How do Venture Capitalists (actually) make investment decisions?
Internal evidence from a private startup accelerator

Tristan Fitzgerald
Texas A&M University
tfitzgerald@mays.tamu.edu

Korok Ray
Texas A&M University
korok@mays.tamu.edu

December 2025

ABSTRACT

Using a proprietary dataset detailing all startup applications, internal judging scores, judges' written comments, and even audio recordings of interviews involving one of the largest VC-backed startup accelerators in the United States, we open the 'black box' of VC investment decision-making. To do so, we study the entire internal VC investment selection process from the evolution of judging scores across interview stages to the voting rules applied for making final portfolio firm decisions. We first find that the evaluation of the *same* startup applicant by the *same* VC partner can differ by up to 25% depending on whether the partner is judging autonomously ('solo evaluation') vs. in collaboration with their colleagues ('group evaluation'). We trace part of these 'within-applicant' judging disagreements to the presence of significant VC judge-startup founder 'homophily' biases that are amplified in solo judging environments. Second, we document the value implications of different judging policies for VC fund performance. We find that group evaluations consistently outperform solo evaluations, while 'single champion' decision rules produce investment outcomes with much higher variance compared to 'consensus'-based decision rules. Finally, we show that one of the key mechanisms explaining the outperformance of group evaluations is that collaborative Q&A sessions encourage judges to ask interview questions and submit written explanations of scores that focus more on startup firm "fundamentals" rather than affinity-based founders' traits.

Keywords: Venture capital, entrepreneurial finance, business accelerators, portfolio firm selection, homophily, optimal voting rules, innovation

JEL Classification: G24, G30, G41, L26, M13

We greatly appreciate the generous time and assistance of our anonymous venture capital fund in supporting this research project. We would also like to thank Thomas Boulton, Juanita Gonzalez-Uribe, Isaac Hacamo, Alexander Montag, Fiona Paine, Alminas Zaldokas (discussants), Espen Eckbo, Alex Edmans, Peter Klein, and Ting Xu as well as all conference and seminar participants at the 2025 Olin Finance Conference at WashU, NFA Annual Meeting (2025), EUROFIDAI-ESSEC Paris Finance Meeting, 2025 JCF Special Issue Conference, 2025 Boca Finance and Real Estate Conference, EFA Annual Meeting (2025), Aggie Longhorn Innovation Conference, Baylor University and Texas A&M University, for their invaluable feedback.

1. INTRODUCTION

All organizations face two critical questions when making corporate decisions: (1) should employees work independently or collaboratively to evaluate alternative options; and (2) what is the optimal 'rule' for aggregating individual employee opinions to make a final corporate decision?[1] Importantly, there is substantial disagreement among academics and practitioners about whether a person makes more rigorous and insightful judgments when assessing possible options individually (i.e., on a 'solo' basis) or as part of a team.[2] Similarly, there is much debate about the relative merits of requiring a consensus among a group of employees before the firm proceeds with a particular course of action vs. only requiring that at least one person advocates for (or 'champions') an idea.[3]

A firm's chosen judging policies and decision rules are especially impactful in the context of corporate investment decisions, where binary investment choices must often be made with limited information, ambiguous judging criteria, and large internal disputes among employees about the firm's optimal investment strategy. This is especially prevalent in the case of venture capital (VC) firms, where VC firm partners from a wide range of personal and professional backgrounds must judge thousands of applications to invest in very high risk, early-stage startup companies in order to select a limited number of startups for multi-million dollar investments (e.g., Ewens, Nanda, and Rhodes-Kroph, 2018; Malenko, Nanda, Rhodes-Kropf, and Sundaresan, 2024).

Unfortunately, despite the critical role of judging procedures in the efficient allocation of corporate resources and broader economic growth, relatively little is known empirically about how organizations develop and synthesize various employees' opinions to make corporate decisions.

[1] For example, in the context of employment/hiring decisions, the firm must consider: (1) should employees conduct separate (one-on-one) interviews with a job candidate or organize a group meeting where several firm employees jointly interview the candidate; and (2) should a firm apply a 'majority voting rule' based on some consensus evaluation across all interviewers or prioritize a certain individual's opinion to select which candidate is offered the job?
[2] See studies comparing the trade-off between the promotion of more diverse ideas and efficient learning within teams (Lazear & Shaw, 2007; Jones, 2021) vs. potential for group think (Janis, 1972) and social loafing (Ringelmann, 1913).
[3] See e.g., Sah and Stiglitz (1986, 1988) and Malenko, Nanda, Rhodes-Kropf, and Sundaresan (2024).

For example, research on VC portfolio firm selection has been severely hampered by the lack of comprehensive data on the entire VC investment process, from (1) observing all startup applicants in VCs' initial investment opportunity set to (2) the internal evaluation of startups by individual VC employees to (3) the process for aggregating differing VC partners' assessments to select portfolio firms (Gonzalez-Uribe, Klinger-Vidra, Wang, and Yin, 2023). As a result, empirical researchers have typically been constrained to attempting to infer key features of the VC selection process using only very limited information on VCs' completed investments into a small subset of startups (see generally, Hall and Lerner, 2010; Da Rin, Hellmann, and Puri, 2013; Lerner and Nanda, 2020).

In this paper, we open the 'black box' of internal corporate project evaluation by studying the confidential dataset of a large startup accelerator program that is run by a prominent U.S. VC firm. This proprietary dataset contains comprehensive information on all startup applicants to our VC's accelerator program (totaling over 7,000 unique startups from 6 different continents and covering all major industries), all internal judging scores and written comments submitted by individual VC firm employees (comprising over 18,500 individual VC judge scores across three separate interview stages), and even all available audio recordings of interviews conducted by our VC firm with accelerator applicants. As a result, our unique research setting allows us to "get inside the corporate decision-making room" and observe how VCs evaluate investment opportunities in practice.

There are many unique benefits of our proprietary VC-backed accelerator dataset and research setting relative to the existing literature. First, our investment setting is both *realistic* and relatively *high stakes*. Specifically, the full-time VC investment professionals in our setting are exclusively focused on maximizing financial returns[4] when deciding how to allocate millions of dollars of

[4] This contrasts with other settings such as not-for-profit accelerators and new venture competitions where startups typically compete for small (non-equity) cash prizes, the startup evaluation criteria may not be 100% commercially focused (e.g., support the local economy, promote social goals etc.), and the judges will often include non-investment professionals such as academics and government representatives.

externally sourced investment capital amongst real startup firms.[5] Second, our sample is *complete* and *comprehensive*. This is because our VC firm provides us with the entire record of all application materials and judging assessments across all accelerator cohorts in a 3-year period. As such, we do not face the omitted variable bias concerns typically encountered by prior researchers.[6] In addition, our rich startup applicant–VC judge–cohort linked dataset allows us to employ novel identification strategies that simultaneously combine startup firm, individual judge, and time fixed effects. Third, our setting exemplifies *individual and group decision-making under high uncertainty*. For example, our setting allows us to study the same VC judge attempting to evaluate the long-term potential of the same startup in alternative judging environments (e.g., solo vs. group interviews). In addition, we can also observe the decision rules used for making final investment decisions and construct counterfactuals for how the VC fund would have performed under alternative investment decision rules. Finally, as discussed in Section 2, our sample is broadly *representative* of the activities of sophisticated early-stage VC investors and the traits of startup firms seeking early-stage funding.

In this paper, we examine two main research questions. First, we identify how individual VC employees evaluate early-stage startup firm potential under different judging conditions. Because our dataset contains multiple judges evaluating multiple firms on a repeated basis across time, we can examine topical questions such as how judge's beliefs are formed and adjusted in solo vs. group decision-making environments as well as whether individual VC judge evaluations are significantly influenced by affinity biases with startup firm founding teams (i.e., judge-founder homophily biases). Second, we assess the impact of different judging environments (solo vs. group evaluations) and scoring aggregation rules (consensus vs. champion methodology) on VC fund performance.

[5] Our empirical setting thus differs from laboratory experiments and simulations that rely upon the subject's evaluation of fake firms/people in hypothetical scenarios.
[6] As outlined in Kerr, Lerner, and Schoar (2014), this is because empirical researchers typically do not have access to all the information available to the focal evaluator (whether a firm or individual) at the time of their scoring assessment.

Using our unique database of startup application data and individual VC employee judging scores, we provide several new insights into the investment selection process of (early-stage) VC investors. We initially show that, despite reviewing the exact same application materials, there is substantial heterogeneity in individual VC judging scores for the same startup applicant. To give a sense of the economic significance of these judging disagreements, if our VC firm had solely relied on the lowest pre-interview judging score for its initial applicant screening, then almost half of eventually selected portfolio firms may not have even progressed to the first-round interview stage!

Interestingly, we find that a key factor behind scoring disagreements for the same applicant is that VC judges who share an affinity-based trait with a startup firm founder (based on having the same gender, ethnicity, and/or prior schooling) tend to consistently provide higher scores to those startup firms. The economic magnitude of these affinity-based preferences is substantial, with 'affiliated' judges submitting 8–11% higher scores for the same startup applicant in the first evaluation stage than 'unaffiliated' judges. These affinity-based preferences tend to be more prevalent in cases where individual judges must rely more heavily on more qualitative 'soft' information, rather than more quantitative 'hard' data, when evaluating a startup's potential.

Crucially, however, the magnitude of these judge-founder affinity-based preferences appears to be context-specific. In particular, we show that these preferences are primarily concentrated in 'solo judge' settings where each VC judge must independently score applicants without having the opportunity to interact with a group of their VC firm colleagues. In contrast, the same individual VC judge that is evaluating the same startup but in a collaborative ('group') judging environment (i.e., joint Q&A sessions involving several VC judges as well as internal post-interview discussions among judges) seem to place relatively more weight on startup firm characteristics like measures of business traction (e.g. number of users and revenues to date) when determining applicant scores.

Motivated by these considerable differences in individual scoring assessments under different judging conditions, we next measure the impact of adopting alternative investment decision-making policies on VC fund performance. We first show that enabling VC employees to jointly interview startup firm founders and collaboratively discuss the relative merits of applicants prior to submitting individual judge scores (i.e., a 'group'-based approach) results in the selection of higher quality portfolio firms relative to more siloed 'solo' judging environments. Furthermore, we outline the relative merits of requiring 'consensus' across a plurality of VC firm judges to make an early-stage VC investment vs. only requiring at least one VC firm partner to 'champion' an investment. We find that 'consensus'-based decision rules are better (on average) at detecting higher quality startups but that a 'champion' rule may be superior at identifying very positive outlier venture opportunities.

Finally, we identify the mechanisms explaining the outperformance of collaborative 'group' evaluations relative to more independent 'solo' judge assessments. We first show that the scoring assessments of 'affiliated' judges who share at least one affinity-based trait with a startup founder are significantly less predictive of future startup firm performance than the scores submitted by 'unaffiliated' judges. We then leverage our unique database of written judge comments and audio interview recordings to provide more direct evidence that collaborative Q&A sessions encourage all judges, especially 'affiliated' judges, to ask interview questions and submit written explanations of scores that focus more on startup firm "fundamentals" rather than founders' pre-existing traits.

These combined findings imply that the judge-founder affinity-based preferences that we observe in solo judge settings are more consistent with the presence of significant judge-founder "homophily biases" rather than superior private information flows, and that establishing a panel of judges to collaboratively assess a startup's potential helps to improve the rigor of the VC selection process by mitigating the influence of personal judge-founder affinity biases.

Our paper contributes to several strands of literature. First, our paper contributes to the ongoing debate about the optimal structure of internal corporate decision-making processes. For example, while several papers have proposed theoretical models that outline the potential costs and benefits of employees' collaboratively assessing project alternatives versus requiring each worker to conduct independent analysis,[7] there have been relatively few papers that have been able to empirically test these predictions in real-world business settings.[8] In contrast, our unique research setting allows us to observe the same judge evaluating the same startup on repeated occasions under different judging conditions, for example 'solo' vs. 'group' decision-making environments. As a result, we can empirically test whether group discussion with colleagues prior to the submission of individual judging scores can systematically improve the rigor and impartiality of the internal VC selection process (for example, by mitigating any VC judge-founder homophily preferences).

Relatedly, our highly detailed proprietary dataset enables us to empirically test the value implications of alternative investment decision rules. On the one hand, most of the existing theory literature has emphasized the benefits of rules that require some level of consensus or agreement across employees before the firm proceeds with a given course of action.[9] On the other hand, there has been a growing collection of theory models that have proposed the use of a 'champions voting rule' (i.e., undertake an investment as long as at least one employee advocates for, or 'champions', the deal) to better capture very positive 'outlier' investment opportunities (e.g., Malenko et al., 2024). One of the key benefits of our very granular judge score records is that we can empirically examine how the VC fund would have performed under different judge score aggregation methods.

[7] See e.g. Alchian and Demsetz (1972); Itoh (1991); Legros and Matthews (1993); Landier, Sraer, and Thesmar (2009); and Harel, Mossel, Strack, and Tamuz (2021).

[8] For empirical studies examining the impact of teamwork on productivity in manufacturing firms, see e.g., Hamilton, Nickerson, and Owan (2003), and Boning, Ichniowski, and Shaw (2007). For a review of laboratory experiments comparing individual vs. group decision-making, see Charness and Sutter (2012).

[9] See e.g., Condorcet (1785); Ladha (1992); Baranchuk and Dvbvig (2009); Csaszar and Eggers (2013); Malenko (2014); and Chakraborty and Yilmaz (2017).

Second, since VCs often view portfolio firm selection as the primary driver of investment success (Gompers, Gornall, Kaplan, and Strebulaev, 2020), there has been much research into various aspects of how VCs select startup firm investments and their associated consequences. For example, prior researchers have used a variety of empirical methods to examine both the absolute and relative importance of the startup firm's business ("the horse") versus the startup management team ("jockey") in driving VC selection decisions and ex post performance.[10] Other studies focus on the impact of angel investors, business accelerators, or government funding on startup firm outcomes using a regression discontinuity design based on variation in consensus judging scores across firms.[11] More recently, Jang and Kaplan (2023) use a dataset provided by an early-stage, U.S. Midwest-focused VC to assess how much skill VCs have in screening startups for investment (based on future startup outcomes) while Gonzalez-Uribe et al. (2023) use data provided by a UK seed-stage, software-focused VC fund to study how the VC due diligence process by itself can promote startup growth (even if those startups are not ultimately selected for VC investment).

Crucially, however, these prior studies take the investor's internal scoring process as given, whereby the development of startup applicant scores by individual judges is typically unobserved. In contrast, we are the first paper (to the best of our knowledge) to forensically examine how these individual VC employee scores are formulated in the first place. As a result, our unique perspective and empirical strategy allows us to open the black box of VC portfolio firm selection by uncovering novel factors that significantly influence the VC selection process. For example, we show that individual judge scores are materially affected by both the judging environment (i.e., 'solo' vs. 'group'-based evaluations) and shared affinity-based traits between a judge and a firm founder.

[10] See e.g., Kaplan and Strömberg (2004); Kaplan, Sensoy, and Strömberg (2009); Bernstein et al. (2017); Gompers et al. (2020); Jang and Kaplan (2023); Lyonnet and Stern (2024).

[11] See Kerr et al. (2014); Gonzalez-Uribe and Leatherbee (2018); and Howell, Rathje, Van Reenen, and Wong (2023).

Third, our paper contributes to the ongoing debate about the prevalence of investor biases in startup funding markets and potential mechanisms for mitigating any such biases. To date, there is much disagreement in the prior literature about the existence and/or magnitude of biased startup evaluations by investors based on the inherited traits of startup founders (for example, 'minority' entrepreneurs that are female or non-White: see generally Ewens, 2023; Zhang, 2023; and Cassel, Lerner, and Yimfor, 2024). Depending on the empirical setting and data structure, some researchers find evidence of significant discrimination against minority entrepreneurs[12] while others find no persistent biases against minorities in investors' evaluations of startup firms and their founders.[13]

In contrast to these prior papers, our unique setting helps us to directly test the magnitude of investor biases in startup funding markets, where we offer more nuanced findings on both the dimensions and the judging environments in which such biases are more prevalent. Importantly, all judges in our setting are asked to evaluate the *same startup* firm/founding team at the *same time* with the *same information set*. Furthermore, we know the identity and the background of each VC judge, including their (randomly generated) level of homophily with startup firm founders, as well as possess much greater insight into the underlying reasoning behind a judge's scores from their written comments and recorded interview activity.[14] Since we have access to all the information available to the focal evaluator at the time of their scoring assessment, we can more precisely identify whether a genuine bias against certain types of individuals drove the observed variation in final scores at each stage of the interview process, separate from any differences in inherent applicant quality and/or judging information sets (see generally, Ewens, 2023).

[12] See e.g., Ewens and Townsend (2020), Fairle, Robb, and Robinson (2022), and Cook, Marx, and Yimfor (2023).
[13] See e.g., Bapna and Ganco (2021), and Gornall and Strebulaev (2022).
[14] This contrasts with prior experimental-type studies that involve the tracking of email responses by investors (see e.g., Bernstein et al., 2017) that tend to suffer from low response rates, unknown characteristics of the focal evaluator and/or unknown reasoning behind observed response behavior, and only capture initial expressions of interest rather than final decisions (Bertrand and Duflo, 2017; Zhang, 2023).

Finally, our paper is related to the literature that examines the role of startup accelerators in promoting the growth of early-stage entrepreneurial ventures. Several papers have documented that startup firms that participate in accelerator programs outperform their non-accelerator counterparts due to the benefits of specialized training and mentorship (e.g., Gonzalez-Uribe and Leatherbee, 2018; Robinson, 2022), external certification (e.g., Cohen, Fehder, Hochberg, and Murray, 2019), and entrepreneurs receiving timely and informative signals on whether to pivot or abandon their business idea (e.g., Yu, 2020). However, this prior literature tends to overlook the first order question of how firms are chosen for accelerator programs in the first place, a vital selection issue given that typically only 1–2% of applicants are accepted into each accelerator cohort. In contrast, we believe that we are the first paper to undertake a detailed investigation into how individuals at accelerator organizations evaluate and score startup applicants in practice, thus providing valuable new insights for entrepreneurs and investors into the accelerator selection process.

2. INSTITUTIONAL SETTING AND DATA

2.1 Research setting

To address the pervasive funding and educational impediments to startup firm development (Gonzalez-Uribe and Leatherbee, 2018), there has been an explosion in the number of so-called "startup accelerator" programs, from a handful of accelerators in 2005 to over 7,000 programs worldwide today (Yang, Kher, and Lyons, 2019). These fixed-term, cohort-based programs are designed to offer startups some upfront funding plus an intensive education/mentorship program to promote startup growth and prepare portfolio firms for raising future financing rounds. Importantly, as part of a broader push by VCs into providing financing for earlier stage startup firms (Ewens, Nanda, and Rhodes-Kropf, 2018; Lerner and Nanda, 2020), VC-backed accelerator programs have become one of the most important sources of funding and business support for early-stage startup firms. For example, it is estimated that over 10% of the $11.5 billion invested in U.S. seed funding rounds in 2023 were made by VC-backed accelerators (see generally, Teare, 2024).

We focus our analysis on a large startup accelerator program run by a prominent U.S.-based venture capital firm. Our focal VC firm specializes in providing seed round funding to early-stage entrepreneurial firms through periodic fixed-term, cohort-based accelerator programs. Our VC firm operates these accelerator programs at several major startup hubs (e.g., New York and California) and it has raised multiple funds totalling more than $100 million. Our VC firm invests across all industries and geographies, although most portfolio companies are located in North America.

Analogous to other for-profit business accelerators worldwide, our VC firm provides program participants with both equity capital and educational/other resources to help promote rapid startup growth. To participate in the accelerator program, startups must complete a series of application forms and interviews with VC firm employees, with startups being evaluated on the relative quality of their product/service and management team. This is a highly competitive selection process whereby only 1–2% of applicants are accepted into any given accelerator cohort. Further details about our VC firm's investment selection process and criteria are given in Sections 2.2 and 3.1.

For selected portfolio firms, our VC firm typically invests $150,000–$200,000 for a 3%–7% ownership stake. In addition, portfolio firms are required to participate in an intensive 12-week education/mentorship program that is designed to assist portfolio firms with product development, operations, corporate strategy, human resource management, marketing, and future fundraising.

In order to facilitate greater information sharing during the academic research process as well as to protect the confidentiality of the highly detailed and commercially sensitive information provided to us, we will not disclose the name of our VC investment firm, details of individual firm employees, or de-anonymized data on applicant firms to this VC investor's accelerator program.

2.2 Description of the accelerator selection process

For each cohort in our sample, our VC firm implemented the following process to select the subset of startup applicants who will be invited to participate in our VC's accelerator program. Please refer to Internet Appendix IA.1 for additional details on each evaluation stage.

2.2.1 Stage 1: Application and initial screening

All startups are required to submit an initial online application. One (junior) employee at our VC Fund is then asked to review each initial application to decide whether a further request for information (called a 'due diligence (DD) pack') is sent to the startup applicant to complete (a process we term the '*initial employee screen*'). This DD pack is a much more detailed questionnaire that asks for the startup's 'pitch deck' and answers to over 50 questions that cover a range of topics relating to the company's product or service, potential market size, business model, competitive landscape, milestones achieved to date (i.e., traction), financial information and other metrics, the background and skills of the management team, and the legal/ownership structure.

2.2.2 Stage 2: Pre-interview assessment

Next, it is expected that all VC firm employees will review and submit scores for each applicant based only on the submitted due diligence pack materials (otherwise referred to as '*pre-interview scores*'). This pre-interview score can range between 0 points (worst) to 100 points (best) and is based on sub-scores given for the market potential of the startup's product/service, the quality of the startup's management team, and the level of traction/customer engagement garnered to date.

Critically, our VC firm adopted a judging policy that emphasized the importance of each VC partner/employee submitting truly independent assessments of applicant quality, especially at the pre-interview stage. To enforce this policy, our VC firm implemented several procedures to ensure that individual VC employees would not communicate with each other and would not observe each other's scores and comments prior to submitting their final pre-interview judging assessments.

Our VC firm will then rank startups from best to worst based solely on their average pre-interview scores and invite the top 10% of all applicants to progress to the next evaluation round.

2.2.3 Stage 3: First round interview

Our VC firm will then invite selected applicants to a 30-minute meeting with all available VC firm partners and employees. At these meetings, the startup's management team will make a short presentation about the company and answer a series of questions from VC firm employees. At the conclusion of each first round interview, all attendees from the VC firm will have an open group discussion about the strengths and weaknesses of the applicant. However, analogous to the pre-interview scoring process, each interviewer is asked to provide a single overall score for each first round candidate based on that judge's holistic assessment of the startup's potential and fit for the accelerator program (otherwise referred to as '*first round interview scores*'). Our VC firm will then rank startups from best to worst based solely on their (weighted) average first round interview scores and invite the top 5% of all applicants to progress to the next evaluation stage.

2.2.4 Stage 4: Second round interview and final selection

For those applicants that pass the first round interview stage, these selected startups will hold a second (and final) 45-minute interview with all available VC firm employees where the startup's management team will make a longer presentation and answer additional questions from VC firm employees about the startup's business. After this interview, all VC firm interviewers have an open group discussion about the relative merits of the applicant and their suitability for investment. Each interviewer will then provide a single overall score for each firm candidate based on that VC judge's holistic assessment of the startup's potential and fit for the accelerator (otherwise referred to as '*second round interview scores*'). Our VC firm will then rank startups based solely on their (weighted) average second round interview scores and make funding offers to startup companies to join the accelerator cohort in order of rank until all available accelerator cohort slots are filled (where our VC firm will prespecify a capacity threshold of 12–16 startup firms for each cohort).

2.2.6 Additional observations on integrity of VC selection process

There are several features of our VC firm's internal process that helps to ensure the integrity of the portfolio firm selection process, including maintaining the independence of individual scoring assessments. For example, at every stage of the judging process, accelerator applicants *do not* know: (a) the number or identity of the judges evaluating their application, (b) their raw scores or relative ranking, and (c) the cut-offs for determining whether they progress to the next stage of the selection process. As such, it seems impossible for applicants to manipulate the ranking process.

Furthermore, we typically observe that each VC employee judge will either submit scores for all startup firms in the remaining application pool or for no applicants at all during each evaluation stage, consistent with our VC firm's stated judging philosophy that all VC employees are expected to judge every remaining applicant in the cohort pool wherever possible.[15] As a result, it seems unlikely that VC judges are strategically self-selecting which applicants to grade.[16]

Finally, it also appears unlikely that any judge could precisely manipulate the final ranking of applicants (for example to help a 'friend' to qualify for the accelerator program). This is because: (a) a relatively large number of VC employees evaluate each startup applicant[17] and each VC judge will often evaluate hundreds of applicants at each interview stage; (b) each judge does not know

[15] For example, at the pre-interview stage, we only found two instances across our entire sample where a judge provided scores for only part of the cohort applicant pool. In both cases it was due to the focal employee providing scores for all startups in one 'wave' of applications but not providing any scores for the other 'wave' of applications in the cohort.
[16] In additional unreported tests, we verify that, in the rare cases where a VC employee did not provide judging scores for the entire cohort, there is no systematic trend in the observable characteristics of scored vs. unscored startups.
[17] For example, the median number of individual judges for each startup applicant in the pre-interview process is 9.

the scores of other judges prior to submitting their own scores;[18] and (c) no judge has any advance notice on the exact cut-off score needed for an applicant to progress to the next interview stage (since it depends on a relative ranking of (weighted) consensus applicant scores).

2.3 Sample outline

Our baseline sample consists of all application materials, judging scores, and other internal VC firm records pertaining to 10 accelerator cohorts between January 2017 and October 2019.[19] To link these records together, we utilize unique, permanent identifiers for each startup firm applicant (since it is possible for the same startup firm to submit an application for different accelerator cohort cycles) as well as unique identifiers for each VC firm employee judge.

Across our entire sample, we observe 9,068 initial applications from 7,004 unique startup firms, 1,562 due diligence packs, 13,518 pre-interview scores, 3,401 first-round interview scores, and 1,663 second-round interview scores provided by 24 distinct VC firm partners/employees. See Figure 1 for a breakdown of the total number and progress of each application across our 10 cohort cycles and Figure 2 for the distribution of primary industry groups of startup firm applicants.

2.4 Variable construction

In this section, we describe the dependent and independent variables used in our baseline analysis. Please see Appendix A for further details on the construction of each of these variables.

2.4.1 Outcome variables

With respect to our portfolio firm selection tests, we focus on the variation in individual VC judges' scores. We use the 'raw' overall pre-interview, first round interview, and second round

[18] Each judge must separately submit an individual scoring spreadsheet for an administrative manager to compile.
[19] The start date of our sample corresponds to the first cohort where our VC firm started to systematically retain all materials and records generated throughout the entirety of the accelerator selection and contracting process.

interview scores submitted by each employee for each startup firm (referred to as the *Pre-interview overall score*, *First round interview score*, and *Second round interview score*, respectively).[20]

2.4.2 Judge-founder affinity-based traits

To assess the possibility that shared affinity-based characteristics between individual judges and firm founders may significantly influence the internal VC scoring process, we construct several measures of judge-founder affinity-based traits. Specifically, we create four separate indicator variables that are equal to one if the focal VC firm judge and at least one of the startup's founders have the same gender (*Shared gender*), have the same ethnicity (*Shared ethnicity*), graduate from the same university (*Shared education*), or worked at the same employer (*Shared employer*), and zero otherwise. To construct each affinity-based variable, we rely upon startup pitch decks (which include pictures and descriptions of each founder), detailed application materials with founder biographies and education/employment profiles like LinkedIn pages, and web searches to identify the gender and ethnicity of each firm founder as well as each founder's full list of prior employers and universities attended. We then use internal VC firm records and LinkedIn profiles to identify the corresponding traits for each VC employee judge.

2.4.3 Controls for other judge-founder overlapping characteristics

To account for the possibility that other shared characteristics/experiences (that are separate from those based on affinity-based traits) may systematically influence individual judging scores, we include dummy variables for whether the focal VC firm judge and at least one of the startup's founders both have graduate degrees (*Shared graduate degree*) or both earnt a university degree

[20] In additional (unreported) robustness tests, we obtain similar results if we use 'scaled' score dependent variables by dividing a judge's raw score in the pre-, first round-, or second round-interview stage by that applicant's average score in the same evaluation round (computed across all 'raw' scores of VC firm judges in that cohort), respectively.

from a top-tier university (*Shared top tier university*). We also control for each judge's years of experience working in the same industry sector as the focal startup (*Shared industry experience*).

2.4.4 Startup firm characteristic (time-varying) control variables

Through our access to a wealth of confidential and detailed information about each startup applicant, we can incorporate several firm characteristics in our regressions that capture the past progress, current status, and future potential of each startup to our VC's accelerator. Specifically, we include the following time varying startup firm controls: *Company stage of development*; *Company age*; *Company's lifetime revenue*; *Number of total users since launch*; *Number of paying users since launch*; *External funding raised to date*; *Company runway*; *Current firm valuation*; *Number of FTE company employees*; as well as *Estimated Serviceable Obtainable Market (SOM)*.

2.4.5 Startup founding team (time-varying) control variables

In addition to the inclusion of an extensive set of time-varying firm-level controls, our regressions also incorporate various (time-varying) measures of the professional experience of startup firm founders. Specifically, we include the following firm founding team control variables: *Average top management team (TMT) experience of company founders*; and *Average years of startup founding experience of company founders*.

2.4.6 VC judge (time-varying) control variables

Our research setting also allows us to include several variables that control for the evolution in the experience of individual VC firm employees. Specifically, we control for the focal judge's *Years of Financial Investment experience* and *Amount of startup firm judging experience*.

2.5 Summary statistics

Table 1 provides the mean, median, and standard deviation of the various characteristics of our baseline sample, alternatively computed across all startup applicants (comprising both accepted and rejected applicants) as well as across only the 120 accepted accelerator applicants.

In Panel A of Table 1, we present summary statistics for the number of individual judges and the range of ‘raw’ and ‘scaled’ judging scores for each startup applicant at each assessment phase. Interestingly, we observe a high amount of disagreement amongst individual VC employees about an applicant’s quality, even when all judges are evaluating the same startup firm. We explore the potential factors influencing the significant variation in individual judging scores in Section 3.

In Panel B of Table 1, we provide information on the characteristics of the startups that apply to participate in our VC firm’s accelerator program. In Panel C of Table 1, we show the traits and experiences of individual VC employees who are responsible for judging potential candidates.

2.6 External validity

Despite the numerous advantages of our confidential dataset, it is reasonable to consider the extent to which our research setting, and its associated findings, can generalize to other investment-related settings, from accelerators more specifically to (early-stage) venture capital investors more generally. We examine this question of external validity from two perspectives.

First, in terms of the VC investment firm that forms the basis of our analysis, we claim that the internal practices and the observed outcomes of our focal VC firm are reasonably representative of other reputable VC investors in this investment space. With respect to the VC internal decision-making process, we first note that our VC firm’s multi-step, group-focused evaluation of each investment opportunity is quite similar to the typical decision-making approach of other VC investors (e.g., Gompers et al., 2020).[21] Furthermore, our VC firm's key criteria for selecting portfolio firms for its accelerator (i.e., the quality of the startup's management team, the potential of the applicant's product/service, and the startup's level of traction/customer engagement garnered to date) aligns with existing survey and experimental evidence finding that other VC investors (including early-stage focused funds) also emphasize the critical importance of the startup's management team, business model, and core product/technology in selecting investments (Bernstein et al., 2017; Gompers et al., 2020).[22]

With respect to observed outcomes, our focal VC firm is widely recognized as one of the most prolific and accomplished seed investors in the United States. For example, a 2021 Beta Boom study of over 3,000 accelerators worldwide found that our sample VC firm ranked in the top 5% in terms of exit performance (Paluch, 2021), while the average fund internal rate of return (IRR) through to the end of 2024 is in excess of 25%. While our focal VC firm does appear to exhibit above-average return performance (particularly amongst earlier-stage investment funds), we argue that any bias that we have towards studying the selection and contracting outcomes of a relatively successful private accelerator program is helpful because we are more likely to identify the underlying methods and practices of relatively sophisticated, value-maximizing startup investors.

Second, in terms of the startup companies that participate in our VC-backed accelerator's selection and contracting process, we argue that the set of startup companies considered by our VC firm are broadly representative of the universe of entrepreneurial companies seeking early-stage investment funding. While obtaining comparable and comprehensive information on accelerator

[21] This is particularly true as it pertains to for-profit accelerator programs, where our VC firm's screening and selection process closely aligns with standard industry practice (see e.g., Yu, 2020).

[22] Interestingly, however, our VC firm does appear to place relatively less weight on the importance of the startup's management team when selecting portfolio investments compared to other early-stage VC investors (see e.g., Jang and Kaplan, 2023). This may help to explain why our VC firm exhibits relatively strong ex post return performance.

applicants is extremely challenging (especially for VC-backed accelerator programs), the Global Accelerator Learning Initiative's (GALI) Entrepreneurship Database Program provides some useful summary statistics about the types of startup companies that typically apply to business accelerators.[23] Importantly, the average age (2.6 years), the average number of full-time equivalent (FTE) employees (3.1), the average revenue ($78,230), and the average amount of equity financing raised ($45,698) by applicants to our VC-backed accelerator program is comparable to the mean figures reported in the full sample of applicants in the GALI Entrepreneurship Database (2.7 years, 3.3 FTE employees, revenues of $70,109, and equity financing of $43,906, respectively).[24] In addition, our VC firm considers all early-stage startup companies irrespective of location and industry because our VC firm is not geographically or industry restricted in its investment mandate. Therefore, based on the observable characteristics of our VC firm and the profile of applicants to our VC-backed accelerator program, we argue that our empirical analysis is highly relevant for understanding the portfolio selection and contracting decisions of early-stage VC investment funds (especially within a private seed accelerator context).[25]

3. VC PORTFOLIO FIRM SELECTION

In this section, we investigate the potential drivers of disagreements among individual VC firm judges in the scoring evaluations of startup firm applicants and consider the circumstances in which any potential judging biases may be amplified or mitigated during the VC selection process.

[23] Through a joint collaboration between Emory University and the Aspen Network of Development Entrepreneurs (ANDE), this database is based on information from over 23,000 startups applying to over 300 accelerator programs worldwide that operated between 2013 and 2019.

[24] Although our sample of startups may display slightly more advanced development in terms of revenue and equity funding raised, it should be noted that the GALI Database includes data from not-for-profit accelerator organizations that generally consider even earlier-stage applicants than those typically considered by for-profit accelerators.

[25] Nevertheless, regardless of the similarities between our VC-backed accelerator program and other early-stage investment vehicles, we acknowledge that each early-stage startup investor will have their own unique characteristics (for example with respect to internal decision-making processes, geographic/industry focus, and portfolio firm outcomes) that may somewhat affect the generalizability of our findings to other related investment contexts.

3.1 Initial empirical observations on VC selection process

From Panel A of Table 1, we make two important initial observations about our VC-backed accelerator's internal portfolio firm selection process. First, there appears to be a large emphasis on consensus decision-making throughout our VC's various interview phases. For example, the median number of judges for each startup applicant in the pre-interview screening stage is 9 judges. This suggests that our VC firm places a high value on obtaining multiple independent perspectives about a startup's potential before formulating a consensus view on whether that applicant should proceed to the next phase of the internal evaluation process (see generally, Da and Huang, 2020).

Second, however, we also observe a very high amount of disagreement amongst individual VC employees about an applicant's quality and future potential, even when all judges are assessing the exact same startup firm. In statistical terms, when grading applicants in the pre-interview stage out of a total of 100 points, the standard deviation of internal judging scores for the same startup is 17% with an interquartile range of 29% (compared to a median pre-interview score of 53).

To express the real-world importance of this substantial individual judge score variability in more concrete economic terms, let us focus on the set of 120 eventually selected portfolio firms invested in by our VC firm. If our VC firm had relied on the bottom or lowest pre-interview judging score for its initial applicant screening rather than an average-based scoring system, then almost half of these eventually selected portfolio firms may not have even been invited to the first round interview stage and thus never participated in our VC firm's startup accelerator program at all!

3.2 Empirical methodology

Given these initial observations from our internal judging score data, we believe that a natural question to ask is what drives this substantial variation in individual VC employee judging scores? Despite the rapidly growing literature that relies upon the level or variation in *consensus* judging scores to examine the effect of various investor treatments on startup firm outcomes,[26] there has been no empirical research (to the best of our knowledge) that attempts to systematically understand and explain what shapes the *individual* judging scores that underlies any consensus assessments.

The first critical point to note in our research setting is that startup firm characteristics and/or differences between the information sets of individual VC judges *cannot* explain the high variation in within-startup firm individual VC judging scores. This is because all individual VC judges are tasked with using the exact *same information set* to evaluate the exact *same startup firm applicant*.[27] This means that alternative channels are required to explain the observed variation in within-startup firm judging scores, for example the unique traits and experiences of individual VC employees.

However, an intriguing hypothesis that we believe has not received sufficient attention in the prior literature is whether any shared affinity-based characteristics between an individual judge and a startup firm's founders (based on overlapping gender, ethnicity, educational background and/or employment experiences) may explain some of the large within-applicant judging score variation that we observe. While there is a long-standing literature indicating the presence of 'homophily' (namely the tendency of individuals to prefer to work with others who share similar personal or social characteristics) in other VC-related contexts,[28] there is no other empirical study that has undertaken a large-scale analysis of the relative importance of "judge-founder affinity-based preferences" in driving large differences in individual VC judge evaluations. We initially use the

[26] See e.g., Kerr et al. (2014); Gonzalez-Uribe and Leatherbee (2018); Howell et al. (2023); Jang and Kaplan (2023). For example, Jung and Kaplan (2023) only use their VC firm's consensus scores for 'Team' quality, 'Market size & competition', 'Product & innovation', and likelihood of successful 'Exit' to examine how useful these sub-scores are in predicting the amount of future funding raised, probability of startup survival, and probability of an IPO/M&A exit.
[27] As discussed in Section 2 and Appendix IA.1, each VC employee is provided with an identical set of Q&A-based application materials, pitch decks, interview records, and other data for each different startup in each judging round.
[28] For example, see Gompers et al. (2016) in the context of the formation of venture capital syndicates as well as mixed evidence on the role of informationally relevant social networks and/or homophily in explaining the observed matching between startup firms and their VC investors (e.g., Hedge and Tumlinson, 2014; Huang, 2023; Garfinkel et al., 2024).

more neutral term "judge-founder affinity-based preferences" and then discuss in Section 5 whether these preferences are more consistent with rational information-based choices or personal biases.

To examine this variation in individual judging scores more formally, we run the following ordinary least squares (OLS) regression specification where our primary outcome variable is the overall score given to startup firm applying to cohort by individual VC employee judge :

$$
\begin{aligned}
\_ &= \alpha + \beta \quad \_ + \beta \quad \_ \qquad (1)\\
&+ \beta \quad \_ + \beta \quad \_ \\
&+ \delta \\
&+ \theta \quad + \gamma \\
&+ \varphi \quad - \quad \_ \\
&+ \quad + \quad + \quad + \varepsilon
\end{aligned}
$$

The dependent variable _ is equal to the total score given by judge to startup firm in cohort (i.e., *Pre-interview overall score*, *First round interview score*, and *Second round interview score*, respectively). Our primary independent variables of interest are *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*, respectively, which capture measures of potential affinity between individual judges and applicant firm founders. We also incorporate several additional time-varying controls for various observable startup firm, startup founding team, and VC judge characteristics as well as other judge–founder overlapping traits (see Section 2.4 and Appendix A for further details on variable construction).

There are several advantages afforded by the richness of our multiple judge, multiple startup firm dataset that allows us to cleanly parse out the relative importance of judge-founder affinity based preferences, individual judge traits, and startup attributes in explaining variation in individual VC judging scores. First, by including *Startup firm fixed effects* that keeps the evaluation subject (i.e., the startup applicant) and judges' associated information sets constant, we can specifically identify how the unique experiences and potential judge-founder affinity biases of individual VC employees may affect their evaluation of the *same startup applicant*. Second, by including *Judge fixed effects* that keeps the individual VC employee evaluator (with their leniency and personal taste preferences) constant, we can examine what startup traits and potential judge-founder affinity biases may influence the *same VC judge* in their evaluation of different startup applicants. Finally, by including *Cohort fixed effects* to focus on judge score variability within an accelerator cohort, we can account for any time-related trends at the micro- or macro-level that may impact scoring.

Before continuing, we note that our VC firm did not have a formal randomization process for assigning individual judges to startup applications. However, as discussed in Section 2.2.6, the allocation of VC employees to judge each startup applicant is still somewhat of a random process because our VC firm's widely recognized (and followed) judging policy is that: (a) everyone who is available is expected to participate in judging the current interview round and (b) if an employee is involved in judging an interview round, they are expected to score all applicants in that cohort. By adhering to this policy, our VC selection sample is unlikely to suffer from any meaningful self-selection concerns since individual judges do not seem to have the ability (or even the incentive) to strategically select which accelerator applicants they will evaluate and grade.

3.3 Results with pre-interview scores in 'solo judge' setting

In this first empirical test, our dependent variable is judge 's pre-interview score between 0 and 100 points for a startup applicant in cohort . We reiterate that the key features of this judging setting are that: (a) each judge will evaluate and score accelerator applicants without consulting each other (a so-called '*solo judge*' setting), and (b) for each startup applicant, each judge will base their evaluations on the same set of application information (see Section 2.2 for further details).

In Table 2, we first find that judge-founder affinity-related traits strongly affect individual pre-interview judging scores for the same startup applicant despite our inclusion of a stringent set of fixed effects and control variables. Specifically, *Shared gender*, *Shared ethnicity*, and *Shared education*, all seem to exert a statistically significant influence on the scores that individual VC judges assign to the same startup applicant.[29] In terms of economic significance, if a judge shares the same gender, same ethnicity, or same university background as at least one of the startup founders, these 'affiliated' judges will on average give a 8–11% higher overall pre-interview score than 'non-affiliated' judges for the same applicant firm.[30] Given the closeness of applicant scores around the cut-off threshold for a first round interview invitation in our VC's highly competitive application process, these judge-founder affinity-based preferences can clearly affect a marginal startup's relative ranking and thus their chance of progression through the VC selection process.[31]

More generally, we identify several (observable) startup firm characteristics that VC judges appear to positively value when evaluating the relative merits of accelerator applicants. For example, we find that firm characteristics related to the "fundamentals" of the startup such as having a working prototype of the proposed product or service and demonstrating that the startup business has already attracted paying users appear to be critical data points used by VC firms to (at least initially) identify startup companies with the desired risk-return profile for VC investment. Furthermore, a startup firm having previously raised external capital (whether from angel investors, friends and family etc.) also appears to be viewed quite favorably by individual VC firm evaluators.

[29] In additional robustness tests, we confirm that we obtain very similar results if we use an alternative variable bounded between 0 and 1 that captures the percentage of the startup firm founding team that has the same gender, same ethnicity, shared university experience, or shared prior employer as the focal individual judge, respectively.

[30] Somewhat unsurprisingly, we confirm in unreported tests that the effect of these shared affinity-based traits tends to be concentrated in individual judge's sub-scores for the quality/potential of the startup firm's founder team.

[31] When interpreting the magnitude of this result, it should be noted that if significant judge-founder homophily biases exist at even relatively high performing VC funds, then such biases may be even more prevalent at lower tier VC funds.

Somewhat reassuringly, we find that these 'hard' data metrics appear to be at least four times more influential in affecting individual VC judge scores than judge-founder affinity-based preferences.

Nevertheless, it appears that VCs combine both observable 'hard' data with more intangible 'soft' information about a startup and its founders to identify potential investment opportunities. For example, over 25% of startup applicants invited to the first round interview stage did not have a working prototype nor had a current source of revenue. This implies that VCs' portfolio selection process is a somewhat "part art, part science" process, especially when attempting to evaluate early-stage applicants to a business accelerator based only on submitted application materials.

3.4 Heterogenous treatment effects analysis using pre-interview judging scores

In the next stage of our analysis, we consider the circumstances under which the observed affinity-related preferences in pre-interview judging scores may become more prevalent during the initial stages of the VC selection process. Our hypothesis is that judge-founder affinity-based ties will be more influential in (positively) affecting individual VC judge evaluations when there is a lack of 'hard', quantitative information on which to base an early-stage startup assessment. This is consistent with several theories of discrimination that postulate a greater role for biases when information asymmetries about types or intrinsic quality are heightened (see generally, Simcoe and Waguespack, 2011; Bohren, Imas, and Rosenberg, 2019).

Therefore, we split our sample of startup applicants into several sub-samples based on the level of opacity surrounding the startup firm's business and future prospects. Specifically, on the assumption that it is generally more difficult to evaluate a startup's potential when it does not have an established customer base and limited information exists on the sustainable level of pricing and margins, we first partition startup applicants participating in the pre-interview screening stage into whether or not they are 'pre-revenue' (i.e., a startup firm is 'pre-revenue' if it has not generated any revenue to date from product or service sales). As an alternative measure of firm opacity, we also partition our pre-interview startup firms into whether or not all the firm's founders are 'first-time entrepreneurs' (i.e., all members of the startup firm's founding team have not founded a prior entrepreneurial venture before the current focal startup). This is on the assumption that it is more difficult for a VC investor to evaluate a startup's potential if the applicant firm's founders do not have verifiable experience in creating and managing an entrepreneurial venture.

In Table 3, we find that the effect of judge-founder affinity-based preferences on individual VC judge pre-interview scores appears to be primarily concentrated in situations where there is relatively less "fundamental" information about the startup firm in terms of its existing revenue profile and/or the prior entrepreneurial experience of the firm founders. This suggests that when individual VC firm employees are given more scope to make subjective judgments about a startup's potential due to the lack of 'hard' quantitative data about the focal startup's business model and/or management team, there is a greater chance that individual judges will rely upon affinity-related personal characteristics to help shape their final evaluations of accelerator applicant firms.

3.5 Results with first round and second round interview scores in 'group judge' setting

Up until this point, our only dependent variable has been the pre-interview scores given to accelerator applicants by individual VC employee judges. Importantly, this judging setting involves each judge evaluating a "paper-based" application individually without consulting other judges (in other words, a somewhat siloed, 'solo judge' setting).

In contrast, in the first and second round interview stages of our VC firm's selection process, we often still have the *same* applicant firm being evaluated by the *same* individual VC judge as in the pre-interview round, but the judging setting now has two important differences from the preceding pre-interview round. First, all VC employee judges participate in a joint, in-person 'group interview' where each VC employee listens to an investment pitch from an accelerator applicant and can observe their VC colleagues' questions and interactions with the startup firm founders. Second, after each first and second round interview, VC employees will discuss the focal startup's interview performance and hear each other's perspectives on the startup's long-term potential before submitting their individual first or second round interview score.

Ex ante, it is not clear whether an interactive group interview and VC team discussion setting will mitigate or amplify the role of judge-founder affinity biases in the VC selection process for early-stage startup firms. On the one hand, the salience of shared personal characteristics and overgeneralized stereotypes may be heightened in judges' thought processes with more direct, in-person interactions with startup company founders.[32] Furthermore, the opinions of biased judges (especially more senior managers) may be more strongly transmitted to other firm colleagues in more intimate small group settings due to issues associated with 'social conformity' and 'group think' (see generally, Klocke, 2007; Ishii and Xuan, 2014; Calder-Wang and Gompers, 2021).

On the other hand, however, it is possible that a collaborative interview and team-based discussion of startup applicant firms may instead reduce the influence of idiosyncratic judge-founder affinity biases by re-directing VC employees' attention towards more objective datapoints and issues that are directly relevant for the startup's business model (see generally Gershoni, 2021). For example, a nascent stream of experimental research suggests that group decision-making is less likely to be influenced by behavioral biases, cognitive limitations, and social considerations than individual decision-making (see e.g., Charness and Sutter, 2012; and Maciejovsky, Sutter, Budescu, and Bernau, 2013).

[32] For example, prior studies suggest that female entrepreneurs may be at a greater disadvantage to male entrepreneurs during in-person interviews compared to other VC judging environments due to the differing tone and substantive focus of interviewer questions (e.g., Kanze, Huang, Conley, and Higgins, 2018) as well as expected adherence to certain gender stereotypes (e.g., Bordalo, Coffman, Gennaioli, and Shleifer, 2019; Hu and Ma, 2022).

3.5.1 Baseline specification and results

To test for the prevalence of judge-founder affinity-based preferences in our VC firm's first round and second round interview stages, we re-run the same specification as in Equation (1) but instead use either first-round interview scores (see Table 4) or second-round interview scores (see Table 5) as the dependent variable. Interestingly, all the coefficients on *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*, are now statistically insignificant in both the first round and second round interview phases. This result holds even in the sub-sample of cases where the focal judge in the first or second round stage also evaluated the same startup in the pre-interview stage (see Column 3), meaning that changes in the composition of the VC judge pool is unlikely to explain the differential impact of judge-founder affinity traits across interview stages.[33]

This result implies that the group interview and post-interview discussion process among VC employees appears to have a positive impact on the integrity and "fairness" of the VC selection process, namely by helping to mitigate any undue influence of any judge-founder affinity-based preferences on individual judging scores. For example, consistent with our anecdotal discussions with our VC firm's employees, individual judges seem to appreciate hearing other perspectives on the strengths and weaknesses of each applicant firm and appear to understand the value of having to articulate and justify their key conclusions to their colleagues in a dynamic feedback setting.[34] Our results also point more generally to the potential benefits of a collaborative, team-based evaluation process relative to a more 'siloed', individual decision-making process.

---

[33] Our results in Tables 4 and 5 are also unlikely to be explained by a reduction in test power compared to Table 3 given that we still utilize 3,401 first round and 1,663 second round judge score observations, respectively, and that the number of judges per applicant firm in later in-person interview stages remains similar to the pre-interview stage.

[34] For example, individuals may be encouraged to focus on more "fundamental" business-relevant factors rather than (potentially extraneous) personal characteristics of startup founders, consistent with experimental evidence on the superior performance of teams in strategic laboratory games compared to individuals: see, for example, Laughlin, Bonner, and Miner (2002); Cooper and Kagel (2005); and Feri, Irlenbusch, and Sutter (2010).

3.5.2 Alternative specification controlling for a judge's prior scores for the same applicant

When seeking to understand the drivers of an individual VC judge's first and second round interview scores for a given applicant, it is natural to assume that these later evaluations will be significantly influenced by the focal judge's earlier scores for the same startup applicant in the previous pre-interview phase.[35] However, it is an important and open empirical question as to how much VC firm employees are willing and able to update their assessments on startup firm potential based on additional (dynamic) interactions with startup firm founders and their VC firm colleagues. In other words, how much do "first impressions count" during the internal VC selection process?

As such, we run an alternative specification where we only include the first or second round scores of judges who also evaluated the same applicant in the pre-interview stage. For example:

$$
\begin{aligned}
\_ \quad = \alpha + \beta \quad \_ \qquad (2) \\
+ \beta \quad \_ + \beta \quad \_ \\
+ \beta \quad \_ + \beta \quad \_ \\
+ \beta \quad \_ \times \quad \_ \\
+ \beta \quad \_ \times \quad \_ \\
+ \beta \quad \_ \times \quad \_ \\
+ \beta \quad \_ \times \quad \_ \\
+ \delta \\
+ \theta \quad + \gamma \\
+ \varphi \quad - \quad \_ \\
+ \quad + \quad + \quad + \varepsilon
\end{aligned}
$$

[35] This is analogous to a Bayesian-like framework where the focal judge forms an initial view about the startup firm's potential (as expressed in their overall pre-interview scores) and then updates this prior as new information is procured during the VC firm's entire interview process (see generally, Bhuller and Sigstad, 2024).

The dependent variable is equal to the _ given by judge to startup firm in cohort . _ is equal to the overall score given by judge to the *same* startup firm during the preceding pre-interview evaluation phase in cohort . As such, the coefficient β captures the extent to which judge updates their initial assessment of startup firm after participating in the (group-based) first round interview process.

As before, we include the variables *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*, respectively, which capture measures of potential affinity between individual judges and applicant firm founders. When these affinity-based traits are included in conjunction with the focal judge's prior pre-interview score for that particular startup, these coefficients will indicate whether the focal judge is more or less likely to grant a higher score after the first round interview and group discussion process if they share certain personal traits with that startup founder team (above and beyond any affinity-based biases captured in that judge's pre-interview score).

Furthermore, we also include the interaction of *Pre-interview overall score* and each of these four affinity-based variables. The coefficients on these interaction terms will indicate whether a judge is relatively more or less likely to revise their initial assessment of the focal startup firm after the first round interview process if they share certain personal traits with that startup founder team.

We also include several additional controls for various observable startup firm, startup founding team, and VC judge characteristics in combination with other overlapping judge-founder characteristics, as well as *Judge fixed effects*, *Startup firm fixed effects*, and *Cohort fixed effects*.

We present the results of this alternative specification in Table 6 where we provide several novel insights about the internal VC selection process. First, we show that, while a judge's initial per-interview assessment of a startup clearly has a persistent impact on that judge's first and second round scores for that same applicant, our VC firm's judges are surprisingly open to revising their personal assessment of a startup firm's potential after being given the opportunity to dynamically interact with both the startup firm's founders and their VC firm colleagues. In terms of economic significance, our estimates imply that VC judges will on average submit first round interview scores that are 15–25% different from their own pre-interview score for the same applicant. This suggests that dynamic interviews with startup founders and group discussions with VC firm colleagues play a critical role in shaping a VC employee's finalized assessment of a startup firm's quality/potential.

Second, we find that individual judges are relatively more likely to downwardly revise their assessment of a startup applicant's potential between the pre-interview and first round interview stages if that judge shares an affinity-based trait with the applicant's founder team. In other words, it appears that a judge's first round interview score for a startup firm applicant is more likely to revised downward if that judge had a relatively high pre-interview score for that same startup *and* they also shared a potential affinity-related trait with the startup's founder team. This implies that judges are relatively more likely to revise or update their priors about a startup applicant if that judge is relatively more likely to have formulated an (upwardly biased) initial assessment at the pre-interview stage for that startup firm applicant with whom the judge shares certain affinity-based traits. This result also suggests that interactive meetings with startup firm founders and group discussions with other VC colleagues about their respective startup assessments can help to somewhat "undo" the potentially distortive effect of shared judge-founder affinity-based characteristics on individual pre-interview scores in more isolated, 'solo judge' settings.

4. VALUE IMPLICATIONS OF ALTERNATIVE DECISION-MAKING POLICIES

In this section, we compare different methods in which individual judge scores are aggregated to form final investment decisions and study the value implications of these critical policy choices. First, we measure the impact of a collaborative, group-based decision-making approach on VC investment performance relative to a more independent, individualistic decision-making process. Second, we investigate the net performance impact of adopting "consensus-based" decision-making rules versus implementing a "champion rule" whereby a single enthusiastic proponent of a startup firm's prospects is sufficient to justify investment.

4.1 VC performance outcomes using 'solo' vs. 'group'-based judging process

As discussed in Section 3, our VC firm's pre-interview startup scoring process is based on each judge evaluating a "paper-based" application individually without consulting other judges and then ranking applicants based on the highest average pre-interview scores. In contrast, our VC firm's later round interview process utilizes an interactive group interview and VC team discussion environment to derive a 'consensus' view of each applicant based on average second round scores.

In this subsection, we assess whether it is beneficial to rely upon 'group' interviews and collaborative internal discussions among VC employees before judges submit their individual judging scores. In other words, we compare the performance outcomes of startup firm applicants that would be selected for final VC investment depending on the set of judging rules/scores used (i.e., rank applicants based on 'solo' pre-interview overall scores or 'group' second round interview scores). This empirical analysis not only allows us to evaluate the relative merits of group discussion and feedback on the quality of individual judging scores but also provides us with insight into whether relying on the affinity-based preferences embedded in 'solo' pre-interview scores is incrementally informative in predicting startup firm performance.

We start our analysis by first identifying the most "marginal" startup applicants that could be selected for VC investment depending on the scoring decision rule adopted. Specifically, we create the indicator variable, *Portfolio selection based on second round interview consensus scores*, that is set equal to one if the startup firm applicant: (a) *was* ultimately chosen for investment by our VC firm (based on having a sufficiently high average second round interview score), but (b) *would not have* been selected if applicants were instead ranked based on the highest average pre-interview overall judging scores (i.e., 94 'treated' firms). In contrast, this indicator variable is set equal to zero if the startup firm applicant: (a) *was not* ultimately chosen for investment by our VC firm, but (b) *would have* been selected if applicants were instead ranked based on the highest consensus pre-interview judging scores (i.e., 88 'control' firms). Importantly, we exclude from our analysis any startup applicants that would either: (i) be selected for investment by our VC firm under both decision rules or (ii) be rejected under both decision rules.

We then run the following regression that compares the relative post-application performance of treated vs. control startup firm applicants:

$$= \alpha + \beta \qquad + \qquad + \varepsilon \tag{3}$$

where are various measures that capture different aspects of the focal startup's performance in the time period after their application to our VC firm's accelerator program is either accepted or rejected (see Appendix A for further details). Specifically, we track the firm's survival (*Out of business*), subsequent fundraising activities (*Number of funding rounds post-application* and *Amount of funding raised post-application*), and consequent valuation (*Post-money startup valuation post-application*). We also include *Cohort FEs* to ensure that only firms in the same application cohort are directly compared with one another.

In Table 7 – Panel A, we show that there appears to be substantial benefits to ensuring that judges have the opportunity to jointly interview and collaboratively discuss startup applicants prior to submitting their own individual judging scores. We find that constructing investment portfolios using more 'group'-based 2nd round interview scores leads to the selection of higher quality and more successful startup firms. In particular, ‘treated’ startups are less likely to go out of business and are more likely to raise greater amounts of external capital and obtain higher valuations. Overall, our combined results imply that ‘group’-based decision-making and analysis improves VC investment selection by mitigating personal biases that can arise in ‘solo’ judge settings.

4.2 VC performance outcomes using ‘consensus’ vs. ‘champion’-based decision rules

Another important aspect of our VC firm’s investment decision-making process is that portfolio firms are chosen based on which startups have the highest average second round interview scores (a so-called “consensus” approach). This decision rule implicitly favors the selection of firms that garner a broader range of support across all VC employee judges.

In contrast, Malenko, Nanda, Rhodes-Kropf, and Sundaresan (2024) provide survey evidence that the majority of the top 50 U.S. VC funds use a voting rule where a VC fund will undertake a seed or early-stage investment as long as at least one partner strongly supports the deal, even if others are not bullish on the investment (a so-called “champion” approach). This is primarily justified on the basis that a “champions voting rule” is better able to ‘catch outliers’, namely that the most successful early-stage VC investments may be outstanding on some dimensions but flawed on other dimensions and that consensus decision-making tends to underweight/(overweight) focus on the outstanding/(weaker) aspects of the startup firm’s business.

This dichotomy raises an interesting question as to the relative merits of ‘consensus’ vs. ‘champion’-based decision rules, especially in the context of earlier-stage VC investments. By having access to all the individual internal judging scores of our VC firm, we are uniquely positioned to empirically evaluate the potential benefits and costs of adopting a ‘consensus’ vs. a ‘champion’ investment rule.

Similar to our methodology in Section 4.2, we identify the most "marginal" startup applicants that could be selected for VC investment depending on the decision rule adopted. Specifically, we create the indicator variable, *Portfolio selection based on second round interview consensus scores*, that is set equal to one if the startup firm applicant: (a) *was* ultimately chosen for investment by our VC firm (based on having a sufficiently high *average* second round interview score), but (b) *would not have* been selected if applicants were instead ranked based on the *maximum* second-round interview score received (i.e., 48 'treated' firms).[36] In contrast, this indicator variable is set equal to zero if the startup firm applicant: (a) *was not* ultimately chosen for investment by our VC firm, but (b) *would have* been selected if applicants were instead ranked based on their maximum second-round interview score (i.e., 63 'control' firms). Importantly, we exclude from our analysis any startup applicants that would either: (i) be selected for investment by our VC firm under both decision rules or (ii) be rejected under both decision rules.

Utilizing this comparison sample with the same regression specification in equation (3), we present our empirical results in Table 7 – Panel B. Interestingly, we find that there are both pros and cons associated with using a 'consensus' vs. a 'champion' investment decision rule. On the one hand, 'consensus' based decision rules seem to allow VCs to avoid making relatively more investments in the (value-destroying) left tail of the startup distribution. Specifically, choosing portfolio firms based on average second round interview scores leads to the selection of startups that are less likely to completely fail/go out of business and are more likely to raise more follow-on funding (see columns (1) and (2) respectively). On the other hand, however, there is some evidence in column (4) that a 'champion'-based decision rule can better identify startups that subsequently become very high positive return (i.e., "home run") investments.

---

[36] We use the maximum second round interview score to proxy for a situation where at least one member of our VC firm indicated that they strongly supported (i.e. "championed") making an investment in the focal startup firm.

Our combined evidence implies that the optimal judging aggregation rule may depend on several interrelated aspects of the judging process. First, if individual judges can consistently draw uniquely informative insights about the strengths and weaknesses of a startup applicant separate from their colleagues due to their specialized expertise, then VC funds may be more willing to adopt a champion-based decision rule.[37] Second, the characteristics and circumstances of VC funds themselves may also determine the optimality of consensus vs. champion investment decision rules. For example, larger and more established funds (such as the top 50 VC funds considered in Malenko et al., 2024) may be better placed to follow a champion-based decision-making approach where the greater downside risk is mitigated by the fund's previously established reputation and investor connections. In contrast, relatively younger or less established VC funds may not have the luxury of immediately "swinging for the fences" without first establishing the partners' credibility for making consistently high-quality investments, even at the risk of not investing in a few very high risk, very high reward investment opportunities under a consensus-based decision rule.

5. UNDERLYING MECHANISMS

In this section, we examine the reasons why 'group' evaluations appear to outperform 'solo' evaluations on a variety of ex post-performance metrics. First, we evaluate whether the observed influence of shared judge-founder affinity-based traits on the startup evaluation process stem primarily from behavioral biases or rational information-based beliefs. Second, we utilize the confidential written comments of judges and audio recordings of interviews conducted with startup founders to identify how more collaborative judging environments can affect the reasoning and decision-making process of individual VC firm judges.

[37] This is consistent with the theoretical model of Malenko et al. (2024) whereby the optimality of a champion-based decision rule critically depends not only on the size of the heavily right-skewed distribution of early-stage startup quality and investment returns but also the distinctiveness of different dimensions of startup firm quality as well as each VC partner's ability to separately identify and communicate their informative signals to other VC partners.

5.1 Explanation for role of affinity-based traits in judging process

Given our previously presented VC selection results, a reasonable question to ask is whether the systematically higher pre-interview scores awarded to startup applicants whose founders share the same gender, ethnicity, or university background as the focal judge is due to a rational 'information-based' explanation or a 'homophily-based' preferential bias. On the one hand, it is possible that judge-founder affinity-based ties provide 'affiliated' VC judges with informative signals about the startup founders' underlying quality and facilitate greater information flow between VCs and entrepreneurs during the VC portfolio firm selection process (see generally, Garfinkel et al., 2024). On the other hand, it is plausible that judges have a preference (consciously or unconsciously) to work with others who share similar personal traits or that judges hold an innate belief/prior that startup founders with similar backgrounds to the focal judge have higher intrinsic quality or potential, thus (positively) biasing that judge's evaluation of the startup applicant firm (see generally, Huang, 2023).

While it is somewhat difficult to conclusively distinguish between these two theories in our research setting (for example, because our VC firm aggregates individual assessments to make collective final investment decisions), we argue that our collective evidence is more consistent with the latter 'homophily-based' explanation. First, as discussed in Section 4.1, if individual pre-interview scores based on solo evaluations embed incrementally informative insights into startup applicant founders that are not captured in group-based interview judging settings, then this would imply that final investment decisions based on 'solo' pre-interview judging scores should outperform those supported by 'group'-based second-round interview judging scores. However, we in fact find the opposite result in Table 7 whereby portfolio selection based on (group) second-round interview consensus scores leads to superior ex-post investment outcomes than portfolio selection based on (solo) pre-interview consensus scores.

Second, in our VC firm's pre-interview stage judging process, there is no dynamic interaction or communication between individual judges and startup firm founders, meaning that it is impossible for confidential new information to flow between a 'connected' judge and a startup founder to assist with the focal judge's pre-interview assessment. As such, private information flows are unlikely to explain the large and persistent differences in pre-interview scoring between 'affiliated' and 'unaffiliated' judges.

Third, all our regression estimates with respect to judge-founder affinity-based preferences are conditional on a wide variety of 'fundamental' firm and founder characteristics (including educational experience at a top-tier university) as well as startup firm, VC judge, and time fixed effects. Furthermore, we also explicitly control for other judge–founder overlapping characteristics that arguably capture elements of a judge having better insight about the underlying quality of the startup firm and its founders (for example, both the startup founder and the VC judge having a graduate degree and/or both graduating from a top tier university) and yet we still see a significant incremental effect of affinity-based traits on individual judging scores. As such, any additional information revealed by shared gender, ethnicity, or university education between a VC judge and a founder must be orthogonal to these observable characteristics.

Finally, if a judge did have meaningful "inside" information on a startup firm applicant due to their overlapping personal characteristics with the startup's founders, then an 'affiliated' judge's more positive startup evaluation relative to their peers should continue to persist beyond the pre-interview stage into the first and second round interview phases. However, we find in Table 6 that 'affiliated' judges are instead more likely to *downwardly* revise their first round interview scores relative to their pre-interview scores as compared to ‘unaffiliated’ judging colleagues. This pattern is again inconsistent with a purely rational and enduring information-based explanation.

5.2 Evidence from judges’ written comments

A critical advantage of our research setting is that we can gain unprecedented insight into the underlying rationale for each judge’s submitted scores for each startup applicant at every stage of the internal VC selection process. Specifically, we can observe the written explanation (if any) given by VC firm judges for their applicant scores in the pre-interview, first-round interview, and second-round interview phases. This in turn allows us to better understand what factors each judge is emphasizing in their individual assessment of each startup firm’s accelerator application.

In our empirical analysis, we first identify the set of VC firm judges that are ‘affiliated’ with the focal startup firm founder (i.e., the focal judge and startup firm founder either share the same gender, ethnicity, university institution attended, and/or prior employer) as distinct from an ‘unaffiliated’ judge that does not share any of these overlapping characteristics with the focal startup firm founders. We hypothesize that affiliated judges, owing to their stronger social ties to the focal startup firm founders, will be less inclined to provide a detailed written justification for their (elevated) ‘solo’ pre-interview scores but, if they do, they are relatively more inclined to highlight characteristics associated with the startup’s founding team as justifying their scores.

Next, we run the following regression:

$$
\begin{aligned}
&= \alpha + \beta\ \mathbb{I}\{\quad\}_{\_} + \delta \qquad (4)\\
&\quad + \theta \quad + \gamma \\
&\quad + \varphi \quad - \quad _{\_} \\
&\quad + \quad + \quad + \quad + \varepsilon
\end{aligned}
$$

The dependent variable _ is equal to the observed written comment behavior of judge with respect to startup firm in cohort , specifically:

(1) An indicator variable that is equal to one if the judge provides any written justification for their applicant score, and zero otherwise (i.e., 𝕀{ } _);

(2) The (character) length of written comments provided by the focal judge along with their submitted interview score (i.e., _); and

(3) The percentage of written comments (based on character length) that focus exclusively on the judge's assessment of the startup's founding team (i.e., _).

Our primary independent variable of interest is *Affiliated judge* which captures whether there is potential affinity between an individual judge and an applicant firm founder based on inherited shared traits or experiences. We also incorporate several additional time-varying controls for various observable startup firm, startup founding team, and VC judge characteristics as well as other judge–founder overlapping traits (see Section 2.4 and Appendix A for further details).

As outlined in Table 8 – Panel A, we find evidence consistent with the fact that, in 'solo' judging situations (such as the 'siloed' pre-interview process adopted by our VC firm), affiliated judges are more inclined to be more opaque in stating why they support a particular accelerator candidate and, if prompted, they are more likely to emphasize founding team-related traits as justifying their overall scoring assessment. In contrast, in Table 8 – Panel B, we do not observe any significant differences in the nature of written comments provided by affiliated vs. unaffiliated judges in the first and second round interview stage. We argue that social pressure within firms encourages VC employee judges, especially 'affiliated' judges, to focus more on startup firm "fundamentals" as opposed to the inherited traits/experiences of applicant firm founders.

6. CONCLUSION

Given that venture capital firms are one of the most important sources of funding for aspiring entrepreneurs in the modern knowledge economy, understanding the inner workings of the VC decision-making process remains a critical question for academics, practitioners, and policymakers. However, research on this essential topic remains scarce due to fundamental issues associated with data availability, measurement error, and omitted variable bias. To overcome these pervasive problems, we use a confidential accelerator dataset provided by a prominent U.S. based VC firm to open the black box of VC portfolio firm selection.

We offer novel evidence that at least part of the substantial heterogeneity that we observe in individual judging scores submitted by our VC firm's employees for the same startup applicant can be explained by significant judge-founder affinity biases, which are especially prevalent in solo judge settings where there is a high amount of ambiguity about the startup's potential. However, our empirical results suggest that encouraging judges to interact with each other can at least mitigate such individual biases, thus providing support for the common use of judging panels and other dynamic group judging arrangements in competitive award settings.

REFERENCES

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and W. Kerr. 2018. Innovation, reallocation, and growth. *American Economic Review* 108(11): 3450–3491.

Alchian, A., and H. Demsetz. 1972. Production, information costs, and economic organization. *American Economic Review* 62(5): 777–795.

Bapna, S., and M. Ganco. 2021. Gender gaps in equity crowdfunding: Evidence from a randomized field experiment. *Management Science* 67(5): 2679–2710.

Baranchuk, N., and P. Dvbvig. 2009. Consensus in diverse corporate boards. *Review of Financial Studies* 22(2): 715–747.

Bengtsson, O. 2011. Covenants in venture capital contracts. *Management Science* 57: 1926–1943.

Bernstein, S., Giroud, X., and R. Townsend. 2016. The impact of venture capital monitoring. *Journal of Finance* 71(4): 1591–1622.

Bernstein, S., Korteweg, A., and K. Laws. 2017. Attracting early-stage investors: evidence from a randomized field experiment. *Journal of Finance* 72(2): 509–538.

Bertrand, M., and E. Duflo. 2017. Field experiments on discrimination. In *Handbook of Economic Field Experiments*. Vol 1, ed. Banerjee, A., and E. Duflo, pp. 309–393. Elsevier: Amsterdam.

Bohren, J., Imas, A., and M. Rosenberg. 2019. The Dynamics of Discrimination: Theory and Evidence. *American Economic Review* 109(10): 3395–3436.

Boning, B., Ichniowski, C., and K. Shaw. 2007. Opportunity counts: Teams and the effectiveness of Production Incentives. *Journal of Labor Economics* 25(4): 613–650.

Bhuller, M., and H. Sigstad. 2024. Feedback and Learning: The Causal Effects of Reversals on Judicial Decision-Making. *Review of Economic Studies*, forthcoming.

Bordalo, P., Coffman, K., Gennaioli, N., and A. Shleifer. 2019. Beliefs about gender. *American Economic Review* 109: 739–773.

Calder-Wang, S., and P. Gompers. 2021. And the Children Shall Lead: Gender Diversity and Performance in Venture Capital. *Journal of Financial Economics* 142: 1–22.

Calder-Wang, S., Gompers, P., and K. Huang. 2023. Diversity and Performance in Entrepreneurial Teams. Working Paper, Harvard Business School.

Casamatta, C. 2003. Financing and advising: Optimal financial contracts with venture capitalists. *Journal of Finance* 58(5): 2059–2085.

Cassel, J., Lerner, J., and E. Yimfor. 2024. Racial diversity in private capital fundraising. Working Paper, National Bureau of Economic Research.

Chakraborty, A., and B. Yilmaz. 2017. Authority, consensus, and governance. *Review of Financial Studies* 30(12): 4267–4316.

Charness, G., and M. Sutter. 2012. Groups Make Better Self-Interested Decisions. *Journal of Economic Perspectives* 26(3): 157–176.

Cohen, S., Fehder, D., Hochberg, Y., and F. Murray. 2019. The design of startup accelerators. *Research Policy* 48(7): 1781–1797.

Condorcet, M. 1785. *Essay on the Application of Analysis to the Probability of Majority Decisions*. Imprimerie Royale: Paris.

Cook, L., Marx, M., and E. Yimfor. 2023. Funding black high-growth startups. Working Paper, Columbia University.

Cooper, D., and J. Kagel. 2005. Are Two Heads Better than One? Team versus Individual Play in Signalling Games. *American Economic Review* 95(3): 477–509.

Cornelli, F., and O. Yosha. 2003. Stage financing and the role of convertible securities. *Review of Economic Studies* 70: 1–32.

Csaszar, F., and J.P. Eggers. 2013. Organizational decision-making: An information aggregation view. *Management Science* 59(10): 2257–2277.

Da, Z., and X. Huang. 2020. Harnessing the Wisdom of Crowds. *Management Science* 66(5): 1847–1867.

Da Rin, M., Hellmann, T., and M. Puri. 2013. A survey of venture capital research. In *Handbook of the Economics of Finance*. Vol 2A: Corporate Finance, ed. Constantinides, G. Harris. M., and R. Stulz, pp. 537–638. Elsevier: Amsterdam.

Decker, R., Haltiwanger, J., Jarmin, R., and J. Miranda. 2014. The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives* 28(3): 3–24.

Ewens, M., Nanda, R., and M. Rhodes-Kropf. 2018. Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics* 128(3): 422–442.

Ewens, M., and R. Townsend. 2020. Are early stage investors biased against women? *Journal of Financial Economics* 135(3): 653–677.

Fairle, R., Robb, A., and D. Robinson. 2022. Black and white: Access to capital among minority-owned start-ups. *Management Science* 68(4): 2377–2400.

Feri, F., Irlenbusch, B., and M. Sutter. 2010. Efficiency Gains from Team-Based Coordination – Large-scale Experimental Evidence. *American Economic Review* 100(4): 1892–1912.

Frame, S., Huang, R., Mayer, E., and A. Sunderam. 2023. The impact of minority representation at mortgage lenders. *Journal of Finance*, forthcoming.

Garfinkel, J., Mayer, E., Strebulaev, I., and E. Yimfor. 2024. Alumni networks in venture capital financing. Working Paper, University of Wisconsin-Madison.

Gershoni, N. 2021. Individual vs. group decision-making: Evidence from a natural experiment in arbitration proceedings. *Journal of Public Economics* 201: 104479.

Gompers, P., Gornall, W., Kaplan, S., and I. Strebulaev. 2020. How do venture capitalists make decisions? *Journal of Financial Economics* 135: 169–190.

Gompers, P., and J. Lerner. 2001. *The money of invention: how venture capital creates new wealth*. Cambridge, MA: Harvard Business School Press.

Gompers, P., Murharlyamov, V., and Y. Xuan. 2016. The Cost of Friendship. *Journal of Financial Economics* 119(3): 626–644.

Gonzalez-Uribe, J., and M. Leatherbee. 2018. The effects of business accelerators on venture performance: Evidence from Start-Up Chile. *Review of Financial Studies* 31(4): 1566–1603.

Gonzalez-Uribe, J., Klinger-Vidra, R., Wang, S., and X. Yin. 2023. Reducing entrepreneurial capability constraints: The broader role of venture capital due diligence. Working Paper, London School of Economics.

Gornall, W., and I. Strebulaev. 2022. Gender, race, and entrepreneurship: A randomized field experiment on venture capitalists and angels. *Management Science*, forthcoming.

Hall, B., and J. Lerner. 2010. The financing of R&D and innovation. In *Handbook of the Economics of Innovation*, ed. Hall, B., and N. Rosenberg, pp. 609–639. Amsterdam: Elsevier.

Hall, R., and S. Woodward. 2010. The burden of the non-diversifiable risk of entrepreneurship. *American Economic Review* 100(3): 1163–1194.

Hamilton, B., Nickerson, J., and H. Owan. 2003. Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy* 111(2): 465–497.

Harel, M., Mossel, E., Strack, P., and O. Tamuz. 2021. Rational groupthink. *Quarterly Journal of Economics* 136(1): 621–668.

Hedge, D., and J. Tumlinson. 2014. Does social proximity enhance business partnerships? Theory and evidence from ethnicity's role in U.S. venture capital. *Management Science* 60(9): 2355–2380.

Hellmann, T. 2006. IPOs, acquisitions, and the use of convertible securities in venture capital. *Journal of Financial Economics* 81: 649–679.

Hodor, I. 2021. Financing innovation with future equity. Working Paper, Monash University.

Howell, S. 2020. Reducing information frictions in venture capital: The role of new venture competitions. *Journal of Financial Economics* 136(3): 676–694.

Howell, S. 2021. Learning from feedback: Evidence from new ventures. *Review of Finance* 25(3): 595–627.

Howell, S., Rathje, J., Van Reenen, J., and J. Wong. 2023. Opening up military innovation: Causal effects of 'bottom-up' reforms to U.S. defense research. Working Paper, New York University.

Hsu, D. 2004. What do entrepreneurs pay for venture capital affiliation? *Journal of Finance* 59(4): 1805–1844.

Hu, A., and S. Ma. 2022. Persuading investors: A video-based study. Working Paper, Yale University.

Huang, C. 2023. Networks in venture capital markets. Working Paper, University of Illinois.

Huang, R., Mayer, E., and D. Miller. 2024. Gender bias in promotions: Evidence from financial institutions. *Review of Financial Studies* 37: 1685–1728.

Ishii, J., and Y. Xuan. 2014. Acquirer-Target Social Ties and Merger Outcomes. *Journal of Financial Economics* 112: 344–363.

Itoh, H. 1991. Incentives to help in multi-agent situations. *Econometrica* 59(3): 611–636.

Jang, Y.S., and S. Kaplan. 2023. Venture capital start-up selection. Working Paper, University of Chicago.

Janis, I. 1972. *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascos*. Houghton Mifflin: Boston.

Janis, I. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascos*. Houghton Miflin: Boston.

Jones, B. 2021. The rise of research teams: Benefits and costs in economics. *Journal of Economic Perspectives* 35: 191–216.

Kanze, D., Huang, L., Conley, M., and T. Higgins. 2018. We Ask Men to Win & Women Not to Lose: Closing the Gender Gap in Startup Funding. *Academy of Management Journal* 61(2): 586–614.

Kaplan, S., Sensoy, B., and P. Strömberg. 2009. Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies. *Journal of Finance* 64(1): 75–115.

Kaplan, S., and P. Strömberg. 2003. Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *Review of Economic Studies* 70: 281–315.

Kaplan, S., and P. Strömberg. 2004. Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *Journal of Finance* 59(5): 2177–2210.

Kerr, W., and R. Nanda. 2011. Financing constraints and entrepreneurship. Handbook of Research on Innovation and Entrepreneurship, 88.

Kerr, W., Nanda, R., and M. Rhodes-Kropf. 2014. Entrepreneurship as experimentation. *Journal of Economic Perspectives* 28(3): 25–48.

Kerr, W., Lerner, J., and A. Schoar. 2014. The consequences of entrepreneurial finance: Evidence from angel financings. *Review of Financial Studies* 27(1): 20–55.

Klocke, U., 2007. How to Improve Decision-Making in Small Groups: Effects of Dissent and Training Interventions. *Small Group Research* 38(3): 437–468.

Ladha, K. 1992. The Condorcet Jury Theorem, Free Speech, and Correlated Votes. *American Journal of Political Science* 36(3): 617–634.

Landier, A., Sraer, D., and D. Thesmar. 2009. Optimal dissent in organizations. *Review of Economic Studies* 76: 761–794.

Laughlin, P., Bonner, B., and A. Miner. 2002. Groups Perform Better than the Best Individuals on Letters-to-Numbers Problems. *Organizational Behavior and Human Decision Processes* 88(2): 605–620.

Lazear, E., and K. Shaw. 2007. Personnel economics: The economist's view of human resources. *Journal of Economic Perspectives* 21: 91–114.

Legros, P., and S. Matthews. 1993. Efficient and nearly-efficient partnerships. *Review of Economic Studies* 68: 599–611.

Lerner, J., and R. Nanda. 2020. Venture capital's role in financing innovation: What we know and how much we still need to learn. *Journal of Economic Perspectives* 34(3): 237–261.

Lyonett, V., and L. Stern. 2024. Machine Learning about Venture Capital Choices. Working Paper, University of Michigan.

Maciejovsky, B., Sutter, M., Budescu, D., and P. Bernau. 2013. Teams Make You Smarter: How Exposure to Teams Improves Individual Decisions in Probability and Reasoning Tasks. *Management Science* 59(6): 1255–1270.

Malenko, A., Nanda, R., Rhodes-Kropf, M., and S. Sundaresan. 2024. Catching Outliers: Committee Voting and the Limits of Consensus when Financing Innovation. Working Paper, Imperial College London.

Malenko, N. 2014. Communication and decision-making in corporate boards. *Review of Financial Studies* 27(5): 1486–1532.

Malmendier, U., and J. Lerner. 2010. Contractibility and the design of research agreements. *American Economic Review* 100: 214–246.

Paluch, S. 2021. The best startup accelerators in 2021. *Beta Boom*. Published 20 January 2021. URL: https://betaboom.com/beta-boom-insights/the-best-startup-accelerators-in-2021-beta-boom-ae2b3629748.

Puri, M., and R. Zarutskie. 2012. On the life cycle dynamics of venture capital and non-venture capital financed firms. *Journal of Finance* 67: 2247–2293.

Ringelmann, M. 1913. Research on animate sources of power: The work of man. *Annals of the National Institute of Agronomy*, 2nd series, 12, 1–40.

Robinson, M. 2022. Factors impacting entrepreneurial success in accelerators: Revealed preferences of sophisticated mentors. *Review of Corporate Finance* 2(3): 617–661.

Schmidt, K. 2003. Convertible securities and venture capital finance. *Journal of Finance* 58(3): 1139–1166.

Simcoe, T., and D. Waguespack. 2011. Status, Quality, and Attention: What's in a (Missing) Name? *Management Science* 57(2): 274–290.

Suh, R., and J. Stiglitz. 1986. The architecture of economic systems: Hierarchies and polyarchies. *American Economic Review* 76(4): 716–727.

Suh, R., and J. Stiglitz. 1988. Committees, hierarchies and polyarchies. *Economic Journal* 98: 451–470.

Teare, G. 2024. "U.S. seed investment actually held up pretty well for the past 2 years. Here's what it means for 2024" URL: https://news.crunchbase.com/seed/us-investment-held-up-forecast-2024/#:~:text=Seed%20trends%20over%20the%20decade&text=At%20the%20market%20peak%20in,years%20of%202019%20and%202020

Ulan, J., and V. Harrison. 2023. Quantifying the success of YC and the Largest Accelerators: Takeaways for VCs, LPs, and Startups. *Pitchbook*: Seattle, United States.

Yang, S., R. Kher, and T. Lyons. 2019. Where do accelerators fit in the venture creation pipeline? Different values brought by different types of accelerators. *Entrepreneurship Research Journal* 8(4): 13.

Yang, M., and Y. Zeng. 2019. Financing entrepreneurial production: Security design with flexible information acquisition. *Review of Financial Studies* 32(3): 819–863.

Yu, S. 2020. How do accelerators impact the performance of high-technology ventures. *Management Science* 66(2): 530–552.

Zhang, Y. 2023. Discrimination in the venture capital industry: Evidence from field experiments. Working Paper, Stockholm School of Economics.

Figure 1: Overview of our VC firm's accelerator deal flow by cohort

This figure shows the total number of applications received by our VC firm's startup accelerator by cohort, and the percentage of applicants that were eliminated from further consideration by our VC firm at the initial screening stage, the pre-interview stage, the first round interview stage, and the second round interview stage, as well as the percentage of applicants that were ultimately accepted into our VC firm's accelerator program.

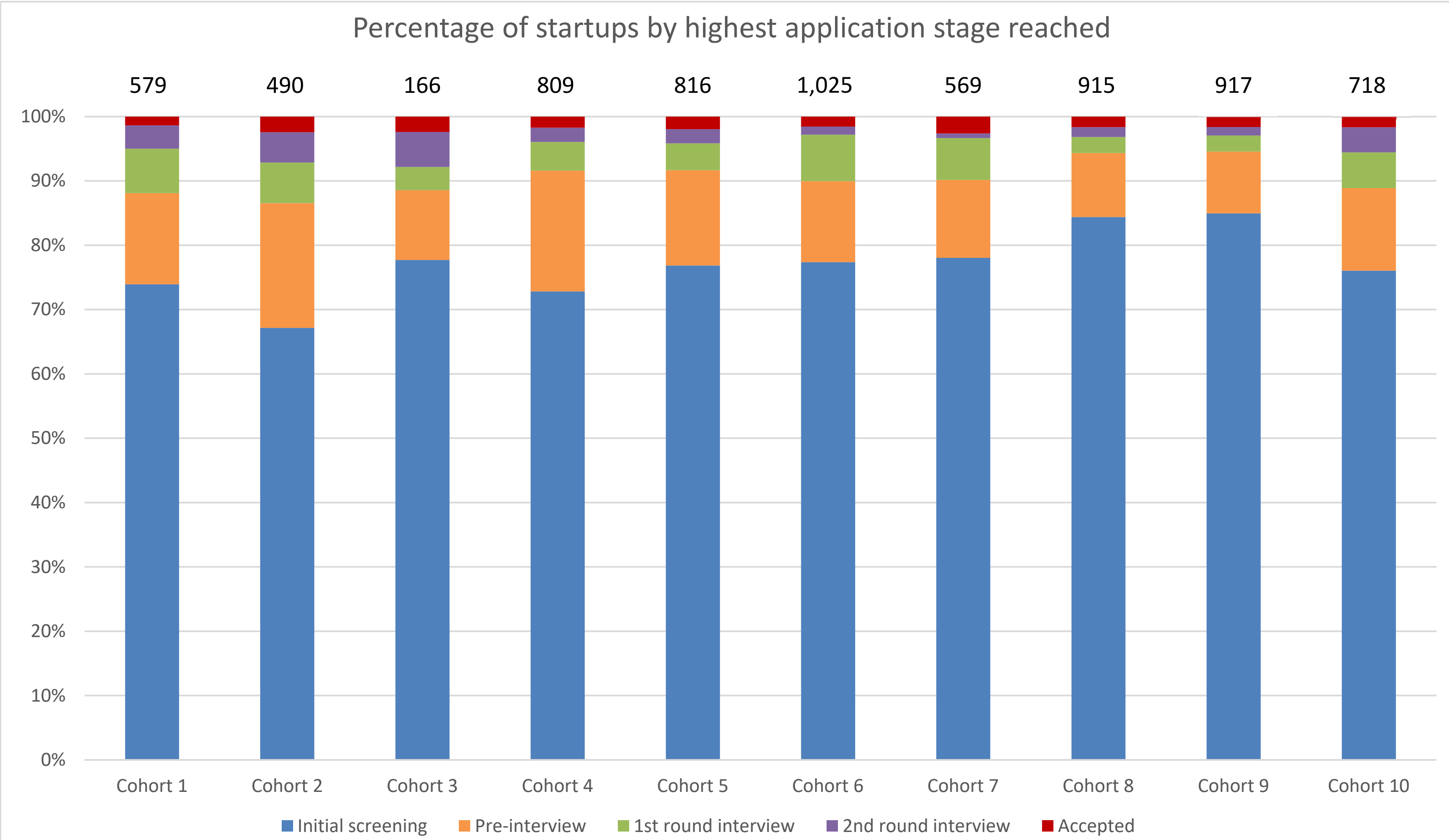

Figure 2: Applicants by industry across all accelerator cohorts

This figure shows the percentage of startup applicants by industry sector across all 10 of our VC firm's accelerator cohorts. The startup firm applicant's industry sector was obtained from answers provided by the startup applicant to the application question, "What is your industry?" Applicants were offered the following choice of industry sectors: (1) Fintech/Financial Services, (2) Biotechnology/Medical Devices/Healthcare, (3) Foodtech/Consumer Food Products, (4) Adtech/Digital Marketing/Media, (5) Internet/Web Service/"Apps"/Software/eCommerce, (6) Hardware/Electronic, (7) Agtech/Energy, (8) Education Technologies/Human Resources, and (9) Other.

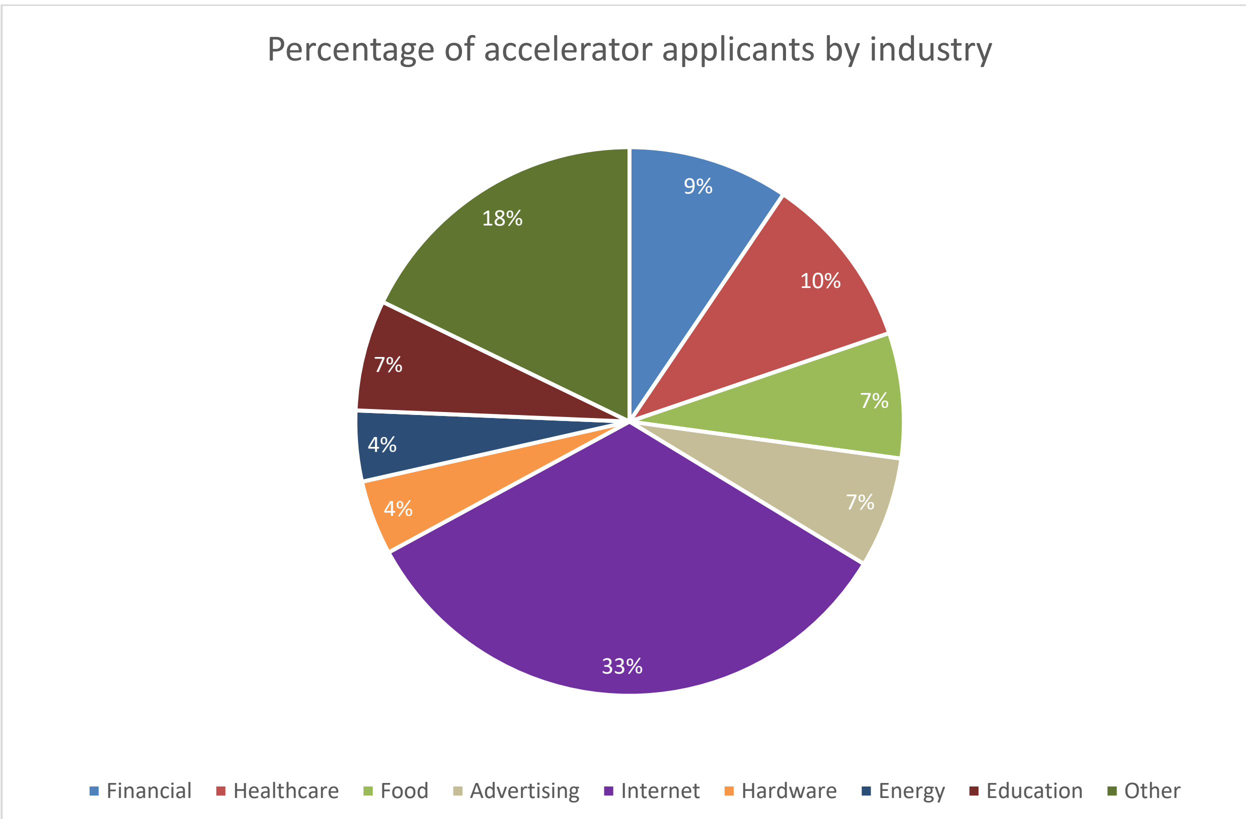

Table 1: Summary statistics

This table reports the summary statistics for the entire sample of startups applying to our VC firm's accelerator program as well as for the sub-sample of startups that are ultimately accepted into the accelerator. Panel A outlines the characteristics of internal VC judging scores for accelerator applicants throughout the VC selection process. Panel B lists the characteristics of startup firms at their time of application. Panel C presents the characteristics of our VC firm's employee judges. Please refer to Appendix A for the definition of all variables listed.

Panel A: Internal VC judge score characteristics

	All applicants			*Accepted applicants only*		
	Mean	Median	Std dev.	Mean	Median	Std. dev.
Scores per applicant in pre-interview stage	8.60	9.00	4.13	8.40	9.00	4.48
Scores per applicant in 1st round interview	6.34	7.00	3.45	6.16	7.00	3.18
Scores per applicant in 2nd round interview	6.41	7.00	3.73	6.09	7.00	3.49
Overall pre-interview judge score	0.53	0.53	0.17	0.67	0.66	0.13
First round interview judge score	0.49	0.46	0.23	0.70	0.71	0.14
Second round interview judge score	0.48	0.48	0.30	0.71	0.72	0.20

Panel B: Startup firm applicant characteristics

	All applicants			*Accepted applicants only*		
	Mean	Median	Std dev.	Mean	Median	Std. dev.
Company age (years)	2.55	1.91	1.62	3.29	2.75	1.94
Company lifetime revenue to date (US$)	78,230	10,012	995,875	401,940	69,434	960,017
External funding raised to date ($US)	45,698	0	201,589	49,772	0	155,587
Company runway (months)	6.54	6.00	7.45	6.90	6.00	4.63
No. of company founders	2.25	2.00	0.88	2.29	2.00	0.87
No. of FTE employees	3.15	3.75	3.74	4.95	4.00	4.33

Panel C: VC employee judge characteristics

	Entire judge sample		
	Mean	Median	Std. dev.
Judge has a Graduate degree	0.46	0.00	0.51
Judge attended a Top tier university	0.50	0.50	0.51
Years of financial investment experience	6.59	3.00	5.88

Table 2: VC selection – Baseline tests using Pre-interview scores

This table presents the results of the ordinary least squares (OLS) regression specification outlined in Equation (1) where the dependent variable is the overall score given by an individual judge for a specific startup firm applicant during the pre-interview judging stage. The judge-founder affinity-based traits are *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*. The list of variables included in the *Startup firm controls*, *Startup founding team controls*, *Judge controls*, and *Controls for other judge–founder overlapping characteristics* vectors, respectively, are described in Sections 2.4 and Appendix A. All regressions include Judge fixed effects, Startup firm fixed effects, and Cohort fixed effects. Robust standard errors (clustered at the startup applicant level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Scoring outcome variable	Overall pre-interview score (1)	Overall pre-interview score (2)
Shared gender	0.04**	0.04**
	(0.02)	(0.02)
Shared ethnicity	0.05***	0.05**
	(0.01)	(0.02)
Shared education	0.06**	0.05*
	(0.03)	(0.03)
Shared employer	0.00	0.00
	(0.03)	(0.03)
Startup firm controls	No	Yes
Startup founding team controls	No	Yes
Judge controls	No	Yes
Controls for other judge–founder overlapping characteristics	No	Yes
Judge FEs	Yes	Yes
Startup firm FEs	Yes	Yes
Cohort FEs	Yes	Yes
Number of observations	13,518	13,518
Adjusted R^2	0.57	0.59

Table 3: VC selection – Heterogenous treatment effects tests using Pre-interview scores

This table presents the results of the ordinary least squares (OLS) regression specification outlined in Equation (1) for various sub-samples of applications where the dependent variable is the overall score given by an individual judge for a specific startup applicant during the pre-interview judging stage. In Columns (1) and (2), our sample is split into startup applicant firms that are pre-revenue versus those that are not pre-revenue. In Columns (3) and (4), our sample is split into firms where all the startup's founders are founding their first ever entrepreneurial venture versus those firms who have founders that are serial entrepreneurs. The judge-founder affinity-based traits are *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*. The independent variables measuring possible discrimination against minority entrepreneurs are *All Female founder team*, *All Black founder team*, and *All Hispanic founder team*. The list of variables included in the *Startup firm controls*, *Startup founding team controls*, *Judge controls*, and *Controls for other judge–founder overlapping characteristics* vectors, respectively, are described in Sections 2.4 and Appendix A. All regressions include Judge fixed effects, Startup firm fixed effects, and Cohort fixed effects. Robust standard errors (clustered at the startup applicant level) are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Scoring outcome variable	Overall pre-interview score	Overall pre-interview score	Overall pre-interview score	Overall pre-interview score
Application subsample	Pre-revenue firms	Not pre-revenue firms	No prior founding experience	Does have prior founding experience
	(1)	(2)	(3)	(4)
Shared gender	0.07**	0.00	0.06*	-0.01
	(0.03)	(0.02)	(0.03)	(0.02)
Shared ethnicity	0.08***	0.01	0.09***	0.00
	(0.02)	(0.02)	(0.03)	(0.03)
Shared education	0.06**	-0.01	0.07**	-0.01
	(0.03)	(0.03)	(0.03)	(0.04)
Shared employer	0.01	-0.01	0.00	-0.01
	(0.03)	(0.02)	(0.04)	(0.04)
Startup firm controls	Yes	Yes	Yes	Yes
Startup founding team controls	Yes	Yes	Yes	Yes
Judge controls	Yes	Yes	Yes	Yes
Controls for other judge–founder overlapping characteristics	Yes	Yes	Yes	Yes
Judge FEs	Yes	Yes	Yes	Yes
Startup firm FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Number of observations	5,399	8,119	9,147	4,471
Adjusted R^2	0.51	0.52	0.46	0.48

Table 4: VC selection – Baseline tests using First and Second round interview scores

This table presents the results of the ordinary least squares (OLS) regression specification outlined in Equation (1) where the dependent variable is the overall score given by an individual judge for a specific startup firm applicant during the first round interview judging stage (Columns (1) and (2)) and the second round interview judging stage (Columns (3) and (4)), respectively. In Columns (1) and (3), all first round and second round interview scores are included in the test sample, irrespective of whether or not an individual employee also judged and scored an accelerator candidate in the pre-interview stage, respectively. In contrast, Columns (2) and (4) only includes the first round and second round interview scores of judges who also submitted an interview score for the same startup applicant in a previous interview round. The judge-founder affinity-based traits are *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*. The list of variables included in the *Startup firm controls*, *Startup founding team controls*, *Judge controls*, and *Controls for other judge–founder overlapping characteristics* vectors, respectively, are described in Sections 2.4 and Appendix A. All regressions include Judge fixed effects, Startup firm fixed effects, and Cohort fixed effects. Robust standard errors (clustered at the startup applicant level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Scoring outcome variable	First round interview score (1)	First round interview score (2)	Second round interview score (3)	Second round interview score (4)
Shared gender	0.01	0.01	0.02	0.01
	(0.03)	(0.03)	(0.04)	(0.06)
Shared ethnicity	0.02	0.02	0.02	0.01
	(0.04)	(0.04)	(0.05)	(0.04)
Shared education	0.00	0.00	0.00	0.01
	(0.05)	(0.05)	(0.04)	(0.05)
Shared employer	-0.01	-0.01	-0.01	-0.02
	(0.04)	(0.04)	(0.04)	(0.05)
Require employee to have judged startup in a previous interview round?	*No*	*Yes*	*No*	*Yes*
Startup firm controls	Yes	Yes	Yes	Yes
Startup founding team controls	Yes	Yes	Yes	Yes
Judge controls	Yes	Yes	Yes	Yes
Controls for other judge–founder overlapping characteristics	Yes	Yes	Yes	Yes
Judge FEs	Yes	Yes	Yes	Yes
Startup firm FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Number of observations	3,401	2,203	1,663	1,152
Adjusted R^2	0.51	0.47	0.49	0.46

Table 5: VC selection – Heterogenous treatment effects tests using First round interview scores

This table presents the results of the ordinary least squares (OLS) regression specification outlined in Equation (1) for various sub-samples of applications where the dependent variable is the overall score given by an individual judge for a specific startup applicant during the first round interview judging stage. In Columns (1) and (2), our sample is split into startup firms that are pre-revenue versus those that are not pre-revenue. In Columns (3) and (4), our sample is split into firms where all the startup's founders are founding their first ever entrepreneurial venture versus those firms who have founders that are serial entrepreneurs. The judge-founder affinity-based traits are *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*. The list of variables included in the *Startup firm controls*, *Startup founding team controls*, *Judge controls*, and *Controls for other judge–founder overlapping characteristics* vectors, respectively, are described in Sections 2.4 and Appendix A. All regressions include Judge fixed effects, Startup firm fixed effects, and Cohort fixed effects. Robust standard errors (clustered at the startup applicant level) are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Scoring outcome variable	First round interview score	First round interview score	First round interview score	First round interview score
Application subsample	Pre-revenue firms	Not pre-revenue firms	No prior founding experience	Does have prior founding experience
	(1)	(2)	(3)	(4)
Shared gender	0.01	0.00	-0.02	-0.04
	(0.04)	(0.04)	(0.03)	(0.02)
Shared ethnicity	0.02	0.01	0.01	0.02
	(0.03)	(0.03)	(0.03)	(0.03)
Shared education	0.04	-0.01	-0.01	-0.03
	(0.06)	(0.04)	(0.03)	(0.04)
Shared employer	0.01	-0.01	0.00	-0.02
	(0.03)	(0.06)	(0.04)	(0.04)
All Female founder team	-0.02	-0.01	-0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.03)
All Black founder team	-0.00	-0.01	-0.03	0.00
	(0.06)	(0.04)	(0.05)	(0.06)
All Hispanic founder team	-0.00	-0.01	-0.02	0.02
	(0.07)	(0.04)	(0.04)	(0.04)
Startup firm controls	Yes	Yes	Yes	Yes
Startup founding team controls	Yes	Yes	Yes	Yes
Judge controls	Yes	Yes	Yes	Yes
Controls for other judge–founder overlapping characteristics	Yes	Yes	Yes	Yes
Judge FEs	Yes	Yes	Yes	Yes
Startup firm FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Number of observations	1,105	2,296	2,607	794
Adjusted R^2	0.45	0.47	0.43	0.45

Table 6: VC selection – Alternative tests controlling for judge's prior score for focal startup

This table presents the results of the ordinary least squares (OLS) regression specification outlined in Equation (2) where the dependent variable is the overall score given by an individual judge for a specific startup firm applicant during the first round interview judging stage. *Pre-interview overall score* is equal to the score that the individual judge gave for the focal startup applicant during their pre-interview evaluations. The judge-founder affinity-based traits are *Shared gender*, *Shared ethnicity*, *Shared education*, and *Shared employer*. The list of variables included in the *Startup firm controls*, *Startup founding team controls*, *Judge controls*, and *Controls for other judge–founder overlapping characteristics* vectors, respectively, are described in Sections 2.4 and Appendix A. All regressions include Judge fixed effects, Startup firm fixed effects, and Cohort fixed effects. Robust standard errors (clustered at the startup applicant level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Scoring outcome variable	First round interview score	First round interview score	Second round interview score	Second round interview score
	(1)	(2)	(3)	(4)
Pre-interview overall score	0.26***	0.22***	0.15***	0.16***
	(0.05)	(0.06)	(0.04)	(0.05)
Shared gender		0.02		0.01
		(0.04)		(0.06)
Shared ethnicity		0.02		0.01
		(0.05)		(0.04)
Shared education		0.00		0.01
		(0.04)		(0.05)
Shared employer		-0.01		-0.02
		(0.04)		(0.05)
Pre-interview overall score × Shared gender		-0.06**		-0.05*
		(0.03)		(0.03)
Pre-interview overall score × Shared ethnicity		-0.07**		-0.05**
		(0.03)		(0.02)
Pre-interview overall score × Shared education		-0.06**		-0.06*
		(0.03)		(0.04)
Pre-interview overall score × Shared employer		0.01		-0.04
		(0.04)		(0.05)
Startup firm controls	Yes	Yes	Yes	Yes
Startup founding team controls	Yes	Yes	Yes	Yes
Judge controls	Yes	Yes	Yes	Yes
Controls for other judge–founder overlapping characteristics	Yes	Yes	Yes	Yes
Judge FEs	Yes	Yes	Yes	Yes
Startup firm FEs	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes
Number of observations	2,203	2,203	1,152	1,152
Adjusted R^2	0.43	0.44	0.42	0.40

Table 7: Startup investment outcomes – Value implications of different portfolio selection rules

This table presents the results of the regression specification outlined in Equation (3) where the dependent variables are various measures capturing the post-investment outcomes of startup firm applicants to our VC firm's accelerator program. Specifically, the startup firm outcome variables in columns (1), (2), (3), and (4) are whether the startup firm is *Out of business*, the *Number of funding rounds post-application*, the *Amount of funding raised post-application*, and *Post-money startup valuation post-application*, respectively. All regressions include Cohort fixed effects. Robust standard errors (clustered at the startup applicant level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

In Panel A, the independent variable *Portfolio selection based on 2nd round interview consensus scores* is an indicator variable that is equal to one if the startup firm applicant *was* ultimately chosen for investment by our VC firm (but *would not have* been selected if applicants were instead ranked by the highest consensus pre-interview overall judging scores), and zero if the startup firm applicant *was not* ultimately chosen for investment by our VC firm (but *would have* been selected if applicants were instead ranked by the highest consensus pre-interview overall judging scores). Note that all firms that would either (i) be selected for investment by our VC firm under both decision rules or (ii) be rejected under both decision rules are excluded from this analysis.

In Panel B, the independent variable *Portfolio selection based on 2nd round interview consensus scores* is an indicator variable that is equal to one if the startup firm applicant *was* ultimately chosen for investment by our VC firm (but *would not have* been selected if applicants were instead ranked by the highest maximum score given by judges during the 2nd round interview process under a so-called "champion" approach), and zero if the startup applicant *was not* ultimately chosen for investment by our VC firm (but *would have* been selected if applicants were instead ranked by the highest maximum score given by judges during the 2nd round interview process under a so-called "champion" approach). Note that all firms that would either (i) be selected for investment by our VC firm under both decision rules or (ii) be rejected under both decision rules are excluded from this analysis.

Panel A: Consensus 2nd round interview score 'portfolios' vs. Consensus pre-interview score 'portfolios'

Scoring outcome variable	Out of business (1)	Number of funding rounds post-application (2)	Amount of funding raised post-application (3)	Post-money startup valuation post-application (4)
Portfolio selection based on 2nd round interview consensus scores	-0.14*** (0.06)	0.22** (0.10)	0.02** (0.01)	0.29* (0.16)
Cohort FEs	Yes	Yes	Yes	Yes
Number of observations	182	182	161	119
Adjusted R^2	0.07	0.06	0.08	0.12

Panel B: Consensus 'portfolios' vs. Champion (maximum) score 'portfolios' using 2nd round interview scores

Scoring outcome variable	Out of business (1)	Number of funding rounds post-application (2)	Amount of funding raised post-application (3)	Post-money startup valuation post-application (4)
Portfolio selection based on 2nd round interview consensus scores	-0.11** (0.05)	0.20* (0.12)	0.01 (0.01)	-0.38* (0.20)
Cohort FEs	Yes	Yes	Yes	Yes
Number of observations	111	111	97	79
Adjusted R^2	0.10	0.12	0.09	0.10

Table 8: Mechanism analysis – Evidence from judges' written comments

This table presents the results of the regression specification outlined in Equation (4) where the dependent variables are various measures capturing the written comment behavior of individual VC firm judges when evaluating a given startup firm applicant. Specifically, the written comment outcome variables in columns (1), (2), and (3) are whether the judge provides any written comment justifying their scoring assessment, and zero otherwise ($\mathbb{I}$*{Written comment}*), the *Comment length*, and the *Relative focus on team*-related characteristics in the judge's written comments, respectively. In Panel A, the sample consists of all pre-interview scores and associated written comments, while in Panel B, the sample consists of all second-round interview scores and associated written comments. All regressions include *Startup firm controls*, *Startup founding team controls*, *Judge controls*, and *Controls for other judge–founder overlapping characteristics* vectors, respectively, are described in Sections 2.4 and Appendix A. All regressions also include Judge fixed effects, Startup firm fixed effects, and Cohort fixed effects. Robust standard errors (clustered at the startup applicant level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Analysis of pre-interview stage written comments ('solo judge' setting)

Written comment outcome variable	$\mathbb{I}${Written comment}	Comment length	Relative focus on team
	(2)	(3)	(4)
Affiliated judge	-0.17***	-0.10***	0.25**
	(0.04)	(0.02)	(0.11)
Startup firm controls	Yes	Yes	Yes
Startup founding team controls	Yes	Yes	Yes
Judge controls	Yes	Yes	Yes
Controls for other judge–founder overlapping characteristics	Yes	Yes	Yes
Judge FEs	Yes	Yes	Yes
Startup firm FEs	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes
Number of observations	13,518	13,518	13,518
Adjusted R^2	0.08	0.09	0.11

Panel B: Analysis of second round interview stage written comments ('group judge' setting)

Written comment outcome variable	$\mathbb{I}${Written comment}	Comment length	Relative focus on team
	(2)	(3)	(4)
Affiliated judge	0.02	-0.01	-0.01
	(0.05)	(0.03)	(0.07)
Startup firm controls	Yes	Yes	Yes
Startup founding team controls	Yes	Yes	Yes
Judge controls	Yes	Yes	Yes
Controls for other judge–founder overlapping characteristics	Yes	Yes	Yes
Judge FEs	Yes	Yes	Yes
Startup firm FEs	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes
Number of observations	1,663	1,663	1,663
Adjusted R^2	0.03	0.05	0.04

Appendix A: Variable definitions

Variable	Description
Panel A: Outcome variables	
Pre-interview overall score	The total score awarded for an individual startup applicant by each judge during our VC firm's pre-interview evaluation process. Scores are standardized across cohorts to be bounded between 0 and 1.
First round interview score	The total score awarded for an individual startup company by each judge after our VC firm conducts a first round interview with the accelerator applicant. Scores are standardized across cohorts to be bounded between 0 and 1.
Second round interview score	The total score awarded for an individual startup company by each judge after our VC firm conducts a second round interview with the accelerator applicant. Scores are standardized across cohorts to be bounded between 0 and 1.
Out of business	An indicator variable that is equal to one if the startup firm is no longer operating and/or is bankrupt, according to the Pitchbook, Crunchbase or CB Insights databases.
Number of funding rounds post-application	The natural logarithm of one plus the number of funding rounds completed by the startup firm in the time period after the startup's initial application to our VC firm's accelerator program, based on the detailed funding round information provided in the Pitchbook, Crunchbase, and CB Insights databases.
Amount of funding raised post-application	The natural logarithm of one plus the cumulative amount of funding (in US$ million) raised by the startup firm in the time period after the startup's initial application to our VC firm's accelerator program, based on the detailed funding round information provided in the Pitchbook, Crunchbase, and CB Insights databases.
Post-money startup valuation post-application	The natural logarithm of the most recent "post-money valuation" of the startup firm reported in the Pitchbook database, using only figures related to funding rounds that occur after the startup's initial application to our VC firm's accelerator program.
Panel B: Judge-founder affinity-based traits and startup founding team inherited characteristics	
Shared gender	An indicator variable equal to one if the individual judge and at least one of the startup firm's founders have the same gender (male or female), and zero otherwise. To identify a person's gender, we start with a startup firm's application materials (including pitch decks and LinkedIn profiles) which almost always includes pictures of each firm founder along with biographies and other work-related descriptions. Typically, the gender of each firm founder is clearly identified from these materials through word use (e.g., written references to 'she', 'he', 'her', 'him', etc.) and/or visual inspection of pictures. For the few remaining ambiguous cases, we use both genderize.io and forebears.io that predict gender based on first names, supplemented with manual web searches using publicly available online sources.

Shared ethnicity	An indicator variable equal to one if the individual judge and at least one of the startup firm's founders have the same ethnicity (White, East Asian, Indian, Middle Eastern, Black/African American, Hispanic/Latino, and Other), and zero otherwise. To identify a person's ethnicity, we use ChatGPT's *Ethnicity Identifier* tool. For each founder, we upload their profile picture and ask this specialized ChatGPT program the following prompt: "based on the attached picture, the person's full name of [insert name], the person's current location of [insert city name, country name], and [if available] the location of the person's undergraduate academic institution (namely [insert university name]), what is the ethnicity and country of origin of this person? Provide a confidence score between 0 and 10 for your predictions, with 0 being the least confident and 10 being the most confident." If the tool identifies the individual as having a mixed ethnic background or the assigned confidence score is less than 9, we set the relevant founder's Ethnicity = "Other". For robustness, we provide the same inputs into ChatGPT's *Ethnicity Guesser* tool and set Ethnicity = "Other" if the two programs disagree on their ethnicity prediction.
Shared education	A dummy variable equal to one if the individual judge and at least one of the startup's founders graduated with a degree from the same university, and zero otherwise. To identify the full list of universities attended (and the associated degrees earned) by each firm founder, we use both application questions asking for the educational background of each member of the firm's current management team as well as education-related information listed in LinkedIn and other similar profiles.
Shared employer	An indicator variable equal to one if the individual judge and at least one of the startup firm's founders worked at the same employer, and zero otherwise. To identify each firm founder's complete list of previous employers prior to founding the focal startup firm, we use both application questions asking for the employment background of each member of the firm's current management team as well as employment-related information listed in LinkedIn and other similar profiles.
Panel C: Other judge-founder overlapping characteristics	
Shared graduate degree	A dummy variable equal to one if both the VC firm judge and at least one of the startup's founders have earned an academic degree after their initial bachelor's degree (namely a Master's degree or a PhD), and zero otherwise.
Shared top tier university	A dummy variable equal to one if both the VC firm judge and at least one of the startup's founders have been granted a degree from a university that is ranked as one of the world's Top 50 best bachelor's degree-granting institutions (according to that year's *Times Higher Education World University Rankings*), and zero otherwise.
Shared industry experience	The natural logarithm of one plus the number of years of experience that the focal VC firm judge has working in the same industry sector as the focal startup company.
Panel D: Startup firm characteristic (time-varying) control variables	
Company stage of development	A dummy variable equal to one if the firm has already publicly launched its product or service, and zero otherwise. For example, a startup firm that is still in the concept or prototype phase will have a value of zero for this indicator variable.
Company age	The natural logarithm of the number of months between the firm's founding date and the date of the startup's application to our VC firm's accelerator program.
Company's lifetime revenue	The natural logarithm of one plus the dollar amount of sales revenue generated by the startup firm's products/services to date.
Number of total users since launch	The natural logarithm of one plus the total number of people that have used the startup's product or service since its launch (if applicable). This variable is set to equal zero if the startup's product or service has not yet been launched.

Number of paying users since launch	The natural logarithm of one plus the total number of people who have paid to use the startup's product or service since its launch (if applicable). This variable is set to equal zero if the startup's product or service has not yet been launched.
External funding raised to date	The natural logarithm of one plus the total amount of capital raised from investors who are not part of the startup's management/founding team (e.g., angel investors, family & friends, governmental entities etc.).
Company runway	The natural logarithm of one plus the number of months left before the startup company exhausts its existing cash reserves (in the absence of any new investment).
Current firm valuation	The natural logarithm of the startup company's (self-reported) current valuation.
Number of FTE company employees	The natural logarithm of one plus the total number of full-time equivalent (FTE) employees working at the startup applicant.
Estimated Serviceable Obtainable Market (SOM)	The natural logarithm of one plus the startup's estimate of the serviceable obtainable market (SOM) for its product or service, defined as the portion of the estimated serviceable addressable market (SAM) that the startup can realistically capture given its business model. We set SOM equal to zero if the startup firm declines to provide an estimate of its SOM due to uncertainty about its preferred target market. Note: SAM is defined as the segment of estimated total addressable market (TAM) realistically targeted by the startup's products and services, where TAM is defined as the total (maximum) revenue opportunity available for a product or service.
Panel E: Startup founding team (time-varying) control variables	
Average TMT experience of company founders	The natural logarithm of one plus the average number of years that each startup co-founder acted in a top management team (TMT)/corporate executive role, including their time at the current applicant company.
Average years of startup founding experience of company founders	The natural logarithm of one plus the average number of years that each co-founder has worked at a startup company that they have created, including their time at the current applicant company.
Panel F: VC judge (time-varying) control variables	
Years of financial investment experience	The natural logarithm of one plus the number of years that the focal VC firm judge has worked in the financial investment sector (including experience gained in venture capital/private equity-focused roles as well as positions in the asset management, investment banking, and management consulting industries.
Amount of startup firm judging experience	The natural logarithm of one plus the total number of separate scoring assessments submitted by the focal judge prior to the current accelerator cohort intake.

The Internet Appendix of

“How do Venture Capitalists (actually) make decisions? Internal evidence from a private startup accelerator”

Appendix IA.1: Additional details on our VC firm's selection and contracting process

In this Appendix, we provide additional information about the entire process that our VC firm employs to select the subset of startup firm applicants who will be invited to participate in our VC's accelerator program as well as how investment contracts are negotiated with accepted applicants.

Stage 1: Application and initial screening

The first step in our VC firm's selection process requires that startups submit an initial online application by the relevant cohort application deadline. This online form asks a series of more basic questions about the company, the firm's founders, the startup's business model, and the company's progress to date. One junior employee at our VC Fund will then review each initial application to make a binary 'Yes' or 'No' determination as to whether a further request for information (called a 'due diligence pack' or 'DD pack') is sent to the startup applicant to complete (a process that we term the '*initial employee screen*'). This DD pack is a much more detailed questionnaire that asks 50+ questions and covers a broad range of subjects relating to the company's product or service, potential market size, business model, competitive landscape, milestones achieved to date (i.e., traction), financial information and other related business metrics, the background and skills of the management team, and the legal/ownership structure. In addition, applicants are also expected to submit 'pitch decks' that provide a more visual representation of the startup's business plan and offer applicants a more open-ended outlet to describe the founding team, the market opportunity, the proposed solution, potential challenges, funding needs etc.

As an aside, the only part of our VC firm's entire investment selection process for which we *do not* have complete information concerns the initial employee screening process. While we observe whether a startup applicant was invited to submit a due diligence pack or not, our VC firm never required their sole junior screener to provide any written justification for their decision nor kept any formal records relating to these decisions. However, during discussions with employees tasked with undertaking this initial screening, it became apparent that the primary purpose of these 'initial screens' was only to filter out obviously poor applications that were clearly not worthy of any further consideration rather than serve as a meaningful part of the VC firm's overall investment selection process. Examples of startups that were rejected at this initial phase were those that did not take the time to fill out many of the basic questions on the initial application form, companies that did not have a functioning website, and/or businesses that were clearly unsuitable for a high-growth orientated accelerator program (e.g., small local businesses with no plans for meaningful expansion). These anecdotal observations are supported empirically in Appendix Table IA.1 where we find that answering every one of our VC's initial online application questions is overwhelming the most important predictor of whether an applicant is asked to submit a DD pack for further evaluation by the entirety of the VC firm's personnel.[38] As such, it appears unlikely that these initial screening decisions will have a material impact on our subsequent analysis and conclusions.

Stage 2: Pre-interview assessment

Once all requested DD packs and pitch decks are received, the widely recognized expectation at our VC firm is that all VC employees who are available to read and evaluate these due diligence materials will submit individual scores for each applicant (otherwise referred to as '*pre-interview scores*'). This pre-interview score can range between 0 points (worst) to 100 points (best) and is based on sub-scores given for the market potential of the startup's product/service, the quality of the startup's management team, and the level of traction/customer engagement garnered to date (with each category assigned roughly equal weights).[39] Given the high volume and wide diversity

[38] Also, our VC firm progressed over 23% of applicants to its pre-interview assessment stage, much higher than the 4% of applicants that the VC firm studied in Jang and Kaplan (2023) chose to "intensively analyze" and formally score.
[39] To protect our VC firm's identity, we cannot disclose the precise weights that they put on each of these three criteria. Thus, our later empirical analysis of pre-interview judging scores will only focus on the total 0-to-100-point scores.

of applicants that each VC employee 'judge' must evaluate within a compressed time period, these pre-interview scores are only meant to be based on the submitted DD packs and pitch decks.[40]

Critically, our VC firm adopted a judging policy that prioritized each VC partner/employee submitting truly independent assessments of applicant quality, especially at the pre-interview stage. Specifically, our VC firm went to great lengths to proactively ensure the independence and integrity of the internal judging process. For example, during the entirety of the pre-interview judging period, employees were physically separated and instructed not to communicate with one another about their personal assessments of startup applicants (e.g., all employees were not allowed to review applications in the office but instead had to work from home during this time period). VC employees would then be required to separately submit an individual scoring spreadsheet to an administrative manager to compile and summarize, thus helping to ensure that no VC judge had access to other employees' scores prior to submitting their own scores and comments.

Next, once all individual pre-interview judging scores are received, a weight is then applied to each individual score. For most cohorts, each pre-interview score (whether from a partner or a full-time employee) is assigned the same weight.[41] Our VC firm then takes an average of the (weighted) judges' scores and ranks startups from best to worst based solely on these weighted average scores. At the beginning of each cohort cycle, our VC firm will prespecify a capacity threshold for how many startups they can include as portfolio firms in the accelerator program (typically 10–12 startups) and how many interview slots they can feasibly accommodate at each

[40] For example, within a two-day period at the start of each cohort cycle, the average number of applications that each VC judge will score in this pre-interview stage is approximately 200 firms. As such, our conversations with VC firm employees suggest that there is simply insufficient time to conduct significant additional research outside of relying on the (extensive) submitted application materials during this pre-interview evaluation process.

[41] In some later cohorts, however, the scores of VC partners received up to double the weight of other VC employees.

stage of the selection process (typically 80–100 startups in the pre-interview stage). Our VC firm then makes first round interview offers in order of rank until all interview slots have been filled.

Stage 3: First round interview

The third step in our VC firm's selection process is that these selected startup applicants would be invited to a 30-minute meeting with all available VC firm partners and employees. These meetings were usually conducted in-person although there were some instances where a video-conferencing call was instead scheduled. At these meetings, key members of the startup's management team will be invited to make a short presentation about the company and answer a series of (impromptu) questions from VC firm employees. At the conclusion of each first round interview, all attendees from the VC firm would have an open group discussion about the strengths and weaknesses of the applicant. However, analogous to the pre-interview scoring process, each interviewer would be required to separately submit an individual scoring spreadsheet after all first round interviews are conducted. For each first round candidate, each interviewer is asked to provide a single overall score based on that VC judge's holistic assessment of the startup's potential and fit for the accelerator program (otherwise referred to as '*first round interview scores*').

Next, once all individual first round interview judging scores are received, a weight is then applied to each individual score based on level of seniority. In this round, all individual scores receive the same weight, with the exception that each VC firm partner's first round interview score was typically given 1.5 times the weight of scores submitted by other full-time VC employees. Our VC firm then takes an average of the (weighted) judges' scores and ranks startups from best to worst based solely on these weighted average first round interview scores. Our VC firm then makes second round interview offers in order of rank until all available interview slots have been filled.

Stage 4: Second round interview and final selection

For applicants that successfully pass the first round interview stage, these shortlisted startup companies will then have a second (and final) 45-minute, in-person interview with all available VC partners and employees. At this second interview, key members of the startup's management team will be invited to make a longer presentation pitch for VC funding and answer a series of additional questions from VC firm employees about the startup's business. At the end of each second round interview, all VC firm interviewers have an open group discussion about the relative strengths and weaknesses of the applicant and their suitability for VC investment. For each second round interview candidate, each interviewer will then be asked to provide a single overall score based on that VC judge's holistic assessment of the startup's potential and fit for the accelerator program (otherwise referred to as '*second round interview scores*'). Analogous to the first round interview process, however, each VC firm interviewer is required to separately submit an individual scoring spreadsheet with associated comments after all second round interviews are conducted.

Next, once all individual second round interview judging scores are received, a weight is then applied to each individual score based on level of seniority. In this round, each VC firm partner's second round interview score was typically given 1.5 times the weight of scores submitted by other full-time VC employees, but all individuals within the same level of seniority received the same scoring weight. Our VC firm then takes an average of the (weighted) judges' scores and ranks startups from best to worst based solely on these weighted average second round interview scores.[42] Our VC firm will then make offers to startup companies to join the accelerator cohort and receive VC firm funding in order of rank until all available accelerator cohort slots are filled.

[42] It should be noted that our VC firm only uses the second-round interview scores to decide whether a startup company is ultimately accepted into the VC's accelerator program (i.e., an applicant's second round interview scores effectively supersede that applicant's pre-interview and first round interview scores).