeds, its gross payoff is $\mathcal{V}$. Its expectation, denoted as υ, is itself a random variable whose realization is unknown *ex ante*, other than its cumulative distribution of $F(\upsilon)$, where $F'(\upsilon) \equiv f(\upsilon) > 0$. We assume that E (but only E) observes the expected-value realization υ at no cost, at the time E makes a project selection choice.
- Depending on the contract's structure, the VC may have the right to override (or "veto") E's preferred project. But to exercise this right effectively, the VC must also be able to observe υ. To do so, we assume the VC has to pay an additional effort cost $k>0$ in order to learn υ. (The content of υ and $\mathcal{V}$ are otherwise assumed not contractible.)

The sequential structure of the game is as follows. In Stage 1, the VC makes a take-it-or-leave-it offer (α, ω) to E. The term $\alpha \in [0,1]$ denotes the VC's share of the project's terminal gross payoff irrespective of project choice (and E thereby gets the complementary (1-α) share). The term $\omega \in \{0,1\}$ represents an indicator variable for whether VC enjoys the right to override E's project choice (i.e., a "veto right"). When $\omega = 1$ the VC has a veto right and pays effort cost k to learn E's signal υ. (When $\omega = 0$ there is no veto right and no sunk effort cost.) E may accept or reject VC's offer. Rejection ends the game, and both parties get their reservation payoffs. If the offer is accepted, the parties move to Stage 2, involving the project choice (*i.e.,* risky or safe). Here, the type of contract entered into makes a significant difference. If $\omega = 0$, E controls project choice made by the firm. If $\omega = 1$, the VC controls project choice (but, as noted above, must pay k to learn υ). After the project is selected (by either E or VC), the game moves to Stage 3, the terminal juncture where the risky project's payoff $\{0, \mathcal{V}\}$ realized (if it was undertaken). E & VC split the payoff of whatever project the firm has pursued in agreed shares $(1-\alpha, \alpha)$, respectively, and the game ends. Figure C1 illustrates the sequence of play described above:

Figure C1: Sequence of Play



The attentive reader will notice that thus far, our framework affords no role for E's demographics (such as gender) to play in the contracting game. And indeed, if limited to the above framework, the model would predict no gender difference in the rate at which similarly situated ventures are funded, the cash flow provisions within such investments, or the control rights retained by the VC investors. Under this stripped-down "no-discrimination" version of the model, any observed differences in funding proclivities, cash flow rights, or control rights would be artifacts not of direct bias/animus by participants, but rather of orthogonal factors causing female and male startups to *not* be similarly situated or distributed.

That said, in the subsections below, we explore how a VC's bias/animus about gender can alter these predictions, considering three distinct model variations. In all versions, we will assume (without loss of generality) that the VC bias/animus proposed manifests only when E's gender "type" $S = F$ (female) and not when $S = M$ (male). We also assume that E's outside options and the quality of E's project do not hinge on E's gender identification.[119]

119. This assumption is not strictly necessary, but it sets a reasonable baseline for comparing how different types of bias/animus by VC affects equilibrium governance dynamics *ceteris paribus*, thus facilitating interpretation and testing. Another implicit assumption below is that k is sufficiently "moderate" in magnitude relative to *ex ante* expected project payoffs (ρ and υ) that VC would always find it profitable to invest and wrest control from E (paying k to learn the realization υ) rather than taking the outside investment return r. Relaxing this assumption complicates our analysis slightly but ultimately changes little.

A. VC Bias About E's Outside Options

We first consider a variation to the baseline model where the VC exhibits *bias in assessing E's outside options in the event of no investment*. This bias can matter because it can cause the VC to make more aggressive demands with funding if it believes that E has no realistic outside alternatives. Specifically, suppose that the VC may underestimate E's outside option to be $\hat{y} = y - \theta$, where $\theta \geq 0$. Larger values of θ connote greater degrees of VC bias. This structure allows us to consider how the VC might act differently in contracting if θ is "high" when $S = F$, but θ takes on a "low" value (normalized here to $\theta = 0$) when $S = M$. Once again, observe that given the assumptions of the model, any belief other than $\theta = 0$ is actuarially flawed. Nevertheless, if the VC exhibits a "naïve" bias, this actuarial mistake will persist uncorrected. The game otherwise proceeds exactly as before.

We seek to characterize perfect Bayesian equilibria to the game, and thus start with Project Selection in Stage 2. Suppose first that $\omega = 0$ (so that there is no VC override in the contract). E will pick the risky project if and only if its expected gains exceed that under the safe project:

$$(1-\alpha) \bullet qv \geq (1-\alpha) \bullet \rho Z \iff v \geq \frac{\rho Z}{q} \tag{1}$$

If ω =1, in contrast, (where the VC has control rights having sunk effort cost k), the VC will pursue the risky project if and only if:

$$\alpha \bullet qv \geq \alpha \bullet \rho Z \iff v \geq \frac{\rho Z}{q} \tag{2}$$

Note that the VC's criterion for pursuing the risky project precisely tracks E's. Consequently, should the game reach this stage, the VC would happily delegate authority to E as a loyal agent, thereby saving VC the superfluous verification costs k. Consequently, in this case the VC would never demand veto rights,and optimally confines its offer to contracts with $\omega = 0$.

Now consider the VC's optimal offer in Stage 1. Assuming (per the reasoning above) that the VC will optimally choose the no-veto contract ($\omega = 0$), E will accept the contract offer $(\alpha, 0)$ with probability:

$$G\left((1-\alpha)\left(\rho Z \bullet F(\tfrac{\rho Z}{q}) + \int_{\rho Z/q}^{\infty} qvf(v)dv\right)\right) \equiv G[(1-\alpha)\Lambda_0], \tag{3}$$

where $\Lambda_0 \equiv \left(\rho Z \bullet F\left(\frac{\rho Z}{q}\right) + \int_{\frac{\rho Z}{q}}^{\infty} q\upsilon f(\upsilon) d\upsilon\right) = E[\max\{\rho Z, q\upsilon\}]$. However, the VC's bias about E's outside option causes the VC to believe (erroneously) that E will accept the contract offer with probability:

$$G[(1-\alpha)\Lambda_0 + \theta] \tag{4}$$

Given this erroneous conjecture, the VC will set α to maximize its *perceived* expected payoff:

$$\Pi \equiv G[(1-\alpha)\Lambda_0 + \theta] \bullet \alpha\Lambda_0 + (1 - G[(1-\alpha)\Lambda_0 + \theta]) \bullet rZ \tag{5}$$

This structure implies the following proposition:

Proposition A: The unique equilibrium of Model A (bias about E's outside options) is characterized as follows:

1. *Funding Incidence: Equilibrium probability of VC funding is strictly decreasing in θ. Thus, as the VC exhibits more bias ($\uparrow\theta$), it will fund less frequently. If VCs manifest more bias for women founders, they will finance women at lower rates than male founders.*
2. *Control Rights: VC's demand for control rights is invariant in θ (VC sets ω=0 regardless of presence / absence of bias).*
3. *Cash Flow Rights: The VC's equity stake is strictly increasing in θ. Greater bias ($\uparrow\theta$) implies larger financial participation rights ($\uparrow\alpha$) for VC.*
   Returns: The VC's expected returns conditional on making an investment are increasing in its bias θ (e.g., against women founders). However, the VC's unconditional returns decrease in θ due to excess investment in the outside project with safe return.

Remarks: Parts (1) and (4) of Proposition A are consistent with existing literature as described above. Part (2) is also consistent with our empirical results. However, prediction (3) is not consistent with our results.

Proof of Proposition A: Part 2 was demonstrated in the text. The remaining components of the proposition come from the implicit function theorem as applied to the VC's and E's objective functions (in (4) and (5) from the above)

The first order conditions associated with the VC's problem are:

$$\frac{\partial \Pi}{\partial \alpha} = \Lambda_0(G[(1-\alpha)\Lambda_0 + \theta] \bullet \Lambda_0 - (\alpha\Lambda_0 - rZ) \bullet \mathrm{g}[(1-\alpha)\Lambda_0 + \theta]) = 0$$

Or equivalently:

$$\frac{G[(1-\alpha)\Lambda_0+\theta]}{g[(1-\alpha)\Lambda_0+\theta]}=(\alpha\Lambda_0-rZ)$$

Note as well that second order conditions for a maximum require:

$$\frac{\partial\Pi^2}{\partial\alpha^2}=-2\Lambda_0 g[(1-\alpha)\Lambda_0+\theta]+\Lambda_0(\alpha\Lambda_0-rZ)g'[(1-\alpha)\Lambda_0+\theta]<0$$

Moreover, observe that:

$$\frac{\partial\Pi^2}{\partial\alpha\partial\theta}=g[(1-\alpha)\Lambda_0+\theta]-(\alpha\Lambda_0-rZ)g'[(1-\alpha)\Lambda_0+\theta]$$

Applying the implicit function theorem, under an optimal contract $(\alpha^*,0)$, it follows that:

$$\frac{d\alpha^*}{d\theta}=-\left.\frac{\frac{\partial\Pi^2}{\partial\alpha\partial\theta}}{\frac{\partial\Pi^2}{\partial\alpha^2}}\right|_{(\alpha^*,0)}=-\frac{g[(1-\alpha)\Lambda_0+\theta]-(\alpha\Lambda_0-rZ)g'[(1-\alpha)\Lambda_0+\theta]}{(-2g[(1-\alpha)\Lambda_0+\theta]+(\alpha\Lambda_0-rZ)g'[(1-\alpha)\Lambda_0+\theta])\Lambda_0}$$

Because the denominator in this expression is also the second order condition, it must be negative at a maximum. After accounting for the leading negative sign, $\frac{d\alpha^*}{d\theta}>0$ if and only if the numerator is positive, or:

$$g[(1-\alpha)\Lambda_0+\theta]-(\alpha\Lambda_0-rZ)g'[(1-\alpha)\Lambda_0+\theta]>0$$

Substituting in from the first order conditions for $(\alpha\Lambda_0-rZ)$ and rearranging, this condition becomes:

$$g[(1-\alpha)\Lambda_0+\theta]^2-G[(1-\alpha)\Lambda_0+\theta]g'[(1-\alpha)\Lambda_0+\theta]>0$$

But we know that since *G(y)* exhibits a monotone hazard rate, the following must hold true for all *y*:

$$\frac{d}{dy}\frac{G(y)}{g(y)}=\frac{(g(y))^2-G(y)g'(y)}{(g(y))^2}>0$$

Thus, for all y it follows that $(g(y))^2-G(y)g'(y)>0$. This establishes part 3, and thus the VC's equilibrium demand for a fractional share (α^*) is increasing in bias. Part 1 immediately follows, recalling that E accepts the contract offer $(\alpha^*,0)$

with probability $G[(1-\alpha^*)\Lambda_0]$, which must be strictly decreasing in θ since $\frac{d\alpha^*}{d\theta} > 0$.

Part 4 follows from noting that conditional on funding, the VC gets expected payoff $\alpha^*\Lambda_0$. Recalling that Λ_0 is invariant in θ, we see that $\frac{d}{d\theta}(\alpha^*\Lambda_0) = \Lambda_0\frac{d\alpha^*}{d\theta} > 0$, and thus the VC's *conditional* payoff and returns are increasing in bias. Now consider the effect of enhanced bias on the VC's expected *unconditional* payoff (rather than perceived returns). The actual VC expected payoff is given by:

$$\Pi_{Actual} \equiv G[(1-\alpha)\Lambda_0]\cdot\alpha\Lambda_0 + (1 - G[(1-\alpha)\Lambda_0])\cdot rZ$$

Following from the analysis above, the first-best choice of α^{fb} that maximizes this expression is characterized by:

$$\frac{G[(1-\alpha^{fb})\Lambda_0]}{g[(1-\alpha^{fb})\Lambda_0]} = (\alpha^{fb}\Lambda_0 - rZ)$$

It is thus clear that $\alpha^{fb} \neq \alpha^*$ whenever $\theta > 0$, and hence VC's choice of α^* rather than α^{fb} reduces unconditional payoffs and returns ex ante.

Finally, note that since only E's type is private information and all information sets for the VC are reached with positive probability, we need not impose additional restrictions on the VC's beliefs about E's type beyond sequential rationality.

Q.E.D. ■

B. VC Bias About E's Managerial Skill

We now consider a variation of our framework involving a manifestation of possible VC bias *not* about E's outside option, but rather about E's managerial skill. In particular, suppose that VC believes that when E' observes signal υ (the project's expected upside value), E will unconsciously over-interpret the signal to be $(\upsilon + \theta)$, where $\theta \geq 0$. As before, larger values of θ reflect larger degrees of VC bias. This framing allows us to consider how the VC might distort its contracting behavior if it assumed that θ is "high" when $S = F$, but treated θ as "low" when $S = M$. Note once again given the other assumptions of the model, any belief other than $\theta = 0$ is actuarially flawed, but the VC exhibits a "naïve" bias and does not correct for this actuarial mistake. Moreover, for the sake of precision we assume VC's bias entails a presumption that E's proclivity to misinterpret the signal is *itself* naïve, and thus E fails to anticipate that proclivity if asked to manage the project.[120]

Characterizing the equilibrium again begins with the Project Selection stage and works backwards. Suppose first that $\omega = 0$ (the VC cannot override E's project choice). Here, the VC's believes E will be too eager to choose the risky project, doing so whenever E's optimistic expected gain exceeds the safe project payoff:

$$(1-\alpha) \bullet q \bullet (v+\theta) \geq (1-\alpha) \bullet \rho Z \iff v \geq \frac{\rho Z}{q} - \theta \tag{6a}$$

If E were "rational," in contrast, E would opt for the risky project only if:

$$(1-\alpha) \cdot q \cdot v \geq (1-\alpha) \cdot \rho Z \iff v \geq \frac{\rho Z}{q} \tag{6b}$$

(Our model, of course, supposes that E *is* rational and thus would—if given the opportunity—follow the criterion in (6b); however, VC's biased *presumption* about E's managerial skill causes the VC to conjecture that E will adhere to (6a)).

If ω =1, in contrast (so that the VC preempts E's judgment but must pay discovery cost k), the VC will pursue the risky project if and only if:

$$\alpha \cdot qv \geq \alpha \cdot \rho Z \iff v \geq \frac{\rho Z}{q} \tag{7}$$

A comparison (6a) to (7) reveals that, for any value of α, whenever $\theta > 0$ the VC believes that E will perceive the risky project to be more attractive than it actually is. In such cases, VC views E as "too eager" to chase the risky project, and thus the VC may think it necessary to retain a veto right—even at the cost k of learning E's signal. (Whether the VC does insist on control in equilibrium, however, will generally depend on the value of k relative to θ, as shown below).

Moving back to the VC's optimal offer at stage 1, the VC's optimal contract offer involves a more complicated choice between governance regimes than in the prior case. For each contingent regime (no veto and veto), the VC will attempt to calibrate the remaining contract terms to maximize its perceived expected payoff, incorporating in both E's probability of acceptance and the VC's expected payoff conditional on acceptance. Consequently, in analyzing the VC's optimal offer (α, ω) and E's willingness to accept it, it is necessary to consider the respective cases where $\omega = 0$ (no VC veto rights) and $\omega = 1$ (VC veto rights).

120. The model developed here can also be applied to situations where $\theta \leq 0$, so that the VC believes (erroneously) that E is unconsciously prone to *under*-interpreting the signal (or equivalently, E is underconfident about generating a successful outcome under the risky project). It is easy to adapt the analysis to accommodate this alternative framing. Doing so, however, yields virtually identical results as in Proposition B below. We therefore concentrate on the $\theta \geq 0$ case here.

Consider first the "no veto" contract with $\omega = 0$. Here, the VC will attempt to incorporate the biased expectation that E will be too eager to pursue the risky project, and that E will not be able to self-correct that inclination. The VC anticipates (here correctly) E will accept the contract only if E perceives the expected value of the contract to exceed E's outside option. This occurs with probability:

$$G\left((1-\alpha)\left(\rho Z \cdot F(\frac{\rho Z}{q}) + \int_{\rho Z/q}^{\infty} qvf(v)dv\right)\right) \equiv G[(1-\alpha)\Lambda(0)], \tag{8}$$

where $\Lambda(0) = \rho Z \bullet F(\frac{\rho Z}{q}) + \int_{\rho Z/q}^{\infty} qvf(v)dv$. It will prove helpful to think of $\Lambda(0)$ as a baseline value for the parameterized function $\Lambda(\theta) \equiv \rho Z \bullet F(\frac{\rho Z}{q} - \theta) + \int_{\frac{\rho Z}{q}-\theta}^{\infty} qvf(v)dv$. Note that $\Lambda'(\theta) = -q\theta \bullet f(\frac{\rho Z}{q} - \theta) < 0$ and that $\Lambda'(\alpha) = 0$.

Given E's acceptance decision, the VC believes (perhaps incorrectly) that E will unwittingly make optimistic errors in managing the project, and the VC will therefore choose α to maximize its expected payoff conditional in anticipation of those expected errors, which is given by:

$$G[(1-\alpha)\Lambda(0)] \cdot \alpha \cdot \Lambda(\theta) + (1 - G[(1-\alpha)\Lambda(0)]) \cdot rZ \tag{9}$$

Notably, the bias variable θ enters only once in the above expression, through $\Lambda(\theta)$, which corresponds to the VC's (biased) assessment of the startup's expected continuation value under E's management.

Now consider the "veto" contract $\omega = 1$. Here, the VC will override E's preferences in managing the project. With E no longer in the driver's seat, the VC conjectures that E will forecast that the VC will make decisions rationally as per expression (7) above. Accordingly, the VC anticipates that E will accept an offer $(\alpha, 1)$ with probability:

$$G[(1-\alpha)\Lambda(0)] \tag{10}$$

Note that this is the same probability as in (8), the "no-veto" case where $\omega = 0$. Indeed, under both scenarios, the VC conjectures that E's forward predictions are actuarially correct, and that E suffers only from faulty judgment at the moment of project choice.

The VC takes the above assessments into account when setting α to maximize its perceived expected payoff:

$$G[(1-\alpha)\Lambda(0)] \cdot \alpha \cdot \Lambda(0) + (1 - G[(1-\alpha)\Lambda(0)]) \cdot rZ - k \tag{11}$$

Observe that in (11), the VC's bias θ plays no role whatsoever. Having chosen to wrest control of the project from E (at cost k), the VC's contract offer no longer depends on marginal variations of θ, since the posited type of bias affects neither the probability that E accepts the VC's offer nor the VC's continuation payoffs afterward. Analyzing the two above cases gives rise to the following result:

Proposition B: The unique equilibrium of Model B (bias about E's managerial skill) is characterized as follows:

1. *Funding Incidence:*
   - *i.* <u>*Contracts with no Veto Rights* (ω = 0)</u>: *The equilibrium probability of VC funding is strictly decreasing in the VC's bias θ. As the VC exhibits more bias (↑θ), it invests marginally less.*
   - *ii.* <u>*Contracts With Veto Rights* (ω = 1)</u>: *The equilibrium probability of VC funding is invariant in the VC's bias θ. As the VC exhibits more bias (↑θ), its likelihood of investing remains constant on the margin, but is lower than in the absence of bias.*
2. *Control Rights: VC demands veto rights when its bias grows sufficiently strong, such that* $\theta \geq \theta_k^* > 0$. *If VC manifests stronger bias intensity against women entrepreneurs, it will also demand veto rights more frequently from women.*
3. *Cash Flow Rights:*
   - *i.* <u>*Contracts with no Veto Rights* (ω = 0)</u>: *The VC's equity stake (α) is strictly increasing in bias θ. Greater bias (↑θ) implies that the VC will demand larger financial participation rights on the margin (↑α).*
   - *ii.* <u>*Contracts With Veto Rights* (ω = 1)</u>: *The equilibrium VC equity stake α is invariant in the VC's bias θ, but is higher than in the absence of bias.*
4. *Returns:*
   - *i.* <u>*Contracts with no Veto Rights* (ω = 0)</u>: *The VC's average returns conditional on funding E are strictly increasing in θ. However, the VC's unconditional returns decline in θ due to excess investment in the outside safe project.*
   - ii. <u>*Contracts With Veto Rights* (ω = 1)</u>: *The VC's average returns (both conditional and unconditional) are invariant on the margin in θ, but are lower than they would be for all* $\theta < \theta_k^*$, *so that a no-veto contract would be optimal.*

Remarks: Part (1) of Proposition B is consistent with existing literature as described above. Part (4) is consistent with existing literature, but only to the extent that VCs offer contracts with no veto rights. Parts (2) and (3) are consistent with our empirical results only to the extent that VCs exhibits sufficient bias against *all types* of E and thus opts for a contract with veto rights regardless of E's type. Consequently, either Part (4) can be reconciled with the literature, or Parts (2) and (3) can, but not both.

Proof of Proposition B: Consider first the "no veto" contract, and the first order conditions for the VC's optimal choice of α in contract $(\alpha, 0)$:

$$\frac{\partial \Pi}{\partial \alpha} = G[(1-\alpha)\Lambda(0)] \cdot \Lambda(\theta) - g[(1-\alpha)\Lambda(0)]\Lambda(0) \cdot (\alpha\Lambda(\theta) - rZ) = 0,$$

or equivalently:

$$\frac{G[(1-\alpha)\Lambda(0)]}{g[(1-\alpha)\Lambda(0)]} = \left(\Lambda(0)\alpha - \tfrac{\Lambda(0)}{\Lambda(\theta)} rZ\right).$$

Note further that:

$$\frac{\partial^2 \Pi}{\partial \alpha \partial \theta} = -g[(1-\alpha)\Lambda(0)] \cdot \Lambda(0)\Lambda'(\theta) > 0.$$

Applying the implicit function theorem to the condition above (and second order conditions) it follows that the VC's optimal choice of α is strictly increasing in θ, or $\frac{d\alpha^*}{d\theta} > 0$. This establishes Part 3(i) of the Proposition.

Under a contract chosen optimally by the VC, the application of the envelope theorem implies:

$$\left.\frac{\partial \Pi^*}{\partial \theta}\right|_{\omega=0} = G[(1-\alpha^*)\Lambda(0)] \cdot \Lambda'(\theta) < 0$$

This establishes that in the space of no-veto contracts, the VC's perceived expected payoff is strictly decreasing in θ. The VC's <u>realized</u> expected payoff *conditional on funding*, however, is given by $\alpha^* \bullet \Lambda(0)$, which is clearly increasing in bias. The VC's unconditional realized expected payoff, however, is given by:

$$\Gamma(\theta) = G[(1-\alpha^*)\Lambda(0)] \cdot \alpha^* \cdot \Lambda(0) + (1 - G[(1-\alpha^*)\Lambda(0)]) \cdot rZ$$

Note that θ enters in the above only through α^*. Differentiating by θ and imposing the first order conditions for α^*, we obtain:

$$\Gamma'(\theta) = \left(\frac{G[(1-\alpha^*)\Lambda(0)]}{g[(1-\alpha^*)\Lambda(0)]} - (\alpha^*\Lambda(0) - rZ)\right) \times \frac{d\alpha^*}{d\theta}\Lambda(0) g[(1-\alpha^*)\Lambda(0)]$$
$$= \left(rZ\left(1 - \tfrac{\Lambda(0)}{\Lambda(\theta)}\right)\right) \times \left(\tfrac{d\alpha^*}{d\theta}\Lambda(0) g[(1-\alpha^*)\Lambda(0)]\right) < 0,$$

where the last inequality is due to the fact that $\Lambda(\theta)$ is strictly decreasing in θ. Consequently, the VC's unconditional realized expected payoff (as well as VC's return) is declining in bias, as reflected in Part 4(i) of the Proposition.

Now consider the "veto" contract $(\alpha, 1)$. Here, the VC will pick the risky project only when the expected payoff from the risky project (as adjusted by VC's bias about E's signal) justifies it. That is, if $\rho Z < q(\upsilon - \theta)$. The VC's expected payoff is therefore given by:

$$G[(1-\alpha)\Lambda(0)]\alpha \cdot E(max\{\rho Z, v\}) + (1 - G[(1-\alpha)\Lambda(0)]) \cdot rZ - k$$

Note that VC bias does not enter into this expression, because the VC is taking control of project management, and both parties believe E can make rational predictions ex ante. The first order conditions for an optimal choice of α imply:

$$\frac{G[(1-\alpha)\Lambda(0)]}{g[(1-\alpha)\Lambda(0)]} = (\Lambda(0)\alpha - rZ)$$

Because the VC's expected payoff is invariant in θ for a veto contract, but it costs the VC k to execute a veto contract, the VC will optimally chose a veto contract only when the VC's anticipated loss in the non-veto contract (analyzed above) exceeds k. Equivalently the veto contract will be chosen at the first point where the VC's bias exceeds a critical value $\theta \geq \theta_k^* > 0$, where θ_k^* is characterized by:

$$\int_0^{\theta_k^*} \frac{\partial \Pi^*}{\partial \theta}\bigg|_{\omega=0} d\theta = -q \int_0^{\theta_k^*} G\big[\big(1-\alpha^*(\theta)\big)\Lambda(0)\big] f\left(\tfrac{\rho Z}{q} - \theta\right) \theta \, d\theta = -k$$

So long as k is sufficiently "moderate" in magnitude, θ_k^* will exist and be finite. Thus, as bias increases it becomes more attractive for the VC to retain control as part of its (perceived) payoff-maximizing contract. This observation establishes Part 2 of the Proposition.

As for Part 3(ii), recall from (11) that the VC's optimal choice of α when $\omega = 1$ is such that it is locally invariant in bias, so that $\frac{\partial \alpha^*}{\partial \theta} = 0$. Having wrested decision control away from E (and paying cost k), the VC has effectively "bought" its way out of having to correct for E's perceived biases. Thus, at least on the margin, small changes in θ have no bearing on the VC's optimal demands for equity compensation, or on the VC's conditional returns. However, the VC's payoff is reduced by k any time the VC opts into the veto contract, and that it would do so only if its expected payoffs exceed the no-veto contract. We can thus conclude that the optimally chosen veto contract (where $\theta \geq \theta_k^*$) yields lower returns than an optimally chosen no-veto contract (where $\theta < \theta_k^*$).

Q.E.D. ■

C. VC-Based Associational Animus

Now consider a situation where the VC has no bias about E's talent or outside options, but instead simply has a taste or preference for doing business with E that depends on E's demographic profile. In particular, suppose that VC suffers disutility of θ simply for entering into a contract depending on E's "type" $S \in \{F, M\}$. A taste disfavoring women would be tantamount to a conditional ordering of $(\theta|F) > (\theta|M)$. In all other respects beyond the VC's tastes, assume it is common knowledge that E's talents and outside options are independent of S. Unlike the above model variations, under a taste-based model it is not possible to characterize the VC's preferences as being "actuarially" accurate or not. (They are simply preferences.)

We seek to characterize perfect Bayesian equilibria to the game, and thus we start with Project Selection in Stage 2. Suppose first that $\omega = 0$ (no veto). E will pick the risky project if and only if:

$$(1-\alpha)\cdot qv \geq (1-\alpha)\cdot \rho Z \iff v \geq \frac{\rho Z}{q} \tag{12}$$

If ω =1, in contrast, (veto but with sunk effort cost k), the VC will pursue the risky project if and only if:

$$\alpha \cdot qv \geq \alpha \cdot \rho Z \iff v \geq \frac{\rho Z}{q} \tag{13}$$

*A*s with Model A, the VC and E have identical criteria for pursuing the risky project. Consequently, in this case the VC would never demand veto rights, and instead optimally confines its offer to contracts with $\omega = 0$. (Indeed, there is some reason to think that the costs to the VC of retaining veto rights may even be higher in the case of associational animus, since vetoing the entrepreneur requires additional interaction between the VC and E.)[121]

Assuming the VC utilizes a no-veto contract, then, the VC will set α to maximize its *perceived* expected payoff:

$$\Pi \equiv G[(1-\alpha)\Lambda_0]\cdot(\alpha\Lambda_0 - \theta) + (1 - G[(1-\alpha)\Lambda_0])\cdot rZ \tag{14}$$

121. One can enrich our model further to give the VC a choice to abandon the project after learning v, preserving a real option to exit unless v is sufficiently high to overcome disutility θ. We conjecture that under such a variation, the VC would be more likely to retain control/veto rights against women founders, specifically to preserve this optionality. While we do not develop this variation here, we observe that the conjectured equilibrium appears inconsistent with our empirical results.

Note that as in (9), the bias parameter θ inters in only one part of (14). Analysis of this problem yields the following proposition:

Proposition C: The equilibrium predictions of Model C (associational animus) are as follows:

1. *Funding Incidence: Equilibrium probability of VC funding is strictly decreasing in θ. Thus, as the VC exhibits more animus ($\uparrow\theta$), it will fund less frequently. If VCs manifest more animus for women founders, they will finance women at lower rates than male founders.*
2. *Control Rights: VC's demand for control rights is invariant in θ; VC sets $\omega = 0$ regardless of presence / absence of animus.*
3. *Cash Flow Rights: Animus (θ) and the VC's equity stake (α) move together. Greater animus ($\uparrow\theta$) implies larger financial participation rights ($\uparrow\alpha$) for VC.*
4. *Returns: Suppose the VC's disutility costs from associational animus are not observable in financial returns. The VC's expected returns conditional on making an investment are increasing in its animus θ (e.g., against women founders). However, the VC's unconditional returns decrease in θ due to excess investment in the outside project with safe return.*

Remarks: Parts (1) and (4) of Proposition C are consistent with existing literature as described above. Part (2) is also consistent with our empirical results. However, prediction (3) is not consistent with our results.

Proof of Proposition C: The proof proceeds through steps nearly identical to those in the proof of Proposition A, and it is therefore omitted here.

Additional Empirical Results

The second portion of this Appendix includes a variety of robustness checks on our empirical results, relating to law firm effects as well as alternative regression specifications.

A. Law Firm Matching Effects

Although the core analysis in the text did not focus on law firm effects, our data set is sufficiently rich to include information about legal representation for the majority of our observations for both the female founder group and their matched counterparts. Overall, we have data on law firms for 75.3% of our matched sample (78.2% of our female-founded firms, and 74.0% of the (unique) male-founded counterparts). We present an overview of our findings below.

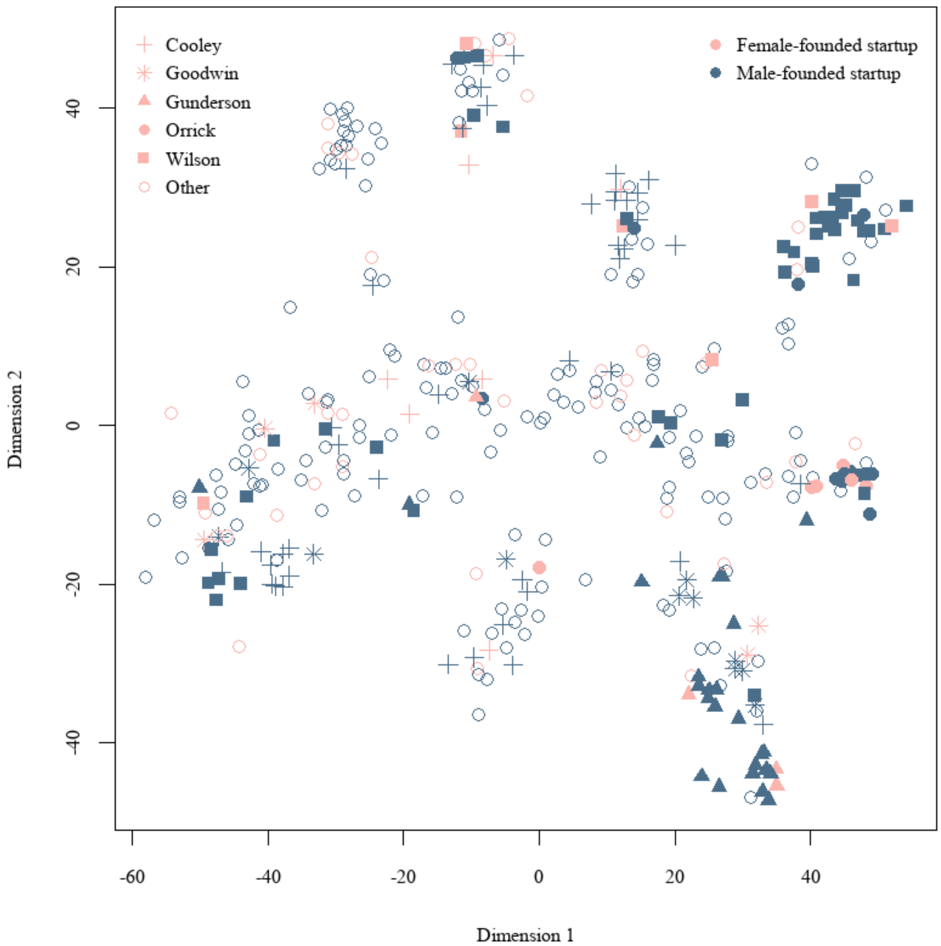
A threshold question—relevant to our inquiry and more generally—is whether governance documents exhibit clustering at the law firm level. This is an important question, since one possible explanation of the semantic distinctions we detect along gender lines (see Part 3 of the main text) may be driven by the ways that founders choose law firms, which in turn tend to follow proprietary templates when assembling governance documents. To investigate this possibility, we recreate Figure 4 from our paper, using only observations where we can identify the law firm for both the female-founded firm and at least one of its three matches is known to us. (If we observe legal representation information on more than one match, we always choose the closest match, resulting in a balanced one-to-one sample (n=400).) Figure C2 illustrates the results, using different symbols to identify each of the top five global venture law firms as reflected in PitchBook's League Tables.[122] We also include a sixth "residual" group representing all other firms not in the top five. The Figure also tracks founder gender, with circular markers denoting female founder teams and triangles representing matches).

It immediately becomes clear that at least four of the top five law firms manifest significant clustering in charters they produce for clients. Specifically, startups advised by Orrick and (to a slightly lesser extent) Wilson Sonsini exhibit highly distinct groups. Most startups advised by Gunderson and Goodwin also cluster together, too, but in a type of "collective" cluster, joined by a small number of other startups. In contrast, the charters of startups advised by Cooley appear to vary more.

At the same time, as in the main text, female- and male-founded charter embeddings appear in different locations, with no apparent tendency to cluster together with companies from the same group. Furthermore, Figure C2 does not reveal any noticeable correlation between the gender of the founders and their selection of a particular law firm. This provides at least some preliminary evidence that startups' law firm choices do not cleanly reveal gender patterns.

122. *See* Jordan Rubio, *Global League Tables*, PITCHBOOK (November 15, 2023), https://pitchbook.com/news/articles/global-league-tables-q3-2023 [https://perma.cc/2MCQ-JGP9].

Figure C2: Two-Dimensional Charter Representations



To get more of a quantitative handle on the law-firm matching question, we also examined whether startups tend statistically to "sort" to different law firms in a manner that tracks gender. Table C1 presents a comparison of female-founded startups and their closest match, once again only where law firm data are available. For each group, the Table illustrates the frequency of assignment to each law firm bin conditional on founder gender (rows therefore sum to 100%):

Table C1: Law Firm Sorting Among Female-Founded Startups and Closest Male-Founded Matches (n=400)

	Cooley	Goodwin	Gunderson	Orrick	Wilson	Other
Female-Founded Firms	14.5%	5.0%	8.0%	6.0%	11.5%	55.0%
Closest Male-Founded Matches	13.5%	4.0%	8.0%	3.5%	15.5%	55.5%

Eyeballing the table, it appears that the two groups of firms sort themselves very similarly across the six law firm bins. And indeed, from a statistical perspective, there is nothing to see here: Our data resoundingly reject the null hypothesis that female-founded firms and their male-founded matches are assigned identically across the tracked law firm bins ($\chi^2(5) = 2.80$; $p = 0.73$). This evidence underscores an assessment that our data do not evince strong—or even weak—law firm effects.

B. Alternative Regression Specifications

Appendix B of the main text presented a series of regressions where select governance provisions are treated as the dependent variable, while founder gender (among other variables) is treated as a control. While this specification simplifies interpretation in certain ways, we reiterate that we do not endeavor to test causal claims in this research (at least yet). Consequently, it may also be interesting to ask whether substantive governance provisions predict founder gender – that is, effectively treating gender as a "left hand" variable in regressions. (This alternative specification also helps square our regression results with our computational text analysis, where we trained a machine learning classifier to predict founder gender using a feature vector of semantic content from charters, which by definition includes the content of substantive governance provisions).[123]

The results of this alternative specification are included in Table C2 below, drawing on the full sample of female founders (Full Sample), the majority female founder (MFF) subsample, and the fully female founder (FFF) subsample. Some of the specifications also add in law firm identifiers—when available—categorized identically to the rubric described above (the omitted category consists of charters where representation was not from a top-5 law firm, our "other" classification in Figure C1).

Our results are largely consistent with Table B1 from Appendix B. On the one hand, there appears to be a consistent negative relationship between participating preferred provisions and female founders, suggesting that

123. We thank Roberta Romano for pointing out this consideration (convincingly).

participating preferred is (ceteris paribus) a counter-indicator of a female founder—an attribute that works to the benefit of female-founded firms. However, nearly all other variables have negligible (and/or inconsistent) predictive coefficients. Note also that the law firm indicator variables appear statistically insignificant from both one another and from the omitted ("other") category, suggesting yet again that there does not appear to be systematic "channeling" of women founded startups into favored firms (See Figure C1 and Table C1 above). A joint *F*-test in specifications [3], [6], and [9] supports this conclusion, strongly failing to reject the null hypothesis that the law firm coefficients are all equal to one another (p = 0.8199 (full sample); p = 0.7619 (MFF); p = 0.7321 (FFF), respectively). All told, we view these results as substantially confirmatory evidence of the empirical findings presented in the main text (and included appendices).

Table C2: Linear Probability Alternative Specification Dependent Variable: Female Founder Team

	[1] Female Founder (Full Sample)	[2] Female Founder (Full Sample)	[3] Female Founder (Full Sample)	[4] Female Founder (MFF Sample)	[5] Female Founder (MFF Sample)	[6] Female Founder (MFF Sample)	[7] Female Founder (FFF Sample)	[8] Female Founder (FFF Sample)	[9] Female Founder (FFF Sample)
ln(Valuation)	0.00328	0.00353	-0.0233	0.00142	0.00167	-0.0368	-0.0292	-0.0273	-0.0553
	(0.0126)	(0.0127)	(0.0228)	(0.0165)	(0.0165)	(0.0318)	(0.0273)	(0.0279)	(0.0872)
ln(Investment)	-0.013	-0.0183	0.0112	-0.0327	-0.0442*	-0.0331	-0.0512	-0.0411	-0.119
	(0.0158)	(0.0186)	(0.0377)	(0.0213)	(0.0251)	(0.0523)	(0.0315)	(0.0386)	(0.0816)
Liquidation Preference	-0.000437	-0.0114	-0.27	-0.2	-0.193	-0.125	-0.425	-0.403	-1.458
	(0.140)	(0.140)	(0.264)	(0.209)	(0.205)	(0.321)	(0.393)	(0.396)	(1.731)
Participating Preferred	-0.109***	-0.117***	-0.181*	-0.106**	-0.111**	-0.214	-0.193**	-0.186**	-0.890*
	(0.0391)	(0.0407)	(0.0997)	(0.0524)	(0.0538)	(0.148)	(0.0813)	(0.0848)	(0.440)
Redemption Rights	0.00834	0.00764	-0.0509	0.0916	0.0897	0.221	0.0126	0.0288	-0.583
	(0.0498)	(0.0502)	(0.110)	(0.0722)	(0.0733)	(0.170)	(0.112)	(0.115)	(0.666)
Contractual Dividends	-0.0268	-0.0266	-0.0524	-0.159**	-0.155**	-0.341	-0.0734	-0.0846	0.821
	(0.0625)	(0.0627)	(0.136)	(0.0645)	(0.0654)	(0.219)	(0.152)	(0.155)	(0.779)
Pay to Play Provision	-0.0404	-0.0397	-0.237	-0.0163	-0.0153	-0.328	0.109	0.111	1.261**
	(0.0664)	(0.0670)	(0.163)	(0.0892)	(0.0905)	(0.230)	(0.165)	(0.165)	(0.571)
Option Pool	0.00245	-0.00174	-0.055	0.00204	-0.000933	-0.0187	-0.0221	-0.0208	-0.165
	(0.0344)	(0.0347)	(0.0807)	(0.0468)	(0.0475)	(0.114)	(0.0830)	(0.0853)	(0.227)
Series_A		0.0105	-0.03		0.0486	0.0289		-0.0329	0.114
		(0.0439)	(0.0927)		(0.0593)	(0.131)		(0.105)	(0.249)
Series_B		0.0247	-0.0195		0.0537	0.0216		-0.0908	0.23
		(0.0567)	(0.114)		(0.0771)	(0.158)		(0.128)	(0.349)
Series_C		0.0926	0.0237		0.0779	-0.0106		0.00604	0.864
		(0.0914)	(0.159)		(0.144)	(0.277)		(0.244)	(0.536)
Series_D_or_greater		-0.0392	0.121		0.0156	0.261		0.0353	0.317
		(0.0966)	(0.201)		(0.139)	(0.300)		(0.261)	(0.796)
Cooley (dummy)			0.0922			0.177*			0.171
			(0.0892)			(0.104)			(0.313)
Goodwin (dummy)			-0.0133			-0.0872			-0.158
			(0.141)			(0.272)			(0.290)
Gunderson (dummy)			0.0261			0.101			0.396
			(0.118)			(0.170)			(0.238)
Orrick (dummy)			0.119			0.222			0.697*
			(0.140)			(0.149)			(0.348)
Wilson (dummy)			0.0529			0.138			-0.0768
			(0.0937)			(0.140)			(0.399)
Observations	943	943	311	515	515	163	184	184	65
R-squared	0.034	0.036	0.101	0.074	0.076	0.196	0.232	0.236	0.447
Adjusted R-squared	-0.003	-0.0056	-0.0355	0.0083	0.0018	-0.0334	0.0812	0.0616	-0.1412
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors clustered at the company level in parantheses. *** p<.01, ** p<.05, * p<.1

About ECGI