Human Capital Breadth and the Financing of Innovative Startups *

Teodor Duevski†

Abstract

I examine how the breadth of venture capital (VC) partners' human capital influences investment selection, startup performance, and innovation. Partners with broader human capital are more likely to lead investments in novel startups with previously unexplored business models and significantly increase their likelihood of major success; however, they underperform when leading non-novel deals. Exploiting plausibly exogenous variation in partner time constraints as a shock to the within-VC firm likelihood of leading a deal, I provide causal evidence for these effects. A theoretical model endogenizes startup creation, partner assignment, and investment to rationalize the empirical findings and provide additional testable predictions. The results highlight the nuanced value of human capital breadth in financing innovation.

JEL Classification: G11, G24, G34

*I am very grateful to Bruno Biais, Thierry Foucault, Johan Hombert, Boris Vallée and Josh Lerner for their continuous guidance and support. I would like to thank Aleksandar Andonov, Malcolm Baker, Scott Baker, Jean-Edouard Colliard, Nicolas Crouzet, Vicente Cuñat, Matthew Denes (discussant), Francois Derrien, Michael Ewens, Quirin Fleckenstein, Xavier Giroud, Arpit Gupta, Ulrich Hege, Sebastian Hillenbrand, Victoria Ivashina, Jessica Jeffers, Pedro Matos, Vincent Maurin, Ramana Nanda, Jacqueline Ng Lane, Vincent Maurin, Ramana Nanda, Aleksei Ovchinnikov, Evren Ors, Maddalena Ronchi, Suproteem Sarkar, Daniel Schmidt, Farzad Saidi, Luis Viceira, Guillaume Vuillemey, Zijun Wang, Moritz Wiedemann (discussant), as well as seminar participants at Harvard Finance Lunch Seminar, 21st WashU Olin Conference (PhD poster), NFA, ENTFIN, Chicago Booth, and HEC Paris Brownbags, for helpful feedback. All errors are my own. This work has benefited from a State grant managed by the Agence Nationale de la Recherche under the Investissements d'Avenir programme with the reference ANR-18-EURE-0005 / EUR DATA EFM. An earlier version of this paper was circulated under "Human Capital Breadth, Portfolio Choice and Performance in Venture Capital".

†HEC Paris, Email: teodor.duevski@hec.edu

Which dimensions of investor human capital are effective for financing early-stage innovation? A long-standing tension in the literature concerns whether investors benefit more from deep specialization or broad, generalist experience. On one hand, research in financial intermediation underscores the benefits of specialization, arguing that deep industry expertise leads to superior outcomes (Gompers et al., 2009; Cressy et al., 2007; Blickle et al., 2023; Spaenjers and Steiner, 2024). On the other hand, the literature in labor economics and finance highlights benefits of breadth-often described as a "jack of all trades" advantage—arguing that individuals with broad, diverse experience are better equipped to adapt to uncertainty and facilitate innovation (Lazear, 2004; Custódio et al., 2013, 2019; Murphy and Zabojnik, 2007). This tension is especially salient in early-stage investing, where limited historical data and untested technologies make evaluation inherently difficult.

Venture Capital (VC) provides a compelling setting to study this trade-off. First, VCs primarily invest in early-stage firms with uncertain prospects and radically novel business models that differ significantly from previously financed startups. Second, the success of this model depends not only on financial capital but also on the human capital of VC investors, who leverage their expertise to identify, support, and scale promising startups.[1] A key unresolved question is whether broader human capital improves the ability to identify and support novel investments, or if deeper specialization yields better outcomes through domain-specific expertise. Ex ante, the answer is ambiguous: specialization may aid in assessing technological feasibility, while breadth may be more useful for evaluating teams, business models, and market potential.[2] In this paper, I address this question both empirically and theoretically.

Empirically, I first show in reduced form that, within VC firms, partners with broader human capital are more likely to lead investments in novel startups—and are significantly more successful when doing so. In contrast, they underperform in non-novel deals. The analysis includes VC firm × year and granular deal-level fixed effects (year, stage, industry, country), allowing comparisons across partners within the same VC firm-year and for otherwise similar deals. To address potential endogeneity in partner–deal assignment, I exploit plausibly exogenous variation in partners' time constraints. Following Abuzov (2024), I define "busyness" as concurrent involvement in a high-value exit, and show that busy partners are significantly less likely to lead new deals. Based on this, I construct a time-varying instrument of the availability of broad-background partners within each VC firm, which shifts the likelihood that a deal is led by a broad vs. narrow partner. Using this approach, I estimate the causal effect of partner breadth and find that it improves the success of novel startups, but decreases it for non-novel ones. I also document two new stylized facts: (i) the share of VC financing allocated to highly novel startups has declined substantially over the past two decades, and (ii) VC investor human capital has become increasingly specialized over the same period.

[1] See for instance: Kaplan and Strömberg (2001); Gompers and Lerner (2002); Casamatta (2003); Kaplan and Strömberg (2004); Chemmanur et al. (2011); Bernstein et al. (2016); Gompers et al. (2020)

[2] Beyond academia, the specialist vs. generalist debate is also of first-order importance in the VC industry. Institutional investors (LPs) allocating capital to VC increasingly weigh whether to back highly specialized funds—able to offer deep sector expertise—or more generalist funds that provide diversification and flexibility. See, e.g., *Wall Street Journal*, "Niche Funds Push the Limits of Specialization" (July 2018) or *Financial Times*, "How to Capitalise on Venture Capital" (March 2010)

These findings hinge on overcoming two core empirical challenges: measurement and identification. Measuring startup novelty ex ante is difficult, and data on VC partners' human capital are not readily available. Moreover, partner–deal matching is non-random: higher-quality startups tend to prefer more experienced VC firms (Hsu, 2004; Sørensen, 2007), and there is substantial skill heterogeneity among partners even within VC firms (Ewens and Rhodes-Kropf, 2015). I address both challenges by constructing new measures of startup novelty and investor breadth, and by implementing an empirical strategy designed to isolate causal effects.

To measure the extent to which a VC-financed startup is novel, I leverage recent advances in the NLP literature and rely on business descriptions of VC-funded startups available in PitchBook.[3] For each startup, I construct an OpenAI LLM-based deterministic embedding vector from the startup's business description.[4] The novelty proxy for a given startup is computed as one minus the maximum cosine-similarity distance between the focal startup and all startups financed by the VC industry in the five years before the focal startup receives its first VC financing.[5] The novelty measure captures how different a given startup's business model is relative to the closest venture financed by the VC industry in the past.

I validate the novelty measure by showing that it (i) is strongly associated with both success and failure, consistent with the high-variance outcomes typical of exploratory innovation, and (ii) predicts future innovation, both in terms of the number of forward patents granted and the number of forward patent citations after VC financing.[6] Importantly, these results hold when including granular VC firm, deal stage, industry, financing year, and country fixed effects. A one-standard-deviation increase in novelty is associated with a 0.8% increase in the probability of failure, a 4.5% increase in the probability of major success, a 7.5% increase in expected patent counts, and a 26% increase in the expected number of forward citations. I also document a steadily declining trend in the average novelty of VC-financed startups over the sample period (2003–2021), driven by a decline in the share of financing allocated to highly novel startups.[7]

To measure human capital breadth of VC partners, I rely on Revelio Labs, a workforce database built from an extensive set of LinkedIn profiles. I match VC partners listed in PitchBook to Revelio Labs and, for the matched subset, collect detailed educational and work histories. Following Custódio et al. (2013), I construct four proxies of breadth: the ratio of distinct job categories, roles, and industries to total employment spells, and the total number of distinct

[3] One concern with using business text descriptions is that startups may pivot from their original business plans. Kaplan et al. (2009) provide evidence that this is very rare, and business plans remain relatively sticky from the time a company receives its first round of venture financing to its public listing. This is also consistent with Paine (2024), who finds that around 87% of words that appear on startups' webpages during their first rounds of financing are also present on their IPO prospectus.

[4] Duevski and Bazaliy (2025) compare the performance of different embedding models and argue that OpenAI-based embedding models achieve superior performance for VC data.

[5] This measure is related to but distinct from Bonelli (2022)'s "backward similarity". A detailed discussion of the relationship between the novelty measure in this paper and other measures is presented in Internet Appendix H.

[6] This novelty measure has several advantages over relying on patent data directly: (i) very few early-stage VC-backed firms patent; (ii) early-stage patenting is concentrated in a few sectors (e.g., life sciences and biotech), limiting generalizability; and (iii) the decision to patent is endogenous to the startup's strategy.

[7] Notice that this trend is not purely mechanical and due to the proliferation of VC activity, since (i) the measure is constructed using a rolling-window look-back and (ii) in robustness analysis, I construct an alternative novelty measure where each focal firm is compared to the same number of VC-financed firms.

educational degrees obtained. I aggregate these into a human capital breadth index using principal component analysis (PCA), and use this index as the main measure in the empirical analysis, with robustness checks based on each individual component. Using this index, I document a decline in human capital breadth among VC partners over the sample period (2003–2021).[8]

First, without claiming to establish causality, I document that within a VC firm, partners with broad backgrounds are more likely to be lead partners on novel deals. A one-standard-deviation increase in a partner's breadth index is associated with a 0.1-standard-deviation increase in deal novelty. This relationship is stronger for investments where the VC firm is a lead investor, for non-syndicated deals, and for deals led by partners earlier in their VC careers. I also find that the interaction between startup novelty and lead partner breadth is positively associated with performance, measured by the likelihood of a major success.[9] However, the baseline coefficient of human capital breadth is negative, implying a negative correlation between human capital breadth and performance for non-novel firms. In these specifications, I include granular fixed effects at the deal stage, industry, year, and VC firm-year levels. This ensures that the observed association is not driven by VC-firm-specific quality, project deal flow, specialization in a particular sector or investment style, or systematic differences in novelty trends across sectors and deal stages. The granular fixed effects also account for time-varying industry and macroeconomic shocks that could influence performance outcomes. For novel (non-novel) firms, a one-standard-deviation increase in human capital breadth is associated with a 1.6% (1.5%) increase (decrease) in the likelihood of a major success.[10] While these results are suggestive, they do not address the endogeneity of partner assignment within VC firms. Broad partners may select themselves into the ex-ante best performing novel firms, but be assigned the hardest non-novel firms which could bias both the baseline human capital breadth estimate downwards and the breadth-novelty interaction upwards.

To study the causal impact of human capital breadth on startups' performance I introduce a novel identification strategy to instrument for the matching between deals and partners. To do so, I first restructure the data following Ewens and Rhodes-Kropf (2015). For each deal made by a VC firm, I construct a set of potential lead partners, defined as partners employed by the same firm within a three-year window around the deal (±3 years). For each potential lead, I incorporate a busyness proxy, following Abuzov (2024), which equals 1 if the partner is involved in a major exit event—either a high-value acquisition or an IPO—within a 90-day window (±90 days) around the focal deal and 0 otherwise. Using this data structure, I provide micro-level evidence that partner-time constraints limit assignment; busy partners are less likely to lead concurrent deals. Quantitatively, being busy reduces the likelihood of leading a concurrent deal by around 1.9%, which is about a 10% relative reduction (compared with 20% unconditional

[8]This pattern is robust to alternative measurement, and a similar downward trend is evident using simpler proxies—such as the changing composition of MBA holders relative to partners with PhDs or STEM degrees.

[9]Following prior literature, I define major success to be an exit via IPO or a high value acquisition. The effect is robust for various thresholds of defining a high value acquisition and also holds when only including IPO exits (Gompers et al., 2020).

[10]In this estimate, novel (non-novel) firms are those with one-standard deviation above (below) mean novelty.

probability of being an investment lead).[11]

These constraints on partner availability form the basis of my identification strategy. When a partner is busy due to exit activity in their existing portfolio, they are less likely to be assigned to new deals, creating plausibly exogenous variation in which partner—broad or specialized—leads a given deal. I exploit this by constructing a shift-share instrument: for each deal, I calculate the average human capital breadth of all available (i.e., non-busy) partners in the VC firm at that time. This instrument varies within VC firm-year, capturing deal-level fluctuations in partner availability, rather than relying on cross-sectional differences across firms. It allows me to isolate the effect of being assigned a broader (vs. more specialized) partner on startup outcomes, mitigating concerns about endogenous partner-deal matching.

The relevance condition is strongly supported by the data, with F-statistics exceeding 44, and is economically meaningful: a one-standard-deviation increase in average breadth availability is associated with a 0.3-standard-deviation increase in the human capital breadth of the selected partner, a strong first-stage relationship. The exclusion restriction is reasonable because variations in partner availability driven by unrelated exit events should influence future startup performance only through their effect on the selection of a broad or narrow-background partner, rather than through other omitted channels.[12] The IV estimates imply that, at very low novelty, an additional standard deviation of breadth reduces the probability of major success by about 11%, representing a negative baseline effect. For firms with novelty of one standard deviation above the mean, a one-standard-deviation increase in human capital breadth leads to an 8.2% increase in the probability of a major exit. The IV estimates are similar in magnitude to the OLS estimates in the same sample.

I conduct several robustness checks. First, I test for heterogeneity by deal lead status. The first-stage relationship is significantly stronger for deals where the VC firm is the lead investor, and the IV estimates are statistically and economically meaningful only in this subsample—consistent with the idea that partner assignment matters more when the partner plays an active role in selection and monitoring. Second, I examine whether shifts in partner availability systematically affect the quality of deal flow by comparing observable characteristics of startups financed during busy versus non-busy periods. I find no significant differences in key variables such as novelty or ex-post performance. Together, these results strengthen the validity of the exclusion restriction and reinforce the interpretation of the IV estimates as causal.[13]

[11] If I focus solely on deals where the VC firm is a lead investor, this magnitude doubles and further increases if I focus on early-stage deals only. This suggests that time constraints of VC partners become more important for deals where the VC firm takes the lead and early-stage investments. This is consistent with the fact that time constraints of partners are more binding when the screening or monitoring effort required by the partner is larger.

[12] This is likely to be valid for early-stage investments, since, as argued by Malenko et al. (2024), individual partners play a vital role in the selection process of startups, and many VC firms apply a champion rule during investment-committee voting. I also provide further evidence that team diversity measured at a fund level does not seem to play a role in the likelihood of financing novelty or future realized performance.

[13] I also conduct a series of additional robustness tests. First, I implement a placebo test by randomly reshuffling the busyness indicator across partners within each deal, while preserving the within-deal distribution of busy and non-busy partners; this generates no significant effects. Second, I test alternative constructions of the busyness proxy, including definitions based solely on IPO events and find consistent results. Third, I vary the time window used to define busyness (±60 and ±45 days), and the estimates remain stable across specifications.

To rationalize the empirical findings and observed secular trends, I develop a three-stage model that endogenizes entrepreneurial entry, the VC firm's partner assignment to projects, and partner screening effort. In the first stage, an entrepreneur chooses whether to produce a high or low quality novel project. In the second stage, the VC firm assigns a partner either specialized or broad, and the partner chooses how much costly effort to exert in screening the project. Screening produces a noisy signal of project quality, with broader partners having lower effort costs for novel projects but also lower fallback payoffs if the project is rejected. In the final stage, surplus is split between the VC and entrepreneur, conditional on project quality and financing. The model generates equilibrium predictions for partner specialization, screening effort, and project acceptance. It rationalizes the empirical finding that broader partners are more likely to finance novel projects and perform better when doing so: lower specialization reduces screening costs and raises acceptance rates and expected returns when the prior for high quality ventures is low. The model also accounts for the long-run decline in financing of novel startups and the rise in partner specialization. When the cost of producing high-quality novel ventures increases (e.g., due to rising complexity or "burden of knowledge" Jones (2009)), fewer entrepreneurs select into high-quality entry, lowering the prior for novel success. In response, VCs optimally assign more specialized partners with higher fallback payoffs, reducing screening effort and further discouraging high-quality entry—creating a feedback loop that drives both trends.[14] A result of the model is that the share of novel ventures financed and specialization co-move in opposite directions.

In sum, my findings underscore the importance of human capital breadth in enabling the financing of novel ventures. This has implications for several stakeholders. For VC firms, it highlights a trade-off in partner selection and assignment: while specialists may deliver better outcomes for incremental innovations, generalists are more effective at evaluating and supporting novel startups. For institutional investors (LPs), the results suggest that due diligence should not focus solely on a fund's investment thesis, but also consider the educational and professional diversity of its partners. For policymakers, the findings offer potential insight into the decline in startup novelty. One explanation, consistent with Bloom et al. (2020), is that as breakthrough opportunities become scarcer, VC firms rationally shift toward specialization. Another view, emphasized by Lerner and Nanda (2020), is that LPs' growing preference for focused investment strategies may limit support for novel, high-uncertainty ventures. In this case, policies that promote cross-sector labor mobility or encourage interdisciplinary education—particularly at the intersection of business and technical fields—may strengthen investors' capacity to identify and back innovative startups. Business schools and executive programs can contribute by embedding such cross-disciplinary training in their curricula.

[14]The equilibrium reacts in a similar way if, instead of increasing the cost of creating high-quality novel ventures, the cost of entry is reduced for entrepreneurs, e.g., as argued by Ewens et al. (2018)

1 Related Literature

I contribute to several strands of literature. First, I contribute to the literature studying the drivers of portfolio choice by venture capitalists and outcomes of funded startups. A substantial body of research underscores the critical role that venture capitalists play in financing and nurturing innovative startups (Kaplan and Strömberg, 2001; Gompers and Lerner, 2002; Casamatta, 2003; Kaplan and Strömberg, 2004; Chemmanur et al., 2011; Bernstein et al., 2016; Gompers et al., 2020). Two primary mechanisms of value creation recur throughout this literature: the ability to attract or select high-potential ventures (Sørensen, 2007; Howell, 2020) and the monitoring that VCs provide post-investment (Hellmann and Puri, 2000; Lindsey, 2008; Bernstein et al., 2016; Ewens and Marx, 2018). Much of the literature on venture capital focuses on how VC firm-level attributes such as reputation, syndication networks, fund size, and the use of data technologies affect startup financing and success (Hochberg et al., 2007; Gompers et al., 2008; Bonelli, 2022). Recent research has begun to zero in on the importance of individual VC partners in startup outcomes. Nahata (2008) links partner experience to investment performance, suggesting that personal track records and industry knowledge play a vital role in building VC reputation. Ewens and Rhodes-Kropf (2015) document a substantial heterogeneity in performance across individual VC partners within the same VC firm. Ewens and Sosyura (2023) provide causal evidence that individual venture capital partners affect startups' outcomes. My primary contribution to this literature is to show that individual VC partners' human capital breadth is particularly important for the selection and monitoring of startups with novel (previously unexplored) business models. I provide causal evidence that breadth enhances performance precisely in high-novelty settings but leads to worse outcomes in more incremental, non-novel investments.

Second, I contribute to the literature on the intersection of labor economics, finance and entrepreneurship by highlighting instances where a broad, generalized skill set can be advantageous. Lazear (2004)'s "Jack of all trades" theory argues individuals with a more balanced skill set are more likely to become entrepreneurs. Lazear (2012) argues that leaders are more likely to be generalists in both their innate characteristics and in their pattern of skill acquisition. In the context of executive leadership, Murphy and Zabojnik (2004) and Murphy and Zabojnik (2007) argue that the shift from firm-specific to general managerial skills has contributed to rising executive compensation and increased competition for top talent. Similarly, Custódio et al. (2013) and Custódio et al. (2019) find that generalist CEOs earn higher salaries, manage more complex firms, and increase R&D investment. Kecht et al. (2025) document an increased demand for generalists CEO, due to their ability to manage uncertainty. I build on and extend this literature in two ways. In terms of measurement, I closely follow the methodology of Custódio et al. (2013) by constructing a human capital breadth index using principal component analysis (PCA) on four proxy variables. I retain the first principal component as the aggregate measure. The proxies include: (i) the number of distinct job categories, (ii) the number of distinct job roles, (iii) the number of distinct industries, each scaled by the total number of employment spells, and (iv) educational breadth, defined as the count of distinct educational degree types (STEM, Social Science and Humanities, IT, Medicine, MBA, and PhD). This index differs from that of Custódio et al. (2013) in several important ways. First, my measure is constructed from the résumés of VC partners prior to their entry into the venture capital industry. As a result, it is a static measure and because career lengths vary across individuals, I normalize the first three work history proxies by the total number of employment spells in order to avoid confounding human capital breadth with overall career duration before entering the VC industry. Second, I exclude the prior-CEO and conglomerate experience dummy variables used by Custódio et al. (2013), as these are proxies specifically tailored to the executive labor market and are less applicable in the context of venture capital. Third, I introduce an educational breadth proxy to capture interdisciplinary academic training as an additional component of VC investors' human capital. [15] I extend this literature by showing that venture capitalists with broader human capital are more likely to identify and finance startups pursuing novel business models and to facilitate their successful exits. However, generalist VC partners perform worse in investments involving more incremental, non-novel ventures. This pattern highlights the context-dependent effectiveness of generalist skills in venture capital and suggests that both generalists and specialists are likely to be necessary for financing innovation: generalists add value in high-novelty settings, while specialists are more effective when evaluating more familiar or incremental opportunities.[16]

Third, I contribute to the literature on performance heterogeneity between specialist and generalist private equity and venture capital (VC) firms. A strand of literature argues that specialist private market intermediaries tend to outperform their diversified counterparts (Cressy et al., 2007; Spaenjers and Steiner, 2024), while Humphery-Jenner (2013) argue that there is a premium for more diversified PE funds. A seminal study by Gompers et al. (2009) finds that industry investment specialization at the partner level is positively associated with performance. I extend this literature in two key ways. First, I introduce a critical distinction between two dimensions of specialization: the breadth of human capital individual VC partners accumulate before entering the VC industry—an inherent personal characteristic—and their investment focus after becoming startup investors. Second, I demonstrate that while broad human capital does not confer an advantage for the average VC-financed firm, it plays a crucial role in supporting early-stage novel projects within a given sector. My findings refine the existing understanding of specialization in venture capital by highlighting the nuanced role of human capital breadth in fostering the financing of novel ventures.

Fourth, I contribute to the literature on limited attention of economic agents. Kacperczyk et al. (2014) study the impact of institutional and individual investors' attention allocation on their trading outcomes. Kempf et al. (2017) investigate the implications of limited attention of institutional investors. Another strand of literature has looked at busy boards and performance

[15] Lazear (2004), for example, incorporates educational background to measure generalist human capital. I proxy educational breadth using the count of distinct degree types listed on CVs, rather than Lazear's course-level breadth measure, since coursework is not observable. Consequently, this proxy captures diversity across degree fields but does not reflect variation within individual degrees.

[16] This broader point—that individuals with different skill profiles contribute value in different contexts—is also reflected in Zambrana and Zapatero (2021), who find that generalist mutual fund managers are better market-timers, whereas specialists are better stock pickers.

outcomes (Fich and Shivdasani, 2006; Field et al., 2013; Falato et al., 2014; Hauser, 2018). In the context of individual VC partners, Abuzov (2024) finds that investments by partners during their IPO engagements underperform. I build on Abuzov (2024) by showing that VC firms internalize this effect and are less likely to allocate concurrent investments to time-constrained partners. This allows me to build a novel identification strategy based on the average human capital breadth of available (non-busy) partners within a VC firm at the time of each deal. The instrument provides plausibly exogenous variation in whether a broad or specialized partner leads a deal, uncovering a new channel of limited attention: by reshaping the allocation of decision-making inside VC firms, partner busyness alters both the financing and performance of novel versus incremental startups.

Fifth, I contribute to the literature that applies text-based natural language processing (NLP) methods to study economic outcomes related to private firms. Kelly et al. (2021) use patent text similarity to improve measures of patent novelty and importance. In the venture capital (VC) literature, Bonelli (2022) apply the TF-IDF methodology from Kelly et al. (2021) to startup business descriptions to measure how "backward-similar" a given startup is, on average, to previously VC-financed startups within the same sector. Building on Bonelli (2022), I use textual data from startups' business descriptions to construct a measure of startup novelty. My approach differs from Bonelli (2022) in several important ways: (i) I do not restrict the comparison set to a specific sector; (ii) I define novelty as $1 - \max(\text{cosine similarity})$ (i.e., the closest-neighbor distance), rather than using a sector-specific average similarity rank; and (iii) I use sentence embeddings (OpenAI's text-embedding-3-small) instead of TF–IDF. I further document the following: (i) a low correlation between my novelty measure and backward similarity, with -0.26 at the sector level, -0.16 at the industry group level, and -0.16 at the industry code level; (ii) in horse-race regressions including both measures, the coefficient on my novelty measure remains stable in magnitude and significance whereas the coefficient on the backwards similarity measure becomes economically smaller and sometimes statistically insignificant; and (iii) divergent predictive patterns: backward similarity is positively associated with failure and negatively with major success, whereas the novelty measure is positively associated with both failure and major success suggesting that the two measures capture distinct dimensions. A methodologically related paper is Lerner et al. (2024), who use a BERT based embedding methods and startups' textual description data from Pitchbook to study relatedness of newly created startups in the developing world to previously funded Chinese startups. Guzman and Li (2023) and Paine (2024) use text from startup webpages from a Way Back Machine to study outcomes related to startup strategic choices and innovation output.

2 Data sources

This paper examines how the human capital breadth of Venture Capitalists (VCs) affects their investment choices, performance, and the innovation output of funded startups. To investigate this, I employ detailed data on VC portfolio allocation, exits of startups funded by VCs, patent applications and citations of VC-funded companies, and the human capital of VC investors. Specifically, I integrate data from several sources: PitchBook, which provides detailed data on VC investments and the subsequent exits of funded startups; a combined dataset from PitchBook and Revelio to construct measures of human capital by VC partners; and USPTO data to evaluate innovation output through patents and citations of VC-funded startups.

2.1 PitchBook

I obtain information on VCs' portfolio choices and performance from PitchBook, accessed via WRDS. The data vendor provides information on deals done by VC firms and VC-financed company characteristics, including textual descriptions, VC investor information, as well as exit types of VC-financed companies. I restrict my main sample to the period between 2003 and 2021, where deal coverage in PitchBook is representative and where business descriptions of VC-financed ventures are readily available (Retterath and Braun, 2020). Garfinkel et al. (2021), for instance, show that during the sample period analyzed in this paper, PitchBook data provides the most comprehensive data coverage when compared with other commercial databases (e.g., Crunchbase or VentureXpert).

Since I focus on investments made by institutional venture capitalists (as opposed to angel investors or corporate venture capital, for instance), I include in my sample deals with the Deal-Class label "Venture Capital" in PitchBook and the following deal type labels: "Seed Round," "Early Stage VC," "Later Stage VC," "Restart - Later VC," and "Restart - Early VC." To obtain a representative sample of VC investors, I also restrict the sample to VC investors who have made at least five investments in different companies over the entire sample period (2003–2021).

For each financed startup, I classify the exit types using the data provided by PitchBook, where I can observe exits up until July 2025, as "IPO" exits, "M&A" exits, and "Major Success." I classify a startup's exit as an "IPO" exit for a given VC investor–portfolio company pair if the given VC firm has exited the company via an IPO. Similarly, I classify the VC investor–portfolio company pair as an "M&A" exit if the given VC investor has exited the company via an M&A.[17] I define a "Major Success" as an outcome if the deal has exited via an IPO or a very profitable acquisition. For most of my tests I define a profitable acquisition as one with an exit value of at least five times total VC invested capital. I run sensitivity checks and robustness analyses for 1x and 3x thresholds (Ewens and Rhodes-Kropf, 2015; Bonelli, 2022). All results are also robust when using only IPO as a measure of successful exit.

2.2 Revelio

To collect detailed educational and job histories of VC partners, I supplement the PitchBook dataset with Revelio. Revelio is a leading workforce database provider that has collected the near-

[17] An exit is classified as an M&A exit if the exit type is labeled as "Merger/Acquisition" or "Merger of Equals" in the exit data provided by PitchBook. To focus on successful M&A exits, I remove exits labeled as "Corporate Divestiture" or "Distressed Acquisition."

universe of LinkedIn profiles. Their resume data includes comprehensive detail on individuals' work and educational histories. The use of Revelio data has been rising in the finance literature (e.g., (Hampole et al., 2025; Dorn et al., 2025)). Since Revelio does not provide an identifier that can be used to match directly with PitchBook data in order to obtain human capital characteristics of VC partners, I utilize the following matching strategy.

First, I match PitchBook VC firms to Revelio firm-level data. To do so, I first clean both Revelio and PitchBook firm names by removing common suffixes (e.g., Ltd, GmbH) and converting the names into lowercase. Then I proceed with matching in four steps. First, I match directly on firm name and website domain—i.e., requiring both firm name and website to be the same in both databases. Then, for the unmatched VC firms, I use information on their headquarters in PitchBook and require that the firm name and the headquarter country be the same.[18] As a third step among unmatched firms, I match only on firm name. Finally, as a fourth step among unmatched firms, I match on website domain and headquarters but not on VC name, and I manually verify the correctness of those matches.

Once I obtain a sample of matched VC firms to Revelio, I use Revelio's jobs data to match the partners available in PitchBook and Revelio directly on first and last name. For the final sample, I require that at least one of the background characteristics (e.g., ethnicity, gender, educational background, work history) for a given partner be available in Revelio.[19]

2.3 USPTO

To measure innovation by VC-financed companies, I supplement the data with deal-level data on patent applications and grants from the United States Patent and Trademark Office (USPTO). The USPTO data also includes patent applications that are still pending, as well as those that have been abandoned, rejected, or canceled. It provides each patent's unique identifier, as well as information on its assignee, its technology class, its application year, and, when applicable, its grant year. I match the VC-financed startups from PitchBook to USPTO data using fuzzy matching, similar to Bernstein et al. (2016).

2.4 Summary Statistics

Table 1 about here.

Table 1 presents summary statistics for the main deal-level sample of the data. Each observation represents a deal–investor–company–lead partner, if the lead partner is available. There are a total of 232,130 new financing rounds during the sample period (2003–2021). These are financing rounds of 92,740 distinct ventures, financed by 8,521 distinct VC investors. Out

[18]Revelio Labs' company data does not contain information on firm headquarters. I construct a proxy for this by assuming that the firm's headquarters is the state or country where the largest number of employees are located which is a reasonable assumption for venture capital firms.

[19]A detailed description of the PitchBook-Revelio matching procedure is outlined in Appendix G.

of those 232,130, for a subset of 81,386, PitchBook also provides the identity of the partner involved in the deal (16,946 distinct partners in total). Out of those 232,130 deals, around 4% have successfully exited to the public market, and around 7% have achieved a major exit. Out of those 16,946 available on PitchBook, I am able to match and retrieve at least one background characteristic from Revelio for 9,880 partners. These partners lead around 47,561 deals, so the match rate is approximately 58%.[20]

In terms of background characteristics of partners assigned to a given deal, around 11% of deals are led by female partners, around 43% of deals are led by partners who have an MBA degree, and only 6% of deals are led by partners who have an advanced PhD degree. 31% of deals are led by partners with a STEM education, and around 63% of deals are led by partners with a Social Science and Humanities education. Around 50% of deals are led by partners who have completed at least one degree at a top educational institution.[21]

Table 2 about here.

Table 2 presents the summary statistics for individual partners i.e., where each observation is a unique partner who has led at least one deal over the sample period. The summary statistics at individual partner level present a similar picture to the deal level summary statistics.

3 Measurement and Stylized Facts

3.1 Measurement of Novelty

To measure the extent to which startups are novel, I rely on startups' business descriptions provided by PitchBook. I draw on recent advances in Natural Language Processing (NLP) and use the textual description of each startup to construct a high-dimensional embedding vector. Specifically, I employ OpenAI's deterministic embedding model text-embedding-3-small, which represents text as a fixed 1,536-dimensional vector. The model is deterministic, meaning that identical input always yields the same output, which is essential for reproducibility in large-scale empirical research.

OpenAI embedding models are trained on a broad mixture of publicly available and licensed text data to capture semantic relationships across diverse domains. As a result, semantically

[20]This is the match rate computed both in terms of overall partner coverage and deal coverage, which means that matched partners are not more or less likely to be leads on investments than unmatched partners suggesting that the sample is representative.

[21]The definition for a top educational institution follows Calder-Wang and Gompers (2021). Specifically, top institutions are: 'Brown University', 'Harvard University', 'Columbia University', 'Cornell University', 'Dartmouth College', 'University of Pennsylvania', 'Princeton University', 'Yale University', 'Duke University', 'Massachusetts Institute of Technology', 'University of Chicago', 'Caltech', 'Stanford University', 'Northwestern University', 'University of California, Berkeley', 'Williams College', 'Cambridge University', 'INSEAD', 'HEC Paris', 'London Business School', 'London School of Economics', 'Oxford University'.

similar texts are positioned close to one another in the embedding space, making the method particularly suitable for identifying related business models and technologies. I use the small variant of the model because it provides substantial computational efficiency while delivering performance comparable to larger alternatives for relatively short business descriptions such as those found in PitchBook. OpenAI embeddings have already been shown to be effective in the analysis of venture capital data (Duevski and Bazaliy, 2025). Importantly, when constructing embeddings from business descriptions, I exclude all firm identifiers (such as the company name) that may introduce look-ahead bias, as documented by He et al. (2025).

I construct my main proxy for the novelty of the startup in the following way. For each deal d for company c made at time t, I define the novelty of startup c at date t as:

$$N_{c,t} = 1 - \max_{j \in (t-5,t)} \text{CosSim}(C_c, C_j), \tag{1}$$

where $\text{CosSim}(C_c, C_j)$ is the cosine similarity between the embedding vector of focal company c and company j, and $\max_{j\in(t-5,t)} \text{CosSim}(C_c, C_j)$ specifies that I take the maximum cosine similarity of the focal company c to any other startup that has received venture financing in the five years prior to year t. In particular, note that according to this definition, $N_{c,t} = 0$ if company c has received venture-backed financing in the past five years. Equation (1) uses a five-year rolling window to avoid mechanical issues related to, for example, the proliferation of VC activity over time.[22]

Intuitively, (1) captures how distinct the business model of company c is from any other company that has received venture financing in the past five years. This proxy evaluates the extent to which the focal company c is novel relative to what the venture capital industry has financed in the past. I term the measure $N_{c,t}$, Novelty (Distance to Closest Firm).

One additional advantage of constructing startup novelty via (1) is that one can explicitly retrieve the closest firm to any given firm receiving VC financing in the past. In Appendix Table A1, I provide the explicit novelty measure for some well-known startups, their percentile ranking in terms of novelty, as well as the firm that is identified to be the closest to them in terms of business description.

3.2 Stylized Facts About Novel Startups

In this subsection, I present several stylized facts about novel startups.

3.2.1 Measure validation: Novelty, Performance and Innovation

In a regression setting, I document two stylized facts about novel startups that also serve as validation for the novelty measure. First, novel startups are more likely to fail or achieve a major exit consistent with high-variance outcomes typical of exploratory innovation. Second, novel startups contribute more to innovation output. The second fact is particularly useful and shows

[22] For example, firms appearing less novel over time simply because the benchmark comparison set is larger.

that even though novelty is constructed from business descriptions—which arguably capture more about the startup's product and business design—they are still correlated with ex-post innovation outcomes and therefore can be used as an ex-ante proxy for the startup's innovation output or social value.

Novelty and Performance: I document that novel startups are more likely to fail or achieve a major success. Specifically, I estimate the following model using a deal-level data structure, where each observation is a deal–investor–company:

$$P_{d,i,t,s,k} = \alpha + \beta N_d + \eta_{i\times t\times s\times k\times c} + \epsilon_{d,i,t,s,k}, \tag{2}$$

where d denotes a deal, i the industry of the deal, t the year the deal is made, s the stage of the deal, k the VC firm financing the deal, and c the country of the financed company. $P_{d,i,t,s,k}$ is a performance outcome indicator, which can be Failure or Major Success. N_d is the novelty of the deal, and $\eta_{i\times t\times s\times k\times c}$ denotes granular industry × time × deal stage × country × VC Firm fixed effects. The coefficient of interest, β, captures the association between deal novelty and deal performance.

Table 4 about here.

The results are presented in Table 4 Panel A. More novel startups are significantly more likely to both fail and achieve a major success. Note that the variation in columns (2) and (4) comes from within an industry × time × deal stage × country × VC firm; that is, we are comparing the performance of more or less novel startups financed by the same investor in the same stage, deal year, industry, and country. This suggests that the results of this stylized fact are not simply driven by certain VC firm characteristics that have been shown to be associated with startup success e.g., VC firm reputation, experience, deal flow or assets under management (Hsu, 2004; Sørensen, 2007; Bhardwaj et al., 2025). Similarly, the performance is not merely driven by time-varying industry or overall economic conditions (e.g., the hotness of the M&A and IPO market in general, industry-specific shocks, and varying country-level economic shocks). In terms of economic magnitude, estimates in columns (2) and (4) imply that a one standard deviation increase in novelty is associated with a 0.8% increase in the probability of failure and a 4.5% increase in the probability of a major success.[23]

I run several robustness analyses to confirm the association between novelty and both failure and major success. In Table A2, I estimate a similar model as in (2), but I split startups into yearly novelty quartiles. The baseline is Novelty (Distance to Closest Firm) = 1, which represents the least novel startups in each year. Estimates in columns (2) and (4) imply that the probability of failure increases by 3% and the probability of a major success exit increases by 9% when moving from the bottom to the top quartile of novelty. In Table IB1, I re-estimate model (2) with an alternative novelty measure defined as one minus the average cosine similarity between the

[23] Calculated as Point Estimate × SD in Novelty (Distance to Closest Firm Measure)

textual description of the startup financed in the deal and the top five closest startups receiving venture capital financing within five years before the deal. In Table IB2, I re-estimate (2) using a sample that includes only firms with first financing rounds between 2018–2021, where the business description likely captures the startup's product description close to its creation. In Table IB3, instead of using Major Success, I re-estimate the model using only IPO as a measure of success. In Table IB2, I also show that novelty positively correlates with both IPO valuation and IPO multiple for the subsample of deals that have exited via the public market.

Novelty and Innovation: To assess the association between novelty and the innovation output of a given startup, I estimate the following model using a deal-level data structure where each observation is a deal–investor–company:

$$I_{d,i,t,s,k} = \alpha + \beta N_d + \eta_{i\times t\times s\times k\times c} + \epsilon_{d,i,t,s,k}, \tag{3}$$

where $I_{d,i,t,s,k}$ is a forward-looking innovation measure. Specifically, for each deal made, I count the number of forward patents (patents granted after the deal is made) and forward citations (citations of patents granted after the deal is made).[24] N_d is the novelty of the deal, and $\eta_{i\times t\times s\times k\times c}$ denotes granular industry × time × deal stage × country × VC Firm fixed effects. The coefficient of interest, β, captures the association between deal novelty and innovation outcomes.

The results are shown in Table 4 Panel B. I use (3) and estimate a Poisson count model (to avoid well-known issues with using the log(1+) model; see Chen and Roth (2024)). Estimated coefficients in column (1)-(4) in Panel B show that startup novelty is positively associated with number of forward patents granted and the number of forward citations. Quantitatively, estimates in column (4) imply that a one standard deviation increase in novelty is associated with a 26% increase in expected forward citations.[25]

I conduct several robustness checks. First, in Table IB5, I re-estimate (3) using an alternative novelty measure defined as one minus the average cosine similarity between the textual description of the startup financed in the deal and the top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. In Table IB6, I use a log(1+) model instead of a Poisson count model for the right hand side variable. In Table A3, I use a novelty quartile measure and show that the effect is monotonic.

3.2.2 Startup Novelty at First Financing Has Declined Over Time

Having established that the novelty measure is predictive of economically meaningful outcomes, I now turn to its evolution over time.

Figure 1 about here.

[24] I adjust the citation number by year and NBER subcategory, as is standard in the literature (e.g., Lerner and Seru (2022)).

[25] Calculated as $e^{3.292\times 0.07} - 1 = 0.26$.

In Figure 1, in the left panel, I plot the distribution of the Novelty (Distance to Closest Firm) measure conditional on $N_{c,t} > 0$, i.e., for startups receiving their first round of venture financing. The median startup's novelty is 0.25. In the right panel of Figure 1, I plot the time evolution of the mean of the Novelty (Distance to Closest Firm) measure and document a decreasing mean novelty over time. Note that these plots are conditional on $N_{c,t} > 0$, so they capture the evolution of novelty for startups receiving their first venture financing, i.e., the pattern is not driven by later financing rounds of the same venture. Note that (1) applies a five-year rolling window to the novelty measure, which means that the decline in novelty presented in Figure 1 is not simply driven by the fact that, as time passes, VC activity proliferates and each benchmark firm is compared to more firms, naturally driving novelty down over time. The sharp observed decline starts around 2010, and quantitatively, the average startup receiving financing in 2020 is around 20% less novel than the average startup receiving a first financing round in the early 2000s.

Figure 2 about here.

To better understand the time evolution of novelty, in the top-left panel of Figure 2, I plot the number of distinct firms financed by the VC industry with positive novelty over time, which shows an increasing trend. In the top-right panel, I plot the raw number of firms with novelty above 0.3—that is, financed firms with a novelty measure above the top quartile for the entire sample—and document that the number of firms financed with relatively high novelty has also been increasing. In the bottom-left graph, I plot the share of firms with a novelty measure above 0.3 and document a declining trend over time. That is, the decline in average novelty observed in Figure 1 is driven by a decline in the share of financed firms with high novelty (even if their absolute number is going up). Finally, in the bottom-right panel, I plot the median novelty measure for the top 100 most novel VC-financed firms in each year and document a declining trend; however, this decline starts later than the general decline in novelty.

Robustness checks. I conduct several robustness checks to address potential measurement concerns. First, to account for the increasing number of comparison firms over time (i.e., the proliferation of VC activity), Figure A1 plots an alternative novelty measure where all focal firms are compared to a fixed number of random firms (300 in the left panel and 500 in the right panel) in each year. The observed decline in novelty remains robust under this alternative construction. Second, in Figure IB1, I construct an alternative novelty measure based on the top five closest firms (rather than just the single closest firm) in terms of business description similarity. This alternative specification also shows a robust decline in novelty over time. Third, in Figure IB2, I use a different PitchBook data vintage (2025) to assess whether changes in business descriptions over time might bias the trend. I find that: (1) fewer than 10% of firms exhibit any meaningful semantic change in business description across vintages; and (2) the decline in novelty is robust when using this alternative data vintage.

Overall, the time trends observed are consistent with the fact that the decline in average financed novelty is primarily driven by a steady decline in the share of VC financing going to high-novelty startups.

3.3 Measurement of Breadth of Human Capital and Stylized Facts

To measure human capital breadth, I rely on the educational and work histories of VC partners obtained from Revelio. For each partner that I match to Revelio, I first obtain the year of their deal as recorded by PitchBook. Then, when constructing their work history, I include only those jobs with an end year prior to their first recorded deal. Using their work histories and educational background, I construct four main proxies for human capital breadth. The first three proxies closely follow (Custódio et al., 2013):

1. Ratio of job categories: This is defined as the ratio between distinct job categories ("job category" variable in Revelio) and the number of employment spells.[26]

2. Ratio of job roles. This is defined as the ratio between distinct job roles ("role k1500" variable in Revelio) and the number of employment spells.

3. Ratio of industries: This is defined as the ratio between the number of distinct industries a partner has worked in and the number of employment spells.

4. Educational breadth: This is a count of the distinct types of degrees an individual has obtained in the past. Distinct types of degrees are: STEM education, Social Science or Humanities Education, IT Education, Medicine, MBA and PhD.

Following Custódio et al. (2013), using PCA, I combine these measures into a single time-invariant index for each individual partner. The PCA weights that form the index are given in the equation below:

$$\text{BreadthIndex}_i = 0.0975\, z_{i,\text{edu}} + 0.5413\, z_{i,\text{ind}} + 0.5795\, z_{i,\text{cat}} + 0.6014\, z_{i,\text{role}} \tag{4}$$

In the PCA, I obtain only one eigenvalue higher than 1 and thus retain only the first principal component, which explains around 60% of the variation (similar to (Custódio et al., 2013)). As expected, all individual components load with the same sign when constructing the measure, which means that they are positively correlated with the underlying concept we are trying to proxy for—namely, human capital breadth. A scree plot of the PCA and the cumulative variance explained is given in Figure IB3. I use the measure defined by (4) in the main test and run robustness analyses for each individual component.

Figure 3 about here.

[26] Job Category is a broad classification of the types of jobs an individual has done into 7 categories: Admin, Finance, Engineer, Scientist, Operations, Marketing, Sales.

In Figure 3, I plot the distribution of the breadth index measure (left panel) and the time evolution of the breadth index measure. I document a decline in average breadth in human capital among VC partners. This trend remains robust when using simpler educational proxies to measure human capital breadth. This is shown in Figure A2. In the top panel of the figure, I plot the share of partners with an MBA degree relative to the share of partners with a PhD or STEM degree only. This share has declined from roughly 80% in the early 2000s to below 70% around 2020. In the bottom-left panel, I plot the share of partners with interdisciplinary education and document a similar decline over time.[27] In the bottom-right panel, I plot the average number of distinct degrees per partner over time and document an average decline.

4 Empirical Results

In this section I present the main empirical findings of the paper. In subsection 4.1, under granular fixed effects I present correlational evidence that (i) lead partner human capital breadth positively with deal novelty (ii) higher human capital breadth is positively associated with deal performance for novel firms, but negatively for non-novel firms. In subsection 4.2, I study the causal effect of human capital breadth on performance for both novel and non-novel firms.

4.1 Correlational evidence: Partner level Human Capital Breadth and Novel Startups

Without claiming causality in this subsection, I present two novel facts that link individual partners' human capital breadth to the selection and performance of novel startups.

4.1.1 Association between Lead Partner Human Capital Breadth and Startup Novelty

I document an association between the human capital breadth of an individual lead partner on a deal and the deal's novelty. To do so, I estimate the following empirical specification at the deal level:

$$N_{j,k,p,t} = \alpha + \beta B_p + X_{t,p} + \eta_{i\times t\times s\times c} + \rho_{t\times j} + \epsilon_{j,k,p,t}, \tag{5}$$

where $N_{j,k,p,t}$ is the deal-level novelty of a deal made by investor j in startup k with lead partner p in year t. B_p is a partner-level human capital breadth index, $X_{t,p}$ represents time-varying partner-level controls measured at the time the deal is made, $\eta_{i\times t\times s\times c}$ denotes industry × deal year × deal stage × country fixed effects, and $\rho_{t\times j}$ is a VC firm × deal year fixed effect. The coefficient of interest is β, which captures the association between the lead partner's breadth index and deal novelty. In all specifications in this section, standard errors are double-clustered at the VC firm and financed company levels.

[27] A partner has interdisciplinary education if they have: 1) an MBA degree and a STEM degree; or 2) Social Science or Humanities education and a STEM degree; 3) an MBA degree and a PhD degree; 4) medical education and an MBA degree.

Table 5 about here.

The results are presented in Table 5. Across all specifications, the lead partner's breadth index is positively associated with startup novelty. The granular fixed effects show that, within a VC firm × deal year, partners with broader human capital lead more novel startups. Intuitively, the coefficient β is estimated by relying on variation in breadth across different partners financing startups of varying novelty within the same VC firm in a given year.

In column (3) of Table 5, which presents the strictest specification, I add VC firm × deal year × partner VC industry entry year fixed effects. In this specification, I compare the novelty of investments made by two partners who differ in their human capital breadth, but who are making investments in the same year while working for the same VC firm and entered the VC industry in the same year. The other high-dimensional fixed effect—industry × deal year × deal stage × country—controls for observable deal characteristics that may be correlated with novelty.

Although the effect we estimate is not strictly causal, the granular fixed effects help rule out several alternative explanations. First, the industry × deal year × deal stage × country fixed effects ensure that the result is not driven by broader partners selecting more novel sectors, investing at times when novelty is generally higher, investing at different stages in the firm's lifecycle, or operating in different countries. Second, the VC firm × year × partner entry year fixed effects ensure that the effect is not simply driven by more novel firms selecting into different types of VC firms—or more experienced VC partners.[28]

The estimates in column (3) imply that a one standard deviation increase in human capital breadth is associated with a 0.1 standard deviation increase in the novelty of the financed deal.

Additional Tests: In Table A4, I estimate specification (5) separately for lead (columns (1) and (4)) and non-lead (columns (2) and (5)) investments. In columns (3) and (6), I include an interaction term between the lead partner's human capital breadth and a dummy, Lead Investment, which equals 1 if the VC firm is a lead investor on the deal.Since lead VC firms typically conduct investment screening and assign partners to boards of directors, the association between human capital and startup characteristics should be stronger when the firm is a lead investor. The results in Table A4 support this. First, the effect size of lead partner human capital breadth on startup novelty is economically larger and statistically more significant for deals where the VC firm is a lead investor (columns (1) and (3)) compared to those where it is not (columns (2) and (4)). Second, the interaction term between lead partner breadth and the lead investor dummy is positive (columns (5) and (6)).

In Table IC1, I re-estimate specification (5) using an alternative novelty measure defined as one minus the average cosine similarity between the textual description of the financed startup and the top five closest startups receiving venture capital financing within five years before the deal, in the same deal stage.In Table IC2, I estimate the same specification but split the sample

[28]The deal year × VC partner entry year controls for the number of years a given partner has spent in the VC industry. VC experience is an additional control capturing the number of deals a partner has completed prior to the focal deal.

into non-syndicated deals (where the VC firm is the sole investor) and syndicated deals (with multiple VC firms). I show that the effect is larger for non-syndicated deals. In Table IC3, I estimate the same specification for each of the four components used to construct the human capital breadth index separately. Finally, in Table IC4, I split the sample based on the time elapsed between the deal year and the partner's industry entry year. I show that the effect is stronger for deals made closer to a partner's entry into the VC industry (columns (1) and (3) versus columns (2) and (4)).

4.1.2 Interaction between Lead Partner Breadth, Deal Novelty, and Performance

I document a positive interaction between a lead partner's human capital breadth and deal novelty on investment performance. Specifically, I estimate the following model at the deal level:

$$P_{j,k,p,t} = \alpha + \beta B_p + \gamma N_{j,k,p,t} + \delta B_p \times N_{j,k,p,t} + X_{t,p} + \eta_{i\times t\times s\times c} + \rho_{t\times j} + \epsilon_{j,k,p,t}, \tag{6}$$

where $P_{j,k,p,t}$ is a deal-level performance outcome for an investment made by VC firm j in startup k with lead partner p at time t. $N_{j,k,p,t}$ is the deal-level novelty measure, B_p is the lead partner's human capital breadth index, and $B_p \times N_{j,k,p,t}$ is their interaction. $X_{t,p}$ includes time-varying partner-level controls measured at the time the deal is made. $\eta_{i\times t\times s\times c}$ denotes industry × deal year × deal stage × country fixed effects, and $\rho_{t\times j}$ is a deal year × VC firm fixed effect.

Table 6 about here.

The results are presented in Table 6. Intuitively, in column (3), the coefficients are estimated using variation in partner breadth and novelty across deals made by the same VC firm, in the same industry, deal stage, and year. The granular fixed effects help rule out the possibility that the observed effect is driven by firm-level heterogeneity—such as VC firms attracting better deal flow while simultaneously hiring more capable partners. These fixed effects also address concerns about time-varying macroeconomic or industry-specific shocks that could affect performance outcomes and indirectly influence the observed relationship with novelty.

The coefficient β on lead partner breadth is negative, indicating that for non-novel deals, broader partners are associated with worse performance compared to narrower partners. In terms of economic magnitude, the estimate in column (3) implies that at zero novelty, a one standard deviation increase in the human capital breadth index is associated with a 5% decrease in the probability of a major success.

Importantly, the interaction term δ between human capital breadth and novelty is positive and statistically significant. This suggests that as novelty increases, the negative baseline effect of breadth reverses. Specifically, for every 0.1 increase in deal novelty, the effect of the breadth index increases by 2%. These estimates imply that the association between breadth and major success is negative for deals below the median novelty level (median = 0.24), but positive for more novel deals.

Additional tests: I conduct several robustness checks to validate these findings.

First, in Table IC5, I re-estimate specification (6) using an alternative novelty measure defined as one minus the average cosine similarity between the textual description of the startup and the five most similar startups funded within five years in the same deal stage. Second, in Table IC7, I estimate the model separately for each component used to construct the human capital breadth index. Third, in Table IC6, I estimate the model separately for lead and non-lead investments. Fourth, in Table IC9, I re-estimate the model using only IPO as a measure of successful exit. Fifth, in Tables IC10 and IC11, I show that the results are robust to using different acquisition value thresholds as cutoffs for defining successful outcomes.

Finally, in Table IC8, I estimate the direct association between human capital breadth and performance by estimating:

$$P_{j,k,p,t} = \alpha + \beta B_p + X_{t,p} + \eta_{i\times t\times s\times c} + \rho_{t\times j} + \epsilon_{j,k,p,t}, \tag{7}$$

I find a clear null result—human capital breadth is not associated with deal performance in general. This strengthens the interpretation that the interaction with novelty is central to understanding when and how human capital breadth matters for investment outcomes.

4.1.3 Additional results: Fund-level diversity and on-the-job VC experience

In Internet Appendix G, I present additional analyses examining whether team-level diversity or on-the-job VC experience are correlated with deal novelty or performance.

In summary, (i) I do not find a significant association between VC team-level diversity (measured as the diversity of partners managing investments within the same fund) and either the fraction of novel projects financed or the performance of those projects; and (ii) I find a positive association between partner-level industry specialization and deal novelty (consistent with Gompers et al. (2009)), as well as a positive association between past financed deal diversity (measured as the average cosine similarity across previously financed deals) and deal novelty.

4.2 Causal impact of Lead Partner Breadth on Performance of Novel Startups

Overall, the two reduced form results highlight (i) the human capital breadth of the lead partner is positively associated with deal novelty (ii) for low novelty deals, higher human capital breadth is negatively associated with performance and for high novelty deals human capital breadth is positively associated with performance. Even with granular fixed effects that isolate deal-level characteristics and the inclusion of VC firm × deal year fixed effects, the deal–lead partner assignment—even within the same VC firm—is endogenous. This can result in a bias in the OLS estimates in specification (6). For example since the lead partner human capital breadth assignment is endogenous, if broader partners are more likely to get the most difficult to execute non-novel deals the baseline estimate β may be downward biased, similarity if due to access and network advantages they get most promising novel deals, the estimate for the interaction effect of partner breadth and deal novelty δ can be upward biased.[29]

The ideal experiment would compare the performance of two startups with the same novelty, one having a broad lead partner and the other a narrow lead partner of the same quality. In the absence of such an experiment, one approach is to utilize an instrumental variable strategy that exogenously shifts the likelihood of a partner being a lead on a specific deal. A natural candidate for such an instrument is the time-varying availability of partners within the same VC firm. Intuitively, if partners face time constraints, during periods of high workload they should be less likely to lead a new investment. One proxy for time-varying partner busyness is the partner's involvement in an exit event of another deal made by the VC firm. Specifically, Abuzov (2024), for instance, finds that partners' deals made during periods when the partner is concurrently involved in an exit event perform worse than deals made by the same partner when the partner is less busy. If VC firms internalize this effect, it would imply that busy partners are ex ante less likely to be assigned as lead partners on deals made during periods of high workload. Thus, a natural shifter for the probability of a partner being a lead on a specific deal is whether the partner is concurrently involved in an exit event for another deal.

4.2.1 Micro-Level Evidence that time constraints limit assignment

Before proceeding with causal estimation and instrument construction, I provide micro-level evidence that time constraints limit partner assignment. Specifically, I show that a partner's busyness reduces the likelihood of that partner being assigned as the lead on a new deal made by the same VC firm.

To test this, I restructure the dataset at the deal level as follows: for each deal d made by VC firm j, I construct a set of potential partners who could have plausibly led the deal. As a baseline, I include all partners who have led at least one deal at the same VC firm within a $[-3, +3]$ year window around the deal date. For each potential partner, I construct a busyness proxy at the time of the deal. Following Abuzov (2024), I define a partner p as *busy* at time t if they are involved in a major exit (via a high-value acquisition or IPO) for another deal in the $(-90, +90)$ day window around the focal deal date t.

Using this structure, I estimate the following model:

$$PartnerChosen_{d,j,p} = \alpha + \beta BusyPartner_{p,t} + X_{p,t} + \eta_d + \rho_j + \epsilon_{d,j,p}, \tag{8}$$

where $PartnerChosen_{d,j,p}$ is an indicator equal to 1 if partner p is chosen to lead deal d at firm j, and 0 otherwise. $BusyPartner_{p,t}$ is an indicator equal to 1 if partner p is busy at time t. $X_{p,t}$ is a set of time-varying partner-level controls. η_d is a deal fixed effect, and ρ_j is a VC firm

[29]The direction of the OLS bias in specification (6) can go either way, the hypothetical bias highlighted in the paragraph is the bias one should worry about i.e., the bias that would make the results statistically significant without a causal interpretation.

fixed effect. Standard errors are clustered at the deal level.

Table 7 about here.

The results are shown in Table 7. In column (1), I estimate the coefficient of interest β for the full sample, and find a statistically significant negative effect.[30] In columns (2) and (3), I split the sample between lead and non-lead investments of the VC firm. If partner time availability matters, it should matter more for investments where the VC firm is the lead investor. This is precisely what we find: the estimated β in column (2) (lead investment sample) is more than twice the size of that in column (3). In column (4), I restrict the sample to investments receiving their first rounds of VC financing (i.e., investments with positive novelty). I find a larger magnitude of the estimated β relative to column (1), since partner time constraints are arguably more important for firms receiving their first VC financing.[31] In columns (5) and (6), I estimate β separately for lead and non-lead investments in the sample of firms receiving first-time VC financing, and show that the magnitude of the coefficient is much higher for investments where the VC firm is a lead investor.

This data structure also allows me to provide further evidence on the association between lead partner human capital breadth and deal novelty. Specifically, since I now have a set of potential partners who could have led each deal, the data structure allows for the inclusion of granular deal and partner-level fixed effects. Specifically, it allows me to estimate the following model:

$$PartnerChosen_{d,j,p} = \alpha + \beta N_d \times B_p + \gamma BusyPartner_{p,t} + X_{p,t} + \eta_d + \rho_j + \sigma_p + \epsilon_{d,j,p}, \quad (9)$$

where N_d is the focal deal's novelty, B_p is the partner's human capital breadth index, and σ_p is a partner fixed effect. The coefficient of interest is β which is the interaction between partner human capital breadth and deal novelty. The specification (9) allows me to include both partner and deal fixed effects which allow me to absorb all deal characteristics that may correlate with specific partner choice as well as all time invariant partner characteristics. Intuitively, the interaction term is estimated by within partner variation across deals of different novelty where the partner is or is not chosen to be a lead partner **and** within deal variation of available partners of different breadth index. Specification (9) also allows me to estimate the busy partner coefficient relying on within partner variation leveraging times where the same partner is busy versus available (all time invariant partner characteristics are controlled for).

Table 8 about here.

[30] The full sample also includes later financing rounds.

[31] Either because screening is more costly or requires more effort for these firms, or because monitoring may be more valuable.

The results are presented in Table 8, and the message of the table is twofold. First, the coefficient β on the interaction between human capital breadth and deal novelty is positive, statistically significant, and economically meaningful. In column (1), which uses the full sample, a one standard deviation increase in both partner breadth and deal novelty raises the likelihood of the partner being assigned as lead to approximately 51%, relative to an unconditional mean of 20%. This interaction effect is stronger in column (2), which restricts to deals where the VC firm is the lead investor, and weaker in column (3), which includes only non-lead investments. Second, the coefficient on the Busy Partner variable remains negative and statistically significant, even after including partner fixed effects. This specification relies on within-partner variation, comparing times when the same partner is busy versus not busy across otherwise similar deals. The estimated magnitude is comparable to that in model (8), reinforcing the interpretation that time constraints reduce the likelihood of being assigned as lead partner.

Additional Tests and Robustness: In Table A5, I perform a placebo test in which the busyness indicator is randomly reshuffled across partners within each deal, preserving the within-deal distribution of busy and available partners. As expected, the placebo variable shows no association with the likelihood of leading a deal, suggesting that the original effect is not driven by chance or mechanical correlations.

In Table ID1, I redefine busyness using only IPO exit events, under the assumption that IPOs impose greater demands on a partner's time. The estimated coefficient remains statistically significant and is larger in magnitude than in the main specification. In Table ID2, I shorten the busyness window to $[-60, +60]$ days and continue to find consistent results. In Table ID3, I apply the same $[-60, +60]$ window but restrict busyness to IPO-related exits only, and again the effect remains robust. Finally, I replicate both tests using an even narrower $[-45, +45]$ day window around the deal date. In Table ID2, I use this window for all exits, and in Table ID3, for IPO-only exits. Across all specifications, the results continue to hold, confirming that the relationship between partner busyness and assignment is not sensitive to the choice of time window or the type of exit used to proxy for workload.

The evidence presented suggests that (i) the time constraints of individual VC partners are significant, and (ii) the estimated effect sizes are economically meaningful and become more pronounced when time constraints are likely to matter most—specifically, in investments where the VC firm acts as the lead investor.

4.2.2 IV Estimates and the causal impact of human capital breadth for novel and non-novel firm performance

To identify the causal effect of breadth, we require a source of variation that shifts partner assignment independently of deal characteristics. Conceptually, we want to compare the performance of the same deal if it were randomly assigned to a broad versus a narrow partner. As discussed in the previous subsection, a natural source of such variation is time-varying partner availability. Table 7 showed that busy partners are less likely to be assigned as lead, conditional on other partner characteristics. I build on this by constructing an instrument that captures the availability of partner breadth at the time a deal is made.

I define the average available breadth index at time t in VC firm j as the sum of breadth indices of non-busy partners employed by VC firm j scaled by the number of non-busy partners employed by firm j at time t. Specifically, I define the average available breadth index of a VC firm j at time t as :

$$AvgB_{j,t} = \frac{\sum_{p\in j} B_p \times I_{p,t}}{\sum_{p\in j} I_{p,t}}, \tag{10}$$

where j denotes a VC firm, p denotes a partner. The sum $p \in j$ is taken over partners who work at VC firm j at time t, B_p is a breadth index measure at a partner level, $I_{p,t}$ is an indicator variable taking a value of 1 if the partner is non-busy at time t. In equation (10) I normalize by the number of available partners, instead the number of total partners to make sure the instrument captures variation in the composition of breadth over time within a VC firm - year, instead of the variation being driven by the overall capacity of the VC firm over time.[32]

Intuitively, at times when broad partners are busy the measure decreases, but at times when broad partners are more available relative to narrow partners the measure increases. A hypothesis is that at the time a deal is made a high average breadth availability should increase the probability a high breadth index partner being assigned to a deal and vice versa. So the average available breadth index is a natural candidate for an instrument of the chosen partner's breadth.

To study the effect of the interaction between breadth and novelty we need another instrument for the breadth × novelty interaction which will be the average available breadth index × deal novelty. Given these two instruments, I estimate the following model via a 2SLS:

$$\textbf{First Stage: } B_{d,p,j,t} = \alpha + AvgB_{j,t} + AvgB_{j,t} \times N_{d,t} + X_{p,t} + \eta_{i\times t\times s} + \rho_j + \epsilon_{d,p,j,t} \tag{11}$$

$$\textbf{First Stage: } N_{d,t} \times B_{d,p,j,t} = \alpha + AvgB_{j,t} + AvgB_{j,t} \times N_{d,t} + X_{p,t} + \eta_{i\times t\times s} + \rho_j + u_{d,p,j,t} \tag{12}$$

$$\textbf{Second Stage: } P_{d,p,j,t} = \alpha + N_{d,t} + \widehat{B_{j,t}} + \widehat{B_{j,t} \times N}_{d,t} + X_{p,t} + \eta_{i\times t\times s} + \rho_j + s_{d,p,j,t}, \tag{13}$$

where in $B_{d,p,j,t}$ is the breadth index partner p chosen for deal d made by investor j at time t. $AvgB_{j,t}$ is the average available breadth at investor j at time t,$AvgB_{j,t} \times N_{d,t}$ is the interaction between deal novelty and the average available breadth, $X_{p,t}$ is a set of time varying chosen partner level controls, $\eta_{i\times t\times s}$ are industry × year × deal stage fixed effects and ρ_j is an investor fixed effect. $P_{d,p,j,t}$ is a performance outcome.

Table 9 about here.

[32]For example if we have 4 partners with breadth index (0.2, 0.4, 0.4, 0.5) and the partner with an index of 0.5 is busy (the broadest partner) if one were to normalize by the number of total partners the average breadth would be 0.25, now if the partners with a breadth index of 0.2 and 0.4 are busy the average available breadth would lower (0.225) which clearly would not satisfy the IV monotonicity assumption.

The results are presented in Table 9. Columns (1) and (2) present the first-stage estimates from specifications (11) and (12). In column (1), the Avg. Available Breadth is a strong predictor of the Breadth Index of the chosen partner, and in column (2) the interaction between the Avg. Available Breadth and Novelty is a strong predictor of the interaction between the Breadth Index and Novelty. The first-stage estimate in column (1) implies that a one standard deviation increase in the average availability of breadth at a given VC firm is associated with an increase of 0.3 standard deviations in the Breadth Index of the partner actually chosen to lead the deal. Column (3) presents the IV estimates from specification (13), where the performance outcome is Major Success. First, the F-statistic for the instrument which is 44.6 passes the conventional weak instrument threshold of 10 (as show in table (3) and explicitly seen in the first stage results presented in columns (1) and (2)).[33] In column (3), the coefficient on the Breadth Index is negative and significant, whereas the coefficient on the interaction between the Breadth Index and Novelty is positive and significant. Column (4) presents OLS estimates for the same sample used in the IV specification.[34]

The effects size estimates imply large and economically meaningful effects: In OLS, a one–standard-deviation increase in breadth reduces major success by about 11 percentage points in incremental firms and raises success by roughly 9 percentage points in highly novel firms. The IV estimates are slightly higher and amplify both tails: a 16 percentage point reduction at low novelty and an 11 percentage point increase at high novelty. One reason for the larger IV estimates could be the fact that the IV identifies the effect among complier firms i.e., firms where the availability of breadth shifts partner assignment. In this setting non-complier deals are (i) the incremental projects that always get a narrow partner (ii) the novel projects that always get a broad partner regardless of partner availability. If for example the non-complier firms (both novel and non-novel) are systematically weaker deals where the leeway in partner assignment is very limited their outcomes could pull the OLS estimates toward zero on both ends of the novelty spectrum.[35]

Discussion of the exclusion restriction: The exclusion restriction underlying the IV is valid

[33]The Stock–Yogo critical values (7.03, 4.58, 3.95, 3.63 for maximal IV size distortions of 10%, 15%, 20%, and 25%, respectively), indicating that weak instruments are not a concern. In addition, weak-instrument robust inference tests (Anderson–Rubin Wald $\chi^2 = 7.31$, $p = 0.026$; Stock–Wright LM S test $\chi^2 = 10.03$, $p = 0.0066$) reject the null of no effect of the endogenous regressors, further supporting instrument relevance.

[34]The construction of the instrument restricts the IV sample. Since the instrument utilizes time variation across different partners in the same VC firm - year, the effect is identified in a subsample of VC firm - year observations where at least two different partners are time constraint at different points in time.

[35]On the low-novelty side, routine incremental projects that always receive a narrow partner tend to have low success probabilities regardless of who leads. Because these weak outcomes are mechanically linked to narrow breadth, OLS observes many low-success/low-breadth data points, which makes narrow partners appear worse on average. This depresses the baseline success rate for narrow partners and reduces the apparent performance gap between narrow and broad partners in incremental projects—attenuating the negative effect of breadth. On the high-novelty side, the most challenging frontier projects are often always staffed with broad partners, but because they are difficult to execute, their success rates are low. These low-success/high-breadth data points dilute the positive effect of breadth at high novelty, since OLS averages them together with the truly causal breadth effects observed among compliers.

only if the instrument influences the likelihood of success solely through partner assignment, and not through other channels. A first concern is that, even within a given year, VC partner activity may correlate with periods of market hotness (e.g., IPO or M&A seasonality), potentially affecting the types of projects selected. If market timing, rather than partner assignment, were driving the results, one would expect this channel to affect both lead and non-lead investments by the VC firm. In Table A6, I estimate the IV separately for lead and non-lead investments and find that the magnitude and statistical significance of both the baseline coefficient on human capital breadth and its interaction with novelty are much smaller for non-lead investments than for lead investments reinforcing the idea that the primary channel is partner assignment. Furthermore, the IV specification includes granular Industry x Year x Country x Stage FE, which absorb time-varying common industry x country shocks which could affect performance.

A second natural concern is that the instrument might proxy for selection on novelty: if broad partners are more inclined to take novel projects, then periods when more broad partners are available might also coincide with more novel deals being chosen. However, what the instrument shifts is who leads among the deals in the pipeline, not the distribution of novelty in the pipeline itself. Because I only observe realized deals, what matters is whether availability systematically changes the novelty of the projects that end up being financed. In Table A7, I show that it does not: average available breadth is uncorrelated with the novelty of chosen projects, both in the full sample and when separating lead and non-lead deals. This pattern is intuitive: the set of projects seeking funding at a given point in time is not driven by partner busyness shocks, but partner availability does shift whether a broad or a narrow partner takes the lead once the deal is selected.

A final concern is that partner busyness might influence the types of startups seeking VC funding, not just assignment. While I cannot observe startups that failed to receive funding, Table A8 shows that financed startups during busy vs. non-busy periods are similar in both novelty and ex-post performance. This suggests that the instrument does not substantially alter the composition of deals receiving funding.[36]

Additional tests: I conduct further robustness checks. In Table ID6, I re-estimate the 2SLS model using only IPO as the measure of success; results are similar. In Table ID9, I vary the time window used to define busyness. In Table ID10, I again vary the time window and restrict the busyness proxy to IPO-related exits. Across all specifications, the results remain consistent and statistically significant.

[36]However, I cannot fully rule out a channel through which partner busyness influences the upstream pool of startups seeking funding.

5 Theoretical Framework

In this section, I present a simple stylized model that endogenizes the venture screening and partner assignment choices of VC firms. I introduce minimal ingredients in order to explain the decision of VC firms to assign either generalist or specialist partners, and how this choice interacts with characteristics of novel ventures. The model focuses on the creation of novel ventures with uncertain quality, which require screening and involve inherently higher informational asymmetries compared to incremental, known ventures.

Roadmap: I model the assignment of a VC partner with specialization γ to screen a novel project with a prior probability π of being high quality. Specialization increases the fallback payoff $R_k(\gamma)$ but also raises the cost of effort, $\frac{1}{2}\gamma e^2$, generating a trade-off between a tighter endogenously determined cutoff s^* and lower screening precision via a reduced optimal effort e^*. I first solve for s^* and e^*, then show that: (i) the likelihood of financing novel projects declines with specialization, and this effect weakens as the prior π increases (Props. 2, 3); (ii) the conditional returns to financed novel projects decline with specialization when the sensitivity of the fallback payoff to specialization is small, and this effect also diminishes as π increases (Props. 4, 5). I then endogenize partner assignment, showing that the optimal specialization γ^* equates the marginal value of the fallback option with the marginal screening and wage costs (Prop. 10), and that optimal specialization decreases with the prior π (Cor. 2). Each result maps directly to the empirical tests in Tables 5, A10, 6, and A12.

In Subsection 5.1, I describe the model, outline the assumptions, and briefly discuss the VC firm's optimization problem. In Subsection 5.2, I present the main cross-sectional implications of the model, discuss the underlying intuition, and test these implications empirically. In Subsection 5.3, I endogenize the VC firm's partner assignment.

5.1 Description of the model

5.1.1 Description and discussion of model assumptions

The model is static. There is one novel project created in the economy, which can pay either $R_h > 0$ if the project is of high quality or $R_l < 0$ if the project is of low quality. There is one VC firm, which is risk-neutral and can invest one unit of capital in the project. The VC firm has a prior on the quality of the novel project, denoted by π, but does not observe the exact project type. The VC firm has a roster of partners labeled by a specialization level $\gamma \in [\gamma_{min}, \gamma_{max}]$, where a higher γ partner is a more specialized partner (a more specialized partner has lower human capital breadth). The VC firm assigns a partner of type γ to screen the novel project. To do so, the partner exerts observable and contractable effort e which incurs a cost $\frac{1}{2}\gamma e^2$. By exerting effort, the partner acquires an informative signal about the project's quality. The observed signal is given by:

$$s = \theta + \epsilon \text{ where } \epsilon \sim N(0, \frac{1}{e}), \tag{14}$$

where $\theta = 1$ if the novel project is of high quality and $\theta = 0$ if the novel project is of low quality, and e is the effort exerted by the partner. The VC firm will find it optimal to invest in the novel project if $s \geq s^*$, where s^* is endogenously determined. If $s < s^*$ the VC firm rejects the novel project and can invest in a fallback (known) venture that yields a return $R_k(\gamma)$ which is a differentiable, concave and increasing function of specialization γ.

Discussion of Assumptions: The framework relies on two assumptions about how partner specialization is related to (i) the returns to fallback, non-novel ventures and (ii) the screening costs of novel ventures.

Assumption 1 (Fallback non-novel venture returns increase with specialization)**.** *$R_k(\gamma)$ is a concave, increasing function of specialization γ.*

Assumption 1 captures the idea that more specialized partners, through accumulated expertise, are better able to extract value from incremental (relatively known) opportunities. For example, if the VC firm invests in biotech ventures, this assumption captures the fact that the more specialized the partner is in biotech, the higher the partner's fallback option on ventures that are close to the existing biotech frontier.

Assumption 2 (Screening costs for novel ventures increase with specialization)**.** *The screening cost for novel ventures $C = \frac{1}{2}\gamma e^2$ increases with specialization γ (i.e., decreases with human capital breadth).*

Assumption 2 stipulates that the costs of screening for novel ventures are lower for more generalist partners. This assumption is empirically motivated. Table A9 provides an empirical justification for this assumption. The results in Table A9 suggest that more novel ventures are relatively more similar to ventures in other sectors than to ventures in the focal venture sector. So, it is reasonable to assume that a more generalist partner, who has worked in multiple sectors, would presumably find it less costly to screen a novel venture.

5.1.2 Optimal novel investment signal threshold and optimal effort

Lemma 1 (Signal cut-off at which a novel project is accepted)**.** *Suppose return parameter values are such that information is valuable i.e., $\pi R_h + (1 - \pi)R_l < R_k(\gamma)$. Then, the cut-off signal at which investment in a novel project becomes acceptable is given by:*

$$s^* = \frac{1}{2} + \frac{\Lambda}{e}, \tag{15}$$

where:

$$\Lambda = ln\left(\frac{(R_k(\gamma) - R_l)(1 - \pi)}{(R_h - R_k(\gamma))\pi}\right) \tag{16}$$

Proof. See Internet Appendix E.

□

Lemma 1 defines the optimal signal threshold required for a novel project to be accepted. Notice that s^* is a function of effort, and through the parameter Λ is a function of γ and π. It can easily be shown that as effort increases s^* decreases. This is because at higher effort levels the signal acquired by the partner is more informative about venture quality so the partner requires a lower threshold in order to finance a novel venture. s^* increases in specialization γ. Higher specialization gives the partner a higher fallback return R_k which raises the investment signal required to accept a novel venture.

Given the signal informativeness e, I define the true positive rate $\alpha(e)$ i.e., the probability a venture is high quality and accepted and the false positive rate $\beta(e)$ i.e., the probability a venture is of low quality and is accepted:

$$\alpha(e) \;=\; \Pr\big[s \geq s^* \;\big|\; \theta = 1\big] \;=\; 1 - \Phi\big((s^* - 1)\sqrt{e}\big) \tag{17}$$

$$\beta(e) \;=\; \Pr\big[s \geq s^* \;\big|\; \theta = 0\big] \;=\; 1 - \Phi\big(s^*\sqrt{e}\big), \tag{18}$$

where $\Phi(\cdot)$ is the standard normal cdf.

The VC expected profit of the VC firm if the VC firm assigns a partner of type γ before paying the partner for her effort cost is then given by:

$$\begin{aligned}\Pi(e) = \;& \underbrace{\pi\,\alpha(e)\,R_h}_{\substack{\text{high quality venture}\\ \text{accepted}}} + \underbrace{(1-\pi)\,\beta(e)\,R_l}_{\substack{\text{low quality venture}\\ \text{accepted}}} \\ & + \underbrace{\pi\,[1-\alpha(e)]\,R_k(\gamma)}_{\substack{\text{high quality venture}\\ \text{rejected}}} + \underbrace{(1-\pi)\,[1-\beta(e)]\,R_k(\gamma)}_{\substack{\text{low quality venture}\\ \text{rejected}}}\end{aligned} \tag{19}$$

The VC firm will assign a partner type γ and will pay the partner a wage $w(e,\gamma)$ to compensate the partner for her effort $\frac{1}{2}\gamma e^2$ and her outside labour market option $u(\gamma)$.[37] The VC firm will pay the partner of type γ a wage to exactly compensate the partner for her effort cost and outside option hence the wage paid by the VC firm $w(\gamma, e)$ is:

$$w(\gamma, e) = \frac{1}{2}\gamma e^2 + u(\gamma) \tag{20}$$

[37] The outside option of the partner is a function of partner type γ but it is determined outside of the model by a competitive labour market. I do not make any parametric assumptions on $u(\gamma)$ except that it is differentiable on the relevant range of γ.

Proposition 1. *Given a partner type γ the VC firm will choose optimal effort to maximize:*

$$\max_{e \geq 0} U(e) \; = \; \underbrace{R_k + \pi\, \alpha(e)\, (R_h - R_k) + (1-\pi)\, \beta(e)\, (R_l - R_k)}_{\Pi(e) \text{ (expected return to VC firm)}} \; - \; \underbrace{\left(\frac{\gamma\, e^2}{2} + u(\gamma) \right)}_{\text{partner's compensation}} \tag{21}$$

Given (π, γ) *the condition below* (22) *defines the optimal level of effort* $e^* > 0$ *which is a maximum of* (21) *and satisfies* $e^* > 2\Lambda$.

$$\pi\, (R_h - R_k)\, \alpha'(e^*) + (1-\pi)\, (R_l - R_k)\, \beta'(e^*) \; = \; \gamma\, e^* \tag{22}$$

Proof. See Internet Appendix E. □

Intuitively, condition (22) equalizes marginal screening benefit (right hand size) to marginal screening cost (left hand side). Marginally increasing effort rises the true positive rate $\alpha(e)$ and lowers the false positive rate $\beta(e)$, since α and β are concave the linear marginal cost (left hand side) will cross the right hand side once for an interior effort defined by (22).

5.2 Main Predictions and Cross-Sectional Empirical Implications

In this subsection I, present four theoretical results, discuss their economic intuition and empirical implications.

Define the likelihood of accepting a novel project at optimal effort for each γ, $L_N(\gamma)$ as:

$$L_N(\gamma) = \pi\alpha(e^*) + (1-\pi)\beta(e^*) \tag{23}$$

Define the expected payoff in the novel sector, conditional on accepting a novel project at optimal effort for each γ, $E_N(\gamma)$ as:

$$E_N(\gamma) = \frac{\pi\alpha(e^*)R_h + (1-\pi)\beta(e^*)R_l}{\pi\alpha(e^*) + (1-\pi)\beta(e^*)} \tag{24}$$

Proposition 2 (The likelihood of financing a novel project decreases with specialization)**.** *Suppose that* $R_k > \frac{R_l + R_h}{2}$*, then at optimal effort the likelihood of accepting a novel project decreases with specialization i.e.,* $\frac{\partial L_N(\gamma)}{\partial \gamma} < 0$.

Proof. See Internet Appendix E. □

The intuition behind Proposition 2 is clear. Higher specialization increases the fallback option and, due to the higher cost of effort, lowers optimal effort. Both of these channels imply that more specialized partners will have a higher acceptance threshold, s^*, for novel projects. This reduces their likelihood of financing a novel project. Proposition 2 implies that, within VC firms, partners with lower specialization are more likely to finance novel projects. This proposition is the theoretical counterpart of the empirical results presented in Table 5, which correlate the lead partner breadth index with the novelty of financed projects and find a positive association between partner human capital breadth and project novelty.

Proposition 3 (The sensitivity of the likelihood of financing a novel project with respect to specialization decreases with the prior π)**.** *Suppose that* $R_k > \frac{R_l+R_h}{2}$*, then at optimal effort we have i.e.,* $\frac{\partial}{\partial \pi}\frac{\partial L_N(\gamma)}{\partial \gamma} > 0$.

Proof. See Internet Appendix E. □

The cross-partial in Proposition 3 states that the negative effect of specialization on the likelihood of accepting a novel project weakens as the prior π improves. The most intuitive way to see this is to examine the effect of an increase in π on the acceptance threshold. When π rises, the log-likelihood term Λ falls, so the sensitivity of the financing threshold $s^* = \frac{1}{2} + \Lambda/e^*$ with respect to effort decreases ($\frac{\partial s^*}{\partial e} = -\frac{\Lambda}{e^2}$.). Empirically, Proposition 3 implies that the difference in the likelihood of acceptance of a novel project between a broad and a specialized agent should be higher at lower π. I test this implication by constructing two sample splits that proxy for (i) deals for which π is lower or higher — early-stage vs. late-stage deals; and (ii) time periods when π is lower or higher — hot VC investment periods vs. normal or cold VC investment periods. Intuitively, early-stage ventures should have a comparatively lower π, and hot VC investment periods have been shown to result in the financing of more experimental ventures, i.e., ventures with lower π (Nanda and Rhodes-Kropf, 2013). I find support for this implication in the data. Table A10 presents the results. The correlation between the partner breadth index and deal novelty is higher in seed or early-stage deals than in late-stage deals (columns (1) and (2)), and in hot investment periods relative to normal or cold investment periods (columns (3) and (4)).[38]

Proposition 4 (Expected return conditional on financing novelty decreases with specialization)**.** *Suppose* $R_k > \frac{R_l+R_h}{2}$*, and the sensitivity of fallback returns with respect to specialization is small, i.e.,* $R_k(\gamma)' \ll \frac{(R_k-R_l)(R_h-R_k)}{R_h-R_l}$ *then the expected return conditional on financing a novel project decreases with the level of specialization,* $\frac{\partial E_N}{\partial \gamma} < 0$.

Proof. See Internet Appendix E. □

The result reflects a trade-off between information acquisition and the investment cut-off. Two forces are at play when specialization γ rises. First, higher γ increases the marginal cost of effort, so the optimal screening effort falls, $e^*(\gamma) \downarrow$. Lower e^* makes the signal noisier, which—holding the cutoff fixed—reduces the quality of accepted novel projects: the true positive rate $\alpha(e^*)$ falls, and the false positive rate $\beta(e^*)$ rises. Second, a higher γ implies a higher fallback return, $R_k(\gamma)$, which, through Λ, raises the investment threshold $s^*(\gamma)$, thereby improving the conditional pool of accepted projects. Proposition 4 states that when the sensitivity of the fallback return with respect to specialization is small, the first channel dominates the second. Proposition 4 implies that the performance of novel ventures should be higher for partners

[38] The evidence is suggestive: the key assumption is that the only difference between early- vs. late-stage ventures and hot vs. non-hot VC markets is the difference in π.

with greater human capital breadth. This is in line with the empirical results presented in Table 6.

The trade-off in Proposition 4 additionally implies that the sensitivity of the expected return conditional on financing novelty with respect to human capital breadth should be lower when the sensitivity of the fallback return with respect to specialization is higher. This suggests the following empirical test: in sectors where the returns of non-novel firms are more sensitive to prior sectoral experience, the difference in expected return conditional on financing novelty between a generalist and a specialist agent should be smaller. I test this implication empirically. First, to proxy for the sensitivity of non-novel venture returns to sectoral experience, I estimate—on the non-novel subsample—a specification in which partner prior sector experience is interacted with sector indicators. From this, I extract the marginal effect of partner sector experience on non-novel outcomes for each sector. The results are presented in Figure A3. Then, to test the main implication, I split the sample and estimate the breadth × novelty interaction for sectors with high vs. low sensitivity to experience. The results are reported in Table A11. This implication is supported by the data: in low-sensitivity sectors, the breadth × novelty interaction is higher than in high-sensitivity sectors (column (2) vs. column (1), and column (4) vs. column (3)).

Proposition 5 (The sensitivity of the conditional return of financing novelty with respect to specialization decreases with the prior π)**.** *Suppose* $R_k > \frac{R_l+R_h}{2}$ *and* $R_k(\gamma)' \ll \frac{(R_k-R_l)(R_h-R_k)}{R_h-R_l}$ *hold, then* $\frac{\partial}{\partial \pi}\frac{\partial E_N(\gamma)}{\partial \gamma} > 0$.

Proof. See Internet Appendix E. □

The cross–partial $\frac{\partial}{\partial \pi}\left(\frac{\partial E_N}{\partial \gamma}\right) > 0$ states that negative effect of γ on returns conditional on financing novelty weakens as the prior π improves. When π rises, the log–likelihood ratio Λ falls and the acceptance cutoff $s^* = \frac{1}{2} + \Lambda/e^*$ moves down, so a given increase in γ produces a smaller tightening of selection. At the same time, a higher π raises the marginal benefit of information in the FOC $\pi(R_h - R_k)\alpha'(e^*) + (1-\pi)(R_l - R_k)\beta'(e^*) = \gamma e^*$, leading the firm to choose a higher e^* at each γ. With a higher baseline e^*, the curvature of the (α, β) frontier implies that a small reduction in precision (induced by a marginal increase in γ) translates into a smaller loss in expected return. Thus both channels: lower s^* and higher e^* at larger π make the adverse impact of specialization on the conditional return less pronounced. Empirically proposition 5 implies that the return difference between a specialist and a generalist conditional on financing novelty should be higher at lower π. I test this implication via the same sample splits used to test proposition 3. The results are presented in Table A12. The difference between the breadth x novelty interaction across sample splits is not statistically significant.

5.3 Optimal Assignment Rule and Time-series Implications

In this section we endogenize the VC assignment rule. We have the following proposition:

Proposition 6 (Optimal assignment rule)**.** *Given π the VC firm maximizes the total expected profit and assigns a partner of type γ s.t.*

$$\max_{\gamma\in[\gamma_{min},\gamma_{max}]} V(\gamma) = \Pi(e^*(\gamma),\gamma) - \frac{1}{2}\gamma e^*(\gamma) - u(\gamma) \tag{25}$$

Given π the condition below defines the optimal partner specialization γ which maximizes (145).

$$R_k(\gamma^*)' q^r = \frac{1}{2} e(\gamma^*)^2 + u'(\gamma), \tag{26}$$

where $q^r = 1 - \pi\alpha(e^*) - (1-\pi)\beta(e^*)$ is the rejection probability.

Proof. See Internet Appendix E. □

The intuition is straightforward: the left-hand side represents the marginal benefit of increasing γ (i.e., assigning a specialist), which corresponds to the marginal return from a fallback deal that the VC firm receives if a novel deal is rejected. The right-hand side (RHS) captures the marginal cost, which includes lower screening efficiency and the marginal wage paid at each level of γ.

Simulating the model: To provide further evidence that the model aligns with the empirical findings, I simulate the full model and show, using the simulated data, that I can recover the signs of both the correlational implications and the IV estimates. To do so, I generate $N = 10{,}000$ random choice situations for different values of π, allowing the VC firm to optimally compute (s, e, γ^*). To fully rationalize my IV strategy, I discretize the γ space and assume that each partner is unavailable with some exogenous probability p_{busy}, which is independent across partners and choice situations. I then re-simulate the model under this constraint. Since, for each choice situation, I observe whether a fallback or novel project was financed, the project payoff, which partners were busy, and the project performance, I can re-estimate the IV specification in equation (13) using the simulated data. The estimates are presented in Table A13, and they show that (i) the coefficient on human capital breadth is negative and significant, and (ii) the interaction between human capital breadth and novelty is positive and significant—fully recovering the empirical IV estimates presented in subsection 4.2. The parameter values and simulation procedure are explained in Internet Appendix A.

5.3.1 Time-series implications

We have the following proposition.

Proposition 7 (Optimal specialization decreases with π)**.** *Assume payoffs are such that $R_k > \frac{R_h+R_l}{2}$ then optimal specialization decreases with π i.e. we have $\frac{\partial\gamma}{\partial\pi} < 0$.*

Proof. See Internet Appendix E. □

Under the specified parameter conditions, an increase in π incentivizes the VC firm to assign a more generalist partner. As π increases, the threshold for project acceptance declines.

Since rejecting a project is typically optimal, a higher π encourages the VC firm to assign a partner capable of generating an informative signal that can raise the project's signal above the threshold. This, in turn, incentivizes the assignment of generalists, as they exert effort at lower cost. Proposition 7 provides a rationale for the negative co-movement between financed novelty and specialization observed in the stylized facts. In Internet Appendix I, I present a model that endogenizes entrepreneurial entry into the novel sector, thereby making π a function of the cost of producing high-quality novelty and the cost of entering entrepreneurship. I show that an increase in the cost of producing high-quality novelty—i.e., a burden-of-knowledge mechanism à la Jones (2009)—directly reduces π, which weakens the VC firm's incentive to assign a generalist partner. This, in turn, creates a feedback loop that depresses π further, as entrepreneurs rationally anticipate VC assignment decisions when deciding whether to enter. A similar mechanism arises if the cost of entry into entrepreneurship declines, as argued by Ewens et al. (2018). In this case, more low-quality entrepreneurs enter, which again lowers π, reducing the VC firm's incentive to assign generalist partners and further depressing π.

6 Conclusion

This paper examines the role of venture capital (VC) partners' human capital breadth in investment selection, startup performance, and innovation outcomes. Empirically, I find that within VC firms, partners with broader backgrounds are more likely to lead investments in novel, high-risk startups. While these partners do not perform better on average average, their involvement in novel ventures significantly increases the likelihood of major success. These patterns are consistent with both selection—where broad-background partners excel at screening novel firms—and potential monitoring—where their engagement enhances firm performance. Exploiting plausibly exogenous variation in partner busyness as a shock to lead-partner assignment, I provide a plausibly causal evidence for these effects. These results highlight the critical role of human capital breadth in financing of novel business ventures and support the role for public policies fostering the development of broad human capital.

Figures

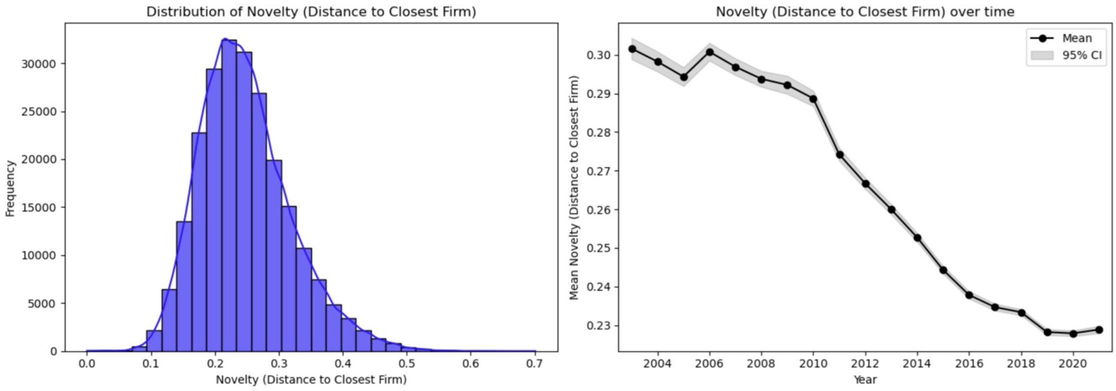

Figure 1: **Distribution of Novelty (Distance to Closest Firm) and time trend of Novelty (Distance to Closest Firm)** Left panel: This figure plots the distribution of the Novelty (Distance to Closest Firm) measure. Right Panel: Time trend of the mean of Novelty (Distance to Closest Firm) measure.

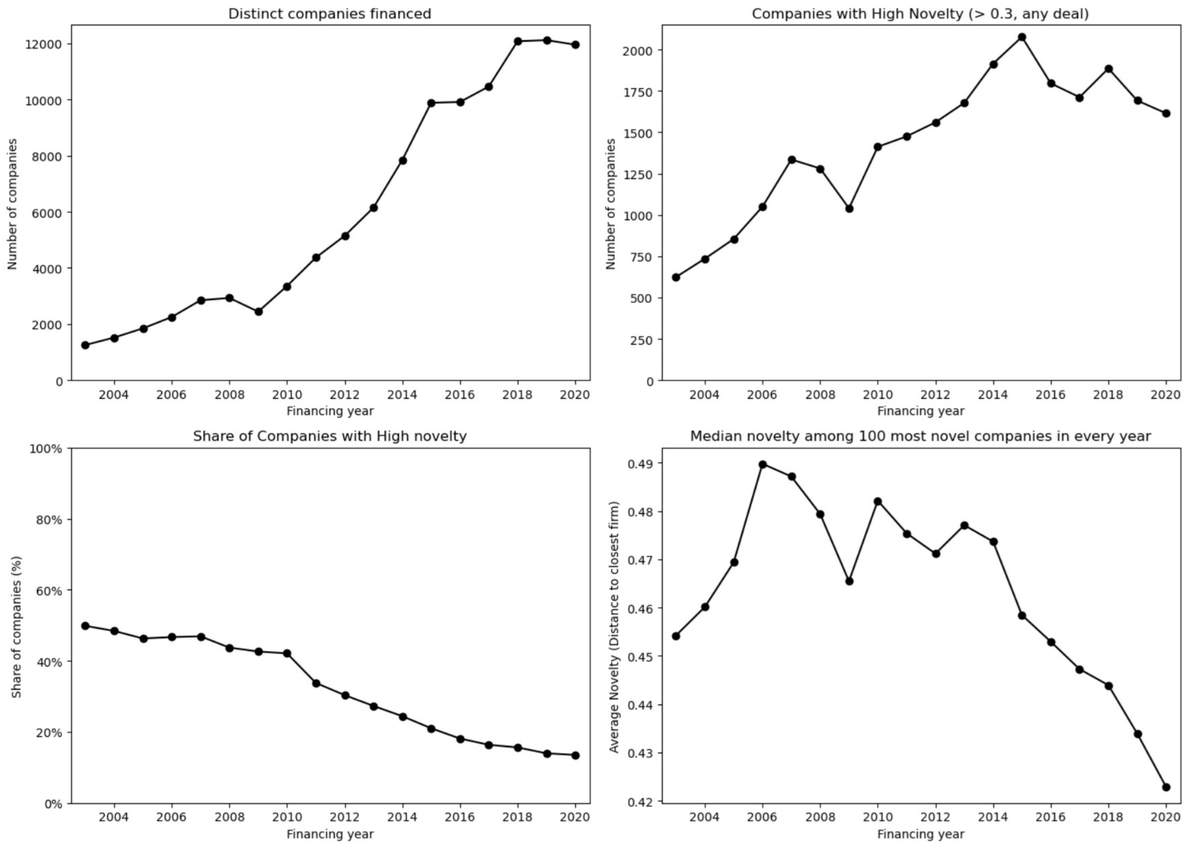

Figure 2: **Stylized facts about novelty of VC financed startups** Top left panel: Total number of newly VC financed; Top right panel: Total number of newly VC financed firms with novelty measure above 0.3 (top 25th percentile in novelty in the overall deal sample); Bottom left panel: Share of newly financed firms with novelty measure above 0.3 (top 25th percentile in novelty in the overall deal sample); Bottom right panel: Time evolution of median novelty among the top 100 most novel financed firms in each year.

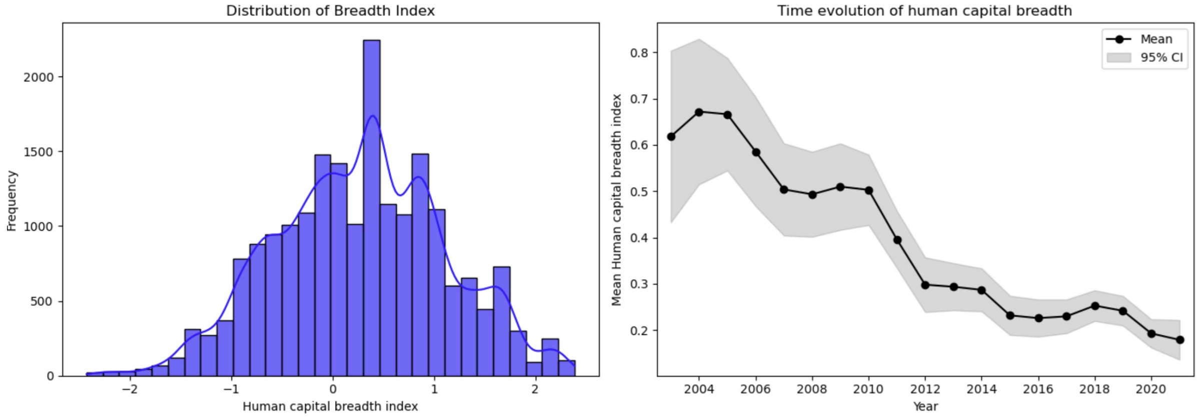

Figure 3: **Distribution of Breadth Index and time trend of Breadth Index** Left panel: This figure plots the distribution of the Breadth Index measure . Right Panel: Time trend of the mean of Breadth Index measure. Graph is done for partners for which we can observe at least 3 jobs prior to VC industry entry.

Tables

Variable	Obs	Mean	Std. Dev.	1%	5%	25%	50%	75%	95%	99%
Novelty (Distance to Closest Firm)	232,130	0.25	0.07	0.11	0.15	0.20	0.24	0.29	0.38	0.45
Novelty (Avg. Distance to Closest Firm)	232,130	0.28	0.07	0.14	0.18	0.23	0.27	0.32	0.41	0.48
IPO Exit	232,130	0.04	0.20	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Major Success	232,130	0.07	0.26	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Failure	232,130	0.35	0.48	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Lead Investment	232,130	0.30	0.46	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Number of Forward Patents	232,130	0.21	2.21	0.00	0.00	0.00	0.00	0.00	0.00	5.00
Number of Forward Citations	232,130	2.76	46.68	0.00	0.00	0.00	0.00	0.00	0.00	39.14
Breadth Index	34,959	-0.03	1.69	-3.14	-3.14	-0.76	0.34	1.21	2.19	2.29
Job categories ratio	28,520	0.59	0.28	0.14	0.20	0.33	0.50	1.00	1.00	1.00
Job roles ratio	28,520	0.83	0.21	0.25	0.43	0.67	1.00	1.00	1.00	1.00
Job industry ratio	28,520	0.71	0.27	0.00	0.25	0.50	0.71	1.00	1.00	1.00
Female	46,870	0.11	0.31	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Partner has MBA	33,250	0.43	0.49	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Partner has PhD	33,250	0.06	0.24	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Partner STEM education	33,250	0.31	0.46	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Partner Social science or Humanities education	33,250	0.63	0.48	0.00	0.00	0.00	1.00	1.00	1.00	1.00
Partner Attended Top School	33,250	0.50	0.50	0.00	0.00	0.00	1.00	1.00	1.00	1.00

Table 1: **Summary Statistics of Deal Level Sample** This table presents summary statistics for the deal level sample. Each observation is a deal. The summary statistics fro partner characteristics are at a deal level, i.e. a mean of 0.11 for the Female indicator means that over the sample period 11% of deals have a lead partner with a female gender.

Variable	Obs	Mean	Std. Dev.	1%	5%	25%	50%	75%	95%	99%
Breadth Index	6,763	0.06	1.51	-3.14	-3.14	-0.57	0.31	1.10	2.10	2.29
Job categories ratio	5,883	0.57	0.27	0.12	0.20	0.33	0.50	0.75	1.00	1.00
Job roles ratio	5,883	0.81	0.21	0.25	0.40	0.67	0.86	1.00	1.00	1.00
Job industry ratio	5,883	0.69	0.26	0.00	0.25	0.50	0.67	1.00	1.00	1.00
Female	9,644	0.13	0.34	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Partner has MBA	6,406	0.37	0.48	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Partner has PhD	6,406	0.06	0.24	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Partner STEM education	6,406	0.28	0.45	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Partner Social science or Humanities education	6,406	0.63	0.48	0.00	0.00	0.00	1.00	1.00	1.00	1.00
Partner Attended Top School	6,406	0.44	0.50	0.00	0.00	0.00	0.00	1.00	1.00	1.00

Table 2: **Summary Statistics for individual partners who have lead at least one investment** This table presents summary statistics for individual partners. Each observation is a unique partner for which the data is available. A mean of 0.13 for the Female indicator here means that 13% of partners who have led at least one deal are female.

Variable	Obs	Mean	Std. Dev.	1%	5%	25%	50%	75%	95%	99%
Novelty (Distance to Closest Firm)	164,762	0.24	0.07	0.11	0.14	0.19	0.24	0.28	0.38	0.45
Novelty (Avg. Distance to Closest Firm)	164,762	0.27	0.07	0.14	0.17	0.22	0.26	0.31	0.40	0.47
IPO Exit	164,762	0.05	0.22	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Major Success	164,762	0.10	0.30	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Failure	164,762	0.24	0.42	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Number of Partners	164,762	12.02	9.28	1.00	2.00	5.00	9.00	18.00	30.00	41.00
Number of Available Partners	164,762	11.09	8.90	0.00	1.00	4.00	9.00	16.00	28.00	38.00
Partner Leads a Deal	164,762	0.21	0.41	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Busy Partner	164,762	0.03	0.18	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Breadth Index	164,762	-0.01	1.72	-3.14	-3.14	-0.75	0.37	1.29	2.19	2.29

Table 3: **Summary Statistics of Choice Model Sample** This table presents summary statistics for the choice model sample. A deal is included if at least 1 partner who could have been a lead partner on the focal deal has either job history or educational history available.

Panel A: Association between startup novelty and likelihood of Failure and Major Success

	100 × Failure		100 × Major Success	
	(1)	(2)	(3)	(4)
Novelty (Distance to Closest Firm)	26.114***	11.102***	64.686***	64.555***
	(2.613)	(3.017)	(3.235)	(4.217)
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00
R^2	0.23	0.52	0.18	0.48

Panel B: Association between startup novelty and innovation outcomes

	Number of Forward Patents		Number of Forward Citations	
	(1)	(2)	(3)	(4)
Novelty (Distance to Closest Firm)	1.717*	1.037	2.727**	3.292***
	(0.920)	(1.250)	(1.152)	(1.194)
Exit Type Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00

Table 4: **Startup Novelty, Exit Outcomes, and Innovation (Panels A–B).** This table presents the association between startup novelty, exit outcomes, and innovation. The main independent variable is *Novelty (Distance to Closest Firm)*, defined as one minus the maximum cosine similarity between the textual description of the financed startup and those of all startups receiving VC financing within five years before the deal in the same stage. *Panel A* reports OLS regressions of exit outcomes on startup novelty. The dependent variables are (i) *Failure*, an indicator equal to 1 if the startup does not go public, is not acquired, and does not receive follow-up financing, and 0 otherwise; and (ii) *Major Success*, an indicator equal to 1 if the startup exits via an IPO or through an acquisition valued at least five times greater than the total venture capital raised, and 0 otherwise. Columns (1) and (3) include Industry × Deal Year × Deal Type × Country FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. *Panel B* reports Poisson regressions of innovation outcomes on startup novelty. The dependent variables are (i) the *Number of Forward Patents* and (ii) the *Number of Forward Citations* received by the financed startup. Exit type controls include indicators for IPO exit, M&A exit, and Failure. Columns (1) and (3) include Industry × Deal Year × Deal Type × Country FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	100 × Novelty (Distance to Closest Firm)			100 × Above Median Novelty		
	(1)	(2)	(3)	(4)	(5)	(6)
Breadth Index	0.162**	0.235*	0.660**	0.908	1.816**	4.027*
	(0.080)	(0.128)	(0.271)	(0.648)	(0.899)	(2.298)
VC Experience	0.032	-0.130	-0.109	0.518	-0.817	0.185
	(0.119)	(0.111)	(0.320)	(0.837)	(0.931)	(2.600)
Partner Industry Experience	-0.314**	0.099	0.346	-3.142**	0.912	5.044**
	(0.157)	(0.189)	(0.260)	(1.299)	(1.538)	(2.168)
Controls	✓	✓	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓	✓	✓
VC Firm × Deal Year FE		✓			✓	
VC Firm × Deal Year × Partner Entry Year FE			✓			✓
Observations	23531.00	23531.00	23531.00	23531.00	23531.00	23531.00
R^2	0.48	0.59	0.66	0.37	0.50	0.58

Table 5: **Association between lead partner's human capital breadth index and startup novelty.** This table reports the results of OLS regressions of deal novelty on the lead partner's human capital breadth. The dependent variable in columns (1)–(3) is *Novelty (Distance to Closest Firm)*, defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. The dependent variable in columns (4)–(6) is an indicator equal to 1 if the deal has above-median novelty. The main independent variable, *Breadth Index*, is defined as the first principal component from a PCA of four human capital breadth proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(6) include Industry × Deal Year × Deal Type × Country FE. Columns (2) and (5) also include VC Firm × Deal Year FE. Columns (3) and (6) include VC Firm × Deal Year × Partner Entry Year FE. All columns include individual-level partner controls: age, sex, and ethnicity. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success		
	(1)	(2)	(3)
Novelty (Distance to Closest Firm)	0.409***	0.431***	0.337***
	(0.079)	(0.081)	(0.084)
Breadth Index	-0.024	-0.034**	-0.052***
	(0.016)	(0.016)	(0.020)
Novelty (Distance to Closest Firm) × Breadth Index	0.133*	0.154**	0.216***
	(0.071)	(0.069)	(0.076)
VC Experience	0.004	0.003	0.014
	(0.004)	(0.005)	(0.012)
Partner Industry Experience	0.014*	-0.007	0.006
	(0.007)	(0.011)	(0.012)
Controls	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓
VC Firm × Deal Year FE		✓	
VC Firm × Deal Year × Partner Entry Year FE			✓
Observations	23531.00	23531.00	23531.00
R^2	0.41	0.55	0.64

Table 6: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success.** The dependent variable in columns (1)–(3) is *Major Success*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least five times greater than the total VC invested capital. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Breadth Index × Novelty (Distance to Closest Firm)* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Partner Leads a Deal					
	(Full Sample)	(Lead)	(Non-Lead)	(Early-Stage)	(Early-Stage Lead)	(Early-Stage Non-Lead)
Busy Partner	-1.730**	-2.814***	-1.399	-2.273**	-3.635***	-1.517
	(0.681)	(1.084)	(0.889)	(0.928)	(1.341)	(1.309)
VC Experience	7.463***	5.712***	8.524***	6.829***	5.182***	8.043***
	(0.117)	(0.196)	(0.147)	(0.157)	(0.246)	(0.207)
Partner Age	-0.094***	-0.096***	-0.094***	-0.079***	-0.043	-0.106***
	(0.014)	(0.023)	(0.018)	(0.018)	(0.028)	(0.025)
Partner Industry Experience	3.078***	2.869***	3.142***	3.125***	2.954***	3.173***
	(0.094)	(0.158)	(0.119)	(0.123)	(0.192)	(0.162)
Controls	✓	✓	✓	✓	✓	✓
Deal FE	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Observations	242751.00	91232.00	151519.00	148916.00	62693.00	86223.00
R^2	0.32	0.31	0.33	0.33	0.31	0.34

Table 7: **Association between partner busyness and the likelihood of a partner leading a deal.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Busy Partner* is an indicator equal to 1 if the partner is involved in exiting a deal via an IPO or a high-value acquisition in the time window (−90, +90) days around the focal deal date. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. All columns include individual-level controls for sex and ethnicity, as well as Deal FE and VC Firm FE. Column (1) reports results for the full deal-level sample. Columns (2) and (3) report results separately for lead and non-lead investments. Column (4) reports results for the restricted sample of early-stage novel investments. Columns (5) and (6) report results separately for lead and non-lead investments within the early-stage novel sample. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Partner Leads a Deal		
	(Full Sample)	(Lead)	(Non-Lead)
Novelty (Distance to Closest Firm) × Breadth Index	4.504***	5.011***	3.535**
	(1.146)	(1.751)	(1.541)
Busy Partner	-1.916**	-3.094**	-1.093
	(0.936)	(1.378)	(1.331)
Controls	✓	✓	✓
Deal FE	✓	✓	✓
VC Firm FE	✓	✓	✓
Partner FE	✓	✓	✓
Observations	148837.00	62666.00	86171.00
R^2	0.43	0.45	0.46

Table 8: **Interaction between human capital breadth and novelty and the likelihood of leading a deal.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Novelty (Distance to Closest Firm) × Breadth Index* is the interaction between a deal's novelty and the partner's human capital breadth. *Busy Partner* is an indicator equal to 1 if the partner is involved in exiting a deal via an IPO or a high-value acquisition in the time window (−90, +90) days around the focal deal date. All columns include individual-level controls for age and experience, as well as Deal FE, VC Firm FE, and Partner FE. Column (1) reports results for the full sample. Columns (2) and (3) report results separately for lead and non-lead investments. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Breadth Index	Breadth Index × Novelty (Distance to Closest Firm)	Major Success	
	(IV-First Stage)	(IV-First Stage)	(IV-Second Stage)	(OLS)
Avg. Available Breadth	0.826*** (0.058)	-0.018 (0.017)		
Avg. Available Breadth × Novelty (Distance to Closest Firm)	-0.565** (0.261)	0.735*** (0.095)		
Breadth Index			-0.163** (0.073)	-0.119** (0.051)
Breadth Index × Novelty (Distance to Closest Firm)			0.841** (0.324)	0.649*** (0.202)
Novelty (Distance to Closest Firm)			0.291 (0.381)	0.488* (0.284)
Controls	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓	✓
Observations	1226.00	1226.00	1226.00	1226.00
R^2	0.87	0.86	0.09	0.53
F-statistic of Instrument			44.60	

Table 9: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup success.** This table reports the results of an instrumental variable regression of startup outcomes on lead partner human capital breadth and its interaction with deal novelty. The dependent variable in Columns (3)–(4) is *Major Success*, an indicator equal to 1 if the startup goes public (IPO) or is acquired at a valuation at least five times greater than the total VC invested capital. The endogenous regressors are the lead partner's *Breadth Index*, constructed as the first principal component of four human capital breadth measures (job category ratio, job role ratio, job industry ratio, and educational breadth count), and the interaction *Breadth Index × Novelty (Distance to Closest Firm)*. *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the focal startup's description and all startups financed in the past five years in the same stage. The instruments are (i) *Avg. Available Breadth*, defined as the average breadth index across all partners at the VC firm who are not busy with a high-value exit event (IPO or acquisition) within $(-90, +90)$ days of the focal deal, and (ii) *Avg. Available Breadth × Novelty (Distance to Closest Firm)*. Column (1) reports the first stage regression for the Breadth Index. Column (2) reports the first stage regression for the interaction Breadth Index × Novelty. Column (3) reports the IV second stage estimates. Column (4) reports OLS estimates in the same sample. All specifications include VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors, reported in parentheses, are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

References

Abuzov, Rustam, “Busy venture capitalists and investment performance,” *Journal of Financial and Quantitative Analysis*, 2024, pp. 1–35.

Bernstein, Shai, Xavier Giroud, and Richard R Townsend, “The impact of venture capital monitoring,” *The Journal of Finance*, 2016, *71* (4), 1591–1622.

Bhardwaj, Abhishek, Abhinav Gupta, Sabrina T Howell, and Kyle Zimmerschied, “Does Fund Size Affect Private Equity Performance? Evidence from Donation Inflows to Private Universities,” Technical Report, National Bureau of Economic Research 2025.

Blickle, Kristian, Cecilia Parlatore, and Anthony Saunders, “Specialization in banking,” Technical Report, National Bureau of Economic Research 2023.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, “Are ideas getting harder to find?,” *American Economic Review*, 2020, *110* (4), 1104–1144.

Bonelli, Maxime, “Data-driven Investors,” 2022.

Calder-Wang, Sophie and Paul A Gompers, “And the children shall lead: Gender diversity and performance in venture capital,” *Journal of Financial Economics*, 2021, *142* (1), 1–22.

Casamatta, Catherine, “Financing and advising: optimal financial contracts with venture capitalists,” *The journal of finance*, 2003, *58* (5), 2059–2085.

Chemmanur, Thomas J, Karthik Krishnan, and Debarshi K Nandy, “How does venture capital financing improve efficiency in private firms? A look beneath the surface,” *The Review of Financial Studies*, 2011, *24* (12), 4037–4090.

Chen, Jiafeng and Jonathan Roth, “Logs with zeros? Some problems and solutions,” *The Quarterly Journal of Economics*, 2024, *139* (2), 891–936.

Cressy, Robert, Federico Munari, and Alessandro Malipiero, “Playing to their strengths? Evidence that specialization in the private equity industry confers competitive advantage,” *Journal of Corporate Finance*, 2007, *13* (4), 647–669.

Custódio, Cláudia, Miguel A Ferreira, and Pedro Matos, “Generalists versus specialists: Lifetime work experience and chief executive officer pay,” *Journal of Financial Economics*, 2013, *108* (2), 471–492.

_ , _ , and _ , “Do general managerial skills spur innovation?,” *Management Science*, 2019, *65* (2), 459–476.

Dorn, David, Florian Schoner, Moritz Seebacher, Lisa Simon, and Ludger Woessmann, “Multidimensional Skills on LinkedIn Profiles: Measuring Human Capital and the Gender Skill Gap,” 2025.

Duevski, Teodor and Viacheslav Bazaliy, “Capital Allocation Concentration Measurement in Venture Capital,” *Available at SSRN 5216676*, 2025.

Ewens, Michael and Denis Sosyura, “Irreplaceable venture capitalists,” Technical Report, Technical report, Working Paper 2023.

_ and Matt Marx, “Founder replacement and startup performance,” *The Review of Financial Studies*, 2018, *31* (4), 1532–1565.

_ and Matthew Rhodes-Kropf, “Is a VC Partnership Greater than the Sum of its Partners?,” *The Journal of Finance*, 2015, *70* (3), 1081–1113.

_ , Ramana Nanda, and Matthew Rhodes-Kropf, “Cost of experimentation and the evolution of venture capital,” *Journal of Financial Economics*, 2018, *128* (3), 422–442.

Falato, Antonio, Dalida Kadyrzhanova, and Ugur Lel, "Distracted directors: Does board busyness hurt shareholder value?," *Journal of Financial Economics*, 2014, *113* (3), 404–426.

Fich, Eliezer M and Anil Shivdasani, "Are busy boards effective monitors?," *The Journal of finance*, 2006, *61* (2), 689–724.

Field, Laura, Michelle Lowry, and Anahit Mkrtchyan, "Are busy boards detrimental?," *Journal of financial economics*, 2013, *109* (1), 63–82.

Garfinkel, Jon A, Erik J Mayer, Ilya A Strebulaev, and Emmanuel Yimfor, "Alumni networks in venture capital financing," *SMU Cox School of Business Research Paper*, 2021, (21-17).

Gompers, Paul A and Josh Lerner, "The money of invention: how venture capital creates new wealth," 2002.

_ , Will Gornall, Steven N Kaplan, and Ilya A Strebulaev, "How do venture capitalists make decisions?," *Journal of Financial Economics*, 2020, *135* (1), 169–190.

Gompers, Paul, Anna Kovner, and Josh Lerner, "Specialization and success: Evidence from venture capital," *Journal of Economics & Management Strategy*, 2009, *18* (3), 817–844.

_ , _ , _ , and David Scharfstein, "Venture capital investment cycles: The impact of public markets," *Journal of financial economics*, 2008, *87* (1), 1–23.

Guzman, Jorge and Aishen Li, "Measuring founding strategy," *Management Science*, 2023, *69* (1), 101–118.

Hampole, Menaka, Dimitris Papanikolaou, Lawrence DW Schmidt, and Bryan Seegmiller, "Artificial intelligence and the labor market," Technical Report, National Bureau of Economic Research 2025.

Hauser, Roie, "Busy directors and firm performance: Evidence from mergers," *Journal of Financial Economics*, 2018, *128* (1), 16–37.

He, Songrun, Linying Lv, Asaf Manela, and Jimmy Wu, "Chronologically Consistent Large Language Models," *arXiv preprint arXiv:2502.21206*, 2025.

Hellmann, Thomas and Manju Puri, "The interaction between product market and financing strategy: The role of venture capital," *The review of financial studies*, 2000, *13* (4), 959–984.

Hochberg, Yael V, Alexander Ljungqvist, and Yang Lu, "Whom you know matters: Venture capital networks and investment performance," *The journal of finance*, 2007, *62* (1), 251–301.

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

Hsu, David H, "What do entrepreneurs pay for venture capital affiliation?," *The journal of finance*, 2004, *59* (4), 1805–1844.

Humphery-Jenner, Mark, "Diversification in private equity funds: On knowledge sharing, risk aversion, and limited attention," *Journal of Financial and Quantitative Analysis*, 2013, *48* (5), 1545–1572.

Jones, Benjamin F, "The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?," *The Review of Economic Studies*, 2009, *76* (1), 283–317.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, "Time-varying fund manager skill," *The Journal of Finance*, 2014, *69* (4), 1455–1484.

Kaplan, Steven N and Per ER Strömberg, "Characteristics, contracts, and actions: Evidence from venture capitalist analyses," *The journal of finance*, 2004, *59* (5), 2177–2210.

— **and Per Strömberg**, "Venture capitalists as principals: Contracting, screening, and monitoring," *American Economic Review*, 2001, *91* (2), 426–430.

— **, Berk A Sensoy, and Per Strömberg**, "Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies," *The Journal of Finance*, 2009, *64* (1), 75–115.

Kecht, Valentin, Alessandro Lizzeri, and Farzad Saidi, "AGING AT THE VERY TOP," 2025.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring technological innovation over the long run," *American Economic Review: Insights*, 2021, *3* (3), 303–320.

Kempf, Elisabeth, Alberto Manconi, and Oliver Spalt, "Distracted shareholders and corporate actions," *The Review of Financial Studies*, 2017, *30* (5), 1660–1695.

Lazear, Edward P, "Balanced skills and entrepreneurship," *American Economic Review*, 2004, *94* (2), 208–211.

— , "Leadership: A personnel economics approach," *Labour Economics*, 2012, *19* (1), 92–101.

Lerner, Josh and Amit Seru, "The use and misuse of patent data: Issues for finance and beyond," *The Review of Financial Studies*, 2022, *35* (6), 2667–2704.

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

— **, Junxi Liu, Jacob Moscona, and David Y Yang**, "Appropriate entrepreneurship? the rise of china and the developing world," Technical Report, National Bureau of Economic Research 2024.

Lindsey, Laura, "Blurring firm boundaries: The role of venture capital in strategic alliances," *The Journal of Finance*, 2008, *63* (3), 1137–1168.

Malenko, Andrey, Ramana Nanda, Matthew Rhodes-Kropf, and Savitar Sundaresan, "Catching outliers: Committee voting and the limits of consensus when financing innovation," *Harvard Business School Entrepreneurial Management Working Paper*, 2024, (21-131).

Murphy, Kevin J and Jan Zabojnik, "CEO pay and appointments: A market-based explanation for recent trends," *American economic review*, 2004, *94* (2), 192–196.

— **and** — , "Managerial capital and the market for CEOs," *Available at SSRN 984376*, 2007.

Nahata, Rajarishi, "Venture capital reputation and investment performance," *Journal of financial economics*, 2008, *90* (2), 127–151.

Nanda, Ramana and Matthew Rhodes-Kropf, "Investment cycles and startup innovation," *Journal of financial economics*, 2013, *110* (2), 403–418.

Paine, Fiona, "Not on Terra FIRRMA: Foreign Investment in US Startups and Innovation," *Available at SSRN 5508418*, 2024.

Retterath, Andre and Reiner Braun, "Benchmarking venture capital databases," *Available at SSRN 3706108*, 2020.

Sørensen, Morten, "How smart is smart money? A two-sided matching model of venture capital," *The Journal of Finance*, 2007, *62* (6), 2725–2762.

Spaenjers, Christophe and Eva Steiner, "Specialization and performance in private equity: Evidence from the hotel industry," *Journal of Financial Economics*, 2024, *162*, 103930.

Zambrana, Rafael and Fernando Zapatero, "A tale of two types: Generalists vs. specialists in asset management," *Journal of Financial Economics*, 2021, *142* (2), 844–861.

Appendix

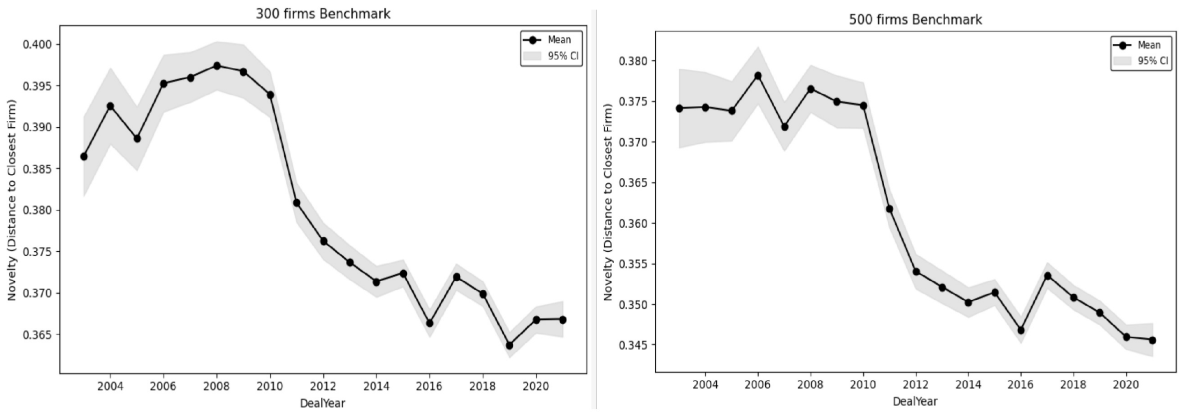

Figure A1: **Time trend of Novelty (Distance to Closest Firm)** Left panel: This panel plots the time trend of mean novelty across VC financed firms over time when novelty of a deal in each year is measured by comparing the deal in each year to a constant number of 300 random firms who have received financing over the last five years prior to the focal deal. Right panel: This panel plots the time trend of mean novelty across VC financed firms over time when novelty of a deal in each year is measured by comparing the deal in each year to a constant number of 500 random firms who have received financing over the last five years prior to the focal deal.

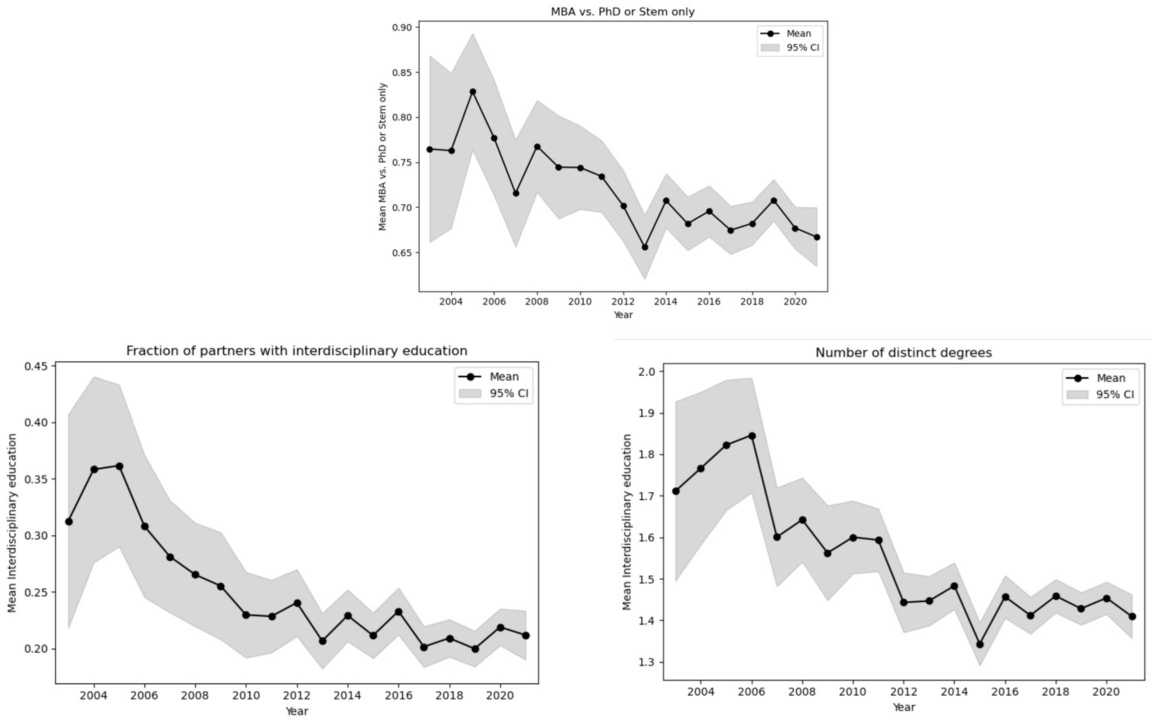

Figure A2: **Time trend of human capital breadth using educational variables to measure human capital breadth** Upper figure: This figure plots the time evolution of the fraction of partners with an MBA degree relative to the fraction of partners who hold either a PhD or are educated in a STEM field over time. Left figure: This figure plots the fraction of partners with an interdisciplinary education over time. Right figure: This figure plots the time evolution of the average number of distinct degrees per partner.

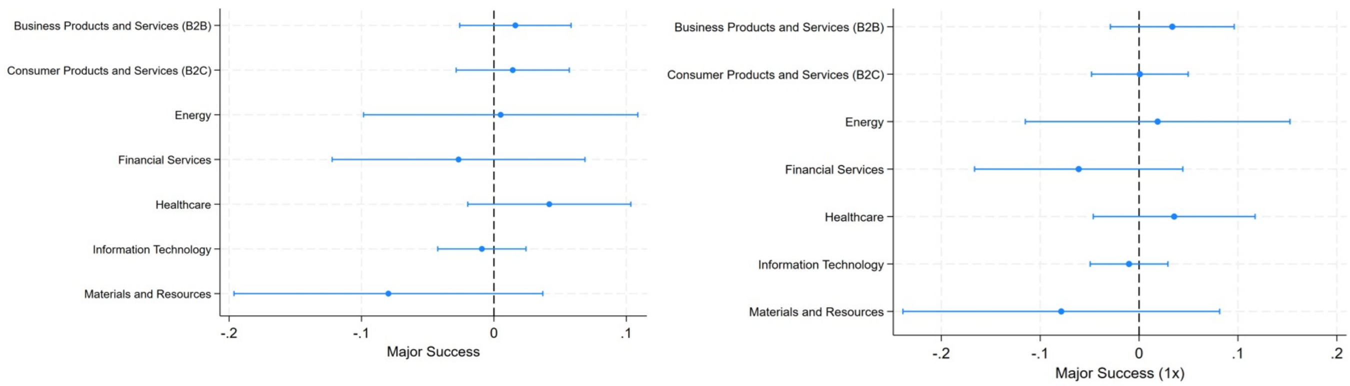

Figure A3: **Sensitivity of non-novel venture returns to past industry experience** This figure plots the OLS estimates of the interaction terms (β_s) of the following regression estimated on a subsample of deals with below median novelty: $P_{j,k,p,t} = \alpha + \sum_s \beta_s \text{PartnerExp}_{p,i} \times \mathbf{1}\{\text{Sector} = i\} + \sum_s \gamma_s \mathbf{1}\{\text{Sector} = i\} + \delta \text{PartnerExp}_{p,i} + X_{t,p} + \eta_{t\times s\times c} + \rho_{t\times j} + \epsilon_{j,k,p,t}$. In the left panel I use major success as an outcome variable. In the right panel I use a less strict definition of major success: major success (1x) as an outcome variable.

Startup	Novelty Measure	Closest Startup	Novelty Percentile
Tesla	0.52	Community Energy	99.9
Facebook	0.51	Mode Media	99.8
Square	0.48	Softgate Systems	99.5
SpaceX	0.48	Zero-G	99.5
Slack	0.48	Octopz	99.4
Airbnb	0.45	Dopplr	99.0
Skype	0.45	Arrival Communications	98.8
Dropbox	0.37	Omnidrive	94.2
Napster	0.27	Deezer	64.1
Revolut	0.19	Wise (Application)	22.9
Instagram	0.18	Pixable	18.8
Stripe	0.13	Moip	2.4

Table A1: **Example of novel and non-novel startups and previously VC funded startups closest to their business model.** This table presents the Novelty (Distance to Closest Firm) of well-known startups and the startup with closest business model them previously financed by the VC industry. Column (1) is the name of the startup. Column (2) is the raw novelty measure. Column (3) is the name of the closest startup and column (4) is a percentile ranking of novelty constructed using the full sample of firms.

	100 × Failure		100 × Major Success	
	(1)	(2)	(3)	(4)
Novelty Quartile (Distance to Closest Firm)=2	2.120***	2.050***	-0.436*	-0.242
	(0.423)	(0.499)	(0.261)	(0.315)
Novelty Quartile (Distance to Closest Firm)=3	4.201***	2.816***	-0.265	-0.102
	(0.439)	(0.524)	(0.281)	(0.341)
Novelty Quartile (Distance to Closest Firm)=4	5.713***	2.948***	8.667***	8.920***
	(0.493)	(0.575)	(0.509)	(0.658)
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00
R^2	0.23	0.52	0.17	0.47

Table A2: **Association between startup novelty quartiles and likelihood of Failure and Major Success.** This table reports the results of OLS regressions of exit outcomes on indicators for the startup's novelty quartile. The independent variables are dummy variables equal to 1 if the financed startup belongs to novelty quartile i (with $i = 2, 3, 4$) within a given year and deal stage, based on the measure *Novelty (Distance to Closest Firm)*, defined as one minus the maximum cosine similarity between the textual description of the financed startup and those of all startups receiving VC financing within the prior five years in the same stage. The omitted category is novelty quartile 1 (least novel deals). The dependent variables are: (i) *Failure*, an indicator equal to 1 if the startup does not go public, is not acquired, and does not receive follow-up financing; and (ii) *Major Success*, an indicator equal to 1 if the startup exits via IPO or through an acquisition valued at least five times greater than the total VC financing raised. Columns (1) and (3) include Industry × Deal Year × Deal Type × Country FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Number of Forward Patents		Number of Forward Citations	
	(1)	(2)	(3)	(4)
Novelty Quartile (Distance to Closest Firm)=2	0.145	-0.063	0.410**	0.316*
	(0.108)	(0.125)	(0.197)	(0.190)
Novelty Quartile (Distance to Closest Firm)=3	0.189	0.068	0.370*	0.315*
	(0.127)	(0.176)	(0.194)	(0.189)
Novelty Quartile (Distance to Closest Firm)=4	0.335**	0.260	0.675***	0.825***
	(0.163)	(0.227)	(0.232)	(0.251)
Exit Type Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00

Table A3: **Association between startup novelty quartiles and innovation outcomes — Poisson count model.** This table reports the results of Poisson regressions of innovation outcomes on deal-level novelty quartiles. The independent variables are novelty quartile dummies indicating whether a startup belongs to the i-th novelty quartile within a given year and deal stage, based on the *Novelty (Distance to Closest Firm)* measure, defined as one minus the maximum cosine similarity between the financed startup's textual description and all startups that received VC financing within the prior five years in the same stage. The omitted category is Novelty Quartile = 1. The dependent variable in columns (1)–(2) is the *Number of Forward Patents*, defined as the total number of patents granted to the firm after the deal date. The dependent variable in columns (3)–(4) is the *Number of Forward Citations*, defined as the total number of forward citations, adjusted by grant year and NBER subcategory, that the firm's patents receive after the deal date. All columns include *Exit Type Controls*, which are dummies for whether the startup exits via IPO, via acquisition, or receives follow-up financing. Columns (1)–(2) include Industry × Deal Year × Deal Type × Country FE. Columns (3)–(4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	100 × Novelty (Distance to Closest Firm)			100 × Above Median Novelty		
	(Lead)	(Non-Lead)	(Interaction)	(Lead)	(Non-Lead)	(Interaction)
Breadth Index	0.509**	0.262	0.157	3.778*	0.985	1.018
	(0.235)	(0.182)	(0.141)	(2.140)	(1.240)	(0.973)
VC Experience	-0.473**	-0.078	-0.131	-1.099	-0.395	-0.818
	(0.204)	(0.174)	(0.110)	(1.737)	(1.546)	(0.923)
Partner Industry Experience	-0.108	0.471	0.097	2.258	0.920	0.887
	(0.433)	(0.301)	(0.189)	(3.700)	(2.375)	(1.536)
Lead Investment=1 × Breadth Index			0.192			1.941*
			(0.146)			(1.153)
Controls	✓	✓	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓	✓	✓
Observations	9519.00	14012.00	23531.00	9519.00	14012.00	23531.00
R^2	0.67	0.65	0.59	0.62	0.55	0.50

Table A4: **Association between lead partner's human capital breadth index and startup novelty: Lead vs. Non - Lead investments** This table reports the results of an OLS regression of deal novelty and lead partner's human capital breadth for lead and non-lead investments. The dependent variable Novelty (Distance to Closest Firm) in columns (1), (2) and (5) is a deal level measure of novelty defined as one minus the maximum of the cosine similarity of the textual description of the startup financed in the deal and all startups receiving venture capital financing within five years before the deal in the same deal stage. The independent variable in columns (3),(4) and (6) is an indicator taking a value of 1 if the deal is a deal with above median novelty. The main independent variable Breadth Index is a measure of human capital breadth which is defined as the first principal component of a PCA from 4 individual human capital breadth proxies: 1) The ratio of distinct job categories to total employment spells 2) The ratio of distinct job roles to total employment spells 3) The ratio of distinct industries worked in to total employment spells 4) Educational breadth count. Lead Investment is an indicator variable equal to 1 if the investment is led by the VC firm. In columns (1) and (3) the sample includes all lead investments by VC firms. In columns (2) and (4) the sample includes all non-lead investments by VC firms. VC experience is a variable defined as the logarithm of one plus the number of deals financed by the lead partner prior to current deal. Partner Industry Experience is a dummy variable taking a value of 1 if the partner has had at least one job in the industry of the venture and 0 otherwise. Columns (1) - (6) include Industry × Deal Year × Deal Type × Financed Company Country FE and Investor × Deal Year FE. All columns include individual level partner controls: age, sex and ethnicity. Standard errors reported in parenthesis are double clustered at the investor and company level. * p<.10; ** p<.05; *** p<.01.

	Partner Leads a Deal					
	(Full Sample)	(Lead)	(Non-Lead)	(Early-Stage)	(Early-Stage Lead)	(Early-Stage Non-Lead)
Busy Partner (Placebo)	0.562 (0.431)	0.520 (0.702)	0.622 (0.550)	1.207** (0.597)	1.299 (0.893)	1.206 (0.813)
VC Experience	7.403*** (0.114)	5.613*** (0.191)	8.476*** (0.144)	6.753*** (0.154)	5.055*** (0.240)	7.993*** (0.203)
Partner Age	-0.094*** (0.014)	-0.096*** (0.023)	-0.094*** (0.018)	-0.078*** (0.018)	-0.043 (0.028)	-0.106*** (0.025)
Partner Industry Experience	3.084*** (0.094)	2.880*** (0.158)	3.147*** (0.119)	3.132*** (0.123)	2.965*** (0.193)	3.178*** (0.162)
Controls	✓	✓	✓	✓	✓	✓
Deal FE	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Observations	242751.00	91232.00	151519.00	148916.00	62693.00	86223.00
R^2	0.32	0.31	0.33	0.33	0.31	0.34

Table A5: **Placebo test of the association between partner busyness and the likelihood of leading a deal.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Busy Partner (Placebo)* is a placebo variable constructed by reshuffling the Busy Partner indicator within each deal while keeping the overall distribution of busy partners unchanged. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the partner prior to the current deal. *Partner Age* is measured in years at the time of the deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. All columns include individual-level controls for sex and ethnicity, as well as Deal FE and VC Firm FE. Column (1) reports results for the full sample. Columns (2) and (3) split the full sample into lead and non-lead investments. Column (4) reports results for the restricted sample of early-stage novel investments. Columns (5) and (6) split this restricted sample into lead and non-lead investments. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Major Success	
	(Lead)	(Non-Lead)
Breadth Index	-0.260***	-0.096
	(0.083)	(0.100)
Breadth Index × Novelty (Distance to Closest Firm)	1.045***	0.311
	(0.292)	(0.528)
Novelty (Distance to Closest Firm)	0.323	0.907
	(0.288)	(0.599)
Total Breadth	0.002	0.000
	(0.002)	(0.002)
Controls	✓	✓
VC Firm FE	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓
Observations	894.00	1328.00
R^2	0.18	0.07
F-statistic of Instrument	42.25	29.10

Table A6: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup success — IV estimates by lead vs. non-lead, controlling for total breadth.** This table reports instrumental variable regressions of startup outcomes on the lead partner's human capital breadth and its interaction with deal novelty. The dependent variable in both columns is *Major Success*, an indicator equal to 1 if the startup goes public (IPO) or is acquired at a valuation at least five times greater than the total VC invested capital. *Breadth Index* is constructed as the first principal component of four human-capital breadth measures (job category ratio, job role ratio, job industry ratio, and educational breadth count). *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the focal startup's description and all startups financed in the prior five years in the same stage. The endogenous regressors *Breadth Index* and *Breadth Index* × *Novelty* are instrumented with *Avg. Available Breadth* and *Avg. Available Breadth* × *Novelty*. All specifications control for *Total Breadth*, the sum of the breadth indices across all partners at the VC firm. Column (1) reports IV second-stage estimates for lead investments; column (2) reports IV second-stage estimates for non-lead investments. Specifications include VC Firm fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Novelty (Distance to Closest Firm)		
	(Full Sample)	(Lead)	(Non-Lead)
Avg. Available Breadth	-0.003	0.001	-0.013**
	(0.003)	(0.005)	(0.006)
Controls	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓
Observations	2222.00	894.00	1328.00
R^2	0.55	0.66	0.61

Table A7: **Correlation between average available breadth and deal novelty.** The dependent variable is *Novelty (Distance to Closest Firm)*, defined as one minus the maximum cosine similarity between the textual description of the focal startup and all startups financed in the previous five years within the same stage. The key independent variable is *Avg. Available Breadth*, defined as the average breadth index across all partners at the VC firm who are not busy with a high-value exit event (IPO or major acquisition) within $(-90, 90)$ days of the focal deal. Controls include partner age, sex, ethnicity, log tenure, and partner industry experience. Column (1) reports results for the full sample, while columns (2) and (3) split the sample into lead and non-lead investments, respectively. All regressions include VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors (in parentheses) are double clustered at the investor and company levels. * p<.10; ** p<.05; *** p<.01.

	Novelty (Distance to Closest Firm)	Major Success	Failure
	(1)	(2)	(3)
Busy	0.002 (0.006)	-0.016 (0.034)	0.036 (0.043)
Controls	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓
Observations	8487.00	8491.00	8491.00
R^2	0.59	0.59	0.65

Table A8: **Balance test on key outcomes for busy vs. non-busy VC Firm periods.** This table reports OLS regressions of three outcomes on the indicator *Busy*, which takes the value 1 if at the time of a deal there is at least one busy partner in the VC firm and 0 otherwise. The dependent variables are: (1) *Novelty (Distance to Closest Firm)*, defined as one minus the maximum cosine similarity between the focal startup and startups financed in the previous five years in the same stage; (2) *Major Success*, an indicator equal to 1 if the startup exits via IPO or acquisition at a value at least five times total VC investment; and (3) *Failure*, an indicator equal to 1 if the startup fails. Controls include partner age, sex, ethnicity, log tenure, and partner industry experience. All regressions include VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors (in parentheses) are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

	Novelty (Distance to Closest Firm)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative Similarity (50 Closest)	-0.875*** (0.003)		-0.874*** (0.004)					
Relative Similarity (100 Closest)		-0.862*** (0.004)		-0.860*** (0.005)				
Relative Similarity Ratio (50 Closest)					-0.037*** (0.012)		-0.070*** (0.016)	
Relative Similarity Ratio (100 Closest)						-0.030** (0.013)		-0.077*** (0.017)
Industry × Deal Year × Deal Type × Country FE	✓	✓			✓	✓		
Industry × Deal Year × Deal Type × Country × VC Firm FE			✓	✓			✓	✓
Observations	151760.00	156251.00	151760.00	156251.00	151760.00	156251.00	151760.00	156251.00
R^2	0.75	0.72	0.82	0.80	0.25	0.25	0.50	0.50

Table A9: **Novelty and relative similarity** This table presents the results of an OLS regression of the Novelty (Distance to Closest Firm) measure on the relative similarity of the focal venture to ventures in the same sector relative to ventures in other sectors. The dependent variable *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the focal startup's description and all startups financed in the prior five years in the same stage. The key independent variables capture how similar the focal startup is to ventures in its own sector relative to ventures in other sectors. *Relative Similarity (50 Closest)* and *Relative Similarity (100 Closest)* are defined as:

$$\frac{\text{Average Similarity to Same-Sector Startups} - \text{Average Similarity to Other-Sector Startups}}{\text{Average Similarity to Same-Sector Startups} + \text{Average Similarity to Other-Sector Startups}},$$

computed using the 50 or 100 most similar startups, respectively. *Relative Similarity Ratio (50 Closest)* and *Relative Similarity Ratio (100 Closest)* are defined as:

$$\frac{\text{Average Similarity to Same-Sector Startups}}{\text{Average Similarity to Other-Sector Startups}}.$$

Standard errors (in parentheses) are double clustered at the VC-firm and company level. Standard errors (in parentheses) are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

	Novelty (Distance to Closest Firm)			
	(Seed or Early Stage)	(Late Stage)	(Hot Market)	(Not Hot Market)
Breadth Index	0.292*	-0.273	0.555*	0.154
	(0.150)	(0.297)	(0.283)	(0.134)
VC Experience	-0.200	0.214	-0.334	-0.082
	(0.126)	(0.306)	(0.246)	(0.117)
Partner Industry Experience	-0.023	-0.007	0.668	-0.019
	(0.235)	(0.529)	(0.481)	(0.205)
Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Observations	16832.00	6699.00	4361.00	19170.00
R^2	0.59	0.69	0.59	0.59

Table A10: **Association between lead partner's human capital breadth index and startup novelty: Early vs. Late Stage and Hot vs. Not Hot VC Markets** This table reports the results of OLS regressions of deal novelty on the lead partner's human capital breadth for different sample splits. Column (1) presents the results when the Deal Type is Early Stage or Seed Round. Column (2) presents the results when the Deal Type is Later Stage. Column (3) presents the results for Hot Market conditions defined as years where the total VC fundraising is at least one standard deviation higher than the prior five year rolling average (years of hot VC market in the sample are: 2006, 2007, 2014 and 2015). Column (4) presents the results for Non Hot Market conditions. The dependent variable in columns (1)–(4) is *Novelty (Distance to Closest Firm)*, defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. The main independent variable, *Breadth Index*, is defined as the first principal component from a PCA of four human capital breadth proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(6) include Industry × Deal Year × Deal Type × Country FE. Columns (2) and (5) also include VC Firm × Deal Year FE. Columns (3) and (6) include VC Firm × Deal Year × Partner Entry Year FE. All columns include individual-level partner controls: age, sex, and ethnicity. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success			
	(High sensitivity sectors)	(Low sensitivity sectors)	(Healthcare)	(Other sectors)
Novelty (Distance to Closest Firm)	0.632***	0.154*	1.570***	0.221***
	(0.129)	(0.093)	(0.390)	(0.071)
Breadth Index	-0.033	-0.093***	-0.014	-0.058***
	(0.032)	(0.027)	(0.081)	(0.018)
Novelty (Distance to Closest Firm) × Breadth Index	0.164	0.319***	0.143	0.202***
	(0.113)	(0.108)	(0.301)	(0.071)
VC Experience	-0.009	0.010	-0.053	0.012
	(0.013)	(0.020)	(0.043)	(0.011)
Partner Industry Experience=1	0.053**	-0.019	0.063	0.004
	(0.022)	(0.021)	(0.066)	(0.009)
Controls	✓	✓	✓	✓
Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year × Partner Entry Year FE	✓	✓	✓	✓
Observations	10711.00	12822.00	3424.00	20109.00
R^2	0.66	0.56	0.75	0.52

Table A11: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success. High prior industry experience sensitivity sectors vs. low prior experience sensitivity sectors** Column (1) presents the results for high prior industry experience sensitivity sectors: Healthcare, Energy, Business Products and Services (B2B) and Consumer Products and Services (B2C). Column (2) presents the results for low prior industry experience sensitivity sectors: Materials and Resources, Information Technology and Financial Services. Column (3) presents the results for the Healthcare sector - the sector with the highest prior industry experience sensitivity. Column (4) presents the estimates for all other sectors except Healthcare. The dependent variable in columns (1)–(4) is *Major Success*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least five times greater than the total VC invested capital. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Breadth Index × Novelty (Distance to Closest Firm)* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success			
	(Seed or Early Stage)	(Late Stage)	(Hot Market)	(Not Hot Market)
Novelty (Distance to Closest Firm)	0.280***	0.843***	0.832***	0.324***
	(0.087)	(0.221)	(0.198)	(0.079)
Breadth Index	-0.034**	-0.033	-0.046	-0.032*
	(0.017)	(0.037)	(0.038)	(0.017)
Novelty (Distance to Closest Firm) × Breadth Index	0.162**	0.140	0.156	0.153**
	(0.072)	(0.154)	(0.151)	(0.071)
VC Experience	0.002	0.002	-0.006	0.005
	(0.005)	(0.013)	(0.011)	(0.004)
Partner Industry Experience=1	-0.001	-0.034	-0.020	-0.005
	(0.013)	(0.032)	(0.023)	(0.011)
Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Observations	16832.00	6699.00	4361.00	19170.00
R^2	0.55	0.66	0.54	0.56

Table A12: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success. Early vs. Late Stage and Hot vs. Not Hot VC Markets** Column (1) presents the results when the Deal Type is Early Stage or Seed Round. Column (2) presents the results when the Deal Type is Later Stage. Column (3) presents the results for Hot Market conditions defined as years where the total VC fundraising is at least one standard deviation higher than the prior five year rolling average (years of hot VC market in the sample are: 2006, 2007, 2014 and 2015). Column (4) presents the results for Non Hot Market conditions. The dependent variable in columns (1)–(4) is *Major Success*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least five times greater than the total VC invested capital. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Breadth Index* × *Novelty (Distance to Closest Firm)* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Performance	
	(IV)	(OLS)
Breadth Index	-0.184***	-0.214***
	(0.003)	(0.003)
Novelty × Breadth Index	0.762***	0.833***
	(0.231)	(0.147)
Novelty	3.928***	3.920***
	(0.212)	(0.209)
Observations	10,000	10,000
R^2	0.13	0.13
F-statistic of Instrument	601.30	

Table A13: **IV and OLS Estimates on Simulated Data** This table reports the results of an instrumental variable regression and an OLS regression of startup outcomes on lead partner human capital breadth and its interaction with deal novelty using the simulated dataset described in Appendix A. The simulation draws $N = 10{,}000$ deals. Payoffs are $R_h = 15$ for high-type novel projects and $R_\ell = -8$ for low-type novel projects; the fallback payoff is $R_k(\gamma) = 5 + 2\sqrt{\gamma}$. Effort precision is $\tau = \kappa e$ with $\kappa = 1$ and cost $\frac{1}{2}\gamma e^2$. The entrepreneur's share of a successful high-type project is $\epsilon = 0.1$. Partner specializations are drawn from $\Gamma = \{0.05, \ldots, 0.60\}$ (baseline $M = 5$ partners). The cost of hiring a partner is $u(\gamma) = a_1/\gamma + a_2\gamma^2$ with $a_1 = 0.10$ and $a_2 = 5.0$. The breadth index used in regressions is $B = 1/\gamma$. The prior probability of a high-type project is $\pi \sim U[0.25, 0.50]$, and partners are independently unavailable with probability $p_{\text{block}} = 0.50$. The instrument is the average breadth of available partners, standardized within the estimation sample. Robust standard errors are in parentheses

Internet Appendix

Human Capital Breadth and Financing of Innovative Startups

(not intended for publication)

Internet Appendix A: Description of Simulation

This appendix describes how the simulated dataset is generated and how the OLS and IV specifications are estimated. The underlying model is presented in the main text; this section documents the functional forms, parameter values, assignment mechanics, and estimation steps.

Each deal involves a project that may be of high or low type. The payoff from financing a high-type project is $R_h = 15$, and for a low-type project, $R_\ell = -8$. If the project is not financed, the investor earns a fallback payoff increasing with the partner's specialization γ:

$$R_k(\gamma) = 5 + 2\sqrt{\gamma}.$$

Effort $e \geq 0$ maps to signal precision via $\tau = \kappa e$ with $\kappa = 1$, and incurs cost $\frac{1}{2}\gamma e^2$. The entrepreneur retains share $\epsilon = 0.1$ if a high-type project is financed. Partner specialization γ determines a U-shaped cost of hiring:

$$u(\gamma) = \frac{a_1}{\gamma} + a_2\gamma^2, \quad \text{with } a_1 = 0.10,\ a_2 = 5.0,$$

yielding an interior optimum when screening is active. The human capital breadth index used in regressions is the inverse of specialization: $B = 1/\gamma$. Each project's prior probability of being high type is drawn from $\pi \sim \text{Uniform}[0.25, 0.50]$.

The VC firm has $M = 5$ partners with specializations $\Gamma = \{\gamma_1, \ldots, \gamma_M\}$, evenly spaced over $[0.05, 0.60]$.

The dataset contains $N =$ 10,000 simulated deals. For each deal $i = 1, \ldots, N$, the simulation proceeds as follows:

- Draw $\pi_i \sim \text{Uniform}[0.25, 0.50]$.
- Generate common random numbers reused across counterfactuals: $u_i \sim \text{Uniform}[0, 1]$ (type draw) and $q_i \sim \mathcal{N}(0, 1)$ (signal noise).
- Each partner j is independently unavailable with probability $p_{\text{block}} = 0.5$. If all are unavailable, one is flipped to available at random. Let $\mathcal{A}_i \subseteq \Gamma$ be the available set.
- The instrument is the average breadth of available partners:

  $$Z_i = \frac{1}{|\mathcal{A}_i|} \sum_{\gamma_j \in \mathcal{A}_i} \frac{1}{\gamma_j},$$

  and we standardize it to $Z_i^{\text{std}} = (Z_i - \bar{Z})/s_Z$.
- The VC firm assigns the partner that maximizes expected value:

  $$\gamma_i^* \in \arg\max_{\gamma_j \in \mathcal{A}_i} V(\pi_i, \gamma_j).$$

Record the assigned γ_i^*, its breadth $B_i = 1/\gamma_i^*$, and a "changed-assignment" flag comparing this to the optimal partner under full availability.

- The realized outcome is R_h with probability π_i, and R_ℓ otherwise (Bernoulli draw).
- Standardize breadth to $B_i^{\text{std}} = (B_i - \bar{B})/s_B$ and construct interaction terms: $N_i B_i^{\text{std}}$ and $N_i Z_i^{\text{std}}$.

The resulting dataset contains, for each deal, the assigned partner breadth, the constructed instrument, deal novelty, outcome, and all relevant interaction terms for estimating OLS and IV specifications.

Internet Appendix B: Stylized Facts Robustness

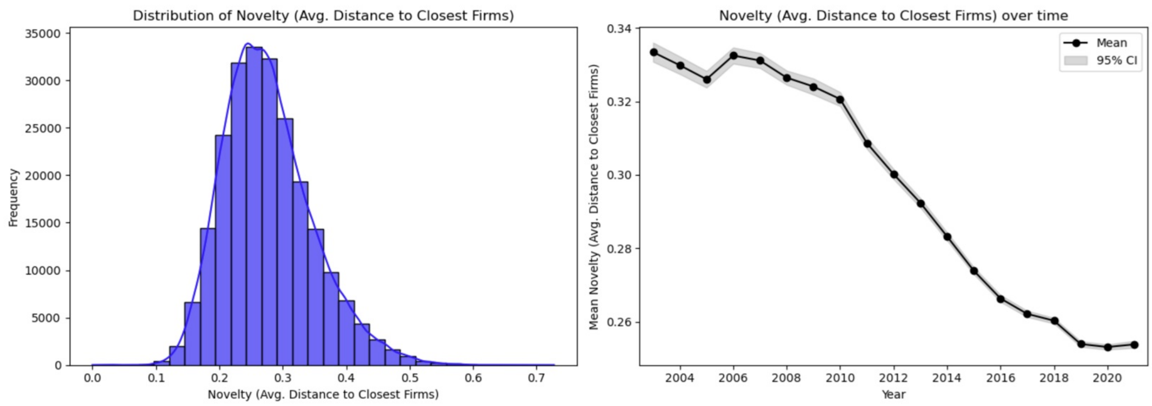

Figure IB1: **Distribution of Novelty (Avg. Distance to Closest Firm) and time trend of Novelty (Avg. Distance to Closest Firm)** Left panel: This figure plots the distribution of the Novelty (Avg. Distance to Closest Firm) measure. Right Panel: Time trend of the mean of Novelty (Avg. Distance to Closest Firm) measure.

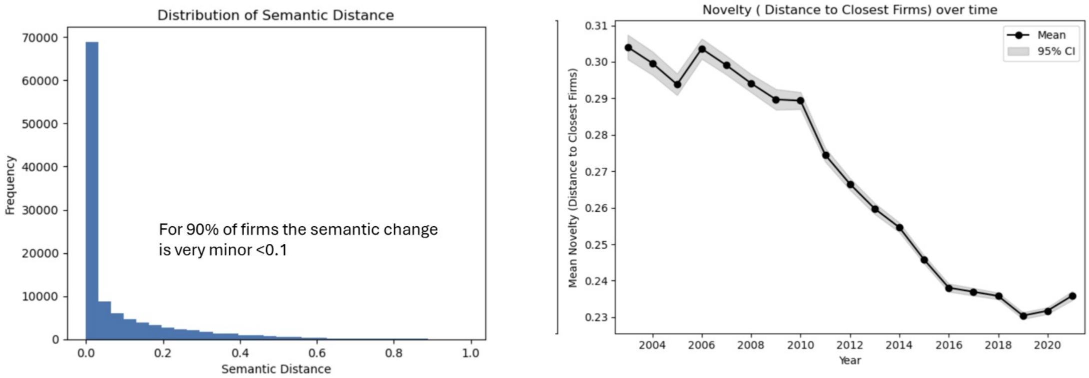

Figure IB2: **Time trend of Novelty (Distance to Closest Firm)** Left panel: This panel plots the time trend of mean novelty across VC financed firms over time when novelty of a deal in each year is measured by comparing the deal in each year to a constant number of 300 random firms who have received financing over the last five years prior to the focal deal. Right panel: This panel plots the time trend of mean novelty across VC financed firms over time when novelty of a deal in each year is measured by comparing the deal in each year to a constant number of 500 random firms who have received financing over the last five years prior to the focal deal.

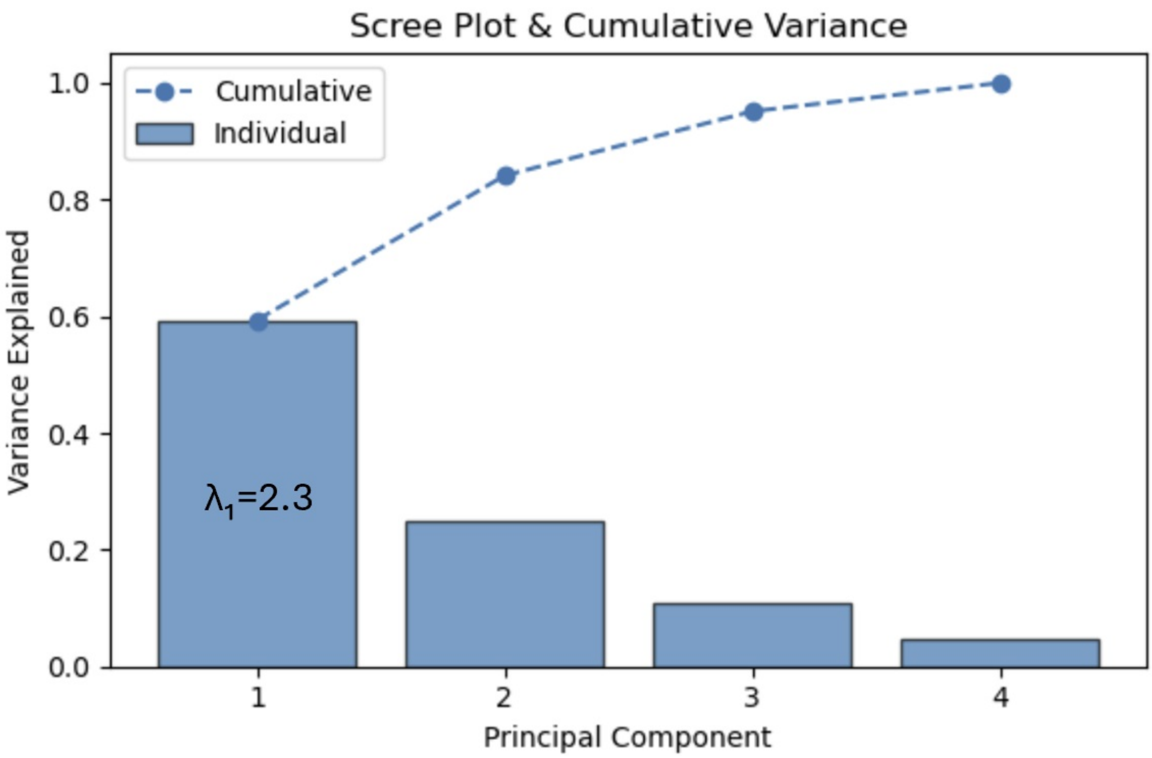

Figure IB3: **Principal component analysis PCA scree plot** Scree plot of PCA

	100 × Failure		100 × Major Success	
	(1)	(2)	(3)	(4)
Novelty (Avg. Distance to Closest Firm)	32.309*** (2.807)	14.550*** (3.172)	67.464*** (3.418)	68.709*** (4.505)
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00
R^2	0.23	0.52	0.18	0.48

Table IB1: **Association between startup novelty and likelihood of Failure and Major Success — robustness to an alternative novelty measure.** This table reports the results of OLS regressions of exit outcomes on deal-level novelty. The independent variable *Novelty (Avg. Distance to Closest Firm)* is defined as one minus the average cosine similarity between the textual description of the financed startup and its five most similar prior startups receiving VC financing within five years in the same stage. The dependent variables are: (i) *Failure*, an indicator equal to 1 if the startup does not exit via IPO, is not acquired, and does not receive follow-up financing; and (ii) *Major Success*, an indicator equal to 1 if the startup exits via IPO or through an acquisition valued at least five times greater than total VC financing received. Columns (1) and (2) include Industry × Deal Year × Deal Type × Country FE. Columns (3) and (4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	100 × Failure		100 × Major Success	
	(1)	(2)	(3)	(4)
Novelty (Distance to Closest Firm)	44.764*** (4.918)	25.285*** (6.223)	28.313*** (3.350)	25.475*** (3.953)
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	92395.00	92395.00	92395.00	92395.00
R^2	0.20	0.51	0.11	0.48

Table IB2: **Association between startup novelty and likelihood of Failure and Major Success — robustness to restricting to firms with first financing rounds between 2018–2021.** This table reports the results of OLS regressions of exit outcomes on deal-level novelty. The independent variable *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the textual description of the financed startup and all prior startups in the same stage receiving VC financing within five years. The dependent variables are: (i) *Failure*, equal to 1 if the startup does not exit via IPO, is not acquired, and does not receive follow-up financing; and (ii) *Major Success*, equal to 1 if the startup exits via IPO or through an acquisition valued at least five times greater than total VC financing received. Columns (1) and (2) include Industry × Deal Year × Deal Type × Country FE. Columns (3) and (4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	100 × Failure		100 × IPO	
	(1)	(2)	(3)	(4)
Novelty (Distance to Closest Firm)	26.114*** (2.613)	11.102*** (3.017)	72.290*** (3.202)	72.889*** (4.403)
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00
R^2	0.23	0.52	0.23	0.53

Table IB3: **Association between startup novelty and likelihood of Failure and IPO exit.** This table reports the results of OLS regressions of exit outcomes on deal-level novelty. The independent variable *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the textual description of the financed startup and all prior startups in the same stage receiving VC financing within five years. The dependent variables are: (i) *Failure*, equal to 1 if the startup does not exit via IPO, is not acquired, and does not receive follow-up financing; and (ii) *IPO*, equal to 1 if the startup exits via an initial public offering. Columns (1) and (2) include Industry × Deal Year × Deal Type × Country FE. Columns (3) and (4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	IPO Valuation		IPO Multiple	
	(1)	(2)	(3)	(4)
Novelty (Distance to Closest Firm)	974.463*** (284.438)	1189.557*** (405.859)	654.215** (323.705)	632.906 (580.750)
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	9086.00	9086.00	7996.00	7996.00
R^2	0.45	0.77	0.22	0.52

Table IB4: **Association between startup novelty and IPO outcomes for a subsample of deals with an IPO exit** This table reports the results of OLS regressions of IPO valuation outcomes on deal-level novelty, conditional on successful IPO exit. The independent variable *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the textual description of the financed startup and all prior startups in the same stage receiving VC financing within five years. The dependent variable in columns (1)–(2) is the *IPO Valuation*, measured at the time of the IPO. The dependent variable in columns (3)–(4) is the *IPO Multiple*, calculated as the ratio of the IPO valuation to the VC firm's investment size. Columns (1) and (3) include Industry × Deal Year × Deal Type × Country FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Number of Forward Patents		Number of Forward Citations	
	(1)	(2)	(3)	(4)
Novelty (Avg. Distance to Closest Firm)	2.212** (0.957)	1.615 (1.371)	3.220*** (1.191)	3.806*** (1.295)
Exit Type Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00

Table IB5: **Association between startup novelty and innovation outcomes — Poisson count model (robustness to an alternative novelty measure).** This table reports the results of Poisson regressions of innovation outcomes on deal-level novelty. The independent variable *Novelty (Avg. Distance to Closest Firm)* is defined as one minus the average cosine similarity between the textual description of the financed startup and the five most similar startups that received VC financing within five years prior to the deal in the same stage. The dependent variable in columns (1)–(2) is the *Number of Forward Patents*, defined as the total number of patents granted to the financed firm after the deal date. The dependent variable in columns (3)–(4) is the *Number of Forward Citations*, defined as the total number of forward citations, adjusted by grant year and NBER subcategory, that the firm's patents receive after the deal date. All columns include *Exit Type Controls*, which are dummies for whether the startup exits via IPO, via acquisition, or receives follow-up financing. Columns (1)–(2) include Industry × Deal Year × Deal Type × Country FE. Columns (3)–(4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Log(1+Number of Forward Patents)		Log(1+Number of Forward Citations)	
	(1)	(2)	(3)	(4)
Novelty (Distance to Closest Firm)	0.047* (0.026)	0.048 (0.032)	0.071 (0.047)	0.084 (0.059)
Exit Type Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓		✓	
Industry × Deal Year × Deal Type × Country × VC Firm FE		✓		✓
Observations	232108.00	232108.00	232108.00	232108.00

Table IB6: **Association between startup novelty and innovation outcomes — log(1+) model.** This table reports the results of deal-level regressions of innovation outcomes on startup novelty. The independent variable *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the financed startup's textual description and all startups that received VC financing within the prior five years in the same stage. The dependent variable in columns (1)–(2) is $\log(1 + \text{Number of Forward Patents})$, where *Number of Forward Patents* is the total number of patents granted to the firm after the deal date. The dependent variable in columns (3)–(4) is $\log(1 + \text{Number of Forward Citations})$, where *Number of Forward Citations* is the total number of adjusted citations (by grant year and NBER subcategory) that the financed firm's patents receive after the deal date. All columns include *Exit Type Controls*, which are dummies for whether the startup goes public, is acquired, or receives follow-up financing. Columns (1)–(2) include Industry × Deal Year × Deal Type × Country FE. Columns (3)–(4) include Industry × Deal Year × Deal Type × Country × VC Firm FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

Internet Appendix C: Human capital Breadth - Novelty Association Additional Robustness Tables

	100 × Novelty (Avg. Distance to Closest Firm)			100 × Above Median Novelty		
	(1)	(2)	(3)	(4)	(5)	(6)
Breadth Index	0.152**	0.235**	0.523*	0.866	2.116**	3.489
	(0.076)	(0.115)	(0.278)	(0.650)	(0.926)	(2.633)
VC Experience	-0.027	-0.141	-0.214	-0.918	-1.522*	-0.937
	(0.116)	(0.103)	(0.305)	(0.846)	(0.871)	(2.319)
Partner Industry Experience	-0.352**	0.059	0.351	-3.702***	1.059	4.738**
	(0.149)	(0.181)	(0.245)	(1.270)	(1.490)	(2.199)
Controls	✓	✓	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓	✓	✓
VC Firm × Deal Year FE		✓			✓	
VC Firm × Deal Year × Partner Entry Year FE			✓			✓
Observations	23531.00	23531.00	23531.00	23531.00	23531.00	23531.00
R^2	0.52	0.62	0.69	0.38	0.52	0.60

Table IC1: **Association between lead partner's human capital breadth index and startup novelty – robustness to an alternative novelty measure.** This table reports the results of OLS regressions of deal novelty on the lead partner's human capital breadth. The dependent variable in columns (1)–(3) is 100 × *Novelty (Avg. Distance to Closest Firm)*, defined as 100 times one minus the average cosine similarity between the textual description of the financed startup and its five closest neighboring startups (based on text embeddings), where the comparison set includes all startups receiving venture capital financing within five years before the deal in the same stage. The dependent variable in columns (4)–(6) is 100 × *Above Median Novelty*, an indicator equal to 100 if the deal's average novelty is above the median and 0 otherwise. The main independent variable, *Breadth Index*, is defined as the first principal component from a PCA of four human capital breadth proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(6) include Industry × Deal Year × Deal Type × Country FE. Columns (2) and (5) also include VC Firm × Deal Year FE. Columns (3) and (6) include VC Firm × Deal Year × Partner Entry Year FE. All regressions include partner-level controls: age, sex, and ethnicity. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	100 × Novelty (Distance to Closest Firm)		100 × Above Median Novelty	
	(Non-Syndicated)	(Syndicated)	(Non-Syndicated)	(Syndicated)
Breadth Index	0.428*	0.167	3.285	0.217
	(0.259)	(0.193)	(2.294)	(1.335)
VC Experience	-0.093	-0.064	-0.273	-0.299
	(0.233)	(0.153)	(1.977)	(1.423)
Partner Industry Experience	0.356	0.203	1.006	0.701
	(0.341)	(0.309)	(3.359)	(2.465)
Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Observations	11963.00	11568.00	11963.00	11568.00
R^2	0.69	0.65	0.62	0.56

Table IC2: **Association between lead partner's human capital breadth index and startup novelty – non-syndicated vs. syndicated investments.** This table reports the results of OLS regressions of deal novelty on the lead partner's human capital breadth, separately for non-syndicated and syndicated investments. The dependent variable in columns (1)–(2) is 100 × *Novelty (Distance to Closest Firm)*, defined as 100 times one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. The dependent variable in columns (3)–(4) is 100 × *Above Median Novelty*, an indicator equal to 100 if the deal's novelty is above the sample median, and 0 otherwise. The main independent variable, *Breadth Index*, is defined as the first principal component from a PCA of four human capital breadth proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1) and (3) restrict the sample to non-syndicated investments (VC firms investing as sole investors), while columns (2) and (4) restrict the sample to syndicated investments (VC firms investing alongside other VCs). All columns include Industry × Deal Year × Deal Type × Country FE and VC Firm × Deal Year FE, as well as partner-level controls: age, sex, and ethnicity. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	100 × Novelty (Distance to Closest Firm)			
	(1)	(2)	(3)	(4)
Job Cat. Ratio	0.866**			
	(0.343)			
Job Role. Ratio		0.437		
		(0.417)		
Job Ind. Ratio			0.594*	
			(0.351)	
Educ. breadth				0.109
				(0.091)
Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Observations	23531.00	23531.00	23531.00	23531.00
R^2	0.62	0.62	0.62	0.62

Table IC3: **Association between individual human capital breadth measures and startup novelty.** This table reports the results of OLS regressions of deal novelty on the four individual components of the human capital breadth index. The dependent variable in columns (1)–(4) is 100 × *Novelty (Distance to Closest Firm)*, defined as 100 times one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. The independent variables are the four proxies used to construct the breadth index: (i) *Job Category Ratio*, the ratio of distinct job categories to total employment spells; (ii) *Job Role Ratio*, the ratio of distinct job roles to total employment spells; (iii) *Job Industry Ratio*, the ratio of distinct industries worked in to total employment spells; and (iv) *Educational Breadth*, a count of distinct educational fields. All columns include Industry × Deal Year × Deal Type × Country FE and VC Firm × Deal Year FE, as well as partner-level controls: age, sex, and ethnicity. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	100 × Novelty (Distance to Closest Firm)		100 × Above Median Novelty	
	(Early-Deals)	(Late-Deals)	(Early-Deals)	(Late-Deals)
Breadth Index	0.547**	0.194	3.981**	0.187
	(0.240)	(0.243)	(1.704)	(2.104)
VC Experience	-0.015	-0.129	-0.361	2.500
	(0.199)	(0.249)	(1.545)	(2.522)
Partner Industry Experience	-0.111	0.060	-2.220	4.819
	(0.333)	(0.363)	(2.717)	(3.022)
Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Observations	14702.00	8829.00	14702.00	8829.00
R^2	0.67	0.62	0.60	0.52

Table IC4: **Association between lead partner's human capital breadth index and startup novelty – early vs. late deals.** This table reports the results of OLS regressions of deal novelty on the lead partner's human capital breadth, separately for early and late deals relative to the partner's entry into the VC industry. The dependent variable in columns (1)–(2) is 100 × *Novelty (Distance to Closest Firm)*, defined as 100 times one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. The dependent variable in columns (3)–(4) is an indicator equal to 1 if the deal has above-median novelty. *Breadth Index* is the first principal component from a PCA of four human capital breadth proxies: (i) ratio of distinct job categories to total employment spells, (ii) ratio of distinct job roles to total employment spells, (iii) ratio of distinct industries worked in to total employment spells, and (iv) educational breadth count. *VC Experience* is the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed startup. Early deals are defined as those made within five years of the partner's VC industry entry; late deals are those made five or more years after entry. All columns include Industry × Deal Year × Deal Type × Country FE and VC Firm × Deal Year FE. All regressions also control for partner age, sex, and ethnicity. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success		
	(1)	(2)	(3)
Novelty (Avg. Distance to Closest Firm)	0.432***	0.452***	0.365***
	(0.081)	(0.082)	(0.086)
Breadth Index	-0.024	-0.041**	-0.062***
	(0.019)	(0.019)	(0.022)
Novelty (Avg. Distance to Closest Firm) × Breadth Index	0.120*	0.162**	0.232***
	(0.073)	(0.073)	(0.082)
VC Experience	0.004	0.003	0.015
	(0.004)	(0.005)	(0.012)
Partner Industry Experience	0.014*	-0.007	0.006
	(0.007)	(0.011)	(0.012)
Controls	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓
VC Firm × Deal Year FE		✓	
VC Firm × Deal Year × Partner Entry Year FE			✓
Observations	23531.00	23531.00	23531.00
R^2	0.41	0.55	0.64

Table IC5: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success — robustness to an alternative novelty measure.** The dependent variable in columns (1)–(3) is *Major Success*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least five times greater than the total VC invested capital. *Novelty (Avg. Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the average cosine similarity between the textual description of the financed startup and the five most similar VC-financed startups within the prior five years in the same deal stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Novelty (Avg. Distance to Closest Firm)* × *Breadth Index* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success	
	(Lead)	(Non-Lead)
Novelty (Distance to Closest Firm)	0.278**	0.434***
	(0.139)	(0.121)
Breadth Index	-0.058*	-0.050**
	(0.030)	(0.025)
Novelty (Distance to Closest Firm) × Breadth Index	0.237*	0.223**
	(0.134)	(0.099)
VC Experience	-0.010	0.007
	(0.006)	(0.009)
Partner Industry Experience	-0.001	-0.009
	(0.020)	(0.020)
Controls	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓
VC Firm × Deal Year FE	✓	✓
Observations	9519.00	14012.00
R^2	0.64	0.61

Table IC6: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success — robustness to lead and non-lead investments.** The dependent variable in columns (1)–(2) is *Major Success*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least five times greater than the total VC invested capital. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Breadth Index × Novelty (Distance to Closest Firm)* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Column (1) reports lead investments; column (2) reports non-lead investments. Both columns include Industry × Deal Year × Deal Type × Country FE and VC Firm × Deal Year FE, as well as individual-level partner controls (age, sex, ethnicity). Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success			
	(1)	(2)	(3)	(4)
Job Cat. Ratio	-0.099* (0.059)			
Novelty (Distance to Closest Firm) × Job Cat. Ratio	0.430* (0.247)			
Job Role. Ratio		-0.125 (0.077)		
Novelty (Distance to Closest Firm) × Job Role. Ratio		0.665** (0.330)		
Job Ind. Ratio			-0.078 (0.061)	
Novelty (Distance to Closest Firm) × Job Ind. Ratio			0.294 (0.256)	
Educ. breadth				-0.038** (0.018)
Novelty (Distance to Closest Firm) × Educ. breadth				0.145* (0.078)
Controls	✓	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Observations	23531.00	23531.00	23531.00	23531.00
R^2	0.55	0.55	0.55	0.55

Table IC7: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success — robustness to individual human capital breadth measures.** The dependent variable in columns (1)–(4) is *Major Success*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least five times greater than the total VC invested capital. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. The independent variable in column (1) is the ratio of distinct job categories to total employment spells. The independent variable in column (2) is the ratio of distinct job roles to total employment spells. The independent variable in column (3) is the ratio of distinct industries worked in to total employment spells. The independent variable in column (4) is the educational breadth count. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(4) include Industry × Deal Year × Deal Type × Country FE and VC Firm × Deal Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success		
	(1)	(2)	(3)
Breadth Index	0.007*	0.004	-0.007
	(0.004)	(0.005)	(0.014)
VC Experience	0.005	0.004	-0.004
	(0.005)	(0.004)	(0.014)
Partner Industry Experience	0.008	-0.009	-0.002
	(0.009)	(0.012)	(0.012)
Controls	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓
VC Firm × Deal Year FE		✓	
VC Firm × Deal Year × Partner Entry Year FE			✓
Observations	36459.00	36459.00	36459.00
R^2	0.37	0.46	0.55

Table IC8: **Association between lead partner human capital breadth and startup success.** The dependent variable in columns (1)–(3) is *Major Success*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least five times greater than the total VC invested capital. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	IPO		
	(1)	(2)	(3)
Novelty (Distance to Closest Firm)	0.467***	0.543***	0.456***
	(0.062)	(0.065)	(0.064)
Breadth Index	-0.044***	-0.050***	-0.060***
	(0.014)	(0.016)	(0.019)
Novelty (Distance to Closest Firm) × Breadth Index	0.211***	0.229***	0.280***
	(0.064)	(0.072)	(0.082)
VC Experience	0.003	0.002	-0.003
	(0.003)	(0.003)	(0.009)
Partner Industry Experience	0.012***	-0.001	0.001
	(0.004)	(0.006)	(0.008)
Controls	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓
VC Firm × Deal Year FE		✓	
VC Firm × Deal Year × Partner Entry Year FE			✓
Observations	23531.00	23531.00	23531.00
R^2	0.52	0.61	0.70

Table IC9: **Association between the interaction of lead partner's human capital breadth and deal novelty and IPO success.** The dependent variable in columns (1)–(3) is *IPO*, an indicator equal to 1 if the startup exits via an initial public offering. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Breadth Index × Novelty (Distance to Closest Firm)* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success (3x)		
	(1)	(2)	(3)
Novelty (Distance to Closest Firm)	0.381***	0.379***	0.289***
	(0.085)	(0.086)	(0.092)
Breadth Index	-0.021	-0.051***	-0.062***
	(0.018)	(0.018)	(0.022)
Novelty (Distance to Closest Firm) × Breadth Index	0.136*	0.220***	0.283***
	(0.076)	(0.076)	(0.086)
VC Experience	0.009*	0.005	0.014
	(0.005)	(0.005)	(0.012)
Partner Industry Experience	0.020**	-0.001	0.016
	(0.008)	(0.012)	(0.013)
Controls	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓
VC Firm × Deal Year FE		✓	
VC Firm × Deal Year × Partner Entry Year FE			✓
Observations	23531.00	23531.00	23531.00
R^2	0.40	0.54	0.64

Table IC10: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success — robustness to an alternative success threshold (3x).** The dependent variable in columns (1)–(3) is *Major Success (3x)*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least three times greater than the total VC invested capital. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Breadth Index × Novelty (Distance to Closest Firm)* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

	Major Success (1x)		
	(1)	(2)	(3)
Novelty (Distance to Closest Firm)	0.252***	0.248***	0.137
	(0.091)	(0.095)	(0.101)
Breadth Index	-0.035*	-0.058***	-0.069***
	(0.019)	(0.020)	(0.025)
Novelty (Distance to Closest Firm) × Breadth Index	0.186**	0.252***	0.311***
	(0.079)	(0.079)	(0.090)
VC Experience	0.011*	0.005	-0.001
	(0.006)	(0.006)	(0.015)
Partner Industry Experience	0.021**	0.009	0.022
	(0.009)	(0.013)	(0.015)
Controls	✓	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓
VC Firm × Deal Year FE		✓	
VC Firm × Deal Year × Partner Entry Year FE			✓
Observations	23531.00	23531.00	23531.00
R^2	0.40	0.55	0.64

Table IC11: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup success — robustness to an alternative success threshold (1x).** The dependent variable in columns (1)–(3) is *Major Success (1x)*, an indicator equal to 1 if the startup goes public or is acquired at a valuation at least equal to the total VC invested capital. *Novelty (Distance to Closest Firm)* is a deal-level measure of novelty defined as one minus the maximum cosine similarity between the textual description of the financed startup and all startups receiving venture capital financing within five years before the deal in the same stage. *Breadth Index* is a measure of human capital breadth defined as the first principal component of a PCA from four individual proxies: (i) the ratio of distinct job categories to total employment spells, (ii) the ratio of distinct job roles to total employment spells, (iii) the ratio of distinct industries worked in to total employment spells, and (iv) an educational breadth count. *Breadth Index × Novelty (Distance to Closest Firm)* is the interaction between deal novelty and the lead partner's human capital breadth. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the lead partner prior to the current deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. Columns (1)–(3) include Industry × Deal Year × Deal Type × Country FE. Column (2) additionally includes VC Firm × Deal Year FE. Column (3) includes VC Firm × Deal Year × Partner Entry Year FE. Standard errors (in parentheses) are double clustered at the VC firm and company levels. * p<.10; ** p<.05; *** p<.01.

Internet Appendix D: Identification Additional Robustness Tables

	Partner Leads a Deal					
	(Full Sample)	(Lead)	(Non-Lead)	(Early-Stage)	(Early-Stage Lead)	(Early-Stage Non-Lead)
Busy IPO Partner	-3.823***	-4.817***	-3.701***	-5.938***	-6.809***	-5.506***
	(0.850)	(1.326)	(1.116)	(1.132)	(1.614)	(1.599)
VC Experience	7.478***	5.713***	8.546***	6.863***	5.189***	8.091***
	(0.116)	(0.194)	(0.145)	(0.156)	(0.244)	(0.205)
Partner Age	-0.093***	-0.095***	-0.094***	-0.078***	-0.042	-0.105***
	(0.014)	(0.023)	(0.018)	(0.018)	(0.028)	(0.025)
Partner Industry Experience	3.080***	2.873***	3.144***	3.127***	2.960***	3.173***
	(0.094)	(0.158)	(0.119)	(0.123)	(0.193)	(0.162)
Controls	✓	✓	✓	✓	✓	✓
Deal FE	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Observations	242751.00	91232.00	151519.00	148916.00	62693.00	86223.00
R^2	0.32	0.31	0.33	0.33	0.31	0.34

Table ID1: **Association between partner busyness and the likelihood of leading a deal — robustness using only IPO events to measure busyness.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Busy IPO Partner* is an indicator equal to 1 if the partner is involved in exiting a deal via an IPO in the time window $(-90, +90)$ days around the focal deal date. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the partner prior to the current deal. *Partner Age* is measured in years at the time of the deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. All columns include individual-level controls for sex and ethnicity, as well as Deal FE and VC Firm FE. Column (1) reports results for the full sample. Columns (2) and (3) split the full sample into lead and non-lead investments. Column (4) reports results for the restricted sample of early-stage novel investments. Columns (5) and (6) split this restricted sample into lead and non-lead investments. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Partner Leads a Deal					
	(Full Sample)	(Lead)	(Non-Lead)	(Early-Stage)	(Early-Stage Lead)	(Early-Stage Non-Lead)
Busy Partner	-1.385*	-2.829**	-0.883	-1.874*	-3.722**	-0.768
	(0.800)	(1.276)	(1.048)	(1.091)	(1.573)	(1.551)
VC Experience	7.438***	5.683***	8.498***	6.798***	5.146***	8.011***
	(0.116)	(0.194)	(0.146)	(0.156)	(0.245)	(0.206)
Partner Age	-0.094***	-0.096***	-0.094***	-0.079***	-0.043	-0.106***
	(0.014)	(0.023)	(0.018)	(0.018)	(0.028)	(0.025)
Partner Industry Experience	3.081***	2.873***	3.145***	3.128***	2.959***	3.176***
	(0.094)	(0.158)	(0.119)	(0.123)	(0.193)	(0.162)
Controls	✓	✓	✓	✓	✓	✓
Deal FE	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Observations	242751.00	91232.00	151519.00	148916.00	62693.00	86223.00
R^2	0.32	0.31	0.33	0.33	0.31	0.34

Table ID2: **Association between partner busyness and the likelihood of leading a deal — robustness to an alternative busyness window.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Busy Partner* is an indicator equal to 1 if the partner is involved in exiting a deal via an IPO or a high-value acquisition in the time window $(-60, +60)$ days around the focal deal date. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the partner prior to the current deal. *Partner Age* is measured in years at the time of the deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. All columns include individual-level controls for sex and ethnicity, as well as Deal FE and VC Firm FE. Column (1) reports results for the full sample. Columns (2) and (3) split the full sample into lead and non-lead investments. Column (4) reports results for the restricted sample of early-stage novel investments. Columns (5) and (6) split this restricted sample into lead and non-lead investments. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Partner Leads a Deal					
	(Full Sample)	(Lead)	(Non-Lead)	(Early-Stage)	(Early-Stage Lead)	(Early-Stage Non-Lead)
Busy IPO Partner	-3.946***	-4.426***	-4.222***	-5.897***	-6.923***	-5.513***
	(1.003)	(1.568)	(1.321)	(1.326)	(1.853)	(1.910)
VC Experience	7.457***	5.674***	8.533***	6.828***	5.146***	8.062***
	(0.115)	(0.193)	(0.145)	(0.155)	(0.243)	(0.204)
Partner Age	-0.093***	-0.095***	-0.094***	-0.078***	-0.042	-0.106***
	(0.014)	(0.023)	(0.018)	(0.018)	(0.028)	(0.025)
Partner Industry Experience	3.081***	2.878***	3.144***	3.128***	2.966***	3.172***
	(0.094)	(0.158)	(0.119)	(0.123)	(0.193)	(0.162)
Controls	✓	✓	✓	✓	✓	✓
Deal FE	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Observations	242751.00	91232.00	151519.00	148916.00	62693.00	86223.00
R^2	0.32	0.31	0.33	0.33	0.31	0.34

Table ID3: **Association between partner busyness and the likelihood of leading a deal — robustness using only IPO events and a narrower busyness window.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Busy IPO Partner* is an indicator equal to 1 if the partner is involved in exiting a deal via an IPO in the time window (−60, +60) days around the focal deal date. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the partner prior to the current deal. *Partner Age* is measured in years at the time of the deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. All columns include individual-level controls for sex and ethnicity, as well as Deal FE and VC Firm FE. Column (1) reports results for the full sample. Columns (2) and (3) split the full sample into lead and non-lead investments. Column (4) reports results for the restricted sample of early-stage novel investments. Columns (5) and (6) split this restricted sample into lead and non-lead investments. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Partner Leads a Deal					
	(Full Sample)	(Lead)	(Non-Lead)	(Early-Stage)	(Early-Stage Lead)	(Early-Stage Non-Lead)
Busy Partner	-1.461	-2.595*	-1.115	-1.756	-3.345*	-0.913
	(0.893)	(1.407)	(1.174)	(1.217)	(1.751)	(1.721)
VC Experience	7.431***	5.663***	8.497***	6.786***	5.119***	8.010***
	(0.116)	(0.194)	(0.145)	(0.156)	(0.244)	(0.205)
Partner Age	-0.094***	-0.096***	-0.094***	-0.078***	-0.043	-0.106***
	(0.014)	(0.023)	(0.018)	(0.018)	(0.028)	(0.025)
Partner Industry Experience	3.082***	2.876***	3.145***	3.130***	2.962***	3.177***
	(0.094)	(0.158)	(0.119)	(0.123)	(0.193)	(0.162)
Controls	✓	✓	✓	✓	✓	✓
Deal FE	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Observations	242751.00	91232.00	151519.00	148916.00	62693.00	86223.00
R^2	0.32	0.31	0.33	0.33	0.31	0.34

Table ID4: **Association between partner busyness and the likelihood of leading a deal — robustness using a narrower busyness window.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Busy Partner* is an indicator equal to 1 if the partner is involved in exiting a deal via an IPO or a high-value acquisition in the time window (−45, +45) days around the focal deal date. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the partner prior to the current deal. *Partner Age* is measured in years at the time of the deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. All specifications include controls for sex and ethnicity, as well as Deal FE and VC Firm FE. Column (1) reports results for the full sample. Columns (2) and (3) split the full sample into lead and non-lead investments. Column (4) reports results for the restricted sample of early-stage novel investments. Columns (5) and (6) split this restricted sample into lead and non-lead investments. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Partner Leads a Deal					
	(Full Sample)	(Lead)	(Non-Lead)	(Early-Stage)	(Early-Stage Lead)	(Early-Stage Non-Lead)
Busy IPO Partner	-3.937***	-4.151**	-4.291***	-5.375***	-5.770***	-5.558***
	(1.117)	(1.721)	(1.472)	(1.487)	(2.107)	(2.101)
VC Experience	7.444***	5.657***	8.520***	6.806***	5.114***	8.047***
	(0.115)	(0.193)	(0.145)	(0.155)	(0.242)	(0.204)
Partner Age	-0.093***	-0.095***	-0.094***	-0.078***	-0.042	-0.106***
	(0.014)	(0.023)	(0.018)	(0.018)	(0.028)	(0.025)
Partner Industry Experience	3.083***	2.879***	3.145***	3.131***	2.968***	3.175***
	(0.094)	(0.158)	(0.119)	(0.123)	(0.193)	(0.162)
Controls	✓	✓	✓	✓	✓	✓
Deal FE	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Observations	242751.00	91232.00	151519.00	148916.00	62693.00	86223.00
R^2	0.32	0.31	0.33	0.33	0.31	0.34

Table ID5: **Association between partner busyness and the likelihood of leading a deal — IPO events only, narrower busyness window.** The dependent variable in all columns is *Partner Leads a Deal*, an indicator equal to 1 if a partner leads a deal and 0 otherwise. *Busy IPO Partner* is an indicator equal to 1 if the partner is involved in exiting a deal via an IPO in the time window $(-45, +45)$ days around the focal deal date. *VC Experience* is defined as the logarithm of one plus the number of deals financed by the partner prior to the current deal. *Partner Age* is measured in years at the time of the deal. *Partner Industry Experience* is a dummy equal to 1 if the partner has had at least one job in the industry of the financed venture, and 0 otherwise. All specifications include controls for sex and ethnicity, as well as Deal FE and VC Firm FE. Column (1) reports results for the full sample. Columns (2) and (3) split the full sample into lead and non-lead investments. Column (4) reports results for the restricted sample of early-stage novel investments. Columns (5) and (6) split this restricted sample into lead and non-lead investments. Standard errors (in parentheses) are clustered at the deal level. * p<.10; ** p<.05; *** p<.01.

	Breadth Index	Breadth Index × Novelty (Distance to Closest Firm)	IPO	
	(IV-First Stage)	(IV-First Stage)	(IV-Second Stage)	(OLS)
Avg. Available Breadth	0.826***	-0.018		
	(0.058)	(0.017)		
Avg. Available Breadth × Novelty (Distance to Closest Firm)	-0.565**	0.735***		
	(0.261)	(0.095)		
Breadth Index			-0.132**	-0.145***
			(0.059)	(0.035)
Breadth Index × Novelty (Distance to Closest Firm)			0.643**	0.701***
			(0.276)	(0.155)
Novelty (Distance to Closest Firm)			0.544*	0.484**
			(0.288)	(0.236)
Controls	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓	✓
Observations	1226.00	1226.00	1226.00	1226.00
R^2	0.87	0.86	0.19	0.65
F-statistic of Instrument			58.79	

Table ID6: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup success — robustness using IPO as the success proxy.** This table reports the results of an instrumental variable regression of startup outcomes on lead partner human capital breadth and its interaction with deal novelty. The dependent variable in Columns (3)–(4) is *IPO*, an indicator equal to 1 if the startup exits via an initial public offering. The endogenous regressors are the lead partner's *Breadth Index*, constructed as the first principal component of four human capital breadth measures (job category ratio, job role ratio, job industry ratio, and educational breadth count), and the interaction *Breadth Index × Novelty (Distance to Closest Firm)*. *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the focal startup's description and all startups financed in the past five years in the same stage. The instruments are (i) *Avg. Available Breadth*, defined as the average breadth index across all partners at the VC firm who are not busy with a high-value exit event (IPO or acquisition) within $(-90, +90)$ days of the focal deal, and (ii) *Avg. Available Breadth × Novelty (Distance to Closest Firm)*. Column (1) reports the first stage regression for the Breadth Index. Column (2) reports the first stage regression for the interaction Breadth Index × Novelty. Column (3) reports the IV second stage estimates. Column (4) reports OLS estimates in the same sample. All specifications include VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors, reported in parentheses, are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

	Breadth Index	Breadth Index × Novelty (Distance to Closest Firm)	Major Success	
	(IV-First Stage)	(IV-First Stage)	(IV-Second Stage)	(OLS)
Avg. Available Breadth	0.873*** (0.081)	-0.013 (0.023)		
Avg. Available Breadth × Novelty (Distance to Closest Firm)	-0.338* (0.182)	0.836*** (0.064)		
Total Breadth	-0.003 (0.002)	-0.001 (0.001)	0.002 (0.001)	0.001 (0.001)
Breadth Index			-0.185*** (0.068)	-0.161*** (0.040)
Breadth Index × Novelty (Distance to Closest Firm)			0.707** (0.318)	0.726*** (0.178)
Novelty (Distance to Closest Firm)			0.517 (0.417)	0.484* (0.264)
Controls	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓	✓
Observations	1232.00	1232.00	1232.00	1232.00
R^2	0.86	0.86	0.09	0.54
F-statistic of Instrument			58.38	

Table ID7: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup outcomes — robustness controlling for total breadth.** This table reports the results of an instrumental variable regression of startup outcomes on lead partner human capital breadth and its interaction with deal novelty. The dependent variable in Columns (3)–(4) is *Major Success*, an indicator equal to 1 if the startup exits via an IPO or an acquisition valued at least five times greater than the total VC capital raised by the company. The endogenous regressors are the lead partner's *Breadth Index*, constructed as the first principal component of four human capital breadth measures (job category ratio, job role ratio, job industry ratio, and educational breadth count), and the interaction *Breadth Index × Novelty (Distance to Closest Firm)*. *Novelty (Distance to Closest Firm)* is defined as one minus the maximum cosine similarity between the focal startup's description and all startups financed in the past five years in the same stage. The instruments are (i) *Avg. Available Breadth*, defined as the average breadth index across all partners at the VC firm who are not busy with a high-value exit event (IPO or acquisition) within $(-90, +90)$ days of the focal deal, and (ii) *Avg. Available Breadth × Novelty (Distance to Closest Firm)*. In addition, all specifications control for *Total Breadth*, defined as the sum of the breadth indices across all partners at the VC firm. Column (1) reports the first stage regression for the Breadth Index. Column (2) reports the first stage regression for the interaction Breadth Index × Novelty. Column (3) reports the IV second stage estimates. Column (4) reports OLS estimates in the same sample. All specifications include VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors, reported in parentheses, are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

	Major Success		
	(Full Sample)	(Lead)	(Non-Lead)
Breadth Index	0.025 (0.021)	0.027 (0.036)	-0.004 (0.047)
Controls	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓
Observations	1232.00	531.00	701.00
R^2	0.01	0.03	0.02
F-statistic of Instrument	352.97	157.34	137.13

Table ID8: **Effect of lead partner's human capital breadth on startup success (IV without novelty interaction).** This table reports the results of an instrumental variable regression of startup performance on the lead partner's human capital breadth. The dependent variable Major Success is an indicator that equals 1 if the startup exits via an IPO or an acquisition at least five times greater than the total VC invested capital. The independent variable Breadth Index is a measure of human capital breadth defined as the first principal component of four proxies for career diversity: (1) the ratio of distinct job categories to total employment spells, (2) the ratio of distinct job roles to total employment spells, (3) the ratio of distinct industries worked in to total employment spells, and (4) an educational breadth count. Breadth Index is instrumented using Avg. Available Breadth, defined as the average breadth index across all partners at the VC firm who are not busy with a high-value exit event in the window $(-90, 90)$ days around the focal deal. Columns (1)–(3) present IV estimates for the full sample, the subsample of lead deals, and the subsample of non-lead deals, respectively. All specifications include controls for partner age, sex, ethnicity, log tenure, and industry experience, as well as VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors, reported in parentheses, are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

	Breadth Index	Breadth Index × Novelty (Distance to Closest Firm)	Major Success	
	(IV-First Stage)	(IV-First Stage)	(IV-Second Stage)	(OLS)
Avg. Available Breadth	0.862*** (0.069)	-0.004 (0.023)		
Avg. Available Breadth × Novelty (Distance to Closest Firm)	-0.692** (0.310)	0.682*** (0.120)		
Breadth Index			-0.097 (0.087)	-0.121* (0.069)
Breadth Index × Novelty (Distance to Closest Firm)			0.343 (0.420)	0.575** (0.285)
Novelty (Distance to Closest Firm)			1.079** (0.530)	0.839** (0.410)
Controls	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓	✓
Observations	809.00	809.00	809.00	809.00
R^2	0.88	0.87	0.09	0.57
F-statistic of Instrument			18.72	

Table ID9: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup outcomes – robustness to alternative busyness window.** This table reports the results of an instrumental variable regression of deal performance on the interaction between lead partner breadth and deal novelty. Column (1) presents the first-stage regression of the instrumented variable Breadth Index on the two instruments Avg. Available Breadth and Avg. Available Breadth × Novelty, where Avg. Available Breadth is the average breadth index across all partners at the VC firm who are not busy with a high-value exit in the window $(-60, 60)$ days around the focal deal. Column (2) presents the first-stage regression of the second instrumented variable Breadth Index × Novelty on the same two instruments. In columns (3) and (4) the dependent variable is Major Success, an indicator equal to 1 if the startup exits via an IPO or an acquisition at least five times greater than the total VC capital raised by the company. Column (3) presents the IV estimates, and column (4) presents the OLS estimates for the same sample. All specifications include controls for partner age, sex, ethnicity, log tenure, and industry experience, as well as VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors, reported in parentheses, are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

	Breadth Index	Breadth Index × Novelty (Distance to Closest Firm)	IPO	
	(IV-First Stage)	(IV-First Stage)	(IV-Second Stage)	(OLS)
Avg. Available Breadth	0.862*** (0.069)	-0.004 (0.023)		
Avg. Available Breadth × Novelty (Distance to Closest Firm)	-0.692** (0.310)	0.682*** (0.120)		
Breadth Index			-0.120 (0.082)	-0.172*** (0.057)
Breadth Index × Novelty (Distance to Closest Firm)			0.477 (0.389)	0.789*** (0.237)
Novelty (Distance to Closest Firm)			0.894* (0.498)	0.582 (0.371)
Controls	✓	✓	✓	✓
VC Firm × Deal Year FE	✓	✓	✓	✓
Deal Stage × Industry × Year × Country FE	✓	✓	✓	✓
Observations	809.00	809.00	809.00	809.00
R^2	0.88	0.87	0.18	0.66
F-statistic of Instrument			18.72	

Table ID10: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup outcomes – robustness to alternative busyness window and IPO as success proxy.** This table reports the results of an instrumental variable regression of deal performance on the interaction between lead partner breadth and deal novelty. Column (1) presents the first-stage regression of the instrumented variable Breadth Index on the two instruments Avg. Available Breadth and Avg. Available Breadth × Novelty, where Avg. Available Breadth is the average breadth index across all partners at the VC firm who are not busy with a high-value exit in the window $(-60, 60)$ days around the focal deal. Column (2) presents the first-stage regression of the second instrumented variable Breadth Index × Novelty on the same two instruments. In columns (3) and (4) the dependent variable is IPO, an indicator equal to 1 if the startup exits via an IPO. Column (3) presents the IV estimates, and column (4) presents the OLS estimates for the same sample. All specifications include controls for partner age, sex, ethnicity, log tenure, and industry experience, as well as VC Firm × Deal Year fixed effects and Deal Stage × Industry × Year × Country fixed effects. Standard errors, reported in parentheses, are double clustered at the VC firm and company level. * p<.10; ** p<.05; *** p<.01.

	Net Multiple					
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of IPO Exits	0.487			0.680*		
	(0.334)			(0.373)		
Fraction of Major Success exits		0.717**			0.899**	
		(0.298)			(0.348)	
Fraction of Acquisition Exits			-0.037			-0.049
			(0.201)			(0.236)
Fund Size	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Industry Controls	✓	✓	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓	✓	✓
VC Firm FE	✓	✓	✓	✓	✓	✓
Vintage Year FE	✓	✓	✓			
Fund Type FE	✓	✓	✓			
Vintage Year × FundType FE				✓	✓	✓
Observations	1458.00	1458.00	1458.00	1458.00	1458.00	1458.00
R^2	0.52	0.52	0.52	0.53	0.53	0.53

Table ID11: **Validation of Exit Measure Used: PitchBook VC Funds with Performance Data from Preqin.** This table reports the results of regressions of realized VC fund performance on the fraction of IPO, Major Success, and Acquisition exits. The dependent variable in columns (1)–(6) is the Net Multiple of a fund, defined as the ratio of the total value of distributions plus unrealized gains to the total capital invested. In columns (1) and (4), Fraction of IPO Exits is defined as the share of a fund's deals that exit via an IPO. In columns (2) and (5), Fraction of Major Success Exits is defined as the share of a fund's deals that exit via either an IPO or an acquisition valued at more than five times the invested capital. In columns (3) and (6), Fraction of Acquisition Exits is defined as the share of a fund's deals that exit via acquisition. Fund Size is the committed size of the fund. All regressions include Industry Controls (capturing the fund's industry composition) and Stage Controls (capturing the fund's stage composition). In columns (1)–(3), Vintage Year FE and Fund Type FE are included. In columns (4)–(6), Vintage Year × Fund Type FE are included. VC Firm FE are included in all specifications. Standard errors, reported in parentheses, are clustered at the VC firm level. * p<.10; ** p<.05; *** p<.01.

Internet Appendix E: Proof from main text

A brief note on notation. Most of the expression in the interned appendix are functions of both e and Λ. When I take a partial derivative of objects with respect to e, I will use a $'$ notation so $\alpha(e)'$ denotes the partial derivative of the true positive rate with respect to effort. When I take a partial derivative with respect to Λ. I use $\alpha_\Lambda(e)$ notation.

Proof of Lemma 1

Proof. We have $s \sim N(\theta, \frac{1}{e})$. Therefore:

$$f(s|\theta = d) = \sqrt{\frac{e}{2\pi}} \exp\left(-\frac{1}{2}e(s-d)^2\right) \tag{27}$$

Next we explicitly calculate the likelihood ratio:

$$\frac{f(s|\theta = 1)}{f(s|\theta = 0)} = \exp\left(\frac{1}{2}e(2s-1)\right) \tag{28}$$

Applying Bayes rule:

$$P(\theta = 1|s) = \frac{f(s|\theta = 1)\pi}{f(s)} \tag{29}$$

Therefore:

$$\frac{P(\theta = 1|s)}{P(\theta = 0|s)} = \frac{\pi}{1-\pi} \exp\left(\frac{1}{2}e(2s-1)\right) \tag{30}$$

Partner invests in unknown project pool if and only if:

$$P(\theta = 1|s)R_h + P(\theta = 0|s)R_l \geq R_k \tag{31}$$

The last inequality can be rewritten as

$$P(\theta = 1|s)R_h + (1 - P(\theta = 0|s))R_l \geq R_k \tag{32}$$

The last equation is equivalent to:

$$P(\theta = 1|s) \geq \bar{p}, \tag{33}$$

where $\bar{p} := \frac{R_k - R_l}{R_h - R_l}$. Therefore the minimal threshold signal is defined by:

$$\frac{P(\theta = 1|s)}{P(\theta = 0|s)} = \frac{\pi}{1-\pi} \exp\left(\frac{1}{2}e(2s-1)\right) \geq \frac{\bar{p}}{1-\bar{p}} \tag{34}$$

We therefore have in log odds

$$ln(\frac{\pi}{1-\pi}) + \frac{1}{2}e(2s^* - 1) = ln(\frac{R_k - R_l}{R_h - R_k}) \tag{35}$$

Hence the cut-off signal is:

$$s^*(e) = \frac{1}{2} + \frac{\Lambda}{e} \tag{36}$$

where

$$\Lambda = ln\left(\frac{(R_k - R_l)(1-\pi)}{(R_h - R_k)\pi}\right) \tag{37}$$

□

Proof of Proposition 1 - Optimal Effort

Proof. By differentiating $U(e)$ with respect to e we obtain the FOC condition. We need to show that optimal effort e^* defined by (44) is indeed a maximum. Given the expressions for the true positive ((141)) and the false positive rate ((142)) we first define:

$$z_\alpha = (s^* - 1)\sqrt{e} = \frac{\Lambda}{e^{1/2}} - \frac{e^{1/2}}{2} \tag{38}$$

$$z_\beta = (s^*)\sqrt{e} = \frac{\Lambda}{e^{1/2}} + \frac{e^{1/2}}{2} \tag{39}$$

Then:

$$\alpha(e) = 1 - \Phi(z_\alpha) \tag{40}$$

$$\beta(e) = 1 - \Phi(z_\beta) \tag{41}$$

Then:

$$\alpha'(e) = -\phi(z_\alpha)\frac{\partial z_\alpha}{\partial e} = \phi(z_\alpha)\frac{2\Lambda + e}{4e^{3/2}} > 0 \tag{42}$$

$$\beta'(e) = -\phi(z_\beta)\frac{\partial z_\beta}{\partial e} = \phi(z_\beta)\frac{2\Lambda - e}{4e^{3/2}} < 0 \tag{43}$$

Intuitively the true positive rate is increasing and the false positive rate is decreasing with a rise in informativeness e. For this to hold we must have in $e^* > 2\Lambda$ which must hold since for any e such that $e < 2\Lambda$ the signal threshold s^* rises above 1 and the true positive rate drops below half and this defines a local minimum.

Now we show that the second order condition is satisfied. We have to show that:

$$\left\{\pi\,(R_h - R_k)\,\alpha''(e) + (1-\pi)\,(R_l - R_k)\,\beta''(e)\right\}\; - \gamma\; \leq 0 \tag{44}$$

Define:

$$H(e) = \pi\,(R_h - R_k)\,\alpha''(e) + (1-\pi)\,(R_l - R_k)\,\beta''(e) \tag{45}$$

We will show that $H(e)$ is negative. First from the expressions for $\alpha'(e)$ and $\beta'(e)$ we have:

$$\alpha''(e) = \phi(z_\alpha)\left(A'(e) + z_\alpha A(e)^2\right), \tag{46}$$

where

$$A(e) = -z'_\alpha \tag{47}$$

Calculating we obtain:

$$\alpha''(e) = \phi(z_\alpha)\frac{(2\Lambda + e)^2(2\Lambda - e) - 4e(6\Lambda + e)}{32e^{7/2}}, \tag{48}$$

Now notice that since $e > 2\Lambda$ all of the terms in the numerator are negative hence the true positive rate is concave with respect to signal informativeness $\alpha(e)'' < 0$. A similar calculation for the false positive rate shows:

$$\beta''(e) = \phi(z_\beta)\frac{(2\Lambda + e)(2\Lambda - e)^2 + 4e(6\Lambda + e)}{32e^{7/2}}, \tag{49}$$

So we have $\beta''(e) > 0$

Since $\alpha''(e) < 0$, we have $\pi\,(R_h - R_k)\,\alpha''(e) < 0$ and since $\beta''(e) > 0$ we have $(1-\pi)\,(R_l - R_k)\,\beta''(e) < 0$ (since $R_l - R_k < 0$ hence, both terms in $H(e)$ are negative, therefore $H(e) < 0$. □

Proof of Corollary - Optimal Effort Comparative Statics

Proof. Define

$$F(\gamma, \pi, e^*) = \left\{(1-\epsilon)\pi\,(R_h - R_k)\,\alpha'(\tau) + (1-\pi)\,(R_l - R_k)\,\beta'(\tau)\right\}\kappa - \gamma\, e^\star = 0 \tag{50}$$

Implicitly differentiating we get:

$$\frac{\partial e^*}{\partial q} = -\frac{F_q}{F_e}, \tag{51}$$

where q is the parameter of the comparative statics and F_e and F_q denote the partial derivatives of F with respect to e and parameter q. Now from the second order condition we have $F_e < 0$ hence:

$$\operatorname{sgn}(\frac{\partial e^*}{\partial q}) = \operatorname{sgn}(F_q) \tag{52}$$

- Comparative statics of optimal effort with respect to γ.

  First by assumption $R_k(\gamma)' > 0$ We have:

$$\frac{\partial \Lambda}{\partial \gamma} = \frac{\partial \Lambda}{\partial R_k} R_k(\gamma)' = \frac{R_h - R_l}{(R_k - R_l)(R_h - R_k)} R_k(\gamma)' > 0 \tag{53}$$

  Denote:

$$Q(\tau, \lambda, \gamma) = \pi(R_h - R_k(\gamma))\alpha(\tau)' + (1-\pi)(R_l - R_k(\gamma))\beta(\tau)' \tag{54}$$

Then applying the product rule:

$$\frac{\partial Q}{\partial \gamma} = -R_k(\gamma)' T(\tau) + M(\tau, \gamma), \tag{55}$$

where:

$$T(\tau) = (1-\epsilon)\pi\alpha(\tau)' + (1-\pi)\beta(\tau)' \tag{56}$$

and

$$M(\tau, \gamma) = (1-\epsilon)\pi(R_h - R_k)\alpha(\tau)'_\Lambda \frac{\partial \Lambda}{\partial \gamma} + (1-\pi)(R_l - R_k)\beta(\tau)'_\Lambda \frac{\partial \Lambda}{\partial \gamma} \tag{57}$$

First we show that $M(\tau, \gamma) < 0$. Explicitly calculating:

$$\alpha(\tau)'_\Lambda = \phi(z_\alpha)\frac{4\tau - 4\Lambda^2 + \tau^2}{8\tau^{3/2}} \tag{58}$$

Similarly:

$$\beta(\tau)'_\Lambda = \phi(z_\beta)\frac{4\tau - 4\Lambda^2 + \tau^2}{8\tau^{3/2}} \tag{59}$$

So we obtain:

$$\alpha(\tau)'_\Lambda = \beta(\tau)'_\lambda e^\Lambda = \beta(\tau)_\lambda \frac{(1-\pi)(R_k - R_l)}{\pi(R_h - R_k)} \tag{60}$$

Plugging in the last equality in the expression for M we obtain:

$$M(\tau, \gamma) = -\frac{\partial \Lambda}{\partial \gamma}\beta'(\tau)_\Lambda (1-\pi)(R_k - R_l)\epsilon < 0 \tag{61}$$

. Hence:

$$F_\gamma = -R_k(\gamma)' T(\tau) + M(\tau, \gamma) - e^*, \tag{62}$$

Not clearly a sufficient (not necessary condition) for a negative F_γ is:

$$T(\tau) > 0 \tag{63}$$

One can show that a sufficient condition for this to be satisfied is:

$$R_k(2-\epsilon) \geq R_h + (1-\epsilon)R_l. \tag{64}$$

- Comparative statics of optimal effort with respect to π.

  We have:

$$F_\pi = (1-\epsilon)(R_h - R_k)\alpha(\tau)' + (R_l - R_k)\beta(\tau)' + \frac{\partial \Lambda}{\partial \pi}\left((1-\epsilon)\pi(R_h - R_k)\alpha(\tau)'_\Lambda + (1-\pi)(R_l - R_k)\beta(\tau)'_\Lambda\right) \tag{65}$$

  Now we have shown in proof of the comparative statics of optimal effort with respect to γ that:

$$\pi(R_h - R_k)\alpha(\tau)'_\Lambda + (1-\pi)(R_l - R_k)\beta(\tau)'_\Lambda < 0 \tag{66}$$

Hence the second term is positive since $\frac{\partial \Lambda}{\partial \pi} < 0$. Now since:

$$(1-\epsilon)(R_h - R_k)\alpha(\tau)' + (R_l - R_k)\beta(\tau)' > 0 \tag{67}$$

since both terms are positive ($R_l - R_k < 0$ and $\beta(\tau)' < 0$)

□

Proof of Proposition 2 - The likelihood of financing a novel project decreases with specialization

Proof. All of the expressions involving effort are evaluated at optimal effort e^*, for brevity I omit the superscript. I use $\exp(x)$ to denote the exponential operator. Explicitly calculating:

$$\frac{\partial L_N(\gamma)}{\partial \gamma} = \frac{\partial}{\partial \gamma}(\pi\alpha(e) + (1-\pi)\beta(e)) \tag{68}$$

Notice that $\alpha(e)$ and $\beta(e)$ depend on γ directly through optimal effort and through Λ which is a function of $R_k(\gamma)$. Computing the derivative we obtain:

$$\frac{\partial L_N(\gamma)}{\partial \gamma} = \frac{\partial e}{\partial \gamma}(\pi\alpha(e)' + (1-\pi)\beta(e)') + \frac{\partial \Lambda}{\partial \gamma}(\pi\alpha_\Lambda(e) + (1-\pi)\beta_\Lambda(e)) \tag{69}$$

Now the second term is clearly negative since $\frac{\partial \Lambda}{\partial \gamma} > 0$ and both $\alpha_\Lambda(e) < 0, \beta_\Lambda(e) < 0$. Hence a sufficient condition (not necessary condition) for negative derivative is:

$$\pi\alpha(e)' + (1-\pi)\beta(e)' > 0. \tag{70}$$

Explicitly calculating:

$$\pi\alpha(e)' + (1-\pi)\beta(e)' = \frac{\pi\phi(z_\alpha)(e+2\Lambda) - (1-\pi)\phi(z_\beta)(e-2\Lambda)}{4e^{\frac{3}{2}}} \tag{71}$$

Now using $\phi(z_\alpha) = \exp(\Lambda)\phi(z_\beta)$ and the functional form of Λ we have:

$$\pi\phi(z_\alpha)(e+2\Lambda) - (1-\pi)\phi(z_\beta)(e-2\Lambda) = (1-\pi)\frac{(R_k - R_l)(2\Lambda + e) - (R_h - R_k)(e - 2\Lambda)}{R_h - R_k} \tag{72}$$

So a sufficient condition is:

$$(R_k - R_l)(2\Lambda + e) - (R_h - R_k)(e - 2\Lambda) = 2R_k e - R_l(2\lambda + e) - R_h(e - 2\Lambda) \geq 0 \tag{73}$$

which is certainly satisfied when:

$$R_k \geq \frac{R_h + R_l}{2} \tag{74}$$

□

Proof of Proposition 3 - Sensitivity of the likelihood of financing novelty with respect to specialization with respect to π

Proof. We have:

$$\frac{\partial L_N(e^*(\gamma))}{\partial \gamma} = \frac{\partial e^*}{\partial \gamma}(\pi\alpha(e)' + (1-\pi)\beta(e)') + \frac{\partial \Lambda}{\partial \gamma}(\pi\alpha_\Lambda(e) + (1-\pi)\beta_\Lambda(e)) \tag{75}$$

Denote:

$$A_1 = \pi\alpha(e)' + (1-\pi)\beta(e)' > 0 \tag{76}$$

$$A_2 = \pi\alpha_\Lambda(e) + (1-\pi)\beta_\Lambda(e), \tag{77}$$

Intuitively A_1 is the marginal effect of increasing signal precision on the acceptance probability. We have:

$$\frac{\partial}{\partial \pi}\frac{\partial L_N(e^*(\gamma))}{\partial \gamma} = A_1\frac{\partial}{\partial \pi}(\frac{\partial e^*}{\partial \gamma}) + \frac{\partial e^*}{\partial \gamma}\frac{\partial A_1}{\partial \pi} + \frac{\partial \Lambda}{\partial \gamma}\frac{\partial A_2}{\partial \pi} \tag{78}$$

We have:

$$\frac{\partial A_2}{\partial \pi} = \alpha_\Lambda - \beta_\Lambda + \frac{\partial e^*}{\partial \pi}\left(\pi\alpha_\Lambda(e)' + (1-\pi)\beta_\Lambda(e)'\right) + \frac{\partial \Lambda}{\partial \pi}\left(\pi\alpha_{\Lambda\Lambda}(e) + (1-\pi)\beta_{\Lambda\Lambda}(e)\right), \tag{79}$$

Now:

$$\alpha_\Lambda - \beta_\Lambda = \frac{\phi(z_\beta) - \phi(z_\alpha)}{e^{\frac{1}{2}}} > 0 \tag{80}$$

Next we have:

$$\frac{\partial e^*}{\partial \pi}\left(\pi\alpha_\Lambda(e)' + (1-\pi)\beta_\Lambda(e)'\right) > 0, \tag{81}$$

since optimal effort increases with π and $\left(\pi\alpha_\Lambda(e)'+(1-\pi)\beta_\Lambda(e)'\right) > 0$ (by explicit calculation). We have:

$$\pi\alpha_\Lambda(e)'+(1-\pi)\beta_\Lambda(e)' = (1-\pi)\frac{R_h - R_l}{R_h - R_k}\beta_\Lambda(e)' = (1-\pi)\frac{R_h - R_l}{R_h - R_k}\phi(z_\beta)\frac{4e - 4\Lambda^2 + e^2}{8e^{\frac{3}{2}}} \tag{82}$$

We have:

$$\pi\alpha_{\Lambda\Lambda}(e) + (1-\pi)\beta_{\Lambda\Lambda}(e) = \pi\frac{z_\alpha\phi(z_\alpha)}{e} + (1-\pi)\frac{z_\beta\phi(z_\beta)}{e} \tag{83}$$

Simplifying and adding the last three terms we get:

$$\frac{\partial A_2}{\partial \pi} = \frac{\phi(z_\alpha)}{\tau^{3/2}} \cdot \frac{R_h - R_\ell}{R_k - R_\ell}\left[\frac{\tau\left(\pi + \frac{1}{2}\right) - \Lambda}{1-\pi} + \left(\frac{\partial e^*}{\partial \pi}\right)\pi\,\frac{e^2 + 4e - 4\Lambda^2}{8}\right]. \tag{84}$$

Hence

$$\frac{\partial \Lambda}{\partial \gamma}\frac{\partial A_2}{\partial \pi} = \frac{\phi(z_\alpha)}{e^{3/2}}\,\frac{(R_h - R_\ell)^2}{(R_k - R_\ell)^2(R_h - R_k)}\,R_k'(\gamma)\,\left[\frac{\tau\left(\pi + \frac{1}{2}\right) - \Lambda}{1 - \pi} + \left(\frac{\partial e^*}{\partial \pi}\right)\pi\,\frac{e^2 + 4e - 4\Lambda^2}{8}\right]. \tag{85}$$

Now since $\tau > 2\Lambda$ we have:

$$\frac{e(\pi + \frac{1}{2}) - \Lambda}{1 - \pi} \geq \frac{2\Lambda(\pi + \frac{1}{2}) - \Lambda}{1 - \pi} = \frac{2\Lambda\pi}{1 - \pi} > 0 \tag{86}$$

Therefore we have:

$$\frac{\partial \Lambda}{\partial \gamma}\frac{\partial A_2}{\partial \pi} > 0 \tag{87}$$

Next we have:

$$\frac{\partial A_1}{\partial \pi} = \alpha(e)' - \beta(e)' + \frac{\partial e^*}{\partial \pi}\Big(\pi\alpha(e)'' + (1 - \pi)\beta(e)''\Big) + \frac{\partial \Lambda}{\partial \pi}\Big(\pi\alpha_\Lambda(e)' + (1 - \pi)\beta_\Lambda(e)'\Big) \tag{88}$$

Now we have:

$$\frac{\partial \Lambda}{\partial \pi}\Big(\pi\alpha_\Lambda(e)' + (1 - \pi)\beta_\Lambda(e)'\Big) = -\,\frac{\phi(z_\alpha)}{8\,e^{3/2}}\left(e^2 + 4e - 4\Lambda^2\right)\frac{R_h - R_\ell}{(R_k - R_\ell)(1 - \pi)} \tag{89}$$

We also have:

$$(\pi\,\alpha'' + (1 - \pi)\,\beta'')\frac{\partial e}{\partial \pi} = \kappa\,\frac{\partial e^*}{\partial \pi}\,\frac{\phi(z_\alpha)}{16\,e^{7/2}}\,\pi\left[A_\alpha + \frac{R_h - R_k}{R_k - R_\ell}\,A_\beta\right], \tag{90}$$

where

$$A_\alpha = 8\Lambda^3 - 4\Lambda^2 e - 2\Lambda e^2 - 24\Lambda e + e^3 + 4e^2, \qquad A_\beta = 8\Lambda^3 + 4\Lambda^2 e - 2\Lambda e^2 - 24\Lambda e - e^3 - 4e^2. \tag{91}$$

and $A_\beta < 0$ if $e > 2\Lambda$. Finally:

$$\alpha' - \beta' = -\phi(z_\alpha)\left[\frac{\Lambda}{2\,e^{3/2}}\,\frac{(R_k - R_\ell)(1 - \pi) - (R_h - R_k)\pi}{(R_k - R_\ell)(1 - \pi)} + \frac{1}{4\,e^{1/2}}\,\frac{(R_k - R_\ell)(1 - \pi) + (R_h - R_k)\pi}{(R_k - R_\ell)(1 - \pi)}\right]. \tag{92}$$

Combining the previous three expressions we get:

$$\begin{aligned}\frac{\partial A_1}{\partial \pi} = \frac{\phi(z_\alpha)}{e^{3/2}}\Bigg\{ -\frac{1}{1 - \pi}\Bigg[&\frac{\Lambda}{2}\Big((1 - \pi) - Q\,\pi\Big) + \frac{e}{4}\Big((1 - \pi) + Q\,\pi\Big) \\ &+ \frac{e^2 + 4e - 4\Lambda^2}{8}\,(1 + Q)\Bigg] + \frac{\kappa\,\pi}{16\,e^2}\left(\frac{\partial e^*}{\partial \pi}\right)\Big(A_\alpha + Q\,A_\beta\Big)\Bigg\},\end{aligned} \tag{93}$$

where $Q = \frac{R_h - R_k}{R_k - R_l}$. Now explicitly we have:

$$1 - \pi - Q\pi = \frac{R_k - \pi R_h - (1-\pi)R_l}{R_k - R_l} > 0, \tag{94}$$

, since we are in the region where information acquisition is valuable. So a sufficient condition for the second term to be negative is:

$$A_\alpha + QA_\beta < 0, \tag{95}$$

Define $m = e/\Lambda > 2$. Then

$$A_\alpha + QA_\beta = \Lambda^3 \left[(m^3 + 2m^2 - 28m + 8) - Q(m^3 + 6m^2 + 20m - 8) \right].$$

Hence,

$$A_\alpha + QA_\beta < 0 \quad \Longleftrightarrow \quad Q > Q^*(m) := \frac{m^3 + 2m^2 - 28m + 8}{m^3 + 6m^2 + 20m - 8}.$$

If $Q > Q^*(m)$ holds then both expressions in the bracket are negative. Therefore $\frac{\partial A_1}{\partial \pi} < 0$. We this have:

$$\frac{\partial e^*}{\partial \gamma}\frac{\partial A_1}{\partial \pi} > 0 \tag{96}$$

Finally we have:

$$A_1 \frac{\partial}{\partial \pi}\frac{\partial e^*}{\partial \gamma} > 0, \tag{97}$$

which proves the claim.

□

Proof of Proposition 4 - The expected return in the novel sector decreases with specialization

Proof. All of the expressions are evaluated at optimal effort. We have:

$$E_N(\gamma) = \frac{\pi\alpha(e)R_h + (1-\pi)\beta(e)R_l}{L_N} \tag{98}$$

Denote:

$$R(\gamma) = \pi\alpha(e)R_h + (1-\pi)\beta(e)R_l \tag{99}$$

$$\frac{\partial E_N}{\partial \gamma} = \frac{\frac{\partial R}{\partial \gamma}L_N - \frac{\partial L_N}{\partial \gamma}R}{L_N^2} \tag{100}$$

Explicitly calculating we can show:

$$\frac{\partial R}{\partial \gamma}L_N - \frac{\partial L_N}{\partial \gamma}R = \pi(1-\pi)(R_h - R_l)\left(\frac{\partial e}{\partial \gamma}(\beta\alpha' - \beta'\alpha) + \frac{\partial \Lambda}{\partial \gamma}(\beta\alpha_\Lambda - \alpha\beta_\Lambda)\right) \tag{101}$$

Hence E_N decreases with γ if and only if:

$$\frac{\partial e}{\partial \gamma}(\beta\alpha' - \beta'\alpha) + \frac{\partial \Lambda}{\partial \gamma}(\beta\alpha_\Lambda - \alpha\beta_\Lambda) < 0 \tag{102}$$

The term:

$$\frac{\partial e}{\partial \gamma}(\beta\alpha' - \beta'\alpha) < 0 \tag{103}$$

is the loss of conditional return due to a reduction in α and an increase in β due to reducing effort. The term:

$$\frac{\partial \Lambda}{\partial \gamma}(\beta\alpha_\Lambda - \alpha\beta_\Lambda) > 0 \tag{104}$$

is the gain of conditional return due to an increase in the threshold s resulting from the increase in γ. The net effect on conditional return is negative iff the loss from effort reduction term is higher than the gain from setting a stricter threshold due to an increase in Λ. The simplest condition under which this is true is if $\frac{\partial \Lambda}{\partial \gamma}$ is low. This is equivalent to:

$$R_k(\gamma)' << \frac{(R_k - R_l)(R_h - R_k)}{R_h - R_l} \tag{105}$$

Intuitively condition (105) states that the marginal return of increasing specialization on fallback return should be much smaller than the project return differentials. □

Proof of Proposition 5 - Sensitivity of return conditional on financing novelty with respect to specialization with respect to π

We use the following notation:

$$X = \pi(1-\pi)(R_h - R_l)\left(\frac{\partial e}{\partial \gamma}(\beta\alpha' - \beta'\alpha) + \frac{\partial \Lambda}{\partial \gamma}(\beta\alpha_\Lambda - \alpha\beta_\Lambda)\right) < 0 \tag{106}$$

From the proof of the previous proposition we have:

$$\frac{\partial E_N}{\partial \gamma} = \frac{X}{L_N^2} \tag{107}$$

Then by the quotient rule:

$$\frac{\partial}{\partial \pi}\frac{\partial E_N}{\partial \gamma} = \frac{\frac{\partial X}{\partial \pi}L_N - 2X\frac{\partial L_N}{\partial \pi}}{L_N^3} \tag{108}$$

Now explicitly calculating:

$$\frac{\partial L_N}{\partial \pi} = \alpha - \beta + \frac{\partial e}{\partial \pi}\big(\pi\alpha' + (1-\pi)\beta'\big) + \frac{\partial \Lambda}{\partial \pi}\big(\pi\alpha_\Lambda + (1-\pi)\beta_\Lambda\big) > 0 \tag{109}$$

, since $\alpha > \beta$. $\pi\alpha' + (1-\pi)\beta' > 0$ and $\pi\alpha_\Lambda + (1-\pi)\beta_\Lambda < 0$, also intuitively the likelihood of accepting a novel project increases with the prior that the novel project is of high quality. This implies:

$$-2X\frac{\partial L_N}{\partial \pi} > 0 \tag{110}$$

Now since from the last proposition $\frac{\partial \Lambda}{\partial \pi} \approx 0$ and the other derivative terms are strictly finite we approximate X via:

$$X \approx \pi(1-\pi)(R_h - R_l)\frac{\partial e}{\partial \gamma}(\beta\alpha' - \beta'\alpha) \tag{111}$$

We have:

$$\begin{aligned}\frac{\partial X}{\partial \pi} &= (1-2\pi)(R_h - R_l)\,\frac{\partial e}{\partial \gamma}\,(\beta\alpha' - \beta'\alpha)\\ &+ \pi(1-\pi)(R_h - R_l)\left(\frac{\partial}{\partial \pi}\frac{\partial e}{\partial \gamma}\right)(\beta\alpha' - \beta'\alpha)\\ &+ \pi(1-\pi)(R_h - R_l)\,\frac{\partial e}{\partial \gamma}\,\frac{\partial}{\partial \pi}(\beta\alpha' - \beta'\alpha).\end{aligned} \tag{112}$$

The second term is clearly positive since the sensitivity of effort with respect to γ declines in π. Let $A = \beta\alpha' - \beta'\alpha$. We differentiate with respect to π and combine terms:

$$\frac{\partial A}{\partial \pi} = \frac{\partial e}{\partial \pi}(\beta\alpha'' - \alpha'\beta' - \alpha\beta'') + \frac{\partial \Lambda}{\partial \pi}(\beta\alpha'_\Lambda + \alpha'\beta_\Lambda - \beta'\alpha_\Lambda - \alpha\beta'_\Lambda) < 0 \tag{113}$$

so all of the term in the expression for $\frac{\partial X}{\partial \pi}$ positive provided $\pi \geq \frac{1}{2}$ which proves the claim.

Proof of Proposition 6 - Optimal Assignment rule

Proof. We have:

$$\frac{dV}{d\gamma} = \frac{\partial V}{\partial e}(e^*(\gamma), \gamma)\frac{de^*}{d\gamma} + \frac{\partial V}{\partial \gamma}(e^*(\gamma), \gamma) \tag{114}$$

By the envelope theorem and since the outside option of the partner does not depend on optimal effort we have $\frac{\partial V}{\partial e}(e^*(\gamma), \gamma) = 0$

Hence the first order condition reads:

$$\frac{\partial \Pi}{\partial \gamma} = \frac{1}{2}e^2 - u(\gamma)' \tag{115}$$

Now for the FOC to define a local maximum we must have:

$$\frac{\partial}{\partial \gamma}\left(\frac{\partial \Pi}{\partial \gamma} - \frac{1}{2}e^2 - u(\gamma)'\right) < 0 \tag{116}$$

Finally we have the FOC for optimal γ:

$$R_k(\gamma^*)' q^r = \frac{1}{2}e(\gamma^*)^2 + u'(\gamma) \tag{117}$$

$$\frac{\partial}{\partial \gamma}\left(R_k(\gamma^*)' q^r - \frac{1}{2}e(\gamma^*)^2 - u'(\gamma)\right) < 0 \tag{118}$$

This is equivalent to:

$$R''_k(\gamma)q^r - e^* \frac{\partial e^*}{\partial \gamma} - u''(\gamma) < 0 \tag{119}$$

If the outside option of the partners is such that $u''(\gamma) == 0$ γ defines a maximum iff:

$$|R''_k(\gamma)|q^r > e^* |\frac{\partial e^*}{\partial \gamma}| \tag{120}$$

□

Proof of Proposition 7 - Optimal Specialization decreases with π

Proof. Define:

$$J(N, \gamma^*, \pi, e^*) = R_k(\gamma^*)' q^r - \frac{1}{2}(e^*)^2 - u'(\gamma) = 0 \tag{121}$$

Then implicit differentiation again gives:

$$\frac{\partial \gamma}{\partial q} = -\frac{J_q}{J_\gamma} \tag{122}$$

From second order condition for a maximum $J_\gamma < 0$ hence:

$$\text{sgn}(\frac{\partial \gamma}{\partial q}) = \text{sgn}(J_q) \tag{123}$$

- Comparative statics of optimal specialization with respect to π. We have:

  $$J_\pi = R_k(\gamma)' \frac{\partial q^r}{\partial \pi} - e^* \frac{\partial e^*}{\partial \pi} \tag{124}$$

  , where

  $$\frac{\partial q^r}{\partial \pi} = -\alpha(\tau) + \beta(\tau) - \frac{\partial \Lambda}{\partial \pi}\left(\pi \alpha(\tau)_\Lambda + (1-\pi)\beta(\tau)_\Lambda\right) \tag{125}$$

  We have:

  $$\frac{\partial \Lambda}{\partial \pi} = -\frac{1}{\pi(1-\pi)} \tag{126}$$

  $$\alpha(\tau)_\Lambda = -\frac{\phi(z_\alpha)}{\tau^{1/2}} \tag{127}$$

  $$\beta(\tau)_\Lambda = -\frac{\phi(z_\beta)}{\tau^{1/2}} \tag{128}$$

  Hence the claim follows since $\frac{\partial q^r}{\partial \pi} < 0$

□

Proof of Theorem I - Existence of equilibrium

Proof. We prove existence via Brouwer's fixed-point theorem. Define the domain $\mathcal{D} \subset R^3$ as:

$$\mathcal{D}[\underline{\pi}, \overline{\pi}] \times [\gamma_{\min}, \gamma_{\max}] \times [0, e_{\max}] \tag{129}$$

where:

- $\pi \in [\underline{\pi}, \overline{\pi}] \subset (0, 1)$ (bounded away from 0 and 1)
- $\gamma \in [\gamma_{\min}, \gamma_{\max}]$ (specialization bounds)
- $e \in [0, e_{\max}]$ (effort bounded by cost)

$\mathcal{D}$ is **compact** and **convex** as a Cartesian product of compact convex intervals.

Define the mapping $T : \mathcal{D} \to \mathcal{D}$ by $T(\pi, \gamma, e) = (\pi', \gamma', e')$ where:

$$e' \arg\max_{e \geq 0} \left[R_k(\gamma) + (1 - \epsilon)\pi\alpha(\tau)(R_h - R_k(\gamma)) + (1 - \pi)\beta(\tau)(R_l - R_k(\gamma)) - \frac{\gamma e^2}{2} \right] \tag{Optimal Effort}$$

$$\gamma' \arg\max_{\gamma \geq 0} \left[\Pi(e', \gamma) - \frac{1}{2}\gamma(e')^2 - u(\gamma) \right] \tag{Optimal Hiring}$$

$$\pi' \frac{1 - \eta^h}{1 - \eta^l} \quad \text{with} \quad \tau' = \kappa e' \tag{Entry Update}$$

where η^l and η^h are computed using (γ', e').

Step 1: Continuity of T

- e' is continuous in (π, γ) by the Implicit Function Theorem applied to the FOC in Proposition 2, since the objective is strictly concave in e.
- γ' is continuous in (π, e') by the Implicit Function Theorem applied to the FOC in Proposition 5, given $R_k(\gamma)$ concave and $u(\gamma)$ differentiable.
- π' is continuous in (γ', e') because:
  - $\alpha(\tau)$ and $\beta(\tau)$ are smooth (normal CDF)
  - $R_k(\gamma)$ is differentiable
  - Entrepreneurial thresholds are rational functions

Thus, T is continuous on $\mathcal{D}$.

Step 2: Self-Mapping $T(\mathcal{D}) \subseteq \mathcal{D}$

- **Effort:** $e' \in [0, e_{\max}]$ since effort cost $\frac{1}{2}\gamma e^2 \to \infty$ as $e \to \infty$, and $e_{\max}$ bounds the solution.

- **Specialization:** $\gamma' \in [\gamma_{\min}, \gamma_{\max}]$ by partner market constraints.
- **Project Quality:** $\pi' \in [\underline{\pi}, \overline{\pi}]$ because:
  - $\eta^l \geq 0$ by $c_l \geq \beta(\tau)b$ (Condition 1)
  - $\eta^h \leq 1$ by $c_h \leq 1 - \lambda + c_l - \beta(\tau)b + \alpha(\tau)(\epsilon(R_h - R_k) + b)$ (Condition 2)
  - $\eta^h \geq \eta^l$ (Condition 3)

Step 3: Fixed Point Exists Since $\mathcal{D}$ is compact and convex, and $T : \mathcal{D} \to \mathcal{D}$ is continuous, by Brouwer's Fixed-Point Theorem, there exists $(\pi^*, \gamma^*, e^*) \in \mathcal{D}$ such that:

$$T(\pi^*, \gamma^*, e^*) = (\pi^*, \gamma^*, e^*)$$

This fixed point satisfies all equilibrium conditions by construction.

Parameter Restrictions:

1. $c_l \geq \beta(\tau)b$ (ensures $\eta^l \geq 0$)
2. $c_h \in \left[\frac{c_l - \beta(\tau)b}{\lambda} + \alpha(\tau)(\epsilon(R_h - R_k) + b), 1 - \lambda + c_l - \beta(\tau)b + \alpha(\tau)(\epsilon(R_h - R_k) + b)\right]$ (ensures $\eta^h \in [\eta^l, 1]$)
3. $R_k > \frac{R_h + R_l}{2}$ (for Proposition 3)
4. $R_h \geq |R_l|$ (for Proposition 4)
5. $(2 - \epsilon)R_k \geq R_h + (1 - \epsilon)R_l$ (for Corollary 1 and 2)

□

Internet Appendix F: Data Construction

Matching Pitchbook-Revelio Labs

To construct the link between PitchBook investors and Revelio Labs firms, I implement a structured multi-step matching procedure designed to maximize accuracy while minimizing false positives. The procedure begins with a systematic harmonization of identifiers. Specifically, I standardize organization names (PitchBook: InvestorName; Revelio: company) by removing legal suffixes such as Inc., Ltd., or GmbH, stripping bracketed numeric tags, collapsing multiple whitespaces, and converting all characters to lowercase. I also parse website information (PitchBook: Website; Revelio: url) by extracting the registrable domain (e.g., example.com) to reduce noise from URL extensions, and I harmonize location data (PitchBook: HQCountry; Revelio: `country_code`).[39]

Once identifiers are standardized, I carry out a sequence of exact-match passes that proceed in descending order of specificity. In the first pass, I match firms on the joint pair of (cleaned name, website domain), which provides the highest likelihood of uniquely identifying the same entity across datasets. In the second pass, I consider the remaining unmatched records and implement exact matches on the pair (cleaned name, country), conditional on non-missing location data. In the third pass, I restrict attention to unmatched firms with non-missing website information and implement exact matches on the pair (website domain, country). Finally, in the fourth pass, I allow exact matches on cleaned firm name alone. After each pass, I remove all successfully matched PitchBook records from the pool of candidates in order to avoid duplicate links. In cases where multiple potential matches arise, I retain only one-to-one links. The overall sample with non-missing partner IDs from Pitchbook contains 6346 distinct investors (distinct InvestorID). Out of these following this procedure I am able to match 4667 investors to a unique Revelio identifier (unique rcid) resulting in an overall one-to-one match rate of around 74%. For the one - to many matches (i.e., cases where one InvestorID from Pitchbook is matched to multiple rcid identifiers in Revelio) I do a manual check, based on detailed location data and keep only the correct matches. This results in the matching of additional 306 investors bringing the overall match rate to 78%. For the cases where one InvestorID is correctly matched to multiple rcids (e.g., Austin Ventures with an investor ID 10146-16 matched to Austin Ventures LLC and Austin Ventures LP with rcids 20152010 and 22143640 respectively) I keep both rcids as a correct match.

Next, I use the matched Revelio rcids and I collect the full history of employees for the matched rcids. To link individual partners in PitchBook to employees in Revelio Labs, I implement a name-based exact matching procedure conditional on the firm-level match established in the previous step. I begin by harmonizing all person names to ensure comparability across sources. Specifically, I normalize text encoding, split full names into first and last names using a structured parser, and then clean each component by removing legal suffixes, bracketed numeric

[39] Revelio Labs academic access does not provide the company's headquarters. I use the employment dataset and construct a proxy for headquarters based on the country where the majority of employees are based.

identifiers, and other suffixes and prefixes (e.g., Pitchbook names often contain a title such as PhD, MD or JD). I apply the same cleaning procedure symmetrically to both PitchBook partner names and Revelio employee names. I then merge the two datasets by requiring an exact match on three keys: the firm identifier carried over from the firm-level match (the Revelio rcid), the cleaned first name, and the cleaned last name. This conservative design ensures that a partner is linked to a Revelio record only if the firm affiliation and both name components match exactly. After the merge, I remove duplicates to retain a one-to-one mapping.

Industry Match Revelio - Pitcbhook

Because PitchBook and Revelio use different taxonomies, I align PitchBook's Industry Group labels to Revelio's RICS taxonomy using embedding-based cosine similarity. For each PitchBook group, I compute its cosine similarity to each of the 50 Revelio RICS categories and select the two categories with the highest scores as the mapped matches for that group. Table IF1 shows the mapping along with the raw cosine similarity scores.

PB industry group	RICS (top match)	Sim. (top)	RICS (second)	Sim. (second)
Other Financial Services	Financial Services	0.876	Business Services	0.639
Retail	Apparel Retail	0.823	Retail and Consumer Goods	0.804
Media	Media and Entertainment	0.810	Culture and Entertainment	0.550
Pharmaceuticals and Biotechnology	Pharmaceuticals	0.803	Biotech and Healthcare Services	0.713
Transportation	Logistics and Transportation	0.803	Automotive Services	0.479
Healthcare Services	Healthcare and Wellness Services	0.786	Business Services	0.687
Agriculture	Agricultural Services	0.776	Environmental Services	0.407
Apparel and Accessories	Apparel Retail	0.769	Retail and Consumer Goods	0.536
Commercial Services	Business Services	0.761	Marketing and Advertising Services	0.703
IT Services	Information Technology Services	0.760	IT Consulting Services	0.737
Containers and Packaging	Packaging Services	0.750	Logistics and Transportation	0.430
Other Business Products and Services	Business Services	0.744	Digital Commerce Services	0.593
Commercial Transportation	Logistics and Transportation	0.734	Commercial Aviation	0.595
Services (Non-Financial)	Financial Services	0.732	Business Services	0.662
Restaurants, Hotels and Leisure	Food and Hospitality Services	0.714	Hospitality and Tourism Management	0.662
Other Consumer Products and Services	Retail and Consumer Goods	0.690	Consumer Technology Distribution	0.652
Other Energy	Energy and Resources	0.684	Food and Beverage	0.299
Other Information Technology	Information Technology Services	0.664	IT Consulting Services	0.459
Energy Services	Environmental Services	0.643	Energy and Resources	0.592
Other Healthcare	Healthcare and Wellness Services	0.637	Biotech and Healthcare Services	0.484
Commercial Banks	Financial Services	0.617	Business Services	0.454
Consumer Durables	Retail and Consumer Goods	0.604	Consumer Technology Distribution	0.557
Other Materials	Materials Manufacturing	0.604	Miscellaneous	0.417
Energy Equipment	Energy and Resources	0.601	Electronics Manufacturing	0.390
Commercial Products	Retail and Consumer Goods	0.594	Marketing and Advertising Services	0.590
Healthcare Technology Systems	Biotech and Healthcare Services	0.585	Healthcare and Wellness Services	0.573
Computer Hardware	Electronics Manufacturing	0.577	Industrial Manufacturing	0.447
Consumer Non-Durables	Retail and Consumer Goods	0.572	Consumer Technology Distribution	0.541
Textiles	Materials Manufacturing	0.556	Apparel Retail	0.542
Communications and Networking	Telecommunications Services	0.530	Media and Entertainment	0.406
Healthcare Devices and Supplies	Healthcare and Wellness Services	0.523	Wellness Products	0.479
Semiconductors	Electronics Manufacturing	0.493	Materials Manufacturing	0.351
Construction (Non-Wood)	Engineering and Construction Services	0.491	Materials Manufacturing	0.397
Metals, Minerals and Mining	Energy and Resources	0.477	Materials Manufacturing	0.472
Software	Automation Solutions	0.475	Information Technology Services	0.438
Utilities	Energy and Resources	0.446	Environmental Services	0.437
Insurance	Financial Services	0.434	Legal Services	0.398
Capital Markets/Institutions	Financial Services	0.432	Professional and Trade Associations	0.412
Forestry	Agricultural Services	0.416	Environmental Services	0.403
Exploration, Production and Refining	Energy and Resources	0.404	Industrial Manufacturing	0.403
Chemicals and Gases	Pharmaceuticals	0.378	Environmental Services	0.307

Table IF1: Mapping of PitchBook industry groups to closest RICS categories with cosine similarity scores.

Internet Appendix G: Additional Results

6.1 Individual or team?

The results in the previous subsection highlighted the importance of individual VC partners' human capital breadth in financing novel ventures. In this subsection, I shift the focus to the team level and investigate whether diversity among partners at the fund level is associated with the financing of more novel ventures and with the successful exit of such ventures.

To conduct this analysis, I aggregate individual-level data to the fund level and construct several diversity indices, following a similar methodology to the construction of the human capital breadth index. These fund-level diversity measures capture heterogeneity in partners' prior professional experiences, education and gender composition. Specifically to capture heterogeneity in prior professional experience, I define three proxies: (i) job category diversity, which reflects variation in the functional roles held by partners; (ii) job industry diversity, which captures the range of industries in which partners have previously worked; and (iii) job role diversity, which measures variation in the specific roles held across employment spells. Each index is computed at the fund level. For example, the job role diversity index is defined as:

$$D_f = \frac{\text{Number of distinct roles}_f}{\text{Number of total employment spells}_f} \tag{130}$$

the numerator simply counts the total number of distinct roles held by all partners in the past and the denominator scales this measure by the number of total past employment spells of partners involved in the funds' deals.[40]

I also construct educational and gender diversity categories at a fund level, defined as follows:

$$D_f \ = \ \exp\left(-\sum_{i=1}^{K} \frac{n_{f,i}}{N_f} \ln\left(\frac{n_{f,i}}{N_f}\right)\right), \tag{131}$$

where N_f is the total number of distinct partners in the fund, $n_{f,i}$ is the number of partners in category i and K is the number of categories. For example, in constructing gender diversity $K = 2$ for Male and Female partners. Intuitively, (131) captures converts an entropy diversity measure into effective categories. The measure ranges between 1 and the number of categories, for example if the fund consists only of partners of the female gender the measure is 1 and if the fund is perfectly balanced i.e., 50% of partners are male and 50% of partners are female the effective gender diversity is 2.[41]

Analogously, to the individual level tests, I first test whether fund level diversity is associated with financing more novel ventures. Specifically, I estimate,

$$\text{Frac. Novel}_f = \alpha + \beta D_f + X_f + \rho_{t\times c} + \epsilon_f, \tag{132}$$

[40] This is a fund level analogue to the individual partner human capital breadth.

[41] To construct the effective educational diversity measure I use the following degree types: STEM degree, Social Science or Humanities degree, MBA degree, PhD degree, Medical Doctor degree.

where Frac. Novel$_f$ is the fraction of top quartile novelty firms financed by fund f, D_f is a diversity measure, X_f are fund level controls which include fund size, a dummy for first fund equal to 1 if it is the first fund raised by a given GP as well as controls for industry and stage allocation and $\rho_{t\times c}$ are vintage year $\times$ fund category fixed effects.

Table IG1 about here.

The results are presented in Table IG1. I do not find any economically nor statistically significant relationship between fund level diversity and the fraction of novel firms financed. Next, along similar lines I test whether fund level human capital diversity is associated with performance. I estimate the following specification:

$$\text{Frac. Successfull Exits}_f = \alpha + \beta_1 D_f + \beta_2 \text{Frac. Novel}_f + \beta_3 D_f \times \text{Frac. Novel}_f + X_f + \rho_{t\times c} + \epsilon_f, \tag{133}$$

where Frac. Successfull Exits$_f$ is the fraction of investments exited via an IPO or high value acquisition. The coefficients of interest are β_1, β_2 and β_3.

Table IG2 about here.

The results are presented in Table IG2. Analogous to the individual exit regressions, I find that the fraction of financed novel firms (β_2 estimate) are correlated with the fraction of successful exits, however, I do not find robust evidence that fund level diversity is associated with performance for both funds with a high and low fraction of financed novel firms.

In Appendix Tables IG3 and IG4, I use PCA to reduce dimensionality and construct summary measures to capture fund-level diversity from the individual measures. I do not find robust evidence that fund-level diversity is associated with financing a higher fraction of novel projects, nor do I find evidence that it is associated with stronger performance for novel firms.

6.2 The role of human capital breadth over time

In this subsection, I examine how the relationship between human capital breadth and novelty has evolved over time. The stylized facts point to a gradual decline in both average startup novelty - driven by a lower fraction of high novelty firms financed by the venture capital industry and human capital breadth.

To examine this trend, I re-estimate specification (5) by progressively excluding earlier cohorts of VC-backed deals, thereby focusing on more recent years.

Figure IG1 about here.

Figure IG1 presents the estimated coefficients on human capital breadth. Each point reflects the coefficient from specification (5), estimated on a sample restricted to deals financed after year t. The results indicate a gradual weakening in the relationship between human capital breadth and novelty over time, particularly beginning around 2015.

Next, I examine how the relationship between human capital breadth, novelty, and performance has changed by estimating specification (6) over time, again excluding earlier deals.

Figure IG2 about here.

Figure IG2 shows the estimated coefficients on the interaction between human capital breadth and novelty. As before, each point reflects the estimate from a sample restricted to deals completed after year t. Unlike the previous result, the interaction effect appears stable over time.

6.3 Venture sector heterogeneity

In this section, I present heterogeneity of the reduced form estimates by sector of the financed venture. PitchBook classifies ventures into 7 broad sectoral categories: Business Products and Services (B2B), Consumer Products and Services (B2C), Energy, Financial Services, Healthcare, Information Technology, Materials and Resources. The distribution of deals with positive novelty in each sector is presented in Table IG5. Since the distribution is quite skewed towards certain sectors (e.g., Information Technology), I combine the 7 categories into 3 broader categories (i) Information Technology + Financial Services; (ii) Business Products and Services (B2B) + Consumer Products and Services (B2C); (iii) Healthcare + Energy + Materials and Resources. I first re-estimate specification (5) separately for each of these 3 broader categories separately.

Figure IG3 about here.

The point estimates of β - the coefficient on lead partner breadth and the confidence intervals are presented in Figure IG3. First the coefficient on human capital breadth is positive across all sectors. Second there seems to be heterogeneity in both magnitude and statistical significance of the association across sectors. The magnitude and statistical significance is strongest for deals in Business and Consumer Product category and weaker in the Healthcare, Energy and Materials sector. This is consistent with the fact that presumably the Healthcare, Energy and Material sectors require a higher technical expertise from the partner to access the technological novelty of the startup. These requirements are presumably lower in the Consumer and Business Product and Services Sectors and the IT and Financial Services sectors where one would presumably be more interested in the venture's business potential where the breadth of the partner could arguably matter more.

Figure IG4 about here.

In Figure IG4, I re-estimate (6) and plot the coefficient on the lead partner - breadth novelty interaction for different sectors.

6.4 The role of on the job VC experience on financing novelty

Here, I examine the role of experience of partners acquired on the job i.e., experience acquired through investments made post VC industry entry on financing of novel startups. In particular, I examine the role of (i) Industry specialization post VC industry entry (ii) Deal diversity post VC industry entry. To do so, I construct a Herfindahl-Hirschmann industry specialization measure following Gompers et al. (2009) based on the industry sector of past deals financed by VC partners:

$$HHI_{p,t} = \sum_{j} (\frac{n_{p,j,t}}{N_{p,t}})^2, \tag{134}$$

where p denotes parter, t time, $n_{p,j,t}$ is the number of investments partner p has made from VC industry entry to time t in industry j and $N_{p,t}$ is the total number of investments made by partner p. Similarly, I construct an average deal diversity index defined as:

$$\text{Deal Diversity}_{p,t} = \frac{\sum_{k,j,k\neq j}(1 - CosSim(j,k))}{K_{p,t}}, \tag{135}$$

where k, j denote companies financed by partner p before time t, $CosSim(j, k)$ is the cosine similarity between business model of company j and company k and $K_{p,t}$ is the number of distinct companies financed before year t. I next test whether, partner industry specialization and past deal diversity are correlated with the novelty of the financed focal deal. I estimate:

$$N_{j,k,p,t} = \alpha + \beta_1 HHI_{p,t} + \beta_2 \text{Deal Diversity}_{p,t} + X_{t,p} + \eta_{i\times t\times s\times c} + \rho_{t\times j} + \epsilon_{j,k,p,t}, \tag{136}$$

where the coefficients of interests are β_1 and β_2. Model (136) is estimated on the full set of partners available in Pitchbook data i.e., in this specification I do not require partners to have a matched to Revelio data. Furthermore since for each partner I need to measure past deal diversity, I require partners to have made at least 2 past investments in distinct companies.

Table IG6 about here.

The results are reported in Table IG6. Columns (1) and (2) estimate equation (136) using the full sample, while columns (3) and (4) restrict the analysis to the subsample of investments where the VC firm acts as the lead investor. In columns (1) and (3), I find that both greater industry specialization by the partner after entering the VC industry and greater diversity in previously financed deals are positively associated with the novelty of the focal investment. However, once VC firm fixed effects are included, both the economic magnitude and statistical significance of these relationships decline. This suggests that much of the observed variation is driven by cross-firm differences in investment novelty and specialization, and that the within-firm variation in these characteristics has a more limited association with deal novelty.

Table IG7 about here.

In Table IG7, I re-estimate (5) by including industry specialization post VC entry as well as deal diversity of past financed deals and I show that (i) the baseline effect of human capital breadth remains robust and (ii) post VC entry industry specialization plays a role in selecting most novel firms, highlighting the role of post VC entry industry specialization (Gompers et al., 2009).

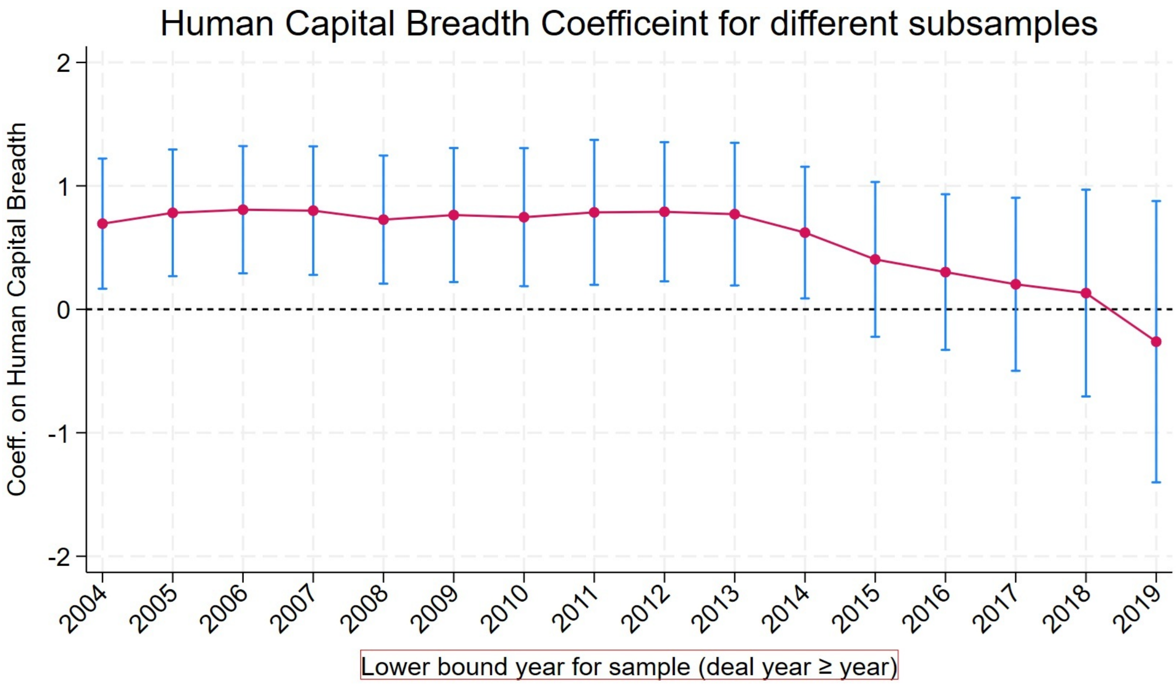

Figure IG1: **Association between Lead Partner Human Capital Breadth and Novelty over time** This figure presents the coefficient β of the association between human capital breadth and novelty estimated via $N_{j,k,p,t} = \alpha + \beta B_p + X_{t,p} + \eta_{i \times t \times s \times c} + \rho_{t \times j} + \epsilon_{j,k,p,t}$ by progressively excluding cohorts of earlier financed VC-deals. For example, the coefficient estimate plotted in year 2010, is the estimated β in the regression excluding deals done before 2010.

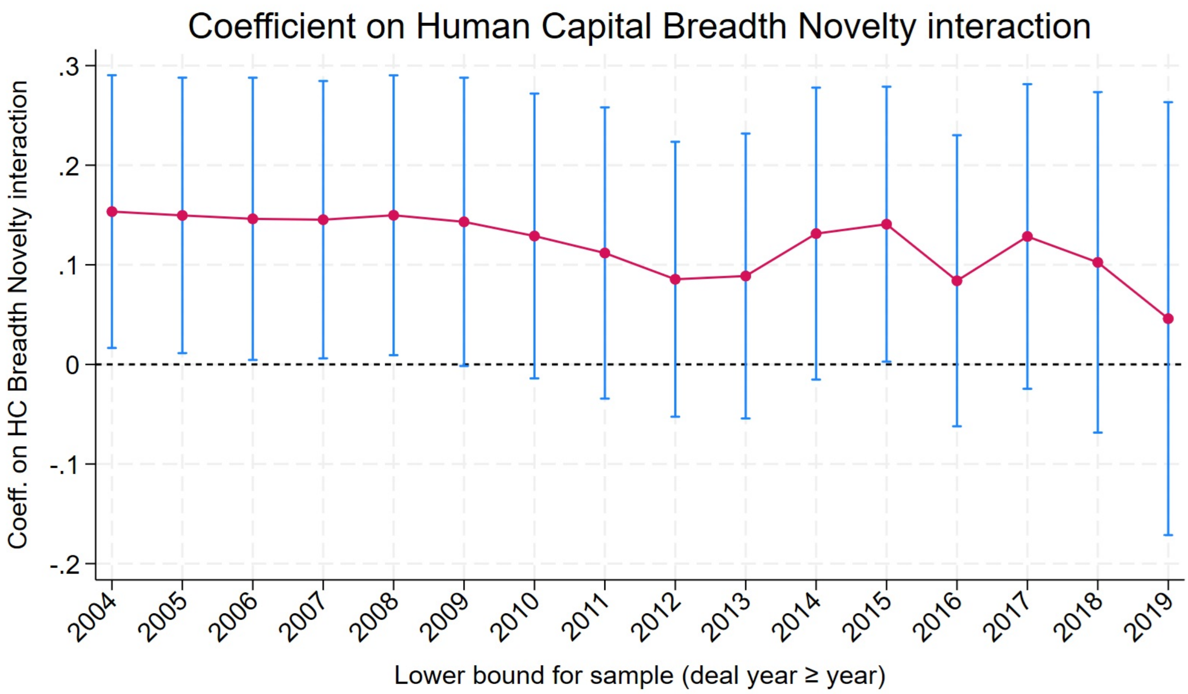

Figure IG2: **Interaction between Lead Partner Human Capital Breadth, Novelty and Investment Performance over time** This figure presents the coefficient δ of the association between the interaction between lead partner breadth and deal novelty and deal performance estimated via $P_{j,k,p,t} = \alpha + \beta B_p + \gamma N_{j,k,p,t} + \delta B_p \times N_{j,k,p,t} + X_{t,p} + \eta_{i \times t \times s \times c} + \rho_{t \times j} + \epsilon_{j,k,p,t}$ by progressively excluding cohorts of earlier financed VC-deals. For example, the coefficient estimate plotted in year 2010, is the estimated δ in the regression excluding deals done before 2010.

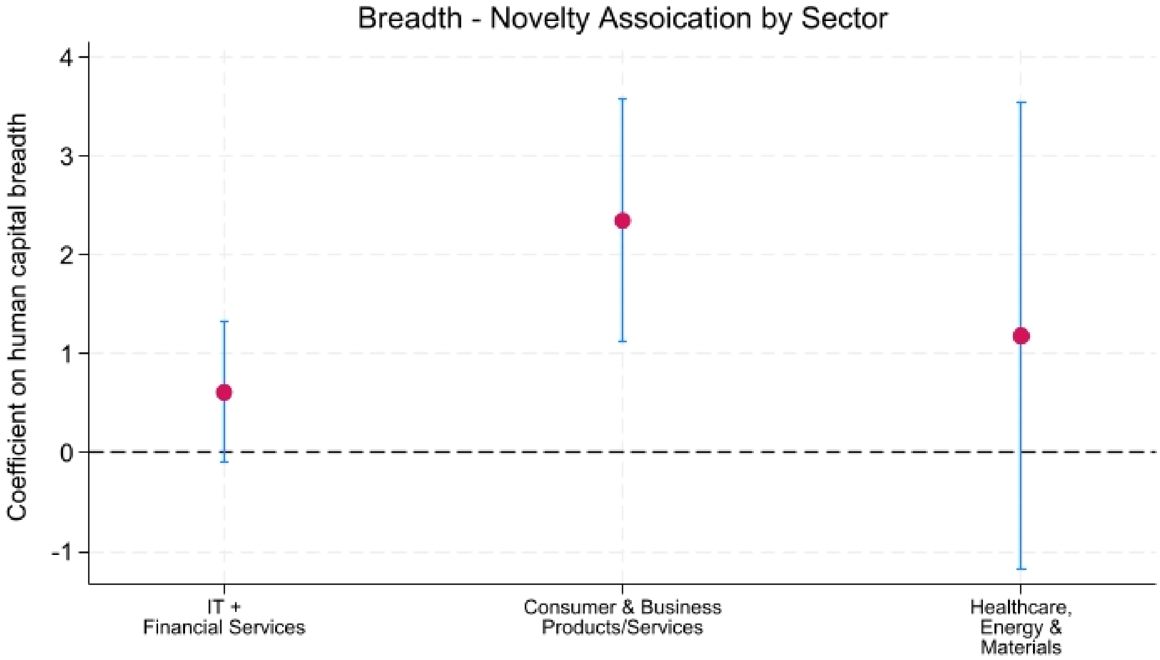

Figure IG3: **Association between lead partner human capital breadth and deal novelty: heterogeneity by sector** This figure presents the coefficient β of the association between human capital breadth and novelty estimated via $N_{j,k,p,t} = \alpha + \beta B_p + X_{t,p} + \eta_{i \times t \times s \times c} + \rho_{t \times j} + \epsilon_{j,k,p,t}$ separately for different broad venture sectors. Leftmost estimate for ventures in the IT and Financial Services sectors. Middle estimate for ventures in Consumer & Business products and services (B2B and B2C) sectors. Rightmost estimate for ventures in the Healthcare, Energy and Materials sectors.

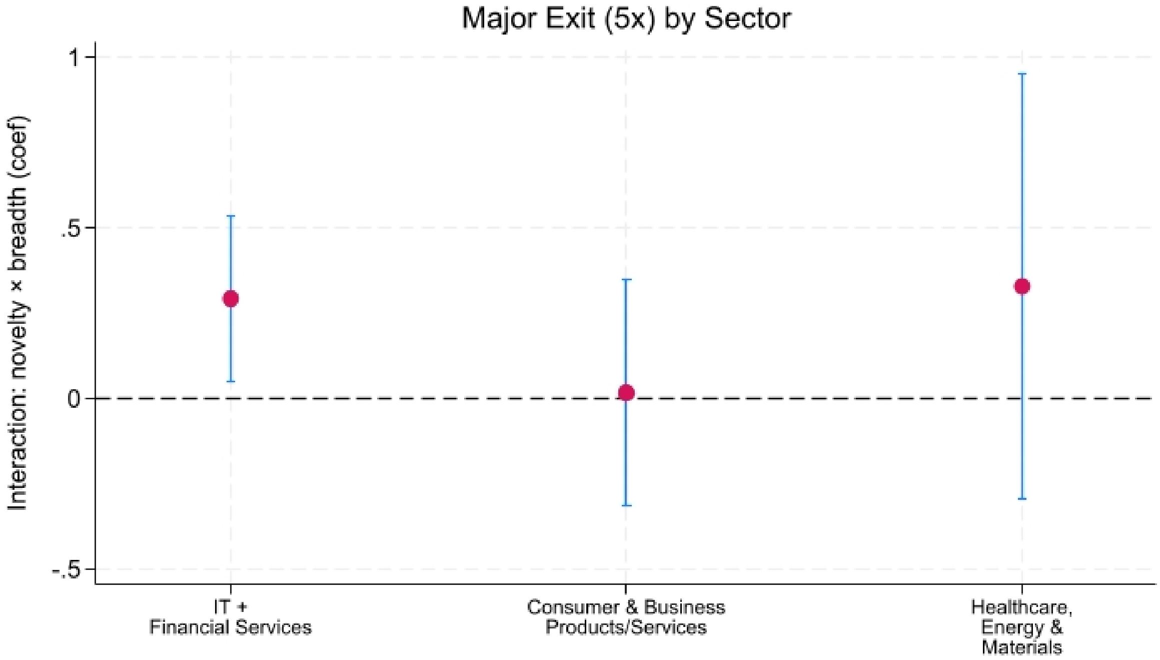

Figure IG4: **Association between the interaction of lead partner human capital breadth and deal novelty and performance: heterogeneity by sector** This figure presents the coefficient δ of the association between the interaction between lead partner breadth and deal novelty and deal performance estimated via $P_{j,k,p,t} = \alpha + \beta B_p + \gamma N_{j,k,p,t} + \delta B_p \times N_{j,k,p,t} + X_{t,p} + \eta_{i \times t \times s \times c} + \rho_{t \times j} + \epsilon_{j,k,p,t}$ separately for different broad venture sectors. Leftmost estimate for ventures in the IT and Financial Services sectors. Middle estimate for ventures in Consumer & Business products and services (B2B and B2C) sectors. Rightmost estimate for ventures in the Healthcare, Energy and Materials sectors.

	Frac. Top Quartile Novelty				
	(1)	(2)	(3)	(4)	(5)
Job Cateogory ratio	0.009 (0.019)				
Job Industry ratio		0.000 (0.022)			
Job Role ratio			-0.005 (0.023)		
Educational diversity index				-0.004 (0.004)	
Gender diversity index					0.001 (0.010)
Fund Size	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
First Fund	-0.013 (0.009)	-0.013 (0.009)	-0.013 (0.009)	-0.018* (0.009)	-0.013 (0.009)
Industry Controls	✓	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓	✓
Vintage Year × FundType FE	✓	✓	✓	✓	✓
Observations	2036.00	2036.00	2036.00	2009.00	2031.00
R^2	0.12	0.12	0.12	0.13	0.12

Table IG1: **Fund level Diversity and Novelty** This table presents of an OLS regression of fund level diversity measures on the fraction of investments in top quartile novelty firms. The dependent variable Frac. Top Quartile Novelty is the fraction of investments in top quartile novelty firms. The independent variable in column (1) Job Category ratio is the ratio between distinct job categories held by all partners in the fund and the total number of past employment spells of all partners in the fund. The independent variable in column (2) Job industry ratio is the ratio between distinct industries worked in by all partners in the fund and the total number of past employment spells of all partners in the fund. The independent variable in column (3) Job Role ratio is the ratio between distinct job roles held by all partners in the fund and the total number of past employment spells by all partners in the fund. The independent variable in column (4) is an educational diversity index at a fund level, which is the exponential of the Shannon entropy computed using the fraction of partners with a given educational degree (STEM education, Social Science or Humanities Education, MBA degree, PhD degree or Medical degree). The independent variable in column (t) is an gender diversity index at a fund level, which is the exponential of the Shannon entropy computed using the fraction of partners with a given gender (Male or Female). Fund Size is the AUM of the fund. First Fund is an indicator variable taking a value of 1 if this is the first fund raised by a VC firm. All columns include Industry Controls, which are separate controls for the fraction of investments made by the fund in each industry sector. All columns include Stage Controls, which are separate controls for the fraction of investments made by the fund in each investment Stage. All columns include Vintage Year × Fund Category fixed effects. Standard errors reported in parenthesis are clustered at the investor level * p<.10; ** p<.05; *** p<.01. Sample constrained on at least 2 partners observable for each fund

	Fraction of Successful Exits				
	(1)	(2)	(3)	(4)	(5)
Frac. Top Quartile Novelty	0.265*** (0.074)	0.396*** (0.110)	0.459*** (0.155)	0.119 (0.085)	0.324*** (0.102)
Job Cateogory ratio	0.008 (0.024)				
Frac. Top Quartile Novelty × Job Cateogory ratio	-0.085 (0.137)				
Job Industry ratio		0.042 (0.027)			
Frac. Top Quartile Novelty × Job Industry ratio		-0.257* (0.150)			
Job Role ratio			0.039 (0.033)		
Frac. Top Quartile Novelty × Job Role ratio			-0.313 (0.195)		
Educational diversity index				-0.005 (0.004)	
Frac. Top Quartile Novelty × Educational diversity index				0.039 (0.028)	
Gender diversity index					0.011 (0.014)
Frac. Top Quartile Novelty × Gender diversity index					-0.075 (0.075)
Industry Controls	✓	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓	✓
Vintage Year × FundType FE	✓	✓	✓	✓	✓
Observations	2036.00	2036.00	2036.00	2009.00	2031.00
R^2	0.33	0.33	0.33	0.33	0.33

Table IG2: **Fund level Diversity, Novelty and Performance.** This table presents of an OLS regression of fund level diversity measures, the fraction of investments in top quartile novelty firms and fund performance. The dependent variable Fraction of Successful Exits is the fraction of deals that have achieved an exit via IPO or high value acquisition (an acquisition with a value of at least five times greater than the total VC amount invested in the company). The independent variable Frac. Top Quartile Novelty is the fraction of investments in top quartile novelty firms. The independent variable in column (1) Job Category ratio is the ratio between distinct job categories held by all partners in the fund and the total number of past employment spells of all partners in the fund. The independent variable in column (2) Job industry ratio is the ratio between distinct industries worked in by all partners in the fund and the total number of past employment spells of all partners in the fund. The independent variable in column (3) Job Role ratio is the ratio between distinct job roles held by all partners in the fund and the total number of past employment spells by all partners in the fund. The independent variable in column (4) is an educational diversity index at a fund level, which is the exponential of the Shannon entropy computed using the fraction of partners with a given educational degree (STEM education, Social Science or Humanities Education, MBA degree, PhD degree or Medical degree). The independent variable in column (t) is an gender diversity index at a fund level, which is the exponential of the Shannon entropy computed using the fraction of partners with a given gender (Male or Female). All columns also include controls of for fund size and first fund. Fund Size is the AUM of the fund. First Fund is an indicator variable taking a value of 1 if this is the first fund raised by a VC firm. All columns include Industry Controls, which are separate controls for the fraction of investments made by the fund in each industry sector. All columns include Stage Controls, which are separate controls for the fraction of investments made by the fund in each investment Stage. All columns include Vintage Year × Fund Category fixed effects. Standard errors reported in parenthesis are clustered at the investor level * p<.10; ** p<.05; *** p<.01. Sample constrained on at least 2 partners observable for each fund

	Frac. Top Quartile Novelty					
	(1)	(2)	(3)	(4)	(5)	(6)
Fund level breadth index PC 1	0.001 (0.003)		0.001 (0.003)	0.000 (0.005)		-0.000 (0.005)
Fund level breadth index PC 2		-0.003 (0.004)	-0.003 (0.004)		-0.003 (0.006)	-0.003 (0.006)
Fund Size	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Industry Controls	✓	✓	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓	✓	✓
VC Firm FE				✓	✓	✓
Vintage Year × FundType FE	✓	✓	✓	✓	✓	✓
Observations	2072.00	2072.00	2072.00	2072.00	2072.00	2072.00
R^2	0.12	0.12	0.12	0.55	0.55	0.55

Table IG3: **Fund level Diversity and Novelty: Diversity measures using Principal Component Analysis** This table presents of an OLS regression of fund level diversity measures on the fraction of investments in top quartile novelty firms. The dependent variable Frac. Top Quartile Novelty is the fraction of investments in top quartile novelty firms. The independent variables Fund level breadth index PC 1, Fund level breadth index PC 2 are the first and second principal component respectively of a fund level breadth index constructed using (1) Job Category ratio: the ratio between distinct job categories held by all partners in the fund and the total number of past employment spells of all partners in the fund, (2) Job industry ratio: ratio between distinct industries worked in by all partners in the fund and the total number of past employment spells of all partners in the fund, (3) Job Role ratio: ratio between distinct job roles held by all partners in the fund and the total number of past employment spells by all partners in the fund, (4) Educational diversity index at a fund level: which is the exponential of the Shannon entropy computed using the fraction of partners with a given educational degree (STEM education, Social Science or Humanities Education, MBA degree, PhD degree or Medical degree), (5) Gender diversity index at a fund level, which is the exponential of the Shannon entropy computed using the fraction of partners with a given gender (Male or Female). Fund Size is the AUM of the fund. All columns include Industry Controls, which are separate controls for the fraction of investments made by the fund in each industry sector. All columns include Stage Controls, which are separate controls for the fraction of investments made by the fund in each investment Stage. All columns include Vintage Year × Fund Category fixed effects. Standard errors reported in parenthesis are clustered at an Investor level * p<.10; ** p<.05; *** p<.01. Sample constrained on at least 2 partners observable for each fund

	Fraction of Successful Exits	
	(1)	(2)
Frac. Top Quartile Novelty	0.221***	0.224***
	(0.038)	(0.043)
Fund level breadth index PC 1	0.002	
	(0.006)	
Frac. Top Quartile Novelty × Fund level breadth index PC 1	0.011	
	(0.028)	
Fund level breadth index PC 2		0.005
		(0.008)
Frac. Top Quartile Novelty × Fund level breadth index PC 2		-0.047
		(0.034)
Industry Controls	✓	✓
Stage Controls	✓	✓
VC Firm FE	✓	✓
Vintage Year × FundType FE	✓	✓
Observations	2072.00	2072.00
R^2	0.68	0.68

Table IG4: **Fund level Diversity, Novelty and Performance: Diversity measures using Principal Component Analysis** This table presents of an OLS regression of fund level diversity measures computed using principal component analysis. The dependent variable Fraction of Successful Exits is the fraction of deals that have achieved an exit via IPO or high value acquisition (an acquisition with a value of at least five times greater than the total VC amount invested in the company). The independent variable Frac. Top Quartile Novelty is the fraction of investments in top quartile novelty firms. The independent variables Fund level breadth index PC 1, Fund level breadth index PC 2 are the first and second principal component respectively of a fund level breadth index constructed using (1) Job Category ratio: the ratio between distinct job categories held by all partners in the fund and the total number of past employment spells of all partners in the fund, (2) Job industry ratio: ratio between distinct industries worked in by all partners in the fund and the total number of past employment spells of all partners in the fund, (3) Job Role ratio: ratio between distinct job roles held by all partners in the fund and the total number of past employment spells by all partners in the fund, (4) Educational diversity index at a fund level: which is the exponential of the Shannon entropy computed using the fraction of partners with a given educational degree (STEM education, Social Science or Humanities Education, MBA degree, PhD degree or Medical degree), (5) Gender diversity index at a fund level, which is the exponential of the Shannon entropy computed using the fraction of partners with a given gender (Male or Female). Fund Size is the AUM of the fund. All columns include Industry Controls, which are separate controls for the fraction of investments made by the fund in each industry sector. All columns include Stage Controls, which are separate controls for the fraction of investments made by the fund in each investment Stage. All columns include Vintage Year × Fund Category fixed effects. Standard errors reported in parenthesis are clustered at an Investor level * p<.10; ** p<.05; *** p<.01. Sample constrained on at least 2 partners observable for each fund

Table IG5: Industry Sector Distribution

Industry Sector	**Freq.**	**Percent**	**Cum.**
Business Products and Services (B2B)	32,694	14.09	14.09
Consumer Products and Services (B2C)	43,828	18.88	32.97
Energy	3,499	1.51	34.48
Financial Services	6,654	2.87	37.34
Healthcare	38,611	16.63	53.98
Information Technology	103,957	44.79	98.77
Materials and Resources	2,865	1.23	100.00
Total	232,108	100.00	

	100 × Novelty (Distance to Closest Firm)			
	(Full Sample)	(Full Sample)	(Lead)	(Lead)
HHI (Past Investments)	0.457**	0.414	0.645*	1.031
	(0.228)	(0.399)	(0.377)	(1.112)
Deal Diversity (Past Investments)	4.116***	1.382	3.551**	2.199
	(0.788)	(1.650)	(1.406)	(4.533)
VC Experience	0.036	-0.110	0.251**	-0.154
	(0.063)	(0.131)	(0.108)	(0.293)
Industry × Deal Year × Deal Type × Country FE	✓	✓	✓	✓
VC Firm × Deal Year FE		✓		✓
Observations	20083.00	20083.00	7607.00	7607.00
R^2	0.40	0.61	0.36	0.67

Table IG6: **Association between lead partner's experience acquired from past deals financed and focal startup novelty** This table reports the results of an OLS regression of deal novelty and measures of lead partner's experience acquired through past financed deals. The dependent variable Novelty (Distance to Closest Firm) in columns (1)-(4) is a deal level measure of novelty defined as one minus the maximum of the cosine similarity of the textual description of the startup financed in the deal and all startups receiving venture capital financing within five years before the deal in the same deal stage. The independent variable HHI (Past Investments) is a Herfindahl-Hirschmann industry specialization measure based on the lead partner's past investments. The independent variables Deal Diversity (Past Investments) is a measure of diversity of past deals financed computed as the average cosine distance between ventures financed by the partner before the focal deal. VC experience is a variable defined as the logarithm of one plus the number of deals financed by the lead partner prior to current deal. Columns (1) - (4) include Industry × Deal Year × Deal Type × Financed Company Country FE. Columns (2) and (4) also include Investor × Deal Year FE. Standard errors reported in parenthesis are double clustered at an Investor and Financed Company level. * p<.10; ** p<.05; *** p<.01.

	100 × Novelty (Distance to Closest Firm)	
	(1)	(2)
Breadth Index	0.356*	1.434**
	(0.209)	(0.706)
HHI (Past Investments)	0.447	2.614*
	(0.640)	(1.522)
Deal Diversity (Past Investments)	-2.116	-8.991
	(2.597)	(10.418)
VC Experience	0.122	0.456
	(0.309)	(1.038)
Partner Industry Experience	0.018	0.805
	(0.388)	(0.675)
Controls	✓	✓
Industry × Deal Year × Deal Type × Country FE	✓	✓
VC Firm × Deal Year × Partner Entry Year FE		✓
Observations	5909.00	5909.00
R^2	0.59	0.75

Table IG7: **Association between lead partner's human capital breadth index and startup novelty controlling for on the VC job experience** This table reports the results of an OLS regression of deal novelty and lead partner's human capital breadth with additional controls related to on the VC job acquired experience. The dependent variable Novelty (Distance to Closest Firm) in columns (1)-(2) is a deal level measure of novelty defined as one minus the maximum of the cosine similarity of the textual description of the startup financed in the deal and all startups receiving venture capital financing within five years before the deal in the same deal stage. The main independent variable Breadth Index is a measure of human capital breadth which is defined as the first principal component of a PCA from 4 individual human capital breadth proxies: 1) The ratio of distinct job categories to total employment spells 2) The ratio of distinct job roles to total employment spells 3) The ratio of distinct industries worked in to total employment spells 4) Educational breadth count. The independent variable HHI (Past Investments) is a Herfindahl-Hirschmann industry specialization measure based on the lead partner's past investments. The independent variables Deal Diversity (Past Investments) is a measure of diversity of past deals financed computed as the average cosine distance between ventures financed by the partner before the focal deal. VC experience is a variable defined as the logarithm of one plus the number of deals financed by the lead partner prior to current deal. Columns (1) include Industry × Deal Year × Deal Type × Financed Company Country FE. Column (2) includes Investor × Deal Year × VC Partner Entry Year FE. All columns include individual level partner controls: age, sex and ethnicity. Standard errors reported in parenthesis are double clustered at an Investor and Financed Company level. * p<.10; ** p<.05; *** p<.01.

Internet Appendix H: Comparison of Novelty Measure to other existing measures

6.5 Comparison with Bonelli (2022)

In this subsection I compare the Novelty (Distance to Closest Firm Measure) to the "backward similarity" measure in Bonelli (2022). The "backward similarity" measure or similarity rank measure in Bonelli (2022) builds on a methodology developed by Kelly et al. (2021) and constructs a measure of backwards similarity of startups based on the distance between the startups' description to all firms financed by the VC in the same industry as the focal startup in a rolling window of five years. The detailed construction of the similarity rank measure is provided by Bonelli (2022) in internet appendix B. I reconstruct Bonelli (2022)'similarity rank measure as described in Bonelli (2022). Two differences are worth noting due to the differences in the main data sources used in this paper and Bonelli (2022) (Pitchbook vs. Crunchbase). (i) Because of the different datasources, the business descriptions of companies would likely be different

(ii) Industry classification and granularity in Pitchbook and Crunchbase data are different which can affect the number of benchmark firms each firm is compared when constructing the similarity rank measure as described by Bonelli (2022). When constructing the Similarity Rank measure I do so separately at the three levels of industry granularities offered by PitchBook: Industry Sector, Industry Group and Industry Code. The brief summary of conclusions from the comparison exercise are the following: (i) There is a relatively low correlation betweeen the Novelty (Distance to Closest Firm) measure ranging from 0.15-0.26 depending on the industry granularity used to construct the benchmark set, (ii) I am able to replicate the stylized facts in Bonelli (2022) with PitchBook data - that is high backward similarity startups are more likely to fail, less likely to achieve a major success and less likely to lead to innovative outcomes. (iii) A horse-race between Novelty (Distance to Closest Firm) and Similarity Rank measure does not change the magnitude nor statistical significance of the coefficient of the Novelty (Distance to Closest) measure (iv) There is a conceptual difference in what the two - measures capture: the similarity rank measure is correlated positively likelihood of failure and negatively with likelihood of major success, however the Novelty (Distance to Closest Firm) measure correlates positively with both failure and major success and this makes it arguably more valid as an ex-ante proxy for startups' risk and captures the idea that financing novelty is riskier.

Table IH1: This table presents a correlation matrix between Novelty (Distance to Closest Firm) measure and Similarity Rank measure in Bonelli (2022) for the 3 industry granularities offered in PitchBook data: Industry Sector, Industry Group and Industry Code.

	Novelty (Distance to Closest Firm)	Sim. Rank (Sector)	Sim. Rank (Group)	Sim. Rank (Code)
Novelty (Distance to Closest Firm)	1.0000			
Sim. Rank (Sector)	-0.2628	1.0000		
Sim. Rank (Group)	-0.1645	0.8478	1.0000	
Sim. Rank (Code)	-0.1609	0.7066	0.8408	1.0000

	100 × Failure		100 × Majors Success		Number of Forward Citations	
	(1)	(2)	(3)	(4)	(5)	(6)
Sim. Rank (Sector)	4.025**	5.224***	-7.116***	-2.151	-1.275	-0.626
	(1.917)	(2.026)	(1.785)	(1.607)	(0.833)	(0.743)
Novelty (Distance to Closest Firm)		20.488**		84.841***		10.585**
		(9.176)		(10.540)		(5.377)
Exit Type Controls					✓	✓
Industry × Deal Year × Deal Type × Country × VC Firm × Lead Partner FE	✓	✓	✓	✓	✓	✓
Observations	81418.00	81382.00	81418.00	81382.00	81418.00	81382.00
R^2	0.71	0.71	0.75	0.76		

Table IH2: **Outcome comparison of Novelty (Distance to Closest Firm) measure with Similarity Rank from (Bonelli, 2022) computed using the least granular PitchBook Industry Sector.** This table reports the results of an OLS regression of various startup outcomes: Failure, Major Success and Number of Forward citations on Similarity Rank measure from Bonelli (2022) and Novelty (Distance to Closest Firm) measure. The independent variable in Columns (1)-(3) is Sim. Rank (Sector) which is similarity rank measure from Bonelli (2022). In Columns (4)-(6) both independent variables Sim. Rank and Novelty (Distance to Closest Firm) are included. All columns include Industry × Deal Year × Deal Type × Country × Investor × Partner FE. Standard errors reported in parenthesis are double clustered at an Investor and Financed Company level. * p<.10; ** p<.05; *** p<.01.

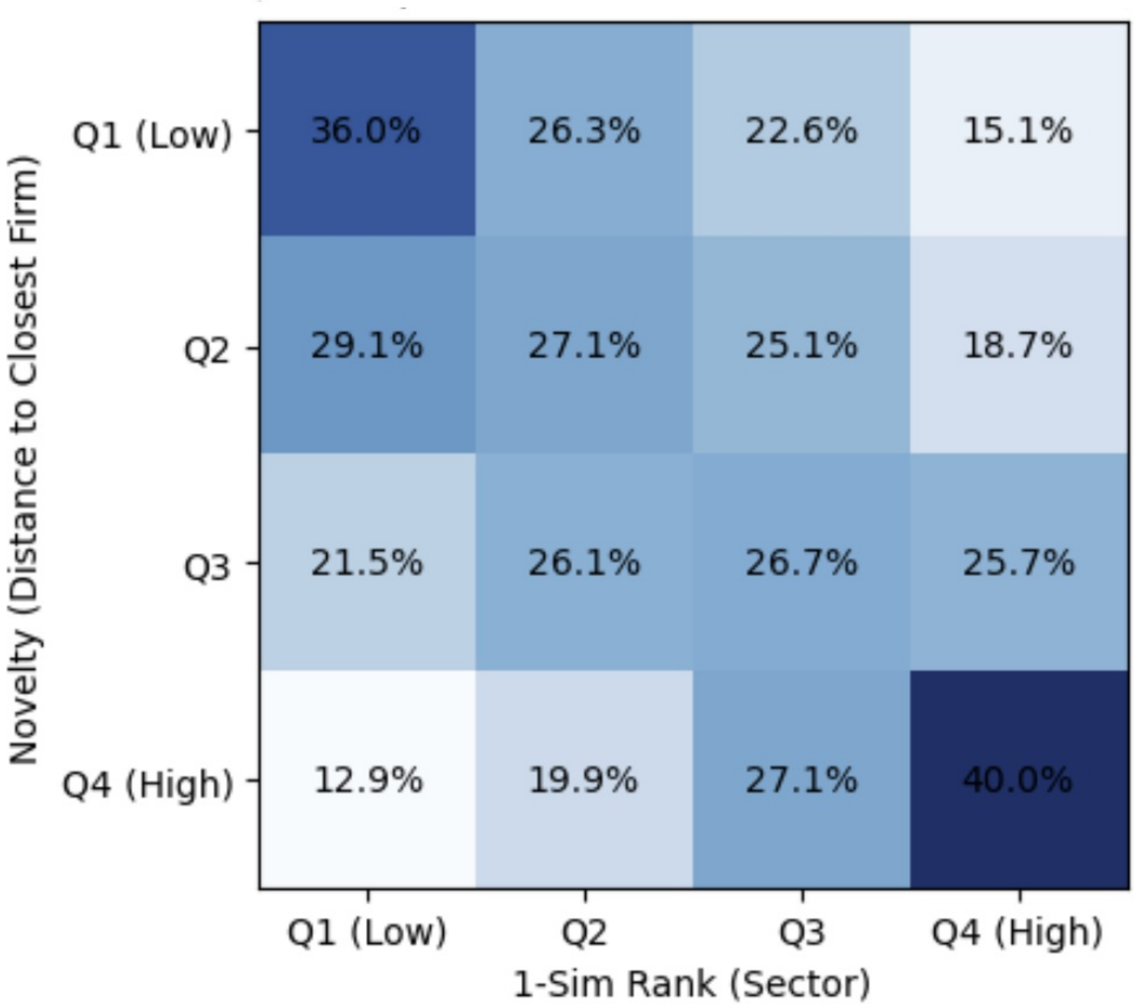

Figure IH1: **Heatmap comparing Novelty (Distance to Closest Firm) measure and 1-Sim. Rank measure from Bonelli (2022)**. This table presents a cross-tabulation of a quartile split of the Novelty (Distance to closest Firm) measure and 1-Sim. Rank measure form Bonelli (2022). Row and Column Percentages sum to 100. For example filed (1,1) represents the fraction of firms in Q1 of Novelty (Distance to Closest Firm) measure that are also in Q1 of the 1-Sim Rank (Sector) measure

Internet Appendix I: Full model and endogenizing entrepreneurial entry

6.6 Model description and timeline

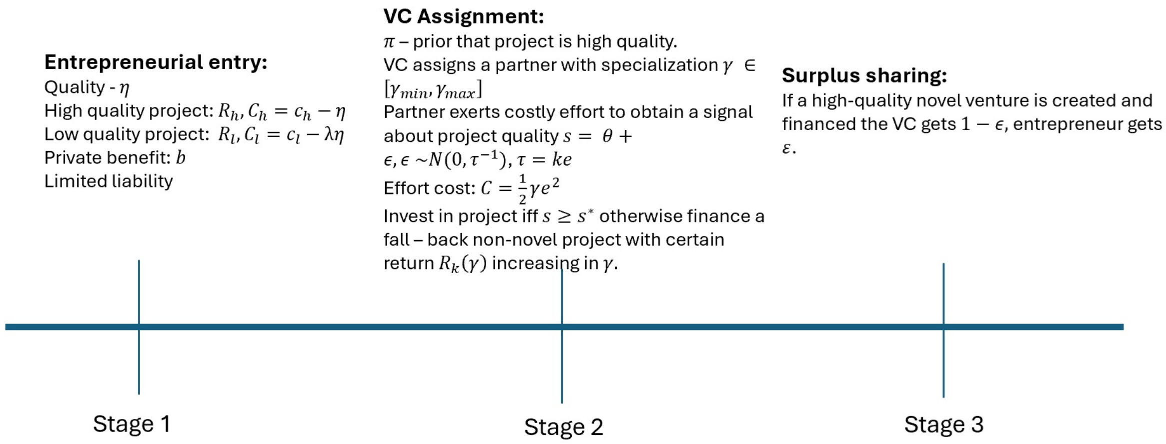

Figure II1: **Timeline of the model** This figure presents the timeline of the model

The timeline of the model is presented in Figure II1 and described as follows:

1. **Entrepreneur Stage**

   (a) An entrepreneur is born with a skill level $\eta \sim \text{Uniform}[0, 1]$.

   (b) Given η, the entrepreneur decides whether to enter. If she enters, she can create a novel venture of either high quality (h) or low quality (ℓ).

   (c) The baseline cost of creating a high- (low-) quality venture for an entrepreneur with no skill ($\eta = 0$) is c_h (c_ℓ), where $c_h > c_\ell$.[42]
   Given skill $\eta > 0$, the cost of creating each type of venture is:

   $$\text{Cost}(h) = c_h - \eta, \qquad \text{Cost}(\ell) = c_\ell - \lambda\eta, \quad 0 < \lambda < 1.$$

   43

   (d) If a project is created and financed, the VC earns $R_h > 0$ if the project is high quality, or $R_\ell < 0$ if it is low quality. The entrepreneur receives an exogenous share ε of the surplus and a non-pecuniary private benefit $b > 0$. The entrepreneur is subject to limited liability.

[42] $c_h > c_\ell$ captures, in reduced form, the idea that producing a high-quality venture requires greater effort (e.g., time or resources).

[43] Entrepreneurial skill reduces the cost of creating both types of ventures, but more so for high-quality ventures ($\lambda < 1$). This assumption is mainly technical and helps derive entry thresholds later in the model.

2. **VC Stage**

   (a) The VC observes all model parameters but not the entrepreneur's specific η. Based on this, the VC forms a prior π over the probability that the novel venture is high quality and assigns a partner with specialization level $\gamma \in [\gamma_{min}, \gamma_{max}]$ to screen the project.[44]

   (b) The partner exerts observable and contractible effort e, which incurs a cost of $\frac{1}{2}\gamma e^2$. The partner receives a signal about project quality:

$$s = \theta + \varepsilon, \qquad \varepsilon \sim \mathcal{N}\left(0, \frac{1}{\tau}\right), \quad \tau = \kappa e.$$

   (c) **Investment rule:** The VC invests in the novel project if $s \geq s^*$, where s^* is endogenously determined. Otherwise, it invests in a fallback (known) venture that yields return $R_k(\gamma)$, a concave increasing function of specialization γ.

3. **Surplus-Sharing Stage**

   (a) After the investment decision, the surplus is split via reduced-form Nash bargaining: the VC receives share $1 - \varepsilon$, and the entrepreneur receives ε, where ε is exogenously given.

6.7 Surplus-sharing stage

We assume that the surplus sharing stage is completely exogenous and if a high quality venture is created and financed the VC gets $1 - \epsilon$ of the surplus (which is $R_h - R_k(\gamma)$) and the entrepreneur gets a fraction ϵ. If a low quality venture is created and financed then since the low quality venture has a negative payoff and the entrepreneur is protected by limited liability the VC bears the full cost of funding such a venture $R_l - R_k(\gamma)$.

6.8 VC stage: Optimal effort and VC assignment rule

Once the surplus-sharing stage is completed, we solve the model backwards. Let π denote the prior probability that the project is of high quality. To screen the project, the VC firm assigns a partner of type γ, who exerts costly effort to acquire a signal about the project's quality. The partner can exert effort e to obtain an informative signal. Effort is costly and incurs a convex cost $C = \frac{1}{2}\gamma e^2$.

We assume that the effort exerted by the partner is observable and contractible. To compensate the partner, the VC firm pays a wage w that covers both the effort cost and the partner's outside opportunity cost $u(\gamma)$, an exogenous parameter that depends on the partner's type.[45]

[44] The VC firm has a roster of partners, each with a different specialization level γ.

[45] The outside option of the partner is treated as exogenous and determined in a competitive labor market for VC partners. In principle, $u(\gamma)$ can be specified as a differentiable function of the partner's specialization level.

6.8.1 Information acquisition and VC payoff

The partner assigned by the VC firm can exert costly effort to acquire an informative signal about the project quality. The observed signal is given by:

$$s = \theta + \epsilon \text{ where } \epsilon \sim N(0, \frac{1}{\tau}), \tag{137}$$

where $\theta = 1$ if the project is of high quality and $\theta = 0$ if the project is of low quality. The parameter τ is the signal informativeness and we assume that it increases with effort e. In particular we assume that $\tau = \kappa e$, where $\kappa > 0$ is a screening efficiency parameter. If the partner decides to not invest in the project the partner can invest in the known pool of ventures where there is a fallback option that returns $R_k(\gamma)$ that depends on the partner type γ.

After observing a signal the VC partner applies a simple threshold rule for investments. The partner accepts the investment if the signal $s \geq s^*$ there s^* is an endogenous threshold above which the signal is valuable. If the signal observed by the partner is below the threshold then the VC firm gets the fallback option $R_k(\gamma)$.

Assumption 3 (Information is valuable)**.** *We assume that for all $\gamma \in [\gamma_{min}, \gamma_{max}]$ it holds:*

$$\pi R_h + (1 - \pi) R_l < R_k(\gamma) \tag{138}$$

Intuitively, (138) states that without information acquisition it is optimal to invest in the fallback option, that is for any partner type γ the fallback option dominates the expected value before acquiring additional information about the type of novel project faced. In particular, since the fall-back option increases with specialization $R_k(\gamma)' > 0$, (138) implies $\pi R_h + (1 - \pi) R_l < R_k(\gamma_{min})$.[46]

Lemma 2 (Signal cut-off at which a novel project is accepted)**.** *The cut-off signal at which investment in a novel project becomes acceptable is given by:*

$$s^* = \frac{1}{2} + \frac{\Lambda}{\tau}, \tag{139}$$

where:

$$\Lambda = ln\left(\frac{(R_k(\gamma) - R_l)(1 - \pi)}{(R_h - R_k(\gamma))\pi}\right) \tag{140}$$

Proof. See Internet Appendix E.

□

Notice that assumption 3 on the value of information in fact guarantees that $\Lambda > 0$. Notice that with $\Lambda > 0$ we have $s^* > \frac{1}{2}$ which implies that in most cases the novel project is rejected

[46]Note that for (138) to hold we need to have $\pi < \frac{R_k(\gamma)_{min} - R_l}{R_h - R_l}$. π is determined endogenously in equilibrium and later we will show that (138) will hold for an interior equilibrium. Intuitively, the prior probability of facing a high quality venture should not be high enough, otherwise screening is not valuable.

and the fallback option is preferred.[47] In this case the following comparative statics regarding the signal threshold are true and trivial to show:

Proposition 8. *The threshold signal satisfies the following comparative statics:* $\frac{\partial s^*}{\partial \pi} < 0$, $\frac{\partial s^*}{\partial (R_k - R_l)} > 0$, $\frac{\partial s^*}{\partial (R_h - R_k)} < 0$.

Proof. Trivial differentiation with respect to parameters of (139). □

The intuition behind proposition 8 is clear. If the prior probability of a good novel venture becomes higher, the signal investment threshold for accepting the project becomes lower. $R_k - R_l$ is the opportunity cost of investing in low quality novel venture, the higher the opportunity cost the higher the threshold for acceptance.[48] $R_h - R_k$ is the payoff difference between the high quality projects and the fallback option, the higher this difference the lower the acceptance threshold.

Given the signal informativeness τ we define the true positive rate $\alpha(\tau)$ i.e., the probability a venture is high quality and accepted and the false positive rate $\beta(\tau)$ i.e., the probability a venture is of low quality and is accepted:

$$\alpha(\tau) \;=\; \Pr\big[s \geq s^\star \;\big|\; \theta = 1\big] \;=\; 1 - \Phi\big((s^\star - 1)\sqrt{\tau}\big) \tag{141}$$

$$\beta(\tau) \;=\; \Pr\big[s \geq s^\star \;\big|\; \theta = 0\big] \;=\; 1 - \Phi\big(s^\star \sqrt{\tau}\big), \tag{142}$$

where $\Phi(\cdot)$ is the standard normal cdf.

6.8.2 Optimal Effort Choice

Proposition 9. *Given a partner type* γ *the VC firm will choose optimal effort to maximize:*

$$\max_{e \geq 0} U(e) \;=\; \underbrace{R_k + (1-\epsilon)\pi\,\alpha(\tau)\,(R_h - R_k) + (1-\pi)\,\beta(\tau)\,(R_l - R_k)}_{\Pi(e)\ \text{(expected return to VC firm)}} \;-\; \underbrace{\left(\frac{\gamma\, e^2}{2} + u(\gamma)\right)}_{\text{partner's compensation}} \tag{143}$$

Given (π, γ) *the condition below* (144) *defines the optimal level of effort* $e^* > 0$ *which is a maximum of* (143) *and satisfies* $\tau^* > 2\Lambda$.

$$\big\{(1-\epsilon)\pi\,(R_h - R_k)\,\alpha'(\tau) + (1-\pi)\,(R_l - R_k)\,\beta'(\tau)\big\}\;\kappa \;=\; \gamma\, e^\star \tag{144}$$

Proof. See Internet Appendix E. □

[47] This result resonates with the fact that in the VC context very few firms seeking financing eventually obtain financing, unlike for instance in the case of more mature projects applying for bank loans.

[48] $R_k - R_l$ is the difference between the fallback in case a venture is rejected and the payoff in case a low quality venture is financed, hence it is the opportunity cost of financing a novel venture and ending up with a low quality venture.

Intuitively, condition (144) equalizes marginal screening benefit (right hand size) to marginal screening cost (left hand side). Marginally increasing effort rises the true positive rate $\alpha(\tau)$ and lowers the false positive rate $\beta(\tau)$, since α and β are concave the linear marginal cost will (left hand side) will cross the right hand side once for an interior effort defined by (144).

As a corollary we have the following comparative statics results:

Corollary 1 (Comparative statics of optimal effort)**.** *Suppose parameter values are such that $R_k(2-\epsilon) \geq R_h + (1-\epsilon)R_l$ then optimal effort e^* is decreasing with respect to γ i.e., $\frac{\partial e^*}{\partial \gamma} < 0$. Optimal effort is increasing with respect to π i.e., $\frac{\partial e^*}{\partial \pi} > 0$.*

Proof. See Internet Appendix E. □

Figure II1 about here.

In Figure II1 I provide a numerical solution and an illustration of the main results of this subsection. Namely, I plot the acceptance probability of novel projects, the conditional return of novel projects and optimal effort exerted as a function of specialization.

Figure II2 about here.

In Figure II2 I simulate the model, by simulating a large number $N = 10000$ of choice situations which are evaluated by partners with different γ and plot the empirical acceptance probability, realized conditional returns if a novel project is financed and full realized returns as well as the average optimal effort exerted and the average threshold for investment in novel project s^* for different values in γ. Once I solve for the optimal assignment rule (below) I will use the simulated data to show that, I can recover the full empirical estimates consistent with my identification strategy.

6.8.3 Optimal Assignment Rule

In this section we endogenize the VC assignment rule. In the previous subsection we optimized effort given γ. Now we solve for the optimal partner type γ. Recall that γ represents the specialization of the partner and we have assumed that the higher the specialization of the partner i.e., the higher the γ the higher the partner's screening cost, but also the higher the fallback option if the project is rejected. Specifically, motivated by the empirical evidence we assume that a generalist has a lower screening cost, and that the specialist has a larger fallback investment. We assume $R_k(\gamma)$ is differentiable, increasing and concave i.e., $R_k'(\gamma) > 0$ and $R_k''(\gamma) < 0$.[49]

We have the following proposition:

[49] One way to interpret this assumption is through ex-ante human capital investment: specialist partners may have spent a significant portion of their careers focusing on a particular sector, accumulating expertise that increases their ability to generate returns from familiar (known) venture types. The returns to these investment are diminishing, hence $R_k(\gamma)$ is concave. By contrast, generalist partners have broader experience across sectors, which lowers their marginal screening cost for evaluating novel ventures.

Proposition 10 (Optimal assignment rule)**.** *Given π the VC firm maximizes the total expected profit and assigns a partner of type γ s.t.*

$$\max_{\gamma \in [\gamma_{min}, \gamma_{max}]} V(\gamma) = \Pi(e^*(\gamma), \gamma) - C(e^*(\gamma), \gamma) - u(\gamma) \tag{145}$$

Given π the condition below defines the optimal partner specialization γ which maximizes (145).

$$R_k(\gamma^*)' q^r = \frac{1}{2} e(\gamma^*)^2 + u'(\gamma), \tag{146}$$

where $q^r = 1 - (1 - \epsilon)\pi\alpha(\tau) - (1 - \pi)\beta(\tau)$ is the rejection probability.

Proof. See Internet Appendix E. □

The intuition is straightforward, the LHS is the marginal benefit of raising γ (hiring a specialist) which is the marginal return of a fall back deal which the firm gets in case the deal from the unknown sector gets rejected. The right hand side is the marginal cost which is lower screening and the marginal wage to be paid at each γ.

Corollary 2 (Comparative statics of optimal specialization with respect to π)**.** *Assume payoffs are such that $(2 - \epsilon)R_k \geq R_h + (1 - \epsilon)R_l$ then optimal specialization decreases with π i.e. we have $\frac{\partial \gamma}{\partial \pi} < 0$.*

Proof. See Internet Appendix E. □

Under the parameter conditions specified when π increases the VC firm is incentivized to assign a more generalist partner. Intuitively when π increases the threshold for acceptance a project drops, since in most cases it is optimal to reject a project an increase in π incentives the VC firm to assign a partner that can generate an informative signal that can bring the signal for the project above the threshold - this incentivizes the hiring of a generalists since generalists have a lower cost of effort.

6.8.4 Entrepreneur

The previous subsections endogenized investment, partner assignment and optimal screening effort. This is sufficient to obtain and rationalize the main empirical cross-sectional findings in the model. Here, I close the model and endogenize π. This will be particularly useful to analyse and speak to the stylized facts relating the decrease in financed novelty and the increase in VC specialization.

In the first period of the model an entrepreneur is born with a skill level $\eta \sim F(\eta)$ where $F(\eta)$ is a CDF. The VC firm does not know the skill of the entrepreneur, but knows the distribution of entrepreneurial skill. The entrepreneur knows her skill. If the entrepreneur is sufficiently skilled to enter the market she can either work to produce a high quality venture or a low quality venture. The cost of producing a high quality venture is c_h. The cost of producing a low quality venture is c_l, where $c_h > c_h$. If the entrepreneur enters the market, produces a venture and the venture gets VC financing in the later periods the entrepreneur gets an additional private benefit of entrepreneurship b.[50].The entrepreneur's skill η can help the entrepreneur lower the cost of working on a given venture via. :

$$c_h(\eta) = c_h - \eta \tag{147}$$

$$c_l(\eta) = c_l - \lambda\eta \tag{148}$$

where $\lambda < 1$.[51]

The expected utility of producing a low quality venture for a type η entrepreneur is :

$$U_l(\eta) = \beta(\tau)b - c_l + \lambda\eta \tag{149}$$

In case the venture gets financing payoff is the private benefit b. Hence the expected payoff in case the entrepreneur produces a low quality venture is the probability of financing (probability of a false positive error by the VC) $\beta(\tau)$ times the private benefit b. The net cost the entrepreneur of skill η needs to pay is $c_l - \lambda\eta$. The utility for producing a high quality venture is:

$$U_h(\eta) = \alpha(\tau)(\epsilon(R_h - R_k) + b) - c_h + \eta \tag{150}$$

In this case the probability of financing is the true positive rate $\alpha(\tau)$ and the entrepreneur gets a fraction ϵ of the surplus plus her private benefit b.

Now we can characterize the entry conditions for the entrepreneur. The entrepreneur produces a high quality venture whenever:

$$U_g(\eta) \geq U_b(\eta) \tag{151}$$

From the last equation we define the quality choice threshold:

$$\eta^h = \frac{(c_h - c_l) - (\alpha(\tau)(\epsilon(R_h - R_k) + b) - \beta(\tau)b)}{1 - \lambda} \tag{152}$$

So the entrepreneur chooses to produce a good venture whenever she is born with $\eta \geq \eta^h$. The condition for entry into entrepreneurship is determined by a participation constant for working on a low quality venture:

$$U_b(\eta) \geq 0 \tag{153}$$

This defines the entry into entrepreneurship threshold

$$\eta^l = \frac{c_l - \beta(\tau)b}{\lambda}. \tag{154}$$

[50]I assume that this private benefit is constant and does not vary by type of venture

[51]The assumption that skill helps reduce the cost of working on a high quality venture more than it helps to reduce the cost of working on a low quality venture to make sure that there are no-overlapping entry thresholds.

whereby if the entrepreneur is born with an $\eta \geq \eta^l$ she enters entrepreneurship.

Hence the probability of entry into entrepreneurship assuming $\eta \sim Uniform[0, 1]$ is given by:

$$1 - \eta^l \tag{155}$$

The probability of entry into entrepreneurship and producing a high quality venture is then:

$$1 - \eta^h, \tag{156}$$

so π is determined by:

$$\pi = \frac{1 - \eta^h}{1 - \eta^l} \tag{157}$$

The following conditions need to hold:

1. High entry threshold higher than low entry threshold $\eta^h \geq \eta^l$
2. Positive low entry threshold $\eta^l \geq 0$.
3. Consistency $\eta^h \leq 1$.

It is obvious that if 3. and 1. hold then also $\eta^l \leq 1$. This defines conditions on parameter values:

$$c_l \geq \beta(\tau)b \tag{158}$$

$$1 - \lambda + c_l - \beta(\tau)b + \alpha(\tau)(\epsilon(R_h - R_k) + b) \geq c_h \geq \frac{c_l - \beta(\tau)b}{\lambda} + \alpha(\tau)(\epsilon(R_h - R_k) + b) \tag{159}$$

6.9 Equilibrium Definition and Existence

We now define the equilibrium of the model and prove its existence. The equilibrium consists of three endogenous variables: the fraction of high-quality projects π, the partner specialization level γ, and the screening effort e. These must satisfy the following conditions simultaneously:

Equilibrium A tuple (π^*, γ^*, e^*) constitutes an equilibrium if:

1. **Optimal Effort:** Given (γ^*, π^*), effort e^* satisfies the first-order condition:

$$\left\{(1 - \epsilon)\pi^*(R_h - R_k)\alpha'(\tau^*) + (1 - \pi^*)(R_l - R_k)\beta'(\tau^*)\right\}\kappa = \gamma^* e^* \tag{160}$$

   where $\tau^* = \kappa e^*$.

2. **Optimal Hiring:** Given π^*, partner type γ^* satisfies:

$$R_k'(\gamma^*)q^r(\tau^*) = \frac{1}{2}(e^*)^2 + u'(\gamma^*) \tag{161}$$

   where $q^r(\tau^*) = 1 - (1 - \epsilon)\pi^*\alpha(\tau^*) - (1 - \pi^*)\beta(\tau^*)$ is the rejection probability.

3. **Entrepreneurial Entry:** The entry threshold η^l and quality threshold η^h satisfy:

$$\eta^l = \frac{c_l - \beta(\tau^*)b}{\lambda} \tag{162}$$
$$\eta^h = \frac{(c_h - c_l) - [\alpha(\tau^*)(\epsilon(R_h - R_k) + b) - \beta(\tau^*)b]}{1 - \lambda} \tag{163}$$

with corresponding project mass and quality fraction:

$$\pi^* = \frac{1 - \eta^h}{1 - \eta^l} \tag{164}$$

Theorem 1. *Under the model assumptions and parameter restrictions, an equilibrium* (π^*, γ^*, e^*) *exists.*

Proof. See Internet Appendix E. □

6.10 Interpreting stylized facts through the lens of the model

n this subsection, I analyse the equilibrium and aim to interpret the stylized facts presented in the paper through the lens of the model. The main stylized facts documented in the paper are:

- A decline in financed novelty over time.
- An increase in specialization by VC firms over time.

I show that both of these facts can be simultaneously explained by analysing how the equilibrium defined in the paper responds to c_ℓ and c_h, which represent the entry costs for working on low- and high-quality ventures, respectively.

6.10.1 One-time increase in c_h — (Working on high-quality novelty is becoming harder over time, à la Jones (2009).)

In this subsection, I analyze the likely new equilibrium following a one-time exogenous increase in c_h by tracing the propagation of the shock through equilibrium values. I do not prove the statements explicitly for now.

1. Direct effect on π. It is clear that

$$\frac{\partial \eta^h}{\partial c_h} = \frac{1}{1 - \lambda} > 0. \tag{165}$$

   Since η^ℓ does not explicitly depend on c_h, it follows that initially $\uparrow c_h \rightarrow \downarrow \pi$.

2. Effect on optimal effort. According to the corollary, when π declines, optimal effort e also declines: $\uparrow c_h \rightarrow \downarrow \pi \rightarrow \downarrow e$.

3. Effect on signal informativeness. As signal precision declines, the true positive rate α decreases and the false positive rate β increases: $\uparrow c_h \rightarrow\downarrow \pi \rightarrow\downarrow e \rightarrow\downarrow \alpha(\tau), \uparrow \beta(\tau)$.

4. Since β is increasing, this decreases η^ℓ, further reducing π.

5. The lower π increases specialization: $\downarrow \pi \rightarrow\uparrow \gamma$.

6. The increase in γ raises the signal threshold for project acceptance, which reduces both α and β, although the overall effect on β remains positive due to the decline in effort.

The above steps outline that the new equilibrium features a lower π, a higher γ, and lower effort e. Intuitively, the higher cost of working on high-quality ventures directly reduces the prior probability that a given project is of high quality. This in turn weakens the VC's screening incentives and increases the value of the fallback option (since a project is more likely to be rejected). Consequently, entrepreneurs have lower incentives to produce high-quality projects—i.e., the new equilibrium level of π is further depressed beyond the direct effect of the higher cost of creating good ventures.

The feedback loop described here provides one interpretation of the empirical findings of the paper, which document a decline in average financed novelty π and an increase in VC specialization over time. Thus, one way to rationalize these empirical trends through the lens of the model is to view them as the result of a one-time upward shift in the cost of producing high-quality novel ventures.

6.10.2 One-time decrease in c_ℓ — (Working on low-quality novelty is becoming easier over time à la Ewens et al. (2018))

Now suppose there is a one-time decrease in the cost of producing a low-type venture, c_ℓ. In this subsection, I analyze the likely new equilibrium following a one-time exogenous shock to c_ℓ by tracing the propagation of the shock through equilibrium values. I do not prove the statements explicitly for now.

1. Direct effect on π. It is clear that

$$\frac{\partial \eta^h}{\partial c_\ell} = \frac{-1}{1-\lambda} < 0. \tag{166}$$

Similarly,

$$\frac{\partial \eta^\ell}{\partial c_\ell} = \frac{1}{\lambda} > 0. \tag{167}$$

Therefore, the direct effect of a one-time decline in c_ℓ is that the skill threshold for working on a good venture increases, while the skill threshold for working on a bad venture decreases. Both of these effects push π down. Hence, $\downarrow c_\ell \rightarrow\downarrow \pi$.

2. Effect on optimal effort. According to the corollary, when π declines, optimal effort e also declines: $\downarrow c_\ell \rightarrow\downarrow \pi \rightarrow\downarrow e$.

3. Effect on signal informativeness. As signal precision declines, the true positive rate α decreases and the false positive rate β increases: $\downarrow c_\ell \rightarrow \downarrow \pi \rightarrow \downarrow e \rightarrow \downarrow \alpha(\tau), \uparrow \beta(\tau)$.

4. Since β is increasing, this decreases η^ℓ, and the decline in α increases η^h, further depressing π.

5. The lower π increases specialization: $\downarrow \pi \rightarrow \uparrow \gamma$.

6. The increase in γ raises the signal threshold for project acceptance, which reduces both α and β, although the overall effect on β remains positive due to the decline in effort.

Therefore, a one-time decrease in c_ℓ in this model generates similar patterns for the new equilibrium as the one-time increase in c_h.

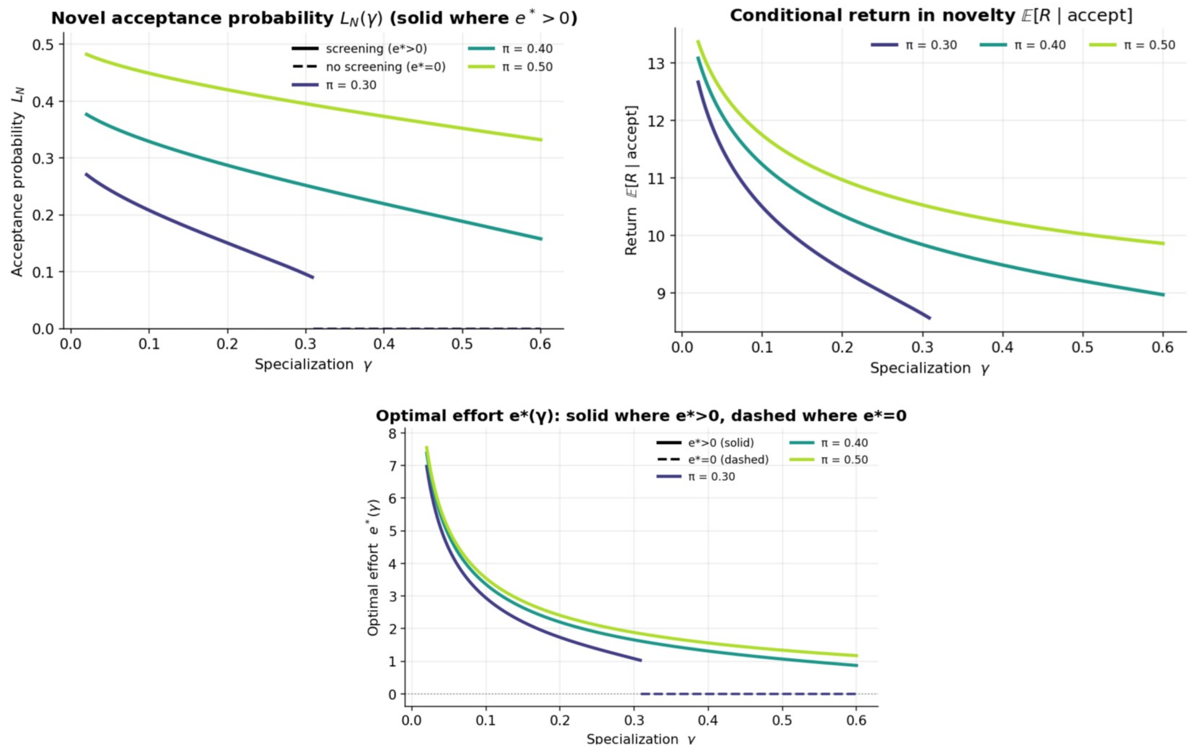

Figure II1: **Numerical solution of the model: Plots of acceptance probability of novel projects, conditional return of novel projects and optimal effort as a function of specialization.** This figure plots: (i) The novel acceptance probability given by (23) as a function of specialization γ for different values of the prior likelihood of a novel project π (top left) (ii) The conditional return given that a novel project has been financed (top right) (iii) Optimal effort exerted (bottom) for different values of the prior likelihood of a novel project π. Payoffs are $R_h = 15$ for high-type novel projects and $R_\ell = -8$ for low-type novel projects; the fallback payoff is $R_k(\gamma) = 5 + 2\sqrt{\gamma}$. Effort precision is $\tau = \kappa e$ with $\kappa = 1$ and cost $\frac{1}{2}\gamma e^2$. The entrepreneur's share of a successful high-type project is $\epsilon = 0.1$.

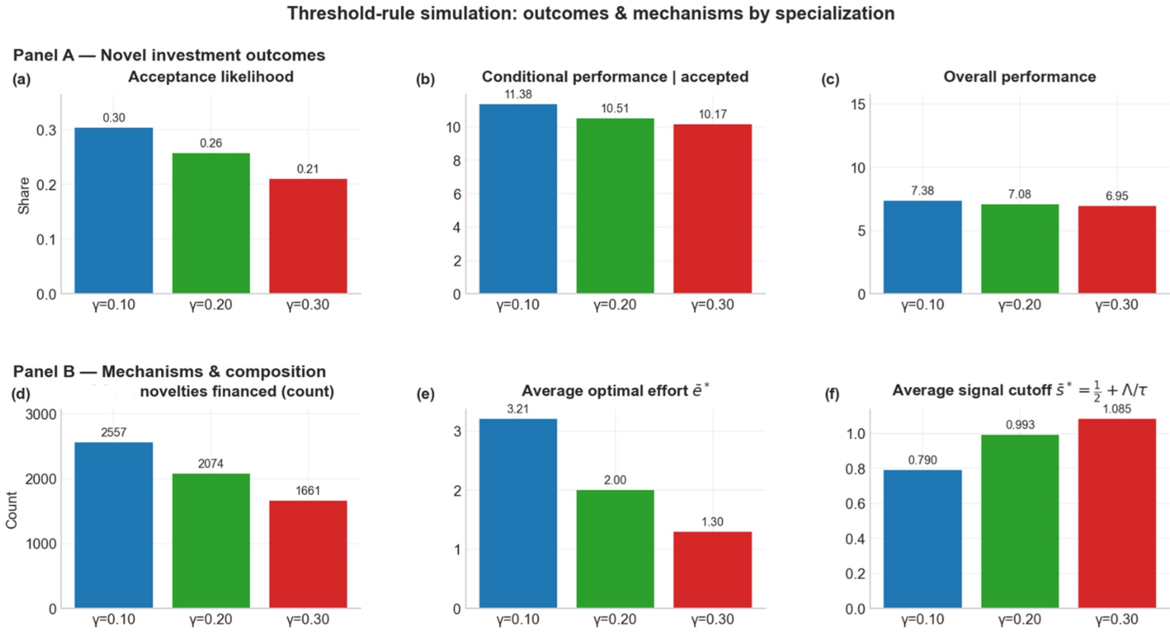

Figure II2: **Outcomes using simulated data for various value of specialization** γ This figure presents model outcomes for on simulated data for three different specialization parameter values. The simulation draws $N = 10{,}000$ deals from a uniform $pi \in [0.25, 0.5]$, where pi is the prior likelihood of a high type novel project. Payoffs are $R_h = 15$ for high-type novel projects and $R_\ell = -8$ for low-type novel projects. The fallback payoff is $R_k(\gamma) = 5 + 2\sqrt{\gamma}$. Effort precision is $\tau = \kappa e$ with $\kappa = 1$ and cost $\frac{1}{2}\gamma e^2$. The entrepreneur's share of a successful high-type project is $\epsilon = 0.1$. Each project arrives and it is evaluated by a partner of different γ so in this simulation, we keep the project simulation parameters constant but vary the project evaluator's specialization level. Panel A, figure a) plots the empirical likelihood of accepting a novel project. Panel A, figure b) plots the realized return on novel projects conditional on novel projects being financed. Panel A, figure c) plots the overall realized return. Panel B) figure d) plots the number of high quality novel projects financed. Panel B) figure e) plots the average effort exerted per project. Panel B) figure f) plots the average signal threshold required to finance a novel project.