Human Capital Breadth, Portfolio Choice and Performance in Venture Capital *

Teodor Duevski†

March 13, 2025

Abstract

I study how the breadth of venture capital (VC) partners' human capital influences investment selection, startup performance, and innovation. Within VC firms, partners with broader backgrounds are more likely to lead novel, high-risk investments, which have a higher likelihood of breakthrough success or failure. On average, these partners underperform, but when leading novel deals, they significantly increase the chances of major success. These results are consistent with both selection—where broad-background partners are skilled at screening novel firms—and monitoring—where their involvement enhances firm performance. Exploiting plausibly exogenous variation in partner busyness as a shock to lead-partner assignment, I provide causal evidence for these effects. To rationalize these findings, I develop an exploration-exploitation portfolio choice model. Broad-background VCs explore riskier sectors, experience early failures, and finance exceptional startups, while narrow-background VCs remain within their expertise. These results highlight the role of human capital breadth in fostering exploration and financing novel projects.

JEL Classification: G11, G24, G34

*I would like to thank Scott Baker, Bruno Biais, Jean-Edouard Colliard, Nicolas Crouzet, Vicente Cuñat, Michael Ewens, Thierry Foucault, Johan Hombert, Jessica Jeffers, Josh Lerner, Pedro Matos, Ramana Nanda, Aleksei Ovchinnikov, Maddalena Ronchi, Daniel Schmidt, Boris Vallee, and Guillaume Vuillemey, as well as seminar participants at 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.

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

1 Introduction

Venture capital (VC) plays a critical role in fostering innovation and economic growth by financing high-risk, high-reward startups (Lerner, 2000; Samila and Sorenson, 2011; Kaplan and Lerner, 2010; Gornall and Strebulaev, 2021). Beyond providing capital, venture capitalists (VCs) engage in pre-investment screening, structure complex contracts, and actively monitor portfolio companies post-investment (Kaplan and Strömberg, 2001; Gompers and Lerner, 2002; Kaplan and Strömberg, 2004; Chemmanur et al., 2011; Bernstein et al., 2016; Gompers et al., 2020). The success of this model hinges not only on financial resources but also on the human capital of VC investors, who apply their expertise to select, support, and scale promising ventures (Sørensen, 2007; Ewens and Rhodes-Kropf, 2015).

The literature on VC human capital and investment performance reveals a fundamental tension regarding the optimal composition of investors' skill sets. On one hand, research in financial intermediation in private markets underscores the benefits of specialization, arguing that deep industry expertise allows VCs to leverage knowledge and networks for superior investment outcomes (Gompers et al., 2009; Cressy et al., 2007; Spaenjers and Steiner, 2024). Conversely, 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 identify novel opportunities, adapt to uncertainty, and facilitate innovation (Lazear, 2004; Custódio et al., 2013, 2019; Murphy and Zabojnik, 2007). This tension is especially pronounced in venture capital, given that VCs primarily invest in early-stage firms with uncertain prospects, sparse historical data, and radically novel business models that differ significantly from previously financed startups. A key unresolved question is whether broader human capital enhances the ability to identify and nurture novel investments or whether deeper specialization consistently drives superior outcomes through targeted expertise. In this paper, I address this question both empirically and theoretically.

Empirically, I find that within VC firms, partners with broad human capital are more likely to lead investments in novel startups with previously unexplored business models. I also find evidence that, although these partners perform worse on average, they significantly increase the likelihood of IPO success and reduce the likelihood of failure when leading investments in novel firms. In all specifications, I include granular fixed effects for deal stage, year of financing, deal industry, and VC firm. Intuitively, I leverage variation in investment novelty within the same VC firm, year, deal stage, and industry sector, as partners with different levels of human capital breadth finance ventures of varying novelty. To further address endogenous partner-venture matching within VC firms, I exploit plausibly exogenous variation in partners' time-varying busyness (Abuzov, 2019), which serves as a probability shifter for partner assignment to deals—akin to a Bartik (1991)-style instrument. Using this approach, I establish a positive causal impact of human capital breadth on novel firm performance. Moreover, I show that these individual partner-level effects scale up to the VC fund and investor levels. To illuminate the underlying mechanism, I document that while investing outside a VC investor's expertise generally deteriorates performance, backing truly novel ventures—those with previously unexplored business models—raises the likelihood of major success. These findings align with an exploration-exploitation trade-off: although investing beyond an investor's expertise reduces average performance, it can generate significant upside when financing a truly promising firm.

To capture the portfolio allocation decisions of venture capitalists (VCs), I develop a dynamic portfolio choice model based on the multi-armed bandit framework, which allows me to formalize the trade-off between exploration and exploitation in venture capital investments. In this model, each 'arm' represents a distinct sector with varying probabilities of yielding high, intermediate, or low returns. At each instant, the VC chooses a sector to draw from, observes a signal about startup quality, and updates their prior beliefs about the sectoral potential in a Bayesian manner. I differentiate VCs by their human capital backgrounds—broad versus narrow—which shape their approaches to risk and exploration. Broad-background VCs, with diverse industry experience, start with a moderately informed understanding of multiple sectors, encouraging them to explore novel, high-risk sectors that could yield exceptional returns. In contrast, narrow-background VCs, who specialize in a particular domain, hold stronger priors about specific sectors and are more inclined to exploit known opportunities, focusing their investments within familiar industries. This distinction is central to the model, as it influences which sectors are sampled, the types of startups selected, and subsequent performance. Following the optimal strategy, broad-background VCs hold sample the novel sector more frequently, experience early failures, and are more likely to finance exceptional, high-quality startups. Narrow-background VCs, in contrast, tend to stick to sectors within their expertise and experience fewer failures.

To study this question empirically, I combine three primary datasets: PitchBook, Crunchbase, and USPTO. From PitchBook, I collect data on U.S.-based VC firms' investments and startup exits. I gather data on VC partners' human capital from Crunchbase. Additionally, to study innovation among VC-backed startups, I collect data on patent applications and grants and citations from the United States Patent and Trademark Office (USPTO). The central empirical challenge lies in accurately measuring both startup novelty and the breadth of human capital among VC partners.

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.[1] Specifically, for each startup with an available business description, I first construct an OpenAI LLM-based embedding vector from the startup's business description.[2] This allows me to compute the embedding-based distance between any two VC-funded startups using cosine similarity. The novelty proxy for a given startup is then computed as 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.[3] The novelty measure captures how truly novel a given startup's business model is relative to the closest venture financed by the VC industry in the past.

Using this measure, I first present several new stylized facts about novel firms and aggregate novelty trends in the VC industry. First, my novelty measure is strongly correlated with the likelihood of achieving either major success or failure. Second, I document that startup novelty correlates positively with both the number of forward patents granted and the number of forward patent citations (after receiving VC financing), suggesting that the most novel firms tend to be the most innovative. In aggregate, I document a declining trend in the average novelty of VC-financed startups over time. Importantly, I find that for startups at the very top of the novelty distribution, novelty remains stable over time, suggesting that the observed downward

[1] One concern with using business text descriptions is that startups may pivot from their original business plans. However, 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.

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

[3] This measure is related but distinct from Bonelli (2022)'s "backward similarity" measure, which is computed as the average past similarity between a focal startup and what the VC industry has financed in the past. That measure captures startup nicheness and the ability to predict future outcomes from past data.

trend is not driven by the most novel firms.

In the absence of comprehensive data on the complete educational and professional histories of venture capital (VC) partners, I develop a text-based methodology for measuring the breadth of human capital. Leveraging recent advancements in NLP, I first create textual prototypes that capture broad and narrow human capital profiles while intentionally abstracting from specific skills, industries, or investment knowledge. Using OpenAI-generated embeddings, I then quantify each VC partner's similarity to these broad and narrow prototypes by computing the textual similarity distance between these prototypes and the textual descriptions of their career backgrounds. The resulting partner-level index of human capital breadth is the normalized difference between the partner's average distance to broad and narrow prototypes.[4]

I validate this human capital breadth measure by correlating it with observable partner characteristics for a subset of individuals with detailed biographical data. I find that greater human capital breadth is positively associated with having an MBA, immigrant status, and being female, while negatively associated with holding a PhD or an undergraduate degree in computer science. Over time, I document a significant decline in the average breadth of human capital among VC partners. VC partnerships are typically heterogeneous, comprising both specialist and generalist partners. However, newer partnerships are increasingly dominated by specialists, underscoring a systematic shift toward specialization in the venture capital industry.

First, while not claiming to establish causality, I provide evidence that partners with broad human capital are more likely to finance novel firms. Furthermore, I show that novel firms led by partners with broad backgrounds achieve better performance. Using investment-level data, 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 corresponds to a 0.2-standard-deviation increase in deal novelty. I also find that the interaction of startup novelty and lead partner breadth is positively associated with performance, specifically in terms of a higher probability of major success and a lower probability of failure. However, the baseline effect of human capital breadth is negative, suggesting that non-novel firms perform better when led by a partner with a narrower background. In these specifications, I control

[4]This normalization ensures the measure is not mechanically influenced by variations in the length of partners' textual descriptions.

for granular fixed effects at the deal stage, deal industry, deal year, and VC firm 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, as well as systematic differences in novelty trends across sectors and deal stages. Additionally, the granular fixed effects account for time-varying macroeconomic and industry shocks that could influence performance outcomes.

To study the causal impact of human capital breadth on the performance of novel startups, I first restructure the data following Ewens and Rhodes-Kropf (2015). Specifically, for each deal made by a VC firm, I construct a set of potential lead partners, defined as those employed by the same firm within a three-year window around the deal (-3 to +3 years). For each potential lead, I incorporate a busyness proxy, following Abuzov (2019), which equals 1 if the partner is involved in an exit event—either an acquisition or an IPO—within a 90-day window (-90 to +90 days) around the focal deal. This alternative data structure yields two key insights. First, it provides an alternative way to demonstrate that partners with higher human capital breadth are more likely to be selected as leads on novel deals. Specifically, I estimate a positive association between deal novelty, human capital breadth, and the probability of a partner being chosen as the lead.[5] Second, I confirm that when controlling for other partner characteristics, the busyness proxy is negatively correlated with the likelihood of a partner leading an investment, reinforcing its potential validity as an natural instrument.

I leverage this data structure to address a central empirical challenge: the endogenous assignment of VC partners to specific deals within the same firm. Because partners and ventures within a VC firm are not randomly matched, estimating the causal effect of partner human capital breadth on startup performance requires an exogenous source of variation in partner assignment. To address this concern, I exploit idiosyncratic fluctuations in partner availability at the time of investment, constructing an instrument based on time-varying busyness resulting from IPO or acquisition exits in a partner's existing portfolio. These exit events generate plausibly exogenous variation in the set of available partners within a firm at the time of deal-making, thereby influencing the likelihood that a broad- or narrow-background partner is chosen

[5] A key advantage of this data structure is the ability to include deal fixed effects, ensuring that identification comes from variation in human capital breadth among potential leads within the same deal. This controls for all venture-specific unobservable characteristics, while the interaction between novelty and breadth is identified through slope differences across deals.

to lead a particular investment. Crucially, because partner exits are driven by factors unrelated to the specific startups under consideration, the average available human capital breadth at the time of a deal is plausibly exogenous to future startup performance, conditional on VC firm, industry, year, and deal-stage fixed effects. By exploiting this within-firm variation, I mitigate concerns about endogenous matching based on unobserved startup quality or partner-specific characteristics. This conditional exogeneity assumption is reasonable because variations in partner availability—driven by unrelated exit events—should influence startup performance only through their effect on the selection of a broad- or narrow-background partner, rather than through other omitted channels.

The intuition behind this instrument is that, within a given VC firm, the probability of assigning a broad-background (narrow-background) partner to a deal increases (decreases) with the average availability of broad human capital. Importantly, because this variation is measured within VC firms and deal years, it captures differences in partner availability across the timing of deals rather than cross-sectional differences in human capital breadth between VC firms. I use average human capital breadth availability to instrument for the chosen partner's human capital breadth and the interaction of average breadth availability and deal novelty to instrument for the interaction of the chosen partner's breadth and deal novelty. The relevance condition is strongly supported by the data, with FF-statistics exceeding 62, 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, demonstrating a strong first-stage relationship. The IV estimates imply that, at very low novelty, an additional standard deviation of breadth reduces the probability of IPO by about 12.6 percentage points (p.p.), representing a negative baseline effect. Each one-standard-deviation increase in novelty, however, raises the slope on breadth by roughly 3.9 percentage points, steadily offsetting and eventually reversing this effect. A one-standard-deviation increase in novelty above its mean raises the marginal effect of breadth enough to yield a net positive 5.2 percentage point increase in IPO likelihood. Thus, for deals below the median in novelty, human capital breadth is harmful; but once the deal is novel enough, its effect becomes positive.

The partner-level analysis presented thus far demonstrates that within VC firms, partners with broader backgrounds are more likely to lead novel deals and enhance the likelihood of successful exits for these innovative investments. These findings emphasize the role of individual lead partners in both the selection and monitoring of innovative investments.[6] However, one may argue that even if a single partner 'championing' an early-stage investment is sufficient for selection, any positive impact on portfolio company outcomes post-investment may result from the collective efforts of several partners rather than solely the lead partner. Additionally, compensation incentives within VC firms are typically structured at the fund level rather than being tied to individual partner-deal outcomes, meaning that all partners within a fund share responsibility for overall investment performance. To address this issue, I document that the investment-level results extend to the fund level. Specifically, I show that the average human capital breadth index at the partnership level is positively associated with the average deal novelty of the partnership's investments, controlling for granular fund-level characteristics such as fund size, the sectoral and stage composition of investments, and vintage-year-by-country fixed effects. Similar to the partner-level results, I also find a positive interaction between average deal novelty, average human capital breadth at the partnership level, and fund performance.

To model the portfolio choice of venture capitalists and rationalize the empirical findings, I build on the multi-armed bandit framework, which captures the dynamic trade-off between exploration and exploitation.[7] In this model, each 'arm' represents a distinct industry sector with varying probabilities of yielding high (IPO-like), intermediate (M&A-like), and low (failure-like) outcomes, represented by a multinomial distribution (Berry and Fristedt, 1985). Sectors in the model are heterogeneous: established sectors offer stable returns, while novel sectors are more volatile, with a higher probability of both low and high outcomes. At each decision point, the venture capitalist chooses a sector to invest in, receives a startup, and updates her prior about the quality distribution of startups in that sector based on the observed quality of the drawn startup, using Bayesian updating. The VC's objective is to maximize the likelihood of reaching an exogenous threshold return within a finite number of draws.[8]

To study the impact of human capital breadth, I differentiate VCs into broad and narrow types in the model, based on their prior beliefs about the probability distribution of startup

[6]Malenko et al. (2024), for example, provide a theoretical argument and empirical evidence that a single partner 'championing' an early-stage investment is optimal under the 'catching outliers' model in VC.

[7]For a survey of this literature in economics, see, e.g., (Bergemann and Valimaki, 2006). In this paper, I rely on the literature studying discrete-time Dirichlet bandit problems; see, e.g., (Berry and Fristedt, 1985).

[8]The exogenous threshold can be thought of as the minimum return required for the VC to raise a subsequent fund.

quality across sectors. Narrow-background VCs start the investment process with a strong and correct prior about the distribution of startup quality in a single sector—the one in which they specialize. They hold a weak but correct prior about other sectors. Broad-background VCs have a moderately strong and correct prior about all sectors of the economy.

I solve the model numerically using a standard value iteration algorithm and simulate the optimal policy for each type of agent (broad- and narrow-background VCs). Following the optimal policy, I show that broad VCs are more likely to sample multiple sectors and to pick novel ones. They experience more failures early in the investment process, which forces them to explore sectors with skewed returns later on in order to reach the required threshold.[9] In terms of performance, this translates into more failures and more IPO outcomes. Narrow-background VCs are more likely to stick to their sector of specialty, less likely to explore novel sectors, and consequently more likely to achieve intermediate outcomes. Following the optimal policy, the overall performance—measured as the likelihood of reaching the required threshold—of broad and narrow VCs is the same. Broad VCs achieve the threshold with a few IPO exits and many failures, whereas narrow VCs rely on intermediate outcomes.

Overall, my findings highlight the critical role of human capital breadth in fostering exploration and financing novel ventures. The observed decline in aggregate novelty among VC-financed startups over time can be explained by two potential factors. One perspective, consistent with Bloom et al. (2020), suggests that as startup business models become less novel, VC firms adapt by specializing, which leads to a decline in broad expertise among partners. Alternatively, as Lerner and Nanda (2020) argues, the VC industry has increasingly focused on narrower investment scopes, prioritizing ventures that align with institutional investors' preferences and risk profiles. This shift may constrain support for a broader range of innovative opportunities and limit financing for novel business models. From this perspective, my findings suggest that policy interventions supporting the development of human capital breadth could be beneficial. For example, investments in multidisciplinary education—integrating liberal arts, entrepreneurship, and technology coursework into business and finance curricula—could help cultivate the diverse skill sets necessary for evaluating novel startups. Business schools and executive education programs could further reinforce this by promoting cross-disciplinary training for aspiring venture capitalists.

[9] This pattern is akin to findings in the mutual fund tournaments literature (e.g., Kempf and Ruenzi (2008)).

2 Related Literature

My primary contribution is 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. 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). While much of the early literature on venture capital focuses on firm-level attributes such as reputation, syndication networks, and overall fund size (Hochberg et al., 2007; Gompers et al., 2008), recent research has begun to zero in on the partner-level drivers of performance. Studies examining how individual venture capitalists' skills and backgrounds influence deal success frequently emphasize the significance of partner-level human capital. Ewens and Rhodes-Kropf (2015), for example, show that differences among partners within the same VC firm have a substantial impact on investment outcomes. Likewise, Nahata (2008) links partner experience to investment performance, suggesting that personal track records and industry knowledge play a vital role in building VC reputation. 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.

In addition to contributing to the literature on venture capital and portfolio choice, I also build on research in labor economics and finance by highlighting instances where a broad, generalized skill set can be advantageous. Lazear (2004)'s "Jack of all trades" theory argues individuals with more balanced skill set are more likely to become successful entrepreneurship. 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 drive greater innovation. I extend this literature by showing that venture capitalists with broader human capital are more likely to identify and finance novel startups and to facilitate their successful exits.

I also 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.

3 Data sources, Measurement and Stylized Facts

3.1 Data Sources

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

3.1.1 PitchBook

I obtain information on VCs' portfolio choices and performance from PitchBook, obtained 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 2000 and 2021, where deal coverage in PitchBook is representative Retterath and Braun (2020). 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 DealClass 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 (2000–2021).

For each financed startup, I classify the exit types using the data provided by PitchBook as "IPO" exits, "M&A" exits, or "Failure." 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.[10] I define a startup exit as a "Failure" for a given VC investor–portfolio company pair if the VC-funded company has not exited via an IPO or an M&A and has not received any follow-up investment.

3.1.2 Crunchbase

I supplement the PitchBook dataset with Crunchbase. Unlike PitchBook accessed through WRDS, Crunchbase provides extensive information on jobs and employees of VC firms, as well as background information and textual descriptions of VC partners' career trajectories. To obtain background characteristics of VC firm partners from PitchBook and Crunchbase, I first match VC firms from PitchBook to Crunchbase directly on investor names after removing punctuation and converting the names to lowercase characters. If no exact match is found, I employ a fuzzy matching algorithm, where I match investor names conditioned on the same investor headquarters state or country, based on Levenshtein edit distance, similar to (González-Uribe, 2020). I only keep the top 10th percentile of matches based on the distance closeness metric and manually verify the correctness of such matches. If I obtain a reliable match, I match the lead partners recorded in PitchBook to partners recorded in Crunchbase within the same matched investor, based on first and last name, keeping only exact matches to avoid false positives. Whenever available, I collect other background partner-level characteristics from Crunchbase, such as age, gender, and educational background.

[10] 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."

3.1.3 USPTO

To measure innovation by VC-financed companies, I supplement the data with deal-level data on patent applications and grants from the USPTO (United States Patent and Trademark Office). The USPTO 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). Table 2 presents summary statistics for the main deal-level sample.

3.2 Measurement and Stylized Facts

3.2.1 Measurement of Novelty

To measure the extent to which startups are novel, I rely on startups' business descriptions provided by PitchBook. I rely on recent advances in Natural Language Processing (NLP) and use the business text description of a startup to construct an embedding vector using state-of-the-art OpenAI embeddings. Embeddings generated by large language models (LLMs) possess the property that similar texts are represented by vectors that are closest in vector space. OpenAI embeddings are particularly suitable for analyzing VC data, as shown by Duevski and Bazaliy (2025).

I 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:

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

where $CosSim(C_c, C_j)$ is the cosine similarity between the embedding vector of focal company c and company j, $max_{j \in (t-5,t)} CosSim(C_c, C_j)$ specifies that I take the maximum cosine similarity of the focal company c to any other startup that has received venture financing in the past five years prior to year t. In particular notice that according to this definition $N_{c,t} = 0$ if the company c has received venture backed financing in the past five years. Intuitively, (1) captures how distinct is the business model of company c to any other company which 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. This definition is related but different from Bonelli (2022)'s "backward similarity measure," which captures how similar a startup is, on average, to what the venture capital industry has financed in the past five years. In particular the $N_{c,t}$ captures how distinct the focal startup is from the financed firm with the closest business model and arguably it is better suited to identify the true novelty of the business model as opposed to, for example whether the business model is niche. I term the measure $N_{c,t}$, Novelty (Distance to Closest Firm).

3.2.2 Stylized Facts About Novel Startups

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

Figure 1 about here.

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 Novelty (Distance to Closest Firm) measure and document a decreasing mean novelty over time. Notice 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.

Figure 2 about here.

To better understand the time evolution of novelty, in the left panel of Figure 2, I plot the time evolution of the distribution of novelty during my sample period. In the right panel of Figure 2, I plot the evolution of the mean novelty, computed using the top 10 most novel startups in each year. The patterns suggest that even though the average and median novelty have declined over time, the novelty of the most novel startups has remained roughly constant over the sample period. In other words, in each year, there is a fraction of venture-backed startups with very high novelty.

Figure 3 about here.

Figure 4 about here.

In Figures 3 and 4, I split the time trend of mean novelty computed using the full sample, and mean novelty across the most novel firms by various regional classifications based on startup location. First, I document a decline in mean novelty across all regions, and even though the mean novelty of startups in EMEA, and notably the APAC region, is higher on average, the most novel firms seem to be created in North America.

Figure 5 about here.

In Figure 5, I plot the mean IPO rate and the mean M&A rate over time for different quartiles of novelty (Distance to Closest Firm). Most of the IPO exits in my sample period are concentrated in the most novel firms receiving financing each year, while most of the M&A exits are concentrated among the least novel firms.[11]

In a regression setting I document two stylized facts about novel startups. First, novel startups are more likely to fail or achieve a major exit. Second, novel startups contribute more to innovation output.

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:

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

where d denotes a deal, i the industry of the deal, t the year the deal is made, s the stage of the deal and k the investor (VC Firm) financing the deal. $P_{d,i,t,s,k}$ is a performance outcome indicator which can be failure, major Success or IPO Exit. N_d the is novelty of the deal, $\eta_{i \times t \times s \times k}$ denotes granular industry $\times$ time $\times$ deal Stage $\times$ investor fixed effects. The coefficient of interest β captures the association between deal novelty and deal performance.

Table 5 about here.

[11] The decreasing time trend of the IPO rate and M&A rate is, to a large extent, mechanical since firms that receive their first venture round in later periods need more time to exit.

The results are presented in Table 5. More novel startups are significantly more likely to both fail and achieve a major success. Notice that the variation in columns (4)-(6) comes from within an Industry $\times$ Time $\times$ Deal Stage $\times$ Investor; that is, we are comparing the performance of more or less novel startups financed by the same investor in the same stage, deal year, and industry. This suggests that the results of this stylized fact are not simply driven by certain investor characteristics that have been shown to be associated with startup success (i.e., VC firm reputation, experience, deal flow). Similarly, the performance is not merely driven by time-varying industry or overall economic conditions (i.e., the hotness of the M&A and IPO market in general or industry-specific shocks). In terms of economic magnitude, estimates in columns (4) and (6) imply that a one standard deviation increase in novelty is associated with a 0.8% increase in the probability of failure and a 5.6% increase in the probability of an IPO.[12]

Table 6 about here.

In Table 6, 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 (4)-(6) imply that the probability of failure increases by 0.3% and the probability of an IPO exit increases by 11.4% when moving from the bottom to the top quartile of novelty. Intuitively, in columns (4)-(6), I am evaluating the performance of startups financed by the same investor in the same year, industry, and deal stage, relying on variation in startup novelty.

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:

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

where $I_{d,i,t,s,k}$ is a forward 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).[13]

[12]Calculated as Point Estimate $\times$ SD in Novelty (Distance to Closest Firm Measure)

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

Table 7 about here.

The results are shown in Table 7. I use (3) and estimate a Poisson count model (to avoid well-known issues with using the log(1+) model Chen and Roth (2024)). Estimates in column (5) imply that a one standard deviation increase in novelty is associated with a 20.8% increase in expected forward citations.[14]

3.2.3 Measurement of Breadth of Human Capital and Stylized Facts

To construct a human capital breadth index at a VC partner level I rely on textual analysis.

Table 4 about here.

First, to systematically classify career trajectories as narrow or broad, I utilize a text-based embedding approach. We define two prototype sets: one representing narrow career trajectories (specialized career paths within a single domain) and another representing broad career trajectories (experiences spanning multiple functions and industries). The sets of prototypes for narrow and broad career trajectories are shown in Table 4. I then construct OpenAI embeddings of each of these prototypes and compute a centroid for narrow and broad human capital, which is simply the average of the embedding vectors of those prototypes, defined as:

$$c_p = \frac{1}{5} \sum_{i \in p} e_i, \tag{4}$$

where p stands for a prototype, which can be "Narrow" (N) or "Broad" (B), and e_i is the embedding vector of a given prototype. I then compute the cosine similarity of the textual description of the partner's career trajectory to each centroid. My final measure of the human capital breadth index at the partner level is given by:

$$B_i = \frac{CosSim(c_B, v_i) - CosSim(c_N.v_i)}{CosSim(c_B, v_i) + CosSim(c_N.v_i)}, \tag{5}$$

where i stands for a partner, B_i is the breadth index for partner i, and v_i is the embedding vector constructed from the textual description of partner i's background. CosSim(c_B, v_i) is the cosine similarity between partner i's background and the broad centroid, and CosSim(c_N, v_i) is the cosine similarity of partner i's background to the narrow centroid. Intuitively, the breadth

[14]Calculated as 2.7×0.07=0.189, $e^{0.189} = 1.208$.

index measure defined by (5) captures how close a given partner's career trajectory is to a broad prototype relative to a narrow prototype.

Figure 6 about here.

On the right panel of Figure 6, I plot the frequency distribution of the breadth index measure (standardized to have a mean of 0 and a standard deviation of 1). On the left panel, I document a decrease in the mean human capital breadth of partners leading deals in the VC industry.

Correlates of Breadth Index:

Figure 7 about here.

In Figure 7, I split the breadth index measure mean by four partner-level characteristics. On average, PhD graduates have lower human capital breadth than non-PhD graduates (top left). Partners with an MBA degree have a higher breadth index than partners without one. Immigrant background partners have, on average, slightly higher human capital breadth (bottom left), and female partners have higher human capital breadth (bottom right).

Table 8 about here.

In Table 8, I correlate my breadth index measure with various partner characteristics. The breadth index is positively correlated with immigrant background, female gender, and having an MBA degree, and negatively correlated with a PhD degree, age, and having a computer science undergraduate degree. VC partners working for firms based in California, Massachusetts, and New York have a higher breadth index than VC partners working in VC firms headquartered in other U.S. states.

4 Empirical Results

In this section I present the main empirical findings of the paper.

4.1 Partner level Human Capital Breadth and Novel Startups

First without claiming causality, I present two novel facts that link individual partner's human capital breadth to selection and performance of novel startups.

4.1.1 Association between Lead Partner Breadth Index and Startup Novelty

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

$$N_{j,k,p,t} = \alpha + \beta B_p + X_{t,p} + \eta_{i\times t\times s} + \rho_j + \epsilon_{j,k,p,t}, \tag{6}$$

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

Table 9 about here.

The results are presented in Table 9. 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, 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 different novelty within a given VC firm.

4.1.2 Interaction between lead partner breadth and deal novelty and performance

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

$$P_{j,k,p,t} = \alpha + \beta B_p + \gamma N_{j,k,p,t} + \delta B_p \times N_{j,k,p,t} + X_{t,p} + \eta_{i\times t\times s} + \rho_j + \epsilon_{j,k,p,t}, \tag{7}$$

where $P_{j,k,p,t}$ is a deal level performance measure for an investment made by firm j in startup k with lead partner p at time t. $N_{j,k,p,t}$ is the deal novelty and B_p is a lead partner breadth index measure, $B_p \times N_{j,k,p,t}$ is an interaction term between lead partner breadth index and startup novelty. $X_{t,p}$ are time varying partner level controls measured at the time the deal is made, $\eta_{i \times t \times s}$ represents industry $\times$ time $\times$ deal stage fixed effects and ρ_j is an investor fixed effect.

Table 10 about here.

The results are presented in Table 10. Intuitively, in column (3), the coefficients are estimated by relying on variation in the partner breadth index and novelty across deals made by the same VC firm, in the same industry, deal stage, and year. The granular fixed effects rule out a story where the effect is driven by any VC firm-specific factors, such as VC firms attracting better deal flow while simultaneously hiring better partners. The fixed effects also rule out time-varying macroeconomic or industry-wide shocks that may influence certain performance outcomes, making them more or less likely and thus driving up novelty. The economic magnitude of β in column (3) implies that non-novel deals financed by broad partners perform worse than non-novel deals financed by narrow partners. At zero novelty, a one standard deviation increase in the breadth index is associated with a 5% decrease in the probability of an IPO. However, as novelty increases, the positive interaction term implies that the probability of an IPO for deals led by broad partners increases with deal novelty. For each 0.1 increase in novelty, the effect of the breadth index increases by 2%. These estimates, in particular, imply that the association between the breadth index and IPO exit is negative for below-median novelty deals (median of 0.24); however, it turns positive for above-median novelty firms.

Robustness: I conduct several robustness tests. First, in Table A7, I re-estimate specification (6) using an alternative novelty measure based on the average distance to the five closest competitors—the conclusions remain unchanged. In Table A8, I re-estimate (6) by removing the bottom decile of partner background descriptions, and the conclusions remain unchanged. In Table A9, I re-estimate specification (7) using an alternative measure of novelty. In Table A10, I re-estimate (7) by excluding the bottom decile of partner background descriptions in terms of length, and the conclusions remain unchanged. In Table A11, I re-estimate specification (7) by splitting the novelty measure into quartiles of novelty.

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

To study the causal impact of human capital breadth on the performance of novel startups, I first restructure the dataset at the deal level in the following way. For each deal d made by VC firm j, I construct a set of potential partners who could have led this deal. As a baseline, I select all partners who have led deals at the same VC firm within a -3 to +3 year window around the deal.[15] For each partner I construct a busyness proxy at the time of the deal. My proxy for busyness follows Abuzov (2019) and I define a partner p to be busy at time t if the same partner p is involved in an exit via an acquisition or an IPO for another deal in a (-90, +90) days time window around the time of the focal deal t. Using this data structure, I first provide an alternative way to show an association between lead partner breadth and the likelihood of being a lead partner on a novel deal. I estimate the following specification at the deal-investor level:

$$PartnerChosen_{d,j,p} = \alpha + \beta B_p + \gamma B_p \times N_d + Xp,t + \eta_d + \rho_j + \epsilon_{d,j,p}, \tag{8}$$

where $PartnerChosen_{d,j,p}$ is an indicator variable taking a value of 1 if partner p working for investor j leads deal d and 0 otherwise. Xp,t is a set of partner level controls measured at the time the deal is made. B_p denotes the partner breadth index measure N_d denotes the deal novelty and $B_p \times N_d$ is an interaction term between the deal's novelty and the lead partner's breadth index, η_d is a deal fixed effect and ρ_j is a partner fixed effect. Notice that in specification (8) the deal novelty is not included as it is absorbed by the deal fixed effect.

Table 11 about here.

The results are shown in Table 11. The advantage of this data structure is that it allows for the inclusion of a deal fixed effect, which absorbs all underlying deal-level characteristics (size, syndication, year of financing, quality). Estimates in column (4) for the control variables have the expected sign and significance. Partners with more experience are more likely to lead deals, while older partners (controlling for experience) and busy partners are less likely to lead deals. The base coefficient on the breadth index indicates that partners are less likely to lead less novel deals. The interaction terms between breadth and novelty indicate that the higher the novelty of the deal, the higher the likelihood that a broad partner is chosen to lead the deal. Intuitively,

[15] I run several robustness tests varying the timing around the deal; e.g., see Table A14.

the interaction specifies the change in the slope of the breadth index across deals of different novelty levels.

The estimate for the base breadth index coefficient (β) implies that for non-novel deals 1 standard deviation increase of breadth index leads to a 2.5 % decrease in the probability that a partner leads a deal. The interaction term (γ) implies that for high novelty deals - with novelty above 0.3 the coefficient on human capital breadth becomes positive and partners with high human capital breadth are more likely to lead novel deals.[16]

4.2.1 The causal impact of human capital breadth on Novel Startup's performance

In Section 4.1.2, we presented a positive association between the interaction of human capital breadth and deal novelty with performance. Even though specification (7) includes granular VC firm fixed effects and various fixed effects related to startup characteristics (Industry × Year × Deal Stage), the assignment or selection of startups, even within a VC firm, is non-random. For example, if deal quality and partner quality within a VC firm are heterogeneous, one may worry that the interaction effect captures a pure quality-matching story—i.e., better-suited partners for novel firms are matched with the highest-quality novel firms, which drives the increase in the probability of an IPO. To capture the causal effect, one would need an instrument that randomly shifts the assignment of a given partner to a deal within a VC firm. In other words, one would like to compare how the same novel deal would perform given that it is randomly assigned to a broad versus a narrow partner and then compare the outcome.

In this section, I propose an instrument that relies on the time-varying availability of partners within a VC firm. Intuitively, when a given deal is made, some partners will be more available than others, providing a natural shifter in the probability of a deal being assigned to specific partners. The results presented in Table 11 suggest that, controlling for other partner characteristics, busy partners are less likely to be assigned to lead a deal. Hence, a natural instrument for partner assignment is the busyness of the partner. To shift the probability of a partner with high (low) human capital breadth being assigned to a given deal, I propose an instrument that relies on time variation in human capital breadth availability within a given VC firm, called the average available breadth index.

[16] Calculated as Novelty cut-off = $\frac{0.025}{0.084}$

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 :

$$AvgB_{j,t} = \frac{\sum_{p \in j} B_p \times I_{p,t}}{\sum_{p \in j} I_{p,t}}, \tag{9}$$

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. 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 $\times$ novelty interaction which will be the average available breadth index $\times$ deal novelty. Given these two instruments, I estimate the following model via a 2SLS:

$$\textbf{First Stage: } B_{d,p,j,t} = \alpha + AvgB_{j,t} + AvgB_{j,t} \times N_{d,t} + X_{p,t} + \eta_{i \times t \times s} + \rho_j + \epsilon_{d,p,j,t} \tag{10}$$

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

$$\textbf{Second Stage: } P_{d,p,j,t} = \alpha + N_{d,t} + \widehat{B_{j,t}} + \widehat{B_{j,t} \times N}_{d,t} + X_{p,t} + \eta_{i \times t \times s} + \rho_j + s_{d,p,j,t}, \tag{12}$$

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 $\times$ year $\times$ deal stage fixed effects and ρ_j is an investor fixed effect. $P_{d,p,j,t}$ is a performance outcome of a deal which is either IPO exit or failure.

Table 12 about here.

The results are presented in Table 12. Columns (1) and (2) present the first-stage estimates from specifications (10) and (11). 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. Columns (3) and (5) present the IV estimates from specification (12), where the performance outcome is either an IPO Exit or a Failure. First, the F statistic for the instrument passes the weak instrument threshold (as also seen in columns (1) and (2)). 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.

4.3 Fund Level and Investor Level Results

The partner-level analysis presented thus far demonstrates that, within VC firms, partners with broader backgrounds are more likely to lead novel deals and enhance the likelihood of successful exits for these innovative investments. However, investment decisions within VC firms are rarely made by individual partners acting in isolation; instead, these decisions often require approval by multiple partners.[17] Similarly, any positive impact on portfolio company outcomes post-investment may result from the collective efforts of several partners, not solely the lead partner. Additionally, compensation incentives within VC firms are usually structured at the fund level rather than being tied to individual partner-deal outcomes, meaning that all partners within a fund share responsibility for overall investment performance.

In this section, we examine whether these individual partner-level effects extend to the fund and investor levels. Specifically, we investigate whether funds and investors characterized by higher average human capital breadth are systematically more inclined to finance novel deals and whether the average fund or investor human capital breadth enhances performance of novel firms.

Figure 8 about here.

First, in Figure 8, I document several novel stylized facts about the generalist and specialist composition of venture capital partnerships. The top-right panel plots the distribution of the

[17] Malenko et al. (2024) provide a theoretical argument and empirical evidence that shows the importance of individual partners 'championing' a given deal especially in the context of early stage investment.

fraction of specialist partners (partners with a negative Breadth Index) within a fund, comparing funds that made their last investment before 2015 with those that made their first investment after 2015. First, as seen in the figure, very few funds are composed entirely of either specialists or generalists; the median fund has a mixed composition. Second, the distribution of specialists in newly created funds is much higher than in funds that made their last investment before 2015, which aligns with the downward trend in the average human capital breadth in the VC industry, as shown in the right panel of Figure 6. This pattern also holds at the investor level, as shown in the bottom panels of Figure 8.

To test whether average human capital breadth is associated with the average novelty of a deal in the partnership we estimate the following specification at a fund level.

$$\text{Avg}N_i = \alpha + \text{Avg}B_i + \text{Fund Controls}_i + \text{Industry Composition}_i + \text{Stage Composition}_i + \eta_{t\times v\times c} + \epsilon_i, \tag{13}$$

where $\text{Avg}N_i$ is the average novelty of deals made by the fund i, $\text{Avg}B_i$ is the average human capital of partners leading deals in the fund. Industry Composition_i is a set of controls for the industry composition of investments for fund i, Stage Composition_i is a set of controls for the composition of stages in deals of fund $i.\eta_{t\times v\times c}$ are Vintage Year × Fund Type × Fund Country fixed effects.

Table 13 about here.

The results are presented in Table 13. In columns (1) and (2), on the left-hand side, I use the average novelty of the deal (two distinct measures). In column (3), the outcome variable is the fraction of deals that fall in the top quartile of novelty in a given deal year. Across columns (1) and (2), the average breadth index across partners is positively associated with average deal novelty. Furthermore, funds with a higher average breadth finance more deals in the top quartile of novelty.

Table 14 about here.

In Table 14, I show that a similar pattern holds for newly created VC firms. I estimate a similar specification as in (13), but now I average partner breadth and deal novelty over the first six years of investment by newly created VC firms. The significance and economic magnitude of the effect remain similar.

Next, I estimate a model similar to (7) with the variables of interest aggregated at a fund or investor level. Specifically I estimate,

$$\text{Perf}_i = \alpha + \text{Avg}B_i + \text{Avg}N_i + \text{Avg}N_i \times \text{Avg}B_i + \text{ Controls}_i + \eta_{t\times v\times c} + \epsilon_i, \tag{14}$$

, where Perf_i is a fund level performance measure which is either the fraction of fund's i deals that have exited via an IPO or have failed. The other variables are defined as in (13) and I use the same set of controls.

Table 15 about here.

The results are presented in Table 15. In columns (1) and (2), for both the base level and interaction, I use the average deal novelty, whereas in columns (3) and (4), I use the fraction of a fund's deals in the top novelty quartile. Across all columns, the baseline coefficient on the average human capital partnership breadth is negative and significant. This shows that funds with high average human capital breadth but low-novelty deals are less likely to achieve IPO exits and more likely to fail. The interaction term between novelty and human capital breadth is positive and significant in columns (1) and (3) (Fraction of IPO Exits) and negative and significant in columns (2) and (4) (Fraction of Failed Exits), implying that the performance of high-average human capital breadth funds improves with average deal novelty (beyond the baseline novelty term).

Table 16 about here.

In Table 16, I present similar findings when I aggregate outcomes at an investor level for early investors.

5 Theoretical Framework

The empirical results documented thus far can be rationalized by the exploration-exploitation trade-off formalized in the theoretical model developed in this section. In particular, the empirical finding that broader-background VC partners are more likely to lead investments in novel, high-risk ventures and to achieve superior performance outcomes in these cases aligns with the model's prediction that broader-background investors optimally explore novel sectors. The model characterizes broader-background investors as agents who hold moderately strong priors across multiple sectors, prompting them to incur early failures but ultimately enabling the identification of highly successful startups. However, it is important to emphasize that this theoretical interpretation represents just one plausible explanation among several for the empirical patterns observed.

To model the portfolio choice of venture capitalists and rationalize the empirical findings, I draw on the multi-armed bandit framework, which captures the dynamic trade-off between exploration and exploitation (Bergemann and Valimaki, 2006). In this model, each "arm" represents a distinct industry sector with varying probabilities of yielding high, intermediate, or low returns. The model assumes that sectors are heterogeneous: established sectors offer stable but limited returns, while novel sectors are more volatile, with a higher probability of both exceptional successes and failures. At each decision point, the VC must choose the sector in which to invest, balancing between sectors with high expected returns (exploitation) and those with greater uncertainty but higher potential payoffs (exploration). After each investment, the VC can observe a signal about the quality of the startup drawn from that sector and update their belief about sectoral potential, reflecting a Bayesian learning process.

In the model, I differentiate VCs by their human capital backgrounds—broad versus narrow—which impacts their approach to risk and exploration. Broad-background VCs, with diverse industry experience, start with a moderately informed understanding of multiple sectors, encouraging them to explore novel, high-risk sectors that could yield exceptional returns. In contrast, narrow-background VCs, who are specialized in a particular domain, hold stronger priors about specific sectors and are more inclined to exploit known opportunities, focusing their investments within familiar industries. This distinction is central to the model, as it influences both the breadth of portfolio diversification, the types of startups selected, and the subsequent performance.

In this section, I present the model in detail. I begin by describing the venture supply framework, outlining the distribution and quality of ventures across sectors. Next, I define the VC's objective function and solve the model under the assumption of full information, where the VC knows the quality distribution of ventures in each sector. Following this, I introduce the distinction between VC types based on their human capital background: narrow-background VCs are modeled as starting with a strong prior in their area of specialization, whereas broad-background VCs begin with weaker priors across multiple sectors. Finally, I solve for the optimal investment strategy for each VC type, highlighting how these differences in prior knowledge shape their approach to sectoral allocation, likelihood of financing novel projects, and subsequent performance.

5.1 Supply of Ventures

I model the supply of available ventures in the following way. There are N available sectors in the economy indexed with i. In each sector N_i, there are startups of high quality, intermediate quality or low quality to be financed. If the startup is of low quality and is financed it yields a payoff of x_{low} to the investor. A startup of intermediate quality yields a payoff of $x_{int} > x_{low}$ and a startup of good quality yields a payoff of $x_{high} > x_{int}$.[18] Each sector is described by a multinomial distribution which specifies the probability of high, intermediate and low quality startups inside that sector. A draw from sector N_i yields a high quality startup with a probability $P_i(x_{high})$, an intermediate quality startup with a probability $P_i(x_{int})$ and a low quality startup with a probability $P_i(x_{low})$, where $P_i(x_{low})+P_i(x_{int})+P_i(x_{high}) = 1$. The sectors are heterogeneous with respect to the distribution of startup quality but not with respect to the payoffs given a high, intermediate and low quality startups. One of the N sectors is a "novel" sector (i.e. a sector of the economy that has not yet been financed - cite papers to justify this assumption). The novel sector, has a highly skewed distribution of startup quality i.e. sampling from this sector yields a higher chance of encountering a low quality startup than any other sector, but also a higher change of encountering a startup that is of high quality. We label this sector with a special index e and based on the discussion above I assume, $P_e(x_{low}) > P_i(x_{low})$ and $P_e(x_{high}) > P_i(x_{high}) \forall i \neq e$.

Figure 9 about here.

[18] In venture capital investments, a low quality startup is a startup that fails. An intermediate quality startup is a startup that can be exited via an acquisition and a high quality startup is a startup that can be exited via an IPO.

5.2 Investor's objective function and Benchmark strategy

The agent, venture capital investor is assumed to maximize the probability of reaching an exogenous threshold of returns T after n draws. For simplicity I assume that the agent is risk-neutral and has a discount factor of 1. In each draw the choice set of the agent is which urn to draw from. In the baseline problem we assume that the agent is fully informed about the distribution of ventures in each sector and once an agent selects a venture the quality of the venture is observed immediately before the agent takes the next action.[19] This is a recursive dynamic programming problem and at each step we define the value function as follows. Let $V^*(d, R)$ denote the maximum probability of reaching or exceeding a the threshold T where the state variable d denotes the number of draws left and the state variable R denotes the cumulative rewards of the agent in the past $n-d$ draws. The Bellman equation at each state (d, R) is given by:

$$V^*(d, R) = max_{i \in [1,N]} \left(\sum_x P_i(x) V^*(d-1, R+x) \right), \tag{15}$$

where:

- $V(d, R)$ is the maximum probability of reaching or exceeding the threshold T with d draws remaining and accumulated returns R,
- $P_i(x)$ is the probability of receiving payoff x from drawing from urn i,
- $i \in [1, N]$ denotes the choice between sectors.
- x represents the possible payoff outcomes from sector i, where $x \in x_{low}, x_{int}, x_{high}$
- $V(d-1, R+x)$ is the continuation value with $d-1$ draws left and accumulated returns $R+x$.

The boundary condition for the value function is:

$$V(0, R) = \begin{cases} 1, & \text{if } R \geq T, \\ 0, & \text{if } R < T. \end{cases}$$

Figure 10 about here.

[19] The results do not change if the agent observes a signal about the startup quality once the startup is drawn as long as this signal is informative.

I illustrate the decision process of the venture capitalist in the complete information case in figure 10. Notice that this is a Markov process as the future decisions of the agent depend only on the current state realized state. In general, we cannot obtain an explicit analytical solution for the decision process described by equation (15). In the following subsection I solve the model numerically using value function iteration and provide intuition for the optimal policy as well as comparative statics results.

5.2.1 Benchmark strategy results

To solve the model numerically, we need to have the distribution of payoffs of all N sectors of the economy. To highlight the main intuition behind the model, the optimal policy followed and the comparative statics results I solve the model numerically under the simplest possible distributional assumptions that highlights the main-trade-off the agent faces. To do so, I will assume that the economy is composed of three sectors and that the probability distribution in each sector is given by table 1. I have assumed that the payoffs of low, intermediate and high outcome are 0, 1 and 4 respectively. Sector 1 is a sector, with the highest expected value and the highest probability of achieving an intermediate outcome, sector 2 is a dominated sector which should never be chosen and sector 3 is a novel sector that has a lower expected value than sector 1, but a higher probability of achieving a high reward. Following optimal policy under full information the agent will choose which of these 3 sectors to draw from each time. To solve the Bellman equation numerically I use a standard value iteration algorithm and then I simulate the optimal policy and average over the number of simulations.

In figure 11, I plot the likelihood of reaching the required threshold under the optimal policy as a function of the threshold value for different number of draws (n). First, as expected across all threshold values the likelhood of reaching the threshold is higher if the agent draws more times. Second, the likelihood of reaching the required threshold return drops faster for smaller values of n the number of draws.

Figure 11 about here.

Comparative statics with respect to Threshold required on Frequency of Sector Selection under optimal policy The main insight of the benchmark strategy is that keeping payoffs and probabilities of outcomes across each sector fixed, increasing the threshold required incentivizes the agent to take more risk and draw more often from the novel sector where the

Table 1: Reward Distributions for Each Sector

Sector	Payoff	Probability
1 (High Likelihood of Intermediate Outcome)	0	0.7
	1	0.25
	4	0.05
2 (Dominated Sector)	0	0.8
	1	0.2
	4	0.0
3 (Novel Sector)	0	0.8
	1	0.1
	4	0.1

probability of achieving a good outcome is higher. In figure 12, I plot the frequency of selection for each sector when the agent follows the optimal policy. As expected the Dominated sector is never chosen, since the agent can always do better by picking the sector with a high likelihood of an intermediate outcome. For very low values of the threshold the agent chooses to only draw from the high likelihood of intermediate outcome sector. In this case the threshold is low enough and the agent can safely reach the threshold by achieving multiple intermediate successes. As the threshold value increases the agent picks the novel sector more frequently and after a certain threshold prefers to draw more frequently from the novel sector. As shown in figure 12, the cut-off point at which it becomes optimal to draw from the novel sector more frequently is lower as the number of draws decreases. Having less draws means that the agent needs to reach the threshold quicker and this incentivizes the agent to draw more from the novel sector.

Figure 12 about here.

5.3 Investor's objective function under incomplete information

In the previous subsection we considered the benchmark problem where the agent knows the distribution in each sector perfectly. When allocation capital and deciding which startups to finance VC investors do not have access to full information about the potential quality and distribution of rewards within all sectors. The incomplete information setting allows us to clearly draw a distinction between broad vs. narrow human capital VCs. At the start of the investment process, these two VC types will differ regarding their prior beliefs about the distribution of startups across sectors. Specificity, we assume that narrow background VC have a strong and correct prior about the distribution of startup quality in one sector - the sector where they are specialized in. They will have a correct, but very weak prior regarding the distribution of startup quality across other sectors. Broad VCs in turn will have a moderately strong and correct prior across all sectors in the economy. Each VC type will follow her optimal investment strategy and update their beliefs about the distirbution of quality in a Bayesian way after drawing from a specific sector.

5.3.1 Description of the incomplete information setting

Agent's prior and updating rules. Since the outcomes in each sector are assumed discrete the agent will have a Dirichlet prior over sector i described by a vector $\alpha_i = (\alpha_i(x_{low}), \alpha_i(x_{int}), \alpha_i(x_{high}))$.[20] Intuitively, the α parameters capture the past experience of the agent in the following way. An $\alpha_i = (\alpha_i(x_{low}), \alpha_i(x_{int}), \alpha_i(x_{high}))$ vector for sector i would mean that the agent has drawn $\alpha_i(x_{low}) + \alpha_i(x_{int}) + \alpha_i(x_{high})$ times from sector i and among those draws the agent has gotten $\alpha_i(x_{low})$,$\alpha_i(x_{int})$ and $\alpha_i(x_{high})$ intermediate outcomes. Given Dirichlet priors the prior belief for getting a startup of quality x when drawing from sector i startup for the agent who samples sector i will be given by:

$$P_i(x) = \frac{\alpha_i(x)}{\alpha_i(x_{low}) + \alpha_i(x_{int}) + \alpha_i(x_{high})} \tag{16}$$

Suppose the agent draws from sector i and an outcome x_{high} is observed. The posterior belief over the distribution in sector i follows a simple updating rule and is now given by $(\alpha_i(x_{low}), \alpha_i(x_{int}), \alpha_i(x_{high}) + 1)$. [21]

Let $\boldsymbol{\alpha_0}$ be an N dimensional vector, where each component is a vector of the agent's prior about the quality distribution in a given sector.

Similarly to the complete information case, the agent's value function $V(d, R, \boldsymbol{\alpha_{n-d}})$ represents the maximum probability of reaching or exceeding the threshold T with d steps left,

[20] Dirichlet distributions are conjugate priors of the multinomial distribution and follow very simple updating rules. I describe and provide an introduction of dirichlet distributions in the appendix B

[21] Proof in appendix B

given an accumulated reward R and the agent's posterior $\boldsymbol{\alpha_{n-d}}$ after $n-d$ draws. The Bellman Equation new reads:

$$V^*(d, R, \boldsymbol{\alpha_{n-d}}) = \max_{i \in \{1,\ldots,N\}} \left(\mathbb{E}\left[V^*(d-1, R+x_i, \boldsymbol{\alpha_{n-d+1}}) \mid \boldsymbol{\alpha_{n-d}}\right]\right) \tag{17}$$

where x_i represents the potential payoffs from sector i ($x_i \in \{x_{low}, x_{int}, x_{high}\}$) and $\boldsymbol{\alpha_{n-d+1}}$ is the dirichet posterior in with $d-1$ steps remaining after observing the outcome in step $n-d$ in the optimally chosen sector.

The expected value of the function if sector i is chosen can be expressed explicitly as:

$$\mathbb{E}\left[V(d-1, R+x_i, \boldsymbol{\alpha_{n-d+1}}) \mid \boldsymbol{\alpha_{n-d}}\right] = \frac{\alpha_{n-d,i}(x_{low})}{\alpha_{n-d,i}(x_{low}) + \alpha_{n-d,i}(x_{int}) + \alpha_{n-d,i}(x_{high})} V(d-1, R+x_{low}, \boldsymbol{\alpha_{n-d+1}})$$
$$+ \frac{\alpha_{n-d,i}(x_{int})}{\alpha_{n-d,i}(x_{low}) + \alpha_{n-d,i}(x_{int}) + \alpha_{n-d,i}(x_{high})} V(d-1, R+x_{int}, \boldsymbol{\alpha_{n-d+1}})$$
$$+ \frac{\alpha_{n-d,i}(x_{high})}{\alpha_{n-d,i}(x_{low}) + \alpha_{n-d,i}(x_{int}) + \alpha_{n-d,i}(x_{high})} V(d-1, R+x_{high}, \boldsymbol{\alpha_{n-d+1}})$$

where :

$$\alpha_{n-d+1,i} = \begin{cases} (\alpha_{n-d,i}(x_{low})+1, \alpha_{n-d,i}(x_{int}), \alpha_{n-d,i}(x_{high})) & \text{if failure (payoff } x_{low}) \\ (\alpha_{n-d,i}(x_{low}), \alpha_{n-d,i}(x_{int})+1, \alpha_{n-d,i}(x_{high})) & \text{if intermediate quality (payoff } x_{int}) \\ (\alpha_{n-d,i}(x_{low}), \alpha_{n-d,i}(x_{int}), \alpha_{n-d,i}(x_{high})+1) & \text{if high quality (payoff } x_{high}) \end{cases}$$

Figure 13 about here.

The investment process under incomplete information is depicted in figure 13. In the beginning the VC starts with $(d, R, \boldsymbol{\alpha}) = (n, 0, \boldsymbol{\alpha_0})$. Then following optimal policy, which is a solution equation (17), the agent picks a sector to invest in, observes the outcome, updates her prior belief about the distribution of startup quality and follows optimal policy thereafter.

Agent's types: Agents can be of two types, broad or a narrow agent. Intuitively, we would like to capture the fact that a narrow agent is a "specialist" in a given sector, but does not know much about anything outside of this sector, whereas a broad agent knows a "little" bit about all of the sectors in the economy. To capture this we assume that the agents are endowed with a different starting prior α before they start the sector selection process. Both the broad the narrow agent start with a correct prior about the distribution of quality of startups in each sector. The broad and the narrow agent will defer with respect to the strength of those priors. Given a sector $i \neq e$, a specialist in sector i is assumed to have a strong and correct prior over the distribution of outcomes in sector i. He has a very weak and correct prior for all other sectors in the economy. The broad agent has a correct and intermediately strong prior over the distribution of quality over all of the sectors in the economy. The construction of these priors reflect the fact that the specialist has had extensive experience in a given sector prior to becoming a VC founder whereas a generalist has had moderate experience across multiple sectors. We formalize these concepts as follows. Suppose we have a narrow background agent who is a specialist in sector i. Their prior, just before starting becoming a VC fonder and allocating capital will be given by:

$$\alpha_0^n = \begin{cases} \alpha_{0,i} = k(P_i(x_{low}), P_i(x_{int}), P_i(x_{high})) \\ \alpha_{0,j} = (P_j(x_{low}), P_j(x_{int}), P_j(x_{high})), \end{cases}$$

where α_0^n denotes the initial prior for the narrow agent, $P_j(x)$ denotes the probability of outcome x in sector j. $k >> 1$ and describes the strength of the prior of the specialist in his own sector relative to the other sectors in which he does not specialize. The strength of the prior will reflect how much the agent updates his prior about startup distribution after observing an outcome in that sector. The assumption $k >> 1$ captures the fact that any signal received after drawing from sector i has a much lower impact on the specialist agents' posterior beliefs in sector i (where the agent is specialized), then receiving the same signal in any other sector.

The prior of the broad background VC will be the same across all sectors and it will be given by:

$$\alpha_{0,j}^b = m(P_j(x_{low}), P_j(x_{int}), P_j(x_{high})),$$

where $\alpha_{0,j}^b$ is the broad agent's prior over sector j and $m > 1$. The ratio $\frac{k}{m} > 1$ is to be interpreted as the relative strength of the prior a narrow background VC specialized in sector i has over a broad background VC. $m > 1$ is to be interpreted as the relative strength a broad background VC has over a narrow background VC in all other sectors except sector i. To clearly understand the difference between the types of agents it is useful to compute the updated probabilities for the narrow agent and the broad agent in sector i (the sector where the narrow agent is specialized in) after drawing from sector i and observing a startup of for example low quality. The posterior after drawing from sector i and observing a startup of low quality for the narrow agent n is given by:

$$\hat{p}_i^n(x_{low}) = \frac{kP_i(x_{low}) + 1}{k(P_i(x_{low}) + P_i(x_{int}) + P_i(x_{low})) + 1} = \frac{kP_i(x_{low}) + 1}{k + 1} \tag{18}$$

The posterior for the broad agent after the same action and observing a startup of low quality is:

$$\hat{p}_i^b(x_{low}) = \frac{mP_i(x_{low}) + 1}{m(P_i(x_{low}) + P_i(x_{int}) + P_i(x_{low})) + 1} = \frac{mP_i(x_{low}) + 1}{m + 1} \tag{19}$$

Now the difference between the posterior for low quality outcome for the narrow agent in sector i will be:

$$\hat{p}_i^n(x_{low}) - P_i(x_{low}) = \frac{1 - P_i(x_{low})}{k + 1} \tag{20}$$

For the broad agent:

$$\hat{p}_i^b(x_{low}) - P_i(x_{low}) = \frac{1 - P_i(x_{low})}{m + 1} \tag{21}$$

Clearly, since $k > m$ we have $\hat{p}_i^n(x_{low}) - P_i(x_{low}) < \hat{p}_i^b(x_{low}) - P_i(x_{low})$ which means that the narrow agent's posterior differs less than the broad agent's posterior after observing additional signals in the sector where the narrow agent specializes.

5.3.2 Results for each VC type under incomplete information

In the following subsection, I solve for the optimal policy followed by each type of agent (narrow vs. broad) under the same simple distributional assumptions as in the benchmark search strategy section given in table 1. I will consider the behaviour of three types of agents under their optimal policy: (i) Narrow background agent specialized in Sector 1 - the high likelihood of intermediate outcome sector, (ii) Narrow background agent specialized in Sector 2 - the dominated sector, (iii) Broad Background agent. This setting will fully rationalize the three main empirical findings (1) Broad agents are on average more likely to explore and try out more sectors which results in them holding more diversified portfolios (2) broad agents are on average more likely to choose the novel sector (3) Broad agents on average have a higher likelihood of achieving both a high outcome (IPO) and failure.

Figure 14 about here.

In Figure 14, I plot the likelihood of reaching the required threshold under the optimal policy for different agent types and different number of draws (n). First, compared to the benchmark strategy of full information, the likelihood of reaching the threshold is lower for each agent. This is because with limited number of draws the agents cannot learn the exact distribution across all sectors. The top left and right panel depict the success rates for the narrow agent who is a specialist in the sector with high likelihood of intermediate outcome and the narrow agent who is a specialist in the dominated sector. The specialist in the dominated sector has a lower performance on average, since after a failure in the other sectors he updates his prior very strongly and chooses to avoid those sectors and sticks to his dominated sector. The overall performance of the broad agent is shown in the bottom panel in figure 14. The performance of the broad agent is similar to the narrow agent who specializes in the high likelihood of intermediate exit sector, they will however reach the threshold return following their optimal policy through different actions.

Figure 15 about here.

In Figure 15, I plot the sector selection frequency for each type of agent for various values of the return threshold required. The top left panel plots the selection frequency for the narrow agent specializing in the high likelihood of intermediate outcome sector. Since this agent is narrow, but specializes in a good enough sector, the agent decides to stick to his speciality and samples predominately the sector with a high likelihood of intermediate outcome. For larger threshold return required the agent chooses to sometimes explore the novel sector. The top right panel plots the selection frequency for the narrow agent that specializes in the dominant sector. At very low thresholds, this agent chooses to stick to his speciality since this is enough for this agent to reach the threshold. As soon as the threshold increases, this agent becomes aware that it won't be possible to reach the threshold by sticking to his dominated sector so the agent will start exploring other sectors. This agent however updates the probability of outcomes in the other sectors very strongly and is therefore much more likely to switch between sector 1 and sector 3 sampling them at roughly the same frequency. The bottom panel depicts the sampling behaviour of the broad agent. The broad agent updates her probability about statup quality moderately after each draw and following her optimal policy decided to explore the novel sector much more at high values of the threshold. The results of this figure rationalize two of the main empirical findings (1) broad agents hold more diversified portfolios on average (2) broad agents are more likely to explore novel sectors.

Figure 16 about here.

In Figure 16, I show that the sampling sectoral choice under the optimal policy shown in figure 15 translates into outcomes. In the left panel I plot the likelihood of getting an IPO (high reward) outcome for each type of agent following optimal policy. Higher rate of sampling from the novel sector translates into higher likelihood of getting an IPO outcome for the broad agent across all threshold values. Similarly it translates into a higher failure rate as shown in the bottom graph. This finding rationalizes the third main empirical results, the fact that broad human capital VCs have a higher failure rate and a higher IPO rate than narrow human capital VCs.

The incomplete information setting demonstrates how human capital broadness impacts venture capitalists' strategic decisions under incomplete information, offering a detailed mechanism for the observed empirical outcomes. Broad-background VCs, characterized by moderate priors across multiple sectors, are shown to explore more sectors and this results in higher portfolio diversification and a greater likelihood of financing novel, high-risk startups. This approach leads to a higher probability of both high-reward exits (e.g., IPOs) and failures, consistent with their propensity to invest in sectors with uncertain yet high-potential returns. Conversely, narrow-background VCs, with strong priors limited to their domain of expertise, exhibit conservative investment patterns, focusing on familiar sectors with stable, lower-risk outcomes. This specialization reduces their likelihood of exploration, resulting in less diversified portfolios and fewer extreme outcomes, whether successes or failures. These results highlight a fundamental trade-off: broad-background VCs prioritize exploration and financing novelty at the cost of higher variance in outcomes, while narrow-background VCs emphasize exploitation of existing knowledge, achieving more stable outcomes.

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 tend to underperform on 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.

To rationalize these findings, I develop a dynamic multi-armed bandit portfolio choice model, demonstrating that broad-background VCs engage in greater exploration, experience more early failures, and finance exceptional startups, whereas narrow-background VCs focus on their established domains. The model shows that, despite differing investment strategies, both types of VCs achieve similar overall performance, with broad VCs relying on a few highly successful exits and narrow VCs favoring stable, intermediate outcomes.

These results highlight the critical role of human capital breadth in fostering exploration and financing novel business ventures and support the role for public policies fostering the development of broad human capital.

Figures

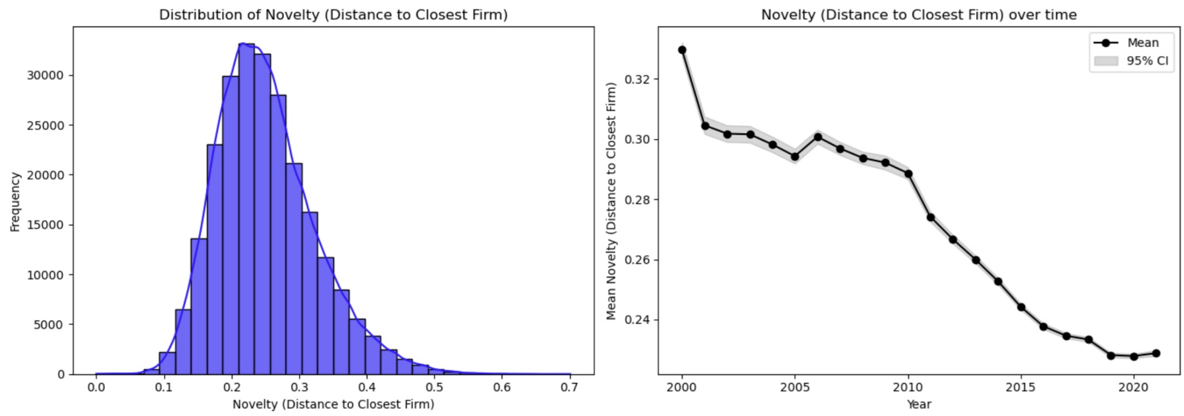

Figure 1: **Distribution of Novelty (Distance to Closest Firm) and time trend of Novelty (Distance to Closest Firm)** Left panel: This figure plots the distribution of the Novelty (Distance to Closest Firm) measure. Right Panel: Time trend of the mean of Novelty (Distance to Closest Firm) measure.

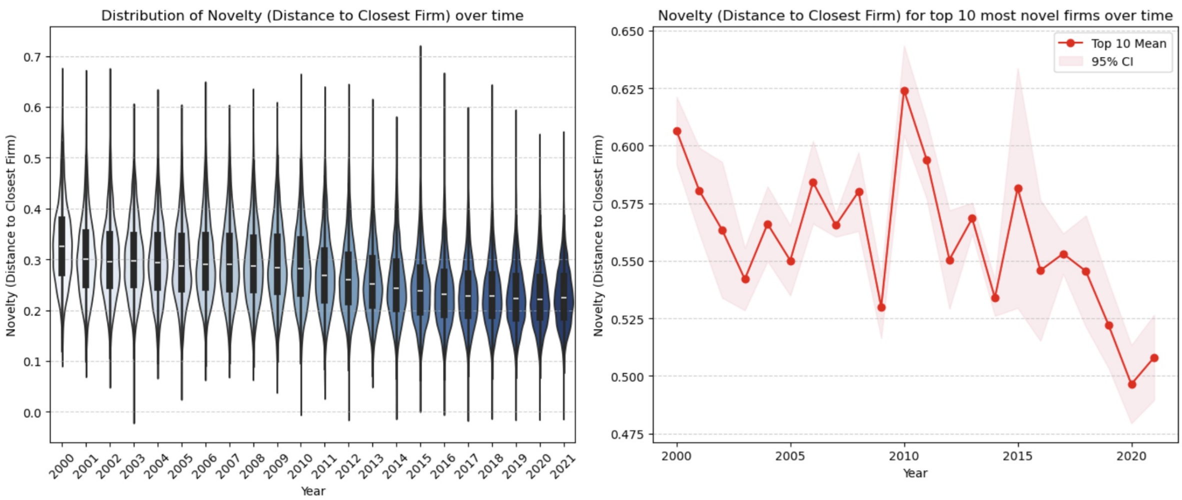

Figure 2: **Evolution of Novelty Distribution over time and the time trend of the mean Novelty for the top 10 most Novel Firm** Left panel: This figure plots the time tend of distribution of Novelty. Left panel: This figure plots the time trend of mean Novelty for the top 10 most novel firms in each year.

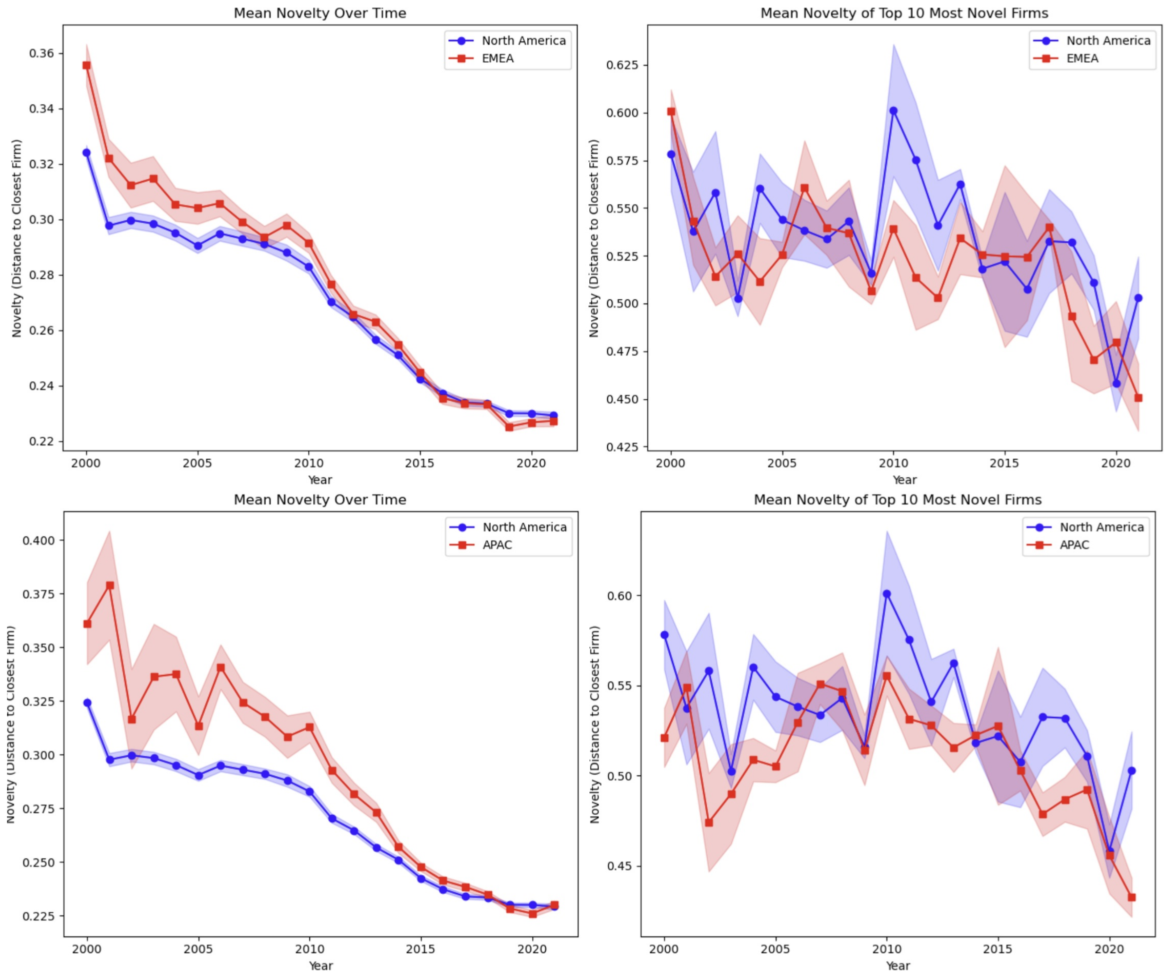

Figure 3: **Evolution of Mean Novelty over time based on region: by startup location** Top - Left panel: This figure plots the time trend of mean Novelty for North America and EMEA regions. Top - Right panel: This figure plots the time trend of mean Novelty of top 10 most novel startups for North America and EMEA regions. Bottom - Left panel: This figure plots the time trend of mean Novelty for North America and APAC regions. Bottom - Right panel: This figure plots the time trend of mean Novelty of top 10 most novel startups for North America and APAC regions.

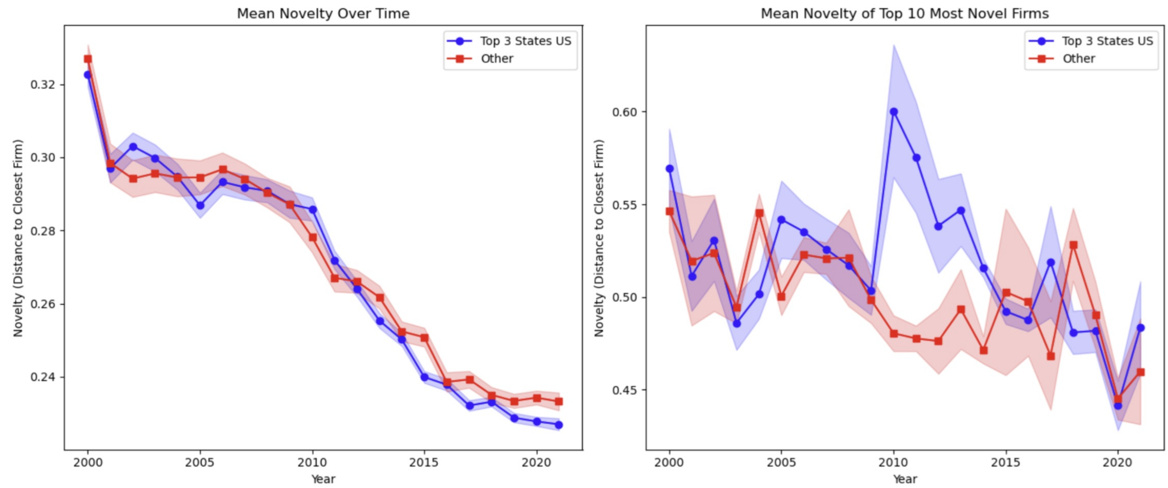

Figure 4: **Evolution of Mean Novelty over time based on region US data: by startup location** Left panel: This figure plots the time trend of mean Novelty for Top 3 US States (California, New York, Massachusetts) and Other US states. Right panel: This figure plots the time trend of mean Novelty of top 10 most novel startups for Top 3 US States (California, New York, Massachusetts) and Other US states.

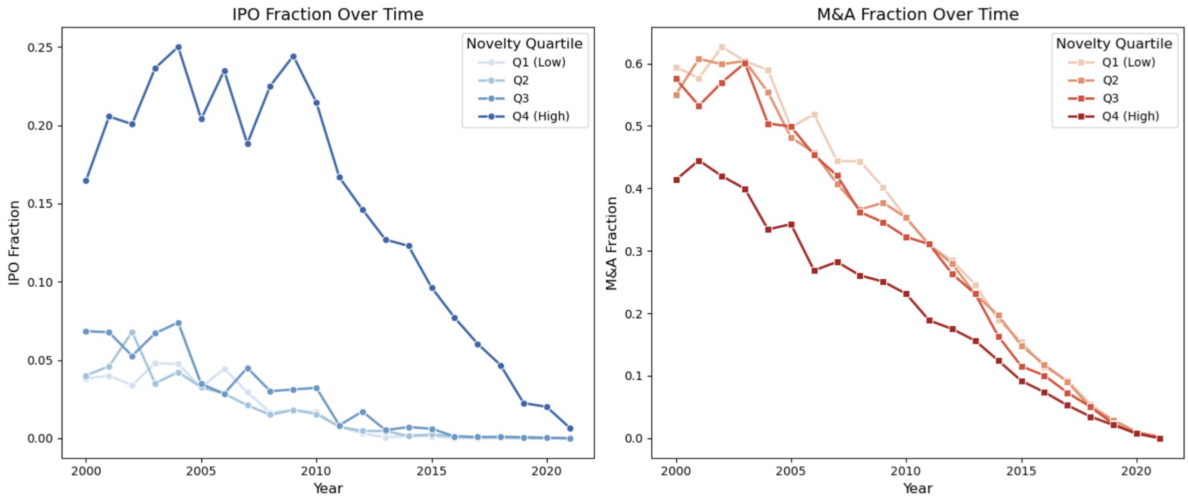

Figure 5: **IPO and M&A rates split by Novelty Quartile** Left panel: This figure plots the time trend of IPO rates for each quartile of Novelty (constructed by splitting the deals in each year by novelty quartile). Right panel: This figure plots the time trend of M&A rates for each quartile of Novelty (constructed by splitting the deals in each year by novelty quartile).

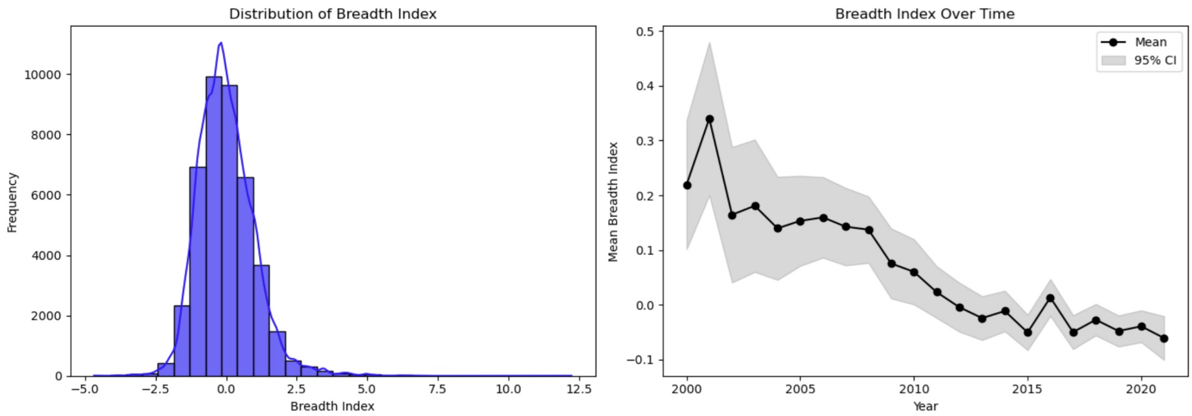

Figure 6: **Distribution of Breadth Index and time trend of Breadth Index** Left panel: This figure plots the distribution of the Breadth Index measure. Right Panel: Time trend of the mean of Breadth Index measure.

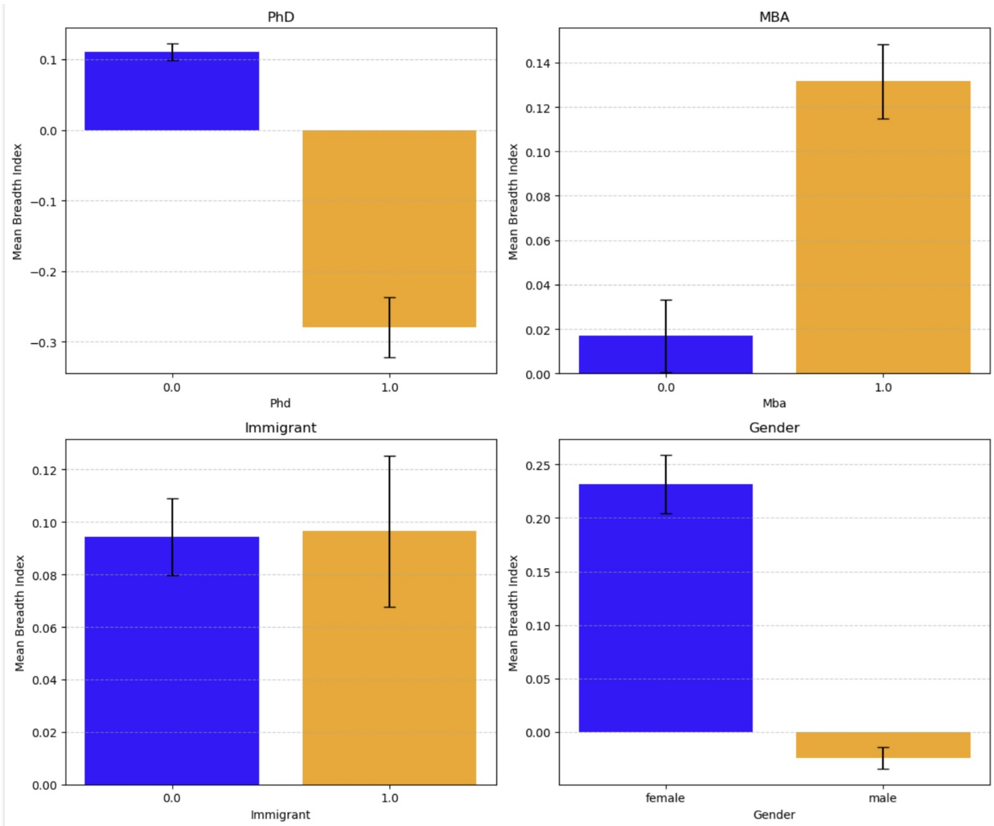

Figure 7: **Split of Breadth Index Mean by Subcategories** This figure plots the split of Mean Breadth Index by various subcategories Top Left: Split by PhD Degree Status. Top Right: Split by MBA Degree status. Bottom left: Split by Immigrant status. Bottom right: Split by gender.

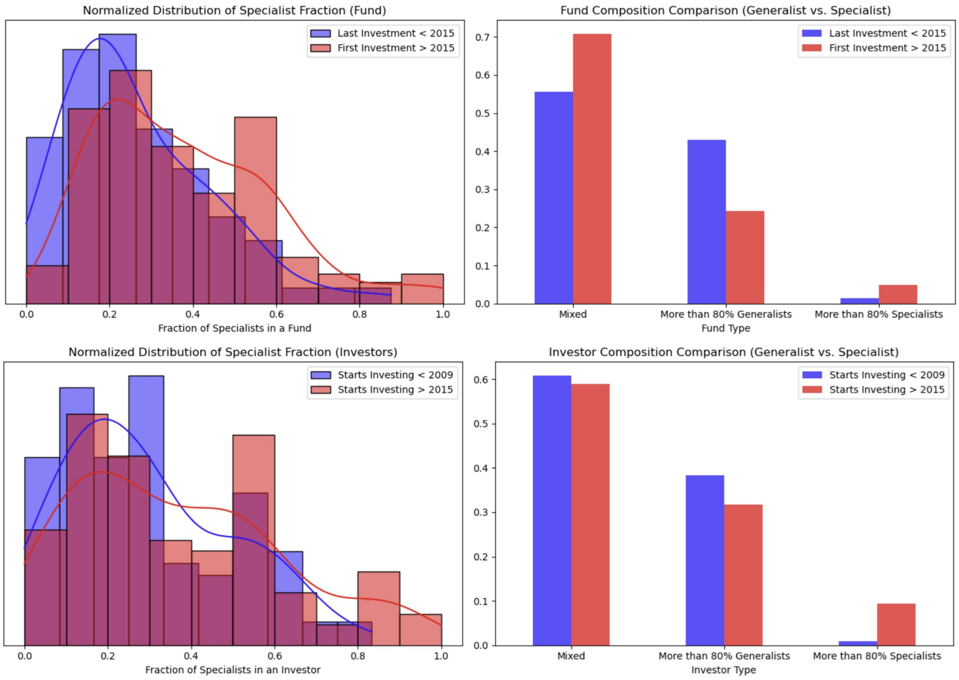

Figure 8: **Distribution of human capital breadth across funds and investors** This figure plots the distribution of human capital breadth across investors and funds. Top right: This figure is a distributional plot of the fraction of partners classified as specialists in each fund for funds that have made their first investment after 2015 (red) and funds that have made their last investment before 2015 (blue). Top left: human capital composition within a fund, blue - funds with last investment made prior to 2015, red - funds with first investment made after 2015. Bottom right: This figure is a distributional plot of the fraction of early partners (partners who have joined the VC firms within 6 years of first investment) classified as specialists in each investor for investors who have made their first investment prior to 2009 (blue) and for investors who have made their first investment after 2015. Bottom left: human capital composition within investor for early partners - blue investors who have started investing before 2009, red-investors who have started investing after 2015.

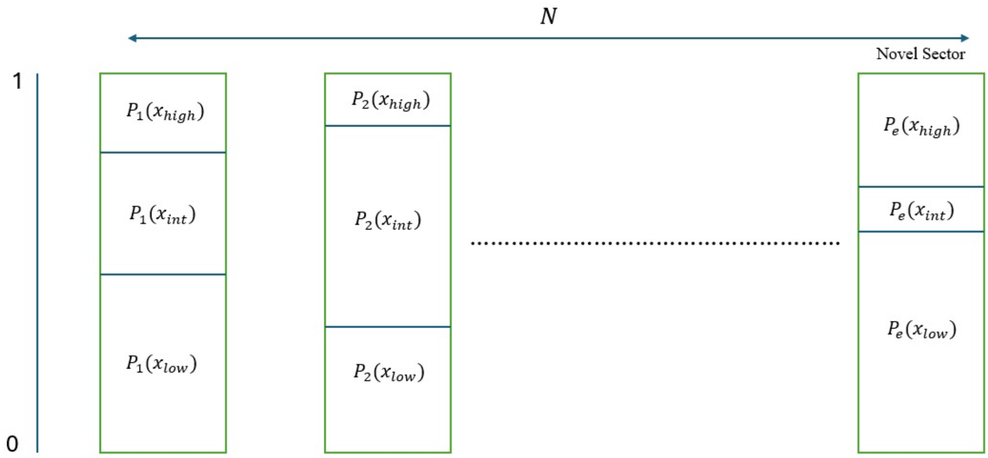

Figure 9: **Supply of Ventures in each of the N sectors:** This figure depicts the supply of ventures in each of the N sectors. $P_i(x_{low})$ denotes the probability of drawing a low quality venture given that the agent draws in sector i. $P_i(x_{int})$ denotes the probability of drawing an intermediate quality venture given that the agent draws in sector i. $P_i(x_{high})$ denotes the probability of drawing an high quality venture given that the agent draws in sector i. For each sector i, $P_i(x_{low}) + P_i(x_{int}) + P_i(x_{high}) = 1$. e indexes the novel sector. The distribution of startups in the novel sector satisfies $P_e(x_{low}) > P_i(x_{low})$ and $P_e(x_{high}) > P_i(x_{high})$ for all other sectors $i \neq e$.

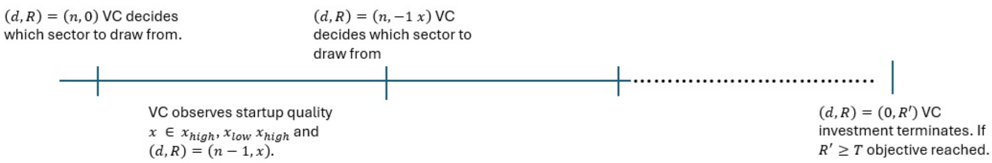

Figure 10: **Value function iteration under complete information:** This figure depicts the sequential draws and value function update for the VC in the complete information case. The VC starts in a $(d, R) = (n, 0)$ state and observes the distribution of startups in each sector. The VC then decides which sector to draw from. The VC draws and observes the outcome moving to a state $(d, R) = (n-1, x)$ and follows optimal policy thereafter. After all n draws are executed the VC either reaches the required threshold, T or does not.

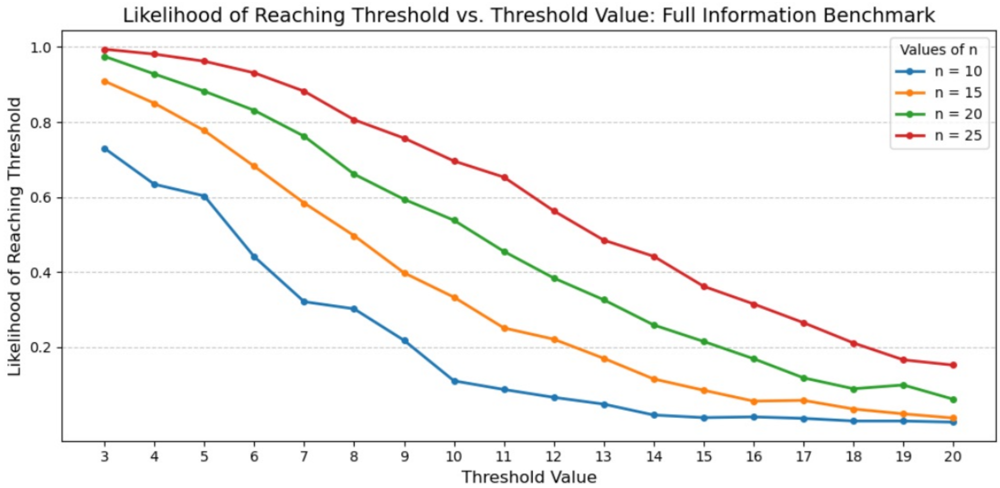

Figure 11: **Likelihood of reaching the required return threshold in the benchmark case for different number of draws.** This figure depicts the likelihood of reaching the threshold as a function of the threshold value for different number of draws. The x-axis depicts the threshold value. They y-axis depicts the number of times the threshold is reached relative to the number of simulations conducted. Simulation parameters: 1000 simulations under optimal policy, sectoral distribution given by table 1. Threshold values ranging from 3-20.

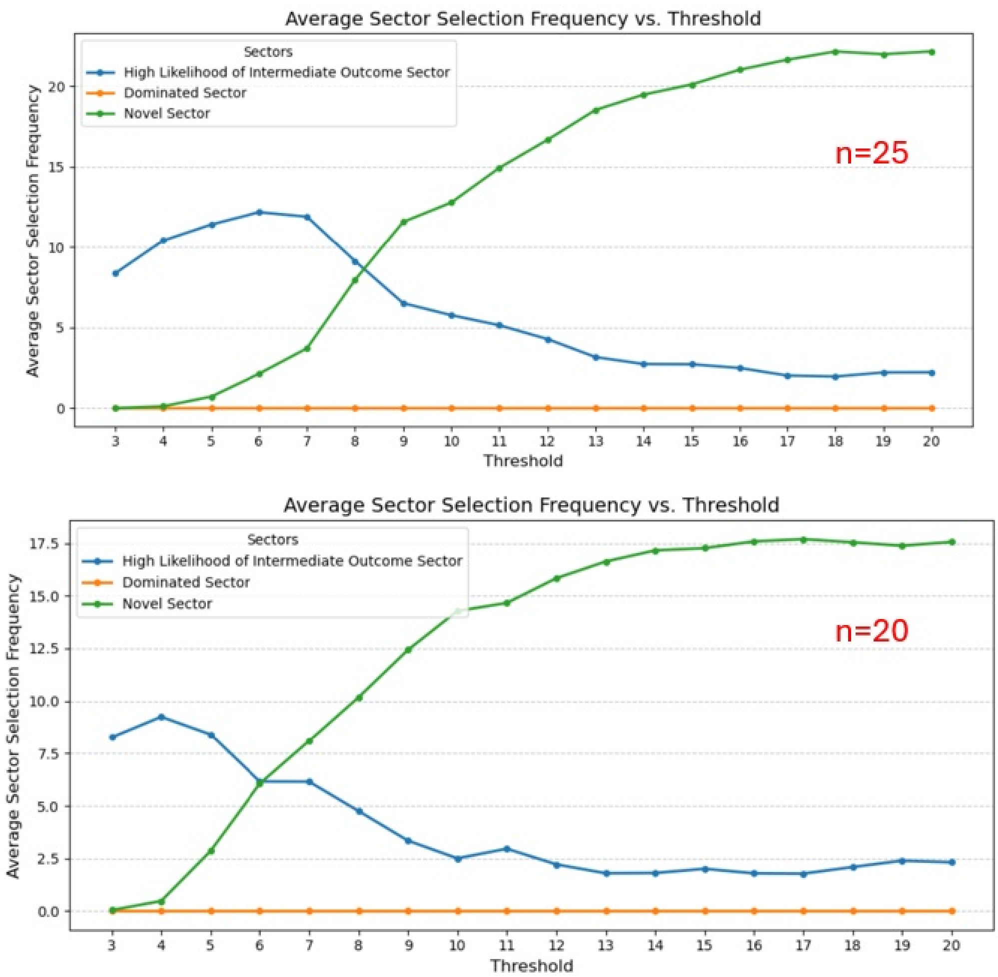

Figure 12: **Frequency of Selection of each sector for different threshold values** This figure depicts the frequency each sector is chosen under the optimal policy as a function of the threshold return required. The x-axis plots the threshold return required. The y-axis plots the number of times a specific sector is chosen under the optimal policy averaged across the number of simulations. Simulation parameters: 1000 simulations under optimal policy, sectoral distribution given by table 1. Threshold values ranging from 3-20.

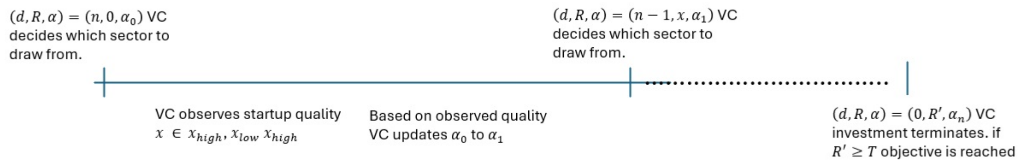

Figure 13: **Investment under incomplete information:** This figure depicts the sequential draws and value function update for the VC in the incomplete information case. The VC starts in a $(d, R, \alpha) = (n, 0, \alpha_0)$ where α_0 denotes the VCs' initial belief for the distribution on quality in each sector i. The VC then decides which sector to draw from. The VC draws and observes x, based on the observed outcome the VC updates the belief α_0 to α_1 the outcome moving to a state $(d, R, \alpha) = (n-1, x, \alpha_1)$ and follows optimal policy thereafter. After all n draws are executed the VC either reaches the required threshold or does not.

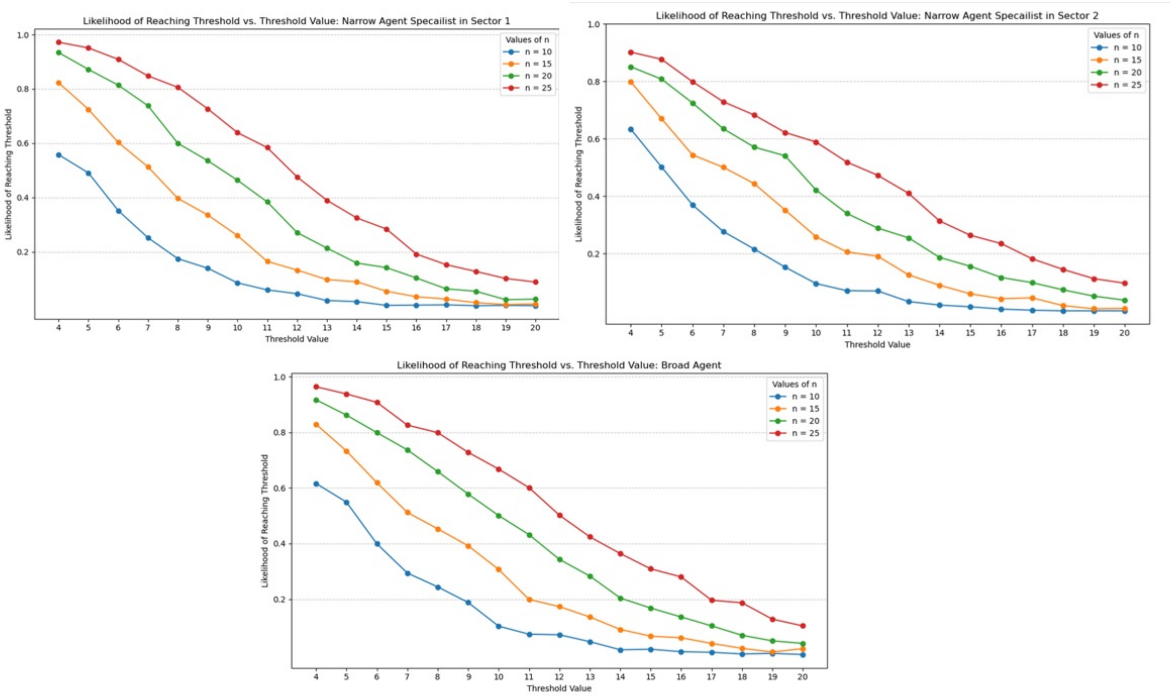

Figure 14: **Likelihood of reaching the required return threshold in the incomplete information case for different types of agents.** This figure depicts the likelihood of reaching the threshold as a function of the threshold value for different number of draws. The x-axis depicts the threshold value. They y-axis depicts the number of times the threshold is reached relative to the number of simulations conducted. Top left depicts the graph for the agent specialized in Sector 1, top right depicts the graph for the agent specialized in sector 2 and bottom graph depicts the graph for the broad agent. Simulation parameters: 1000 simulations under optimal policy, sectoral distribution given by table 1. Threshold values 4, 8, 12, 16, $k = 10$. $m = 5$.

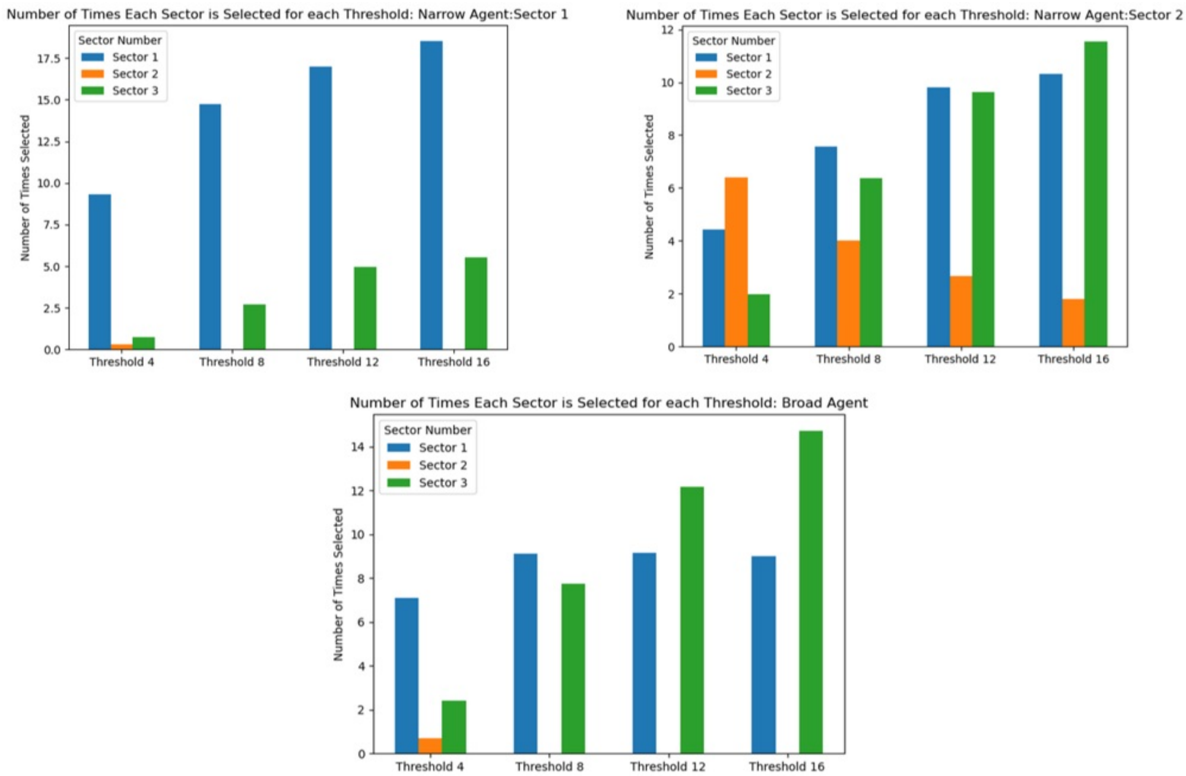

Figure 15: **Number of times each sector is sampled for different agents at different thresholds following optimal policy** This figure plots the number of times each sector is sampled by each agent. Top left depicts the sector sampling under optimal policy for the agent specialized in Sector 1, top right depicts the sector sampling under optimal policy for the agent specialized in sector 2 and bottom graph depicts sector sampling under optimal policy for the broad agent. 1000 simulations under optimal policy, sectoral distribution given by table 1. Threshold values 4, 8, 12, 16,, $k = 10$. $m = 5$.

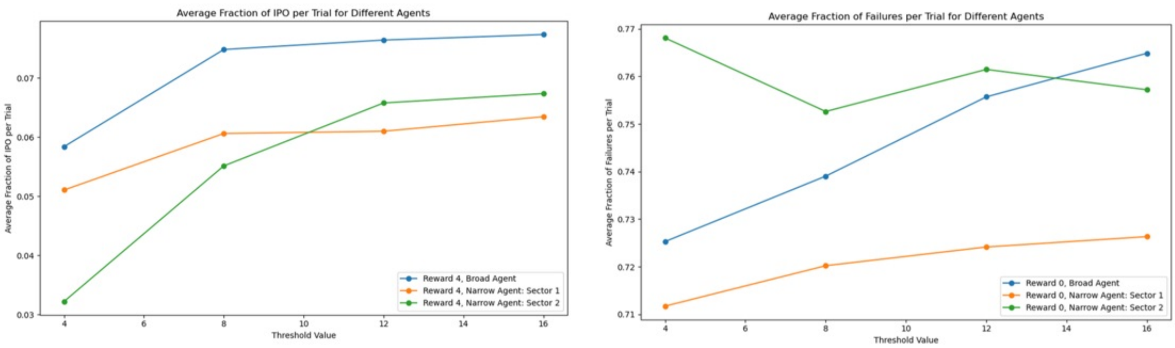

Figure 16: **Likelihood of reaching the required return threshold in the incomplete information case for different types of agents.** This figure depicts the likelihood of reaching the threshold as a function of the threshold value for different number of draws. The x-axis depicts the threshold value. They y-axis depicts the number of times the threshold is reached relative to the number of simulations conducted. Top left depicts the graph for the agent specialized in Sector 1, top right depicts the graph for the agent specialized in sector 2 and bottom graph depicts the graph for the broad agent. Simulation parameters: 1000 simulations under optimal policy, sectoral distribution given by table 1. Threshold values 4, 8, 12, 16, $k = 10$. $m = 5$.

Tables

Variable	Obs	Mean	Std. Dev.	1%	5%	25%	50%	75%	95%	99%
Novelty (Distance to Closest Firm)	241,936	0.25	0.07	0.11	0.15	0.20	0.24	0.29	0.39	0.46
Novelty (Avg. Distance to Closest Firm)	241,936	0.28	0.07	0.15	0.18	0.23	0.27	0.32	0.41	0.48
IPO Exit	241,936	0.04	0.19	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Failure	241,936	0.37	0.48	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Number of Forward Patents	241,936	0.23	2.27	0.00	0.00	0.00	0.00	0.00	0.00	6.00
Number of Forward Citations	241,936	2.94	42.98	0.00	0.00	0.00	0.00	0.00	0.00	52.15
Sum of Forward Claims	241,936	1.62	18.81	0.00	0.00	0.00	0.00	0.00	0.00	36.64
Breadth Index	42,317	0.00	1.01	-1.99	-1.42	-0.68	-0.11	0.57	1.70	3.08
Partner VC Experience	83,637	1.87	1.30	0.00	0.00	0.69	1.79	2.83	4.08	4.91
Female	47,810	0.11	0.31	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Partner has Computer Science Undergraduate	25,801	0.04	0.21	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Partner has MBA	25,801	0.42	0.49	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Partner has PhD	25,801	0.12	0.32	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Partner has Top Degree	25,801	0.35	0.48	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Partner is Immigrant	19,067	0.22	0.41	0.00	0.00	0.00	0.00	0.00	1.00	1.00

Table 2: **Summary Statistics of Deal Level Sample** This table presents summary statistics for the deal level sample

Variable	Obs	Mean	Std. Dev.	1%	5%	25%	50%	75%	95%	99%
Novelty (Distance to Closest Firm)	207,768	0.24	0.07	0.12	0.15	0.19	0.24	0.28	0.38	0.45
Novelty (Avg. Distance to Closest Firm)	207,768	0.27	0.07	0.15	0.17	0.22	0.26	0.31	0.41	0.47
IPO Exit	207,768	0.05	0.21	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Failure	207,768	0.25	0.43	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Number of Partners	207,768	14.47	12.22	1.00	2.00	5.00	11.00	20.00	38.00	53.00
Number of Available Partners	207,768	11.98	11.06	0.00	0.00	4.00	8.00	17.00	34.00	47.00
Partner Leads a Deal	207,768	0.20	0.40	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Busy Partner	207,768	0.08	0.27	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Breadth Index	207,768	0.01	1.02	-1.88	-1.37	-0.68	-0.11	0.57	1.78	3.16

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 a textual description of his or her background.

Category	Prototype Sentences
Narrow	Throughout their professional journey, they have remained dedicated to a single discipline, investing years of study and practice in one area of expertise. Their track record shows minimal shifts in focus or responsibility.
Narrow	This individual's career path reflects a strong focus on mastering one domain, with repeated involvement in projects that reinforce a specific skill set. They rarely venture beyond this established specialty.
Narrow	Over time, they have held roles that closely resemble one another, honing an in-depth approach to a single function. Their résumé reveals a tight concentration on a particular kind of work.
Narrow	From the beginning of their career, this person has operated within a narrowly defined scope. They excel in one niche and show little interest in branching into other roles or disciplines.
Narrow	They built their expertise by continuously refining the same techniques and principles, rarely adopting new methods. Their progression shows incremental specialization rather than expansion into different fields.
Broad	Their professional history spans multiple functions, reflecting a pattern of moving between different responsibilities and challenges. Each step adds another dimension to their growing toolbox of skills.
Broad	By actively seeking roles in distinct areas, this individual has cultivated a broad outlook. Their experience includes learning diverse processes, collaborating with varied teams, and adapting to new contexts.
Broad	Over the course of their career, they've tackled a wide array of problems. Their track record demonstrates an ability to switch between different methodologies, showing breadth in both approach and expertise.
Broad	They are accustomed to pivoting whenever a new opportunity arises, accumulating a mix of experiences that cross traditional boundaries. Their background hints at comfort in tackling varied objectives.
Broad	Their career progression cuts across a range of roles, from technical to strategic. This breadth gives them an expansive viewpoint, enabling them to integrate knowledge from multiple professional domains.

Table 4: **Prototypes of Narrow and Broad Human Capital Backgrounds** This figure describes the prototype sentences used to construct the breadth index measure. Category refers to the category of the prototype which can be Narrow or Broad.

	(1) Failure	(2) Major Success	(3) IPO Exit	(4) Failure	(5) Major Success	(6) IPO Exit
Novelty (Distance to Closest Firm)	0.288***	0.726***	0.792***	0.121***	0.750***	0.802***
	(0.019)	(0.024)	(0.025)	(0.024)	(0.037)	(0.039)
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00
R^2	0.18	0.13	0.17	0.53	0.47	0.52

Table 5: **Association between startup novelty and likelihood of Failure, Major Success and IPO Exit.** This table reports the results of a deal - level regression of Failure, Major Success and IPO Exit on the deal's novelty. The independent variable Novelty (Distance to Closest Firm) is a deal level measure of novelty defined as one - 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. IPO Exit is an indicator variable taking a value of 1 if the firm exits via an IPO. Major Success is an indicator variable taking a value of 1 if the firm exits via an IPO or is acquired at a valuation at least five times greater than the total financing received. Failure is an indicator taking a value of 1 if the firm does not exit via an IPO, is not Acquired or does not receive any follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Failure	(2) Major Success	(3) IPO Exit	(4) Failure	(5) Major Success	(6) IPO Exit
Novelty Quartile (Distance to Closest Firm)=2	0.020***	-0.002*	-0.001	0.020***	0.001	0.002*
	(0.003)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)
Novelty Quartile (Distance to Closest Firm)=3	0.043***	0.002	0.006***	0.028***	0.006***	0.008***
	(0.003)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)
Novelty Quartile (Distance to Closest Firm)=4	0.062***	0.101***	0.110***	0.032***	0.107***	0.114***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00
R^2	0.18	0.12	0.14	0.53	0.47	0.50

Table 6: **Association between startup novelty quartile and likelihood of Failure, Major Success and IPO Exit.** This table reports the results of a deal - level regression of Failure, Major Success and IPO Exit on the deal's novelty quartile. The independent variables are novelty quartile dummies indicating whether a startup belongs to the i the novelty quartile within a given year and deal stage based on the Novelty (Distance to Closest Firm) measure which is a deal level measure of novelty defined as one - 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 omitted category is Novelty Quartile (Distance to Closest Firm) = 1. IPO Exit is an indicator variable taking a value of 1 if the firm exits via an IPO. Major Success is an indicator variable taking a value of 1 if the firm exits via an IPO or is acquired at a valuation at least five times greater than the total financing received. Failure is an indicator taking a value of 1 if the firm does not exit via an IPO, is not Acquired or does not receive any follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Number of Patents	(2) Number Citations	(3) Number of Claims	(4) Number of Patents	(5) Number Citations	(6) Number of Claims
Novelty (Distance to Closest Firm)	1.239***	2.136***	1.648***	1.175	2.787***	1.591*
	(0.401)	(0.482)	(0.468)	(0.793)	(0.994)	(0.880)
Exit Type Controls	✓	✓	✓	✓	✓	✓
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00

Table 7: **Association between startup novelty and innovation outcomes - Poisson count model** This table reports the results of a deal - level regression of innovation outcomes on the deal's novelty. The independent variable Novelty (Distance to Closest Firm) is a deal level measure of novelty defined as one - 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. Number of Patents is the total number of patents granted to the firm after the deal date. Number Citations is the total number of adjusted citations (by grant year and NBER subcategory) that patents of the financed firm have received after the deal date. Number of Claims is total claims in all granted patents after the deal date of the financed firm. All columns include Exit Type controls which are separate controls specifying whether the startup goes public is acquired or receives follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Breadth Index	(2) Breadth Index	(3) Breadth Index	(4) Breadth Index	(5) Breadth Index
Top Degree	-0.009	-0.023	-0.053**	-0.037*	-0.095***
	(0.020)	(0.020)	(0.022)	(0.023)	(0.033)
PhD	-0.292***	-0.284***	-0.233***	-0.215***	-0.211***
	(0.045)	(0.044)	(0.042)	(0.050)	(0.050)
MBA	0.074***	0.092***	0.090***	0.112***	0.058**
	(0.019)	(0.019)	(0.022)	(0.021)	(0.030)
Immigrant	0.070***	0.020	0.062**	0.081**	0.079**
	(0.027)	(0.025)	(0.032)	(0.033)	(0.033)
Age	-0.003**	-0.003**	0.000	-0.002*	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female=1	0.425***	0.431***	0.482***	0.475***	0.477***
	(0.032)	(0.032)	(0.036)	(0.035)	(0.035)
United States=1		-0.083***			
		(0.025)			
Computer Science			-0.125***	-0.135**	-0.139**
			(0.047)	(0.060)	(0.060)
Massachusetts				0.130***	0.132***
				(0.042)	(0.042)
California				0.107***	0.106***
				(0.025)	(0.025)
New York				0.197***	0.199***
				(0.041)	(0.041)
Top Degree × MBA					0.115***
					(0.043)
Country FE	✓				
State FE			✓		
Observations	8867.00	8867.00	6949.00	6949.00	6949.00
R^2	0.07	0.03	0.08	0.04	0.04

Table 8: **Correlates of Human Capital Breadth Index** This figure describes the results of an OLS regression of Breadth Index on partner level correlates. The dependent variable Breadth Index is a measure of human capital breadth defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. Top Degree is an indicator taking a value of 1 if the partner has obtained any degree from a top institution. PhD is an indicator taking a value of 1 if the partner has completed a PhD degree. MBA is an indicator taking a value of 1 if the partner has completed an MBA degree. Age is the Partner's age at entry in the VC industry. Female is an indicator taking a value of 1 if the partner's gender is female. Immigrant is an indicator taking a value of 1 if the partner has immigrant background.Columns (1) and (2) use the full sample, while columns (3)-(6) use the US sample. In column (1) I include Country FE. In column (3) I include US State FE. Heteroscedasticity robust standard errors are in parenthesis. * p<.10; ** p<.05; *** p<.01.

	(1) Novelty (Distance to Closest Firm)	(2) Novelty (Distance to Closest Firm)	(3) Novelty (Distance to Closest Firm)	(4) Novelty (Distance to Closest Firm)
Breadth Index	0.002***	0.002***	0.002**	0.003**
	(0.000)	(0.001)	(0.001)	(0.001)
VC Experience	-0.000	0.001	0.001	0.001
	(0.000)	(0.000)	(0.001)	(0.001)
Controls			✓	✓
Industry × Deal Year × Deal Type FE	✓		✓	
Industry × Deal Year × Deal Type × Investor FE		✓		✓
Observations	42317.00	42317.00	12372.00	12372.00
R^2	0.19	0.30	0.19	0.30

Table 9: **Association between lead partner's human capital breadth index and startup novelty** This table reports the results of an OLS regression of deal novelty and lead partner's human capital breadth. The dependent variable Novelty (Distance to Closest Firm) 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 defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. 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) and (2) include Industry × Deal Year × Deal Type FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (4) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) IPO Exit	(2) IPO Exit	(3) IPO Exit	(4) Failure	(5) Failure	(6) Failure
Breadth Index	-0.060***	-0.058***	-0.050***	0.035***	0.021**	0.010
	(0.007)	(0.008)	(0.019)	(0.009)	(0.009)	(0.020)
Novelty (Distance to Closest Firm)	0.933***	0.958***	1.009***	0.221***	0.139***	0.162**
	(0.035)	(0.037)	(0.075)	(0.034)	(0.036)	(0.065)
Breadth Index × Novelty (Distance to Closest Firm)	0.262***	0.244***	0.197**	-0.101***	-0.062*	-0.001
	(0.034)	(0.035)	(0.078)	(0.031)	(0.032)	(0.073)
VC Experience	0.004***	0.002	0.005*	-0.043***	-0.019***	-0.023***
	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.006)
Controls			✓			✓
Industry × Deal Year × Deal Type FE	✓			✓		
Industry × Deal Year × Deal Type × Investor FE		✓	✓		✓	✓
Observations	42317.00	42317.00	12372.00	42317.00	42317.00	12372.00
R^2	0.21	0.30	0.35	0.20	0.34	0.34

Table 10: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup outcome.** This table reports the results of an OLS regression of startup outcome on the interaction between deal's novelty and lead partner's human capital breadth. The independent variable Novelty (Distance to Closest Firm) 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 Breadth Index is a measure of human capital breadth defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. Breadth Index × Novelty (Distance to Closest Firm) is the interaction between the deal's novelty and the lead partner's human capital breadth index. In columns (1)-(3) the dependent variable IPO Exit is an indicator that equals 1 if the startup goes public. 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) and (2) include Industry × Deal Year × Deal Type FE. In columns (1)-(3) the dependent variable IPO Exit is an indicator that equals 1 if the startup goes public. In columns (4)-(6) the dependent variable Failure is an indicator taking a value 1 if the startup does not go public is not acquired or does not receive follow - up - financing. Columns (1) and (4) include Industry × Deal Year × Deal Type FE. Columns (2), (3), (5) and (6) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (6) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Partner Leads a Deal	(2) Partner Leads a Deal	(3) Partner Leads a Deal	(4) Partner Leads a Deal
Breadth Index	-0.008**	-0.011***	-0.014*	-0.021***
	(0.004)	(0.003)	(0.008)	(0.008)
Breadth Index × Novelty (Distance to Closest Firm)	0.052***	0.054***	0.061**	0.077**
	(0.014)	(0.014)	(0.031)	(0.031)
Partner Age			-0.002***	-0.003***
			(0.000)	(0.000)
Female=1			0.010*	0.011*
			(0.006)	(0.006)
VC Experience			0.058***	0.061***
			(0.002)	(0.002)
Busy Partner			-0.016**	-0.018***
			(0.007)	(0.006)
Deal FE	✓	✓	✓	✓
Investor FE		✓		✓
Observations	207768.00	207768.00	64221.00	64221.00
R^2	0.14	0.17	0.27	0.31

Table 11: **Association between the interaction of lead partner's human capital breadth and deal novelty and the likelihood of a partner leading a deal.** This table reports the results of a regression of the likelihood of leading a deal on partner's human capital breadth and the interaction between deal novelty and partner's human capital breadth. Each observation is a partner-deal-investor. The dependent variable Partner Leads a Deal is an indicator equal to 1 if a partner leads a deal and 0 otherwise. The independent variable Novelty (Distance to Closest Firm) 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 Breadth Index is a measure of human capital breadth defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. Breadth Index × Novelty (Distance to Closest Firm) is the interaction between the deal's novelty and the lead partner's human capital breadth index. In columns (3) and (4) Partner Age is the age of the partner at the time of the deal. Female is an indicator variable equal to 1 if the partner is female and 0 otherwise. 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. Busy Partner is an indicator equal to 1 if the partner leads a different deal that is at the time of the current deal undergoing a public listing (IPO) or acquisition. In columns (1) and (3) we include a Deal FE, in columns (2) and (3) an Deal FE and an Investor FE. Standard errors reported in parenthesis are clustered at a Deal level. * p<.10; ** p<.05; *** p<.01.

	(IV First stage) Breadth Index	(IV first stage) Breadth Index × Novelty (Distance to Closest Firm)	(IV) IPO Exit	(OLS) IPO Exit	(IV) Failure	(OLS) Failure
Avg. Available Breadth	0.308*** (0.053)	-0.144*** (0.016)				
Avg. Available Breadth × Novelty (Distance to Closest Firm)	-0.209 (0.169)	0.857*** (0.063)				
Breadth Index			-0.126*** (0.035)	-0.054** (0.022)	0.003 (0.080)	-0.007 (0.025)
Breadth Index × Novelty (Distance to Closest Firm)			0.556*** (0.123)	0.217** (0.094)	-0.221 (0.151)	0.087 (0.094)
Novelty (Distance to Closest Firm)			0.913*** (0.083)	0.897*** (0.082)	0.152* (0.083)	0.156* (0.081)
Controls FE	✓	✓	✓	✓	✓	✓
Investor FE	✓	✓	✓	✓	✓	✓
Deal Stage × Industry × Year FE	✓	✓	✓	✓	✓	✓
Observations	8324.00	8324.00	8324.00	8324.00	8324.00	8324.00
R^2	0.71	0.69	0.09	0.33	-0.01	0.36
F-statistic of Instrument			62.57		62.57	

Table 12: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup outcomes.** This table presents the results of a instrumental variable regression of deal performance on the interaction between lead partner breadth and deal novelty. Column (1) presents the first stage regressions of the first 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 at a VC firm level across the firm's available partners. Column (2) presents the first stage regression of the second instrumented variable Breadth Index × Novelty on the two instruments Avg. Available Breadth and Avg. Available Breadth × Novelty. Columns (3) and (5) present the IV regression where the dependent variables are IPO Exit and Failure Respectively. Columns (4) and (6) present the OLS estimates of an equivalent model.

	(1) Novelty (Distance to Closest Firm)	(2) Novelty (Avg. Distance to Closest Firm)	(3) Fraction of Top Quartile Novelty Firms
Average Fund Level Breadth Index	0.002** (0.001)	0.002** (0.001)	0.009** (0.004)
Fund Size	-0.000* (0.000)	-0.000* (0.000)	-0.000*** (0.000)
First Fund=1	0.002 (0.001)	0.003** (0.001)	0.005 (0.007)
Industry Controls	✓	✓	✓
Stage Controls	✓	✓	✓
Vintage Year × Country × FundType FE	✓	✓	✓
Observations	7364.00	7364.00	7364.00
R^2	0.63	0.68	0.28

Table 13: **Fund Level Association Between Average Human Capital Breadth and Average Novelty**. This table present the results of a regression of the average novelty of a financed firm by a fund and the average human breadth index within a fund partnership. The dependent variable in column (1) is the average of Novelty (Distance to Closest Firm) for all financed firms by the fund. The dependent variable in column (2) is the average of Novelty (Avg. Distance to Closest Firm) for all financed firms by the fund. The dependent variable in column (3) is the Fraction of Deals that are made in a Top Quartile Novelty Firms. The main independent variable Average Fund Level Breadth Index is an average of the breadth index across all partners that made a deal in the fund. Fund Size is the size of the fund. First Fund is an indicator taking a value of 1 if the Fund is a First Fund raised by a given VC firm. In all columns we include Industry Controls, which are separate controls for the fund's industry composition, Stage Controls which are separate controls for the fund's stage of investment composition. In all columns we include Vintage Year × Country × Fund Type FE. Standard errors reported in parenthesis are clustered at an Investor level * p<.10; ** p<.05; *** p<.01.

	(1) Novelty (Distance to Closest Firm)	(2) Novelty (Avg. Distance to Closest Firm)	(3) Fraction of Top Quartile Novelty Firms
Average Investor Level Breadth Index	0.001** (0.001)	0.002*** (0.001)	0.008** (0.003)
VC Experience	-0.001* (0.001)	-0.001 (0.001)	-0.007* (0.004)
Industry Controls	✓	✓	✓
Stage Controls	✓	✓	✓
Investor Founding Year × Country FE	✓	✓	✓
Observations	2932.00	2932.00	2932.00
R^2	0.55	0.60	0.23

Table 14: **Investor Level Association Between Average Human Capital Breadth and Average Novelty in the yearly years of investor's existence** This table present the results of a regression of the average novelty of a financed firm by an investor within the first 6 years of the investor's existence and the average human breadth index, averaged across partners who have made deals in the first 6 years of the investor's existence. The dependent variable in column (1) is the average of Novelty (Distance to Closest Firm) for all financed firms by the fund. The dependent variable in column (2) is the average of Novelty (Avg. Distance to Closest Firm) for all financed firms by the fund. The dependent variable in column (3) is the Fraction of Deals that are made in a Top Quartile Novelty Firms. The main independent variable Average Investor Level Breadth Index is an average of the breadth index across all partners that made a deals in the investor's first six years of existence. VC experience is the average VC experience across partners in the investor. In all columns we include Industry Controls, which are separate controls for the investor's deal industry composition, Stage Controls which are separate controls for the investor's stage of investment composition. In all columns we include Investor Founding Year × Country FE. Standard errors reported in parenthesis are clustered at a country × Founding Year level * p<.10; ** p<.05; *** p<.01.

	(1) Fraction of IPO Exits	(2) Fraction of Failed Exits	(3) Fraction of IPO Exits	(4) Fraction of Failed Exits
Average Fund Level Breadth Index	-0.101*** (0.023)	0.096*** (0.029)	-0.011** (0.005)	0.016** (0.007)
Novelty (Distance to Closest Firm)	1.182*** (0.117)	0.099 (0.158)		
Average Fund Level Breadth Index × Novelty (Distance to Closest Firm)	0.453*** (0.096)	-0.386*** (0.112)		
Fraction of Top Quartile Novelty Firms			0.175*** (0.018)	0.030 (0.028)
Average Fund Level Breadth Index × Fraction of Top Quartile Novelty Firms			0.115*** (0.026)	-0.084*** (0.028)
Fund Size	0.000*** (0.000)	-0.000 (0.000)	0.000*** (0.000)	-0.000 (0.000)
First Fund=1	-0.014*** (0.004)	0.008 (0.009)	-0.015*** (0.004)	0.009 (0.008)
Industry Controls	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓
Vintage Year × Country × FundType FE	✓	✓	✓	✓
Observations	7364.00	7364.00	7364.00	7364.00
R^2	0.54	0.61	0.52	0.61

Table 15: **Human Capital Breadth, Firm Novelty and Fund Performance** This table reports the results of a regression of fund performance measures on the interaction between the average novelty of a financed firm and the average human capital breadth in a fund partnership. The dependent variable in columns (1) and (3) Fraction of IPO Exits is the fraction of IPO exited deals relative to the total number of deals made by the fund. The dependent variable in columns (2) and (4) Fraction of Failed Exits of deals who have failed (have not received follow up financing or have not exited via an IPO or an Acquisition) relative to the total number of deals made by the fund. In columns (1) and (2) Novelty (Distance to Closest) firm is the average novelty across all deals financed by the fund. In columns (3) and (4) Fraction of Top Quartile Novelty Firms the fraction of deals that are in top quartile of novelty. In columns (1)-(4) Average Fund Level Breadth Index is the average breadth index, averaged across partners who have made deals in the fund. Fund Size is the size of the fund. First Fund is an indicator taking a value of 1 if the Fund is a First Fund raised by a given VC firm. In all columns we include Industry Controls, which are separate controls for the fund's industry composition, Stage Controls which are separate controls for the fund's stage of investment composition. In all columns we include Vintage Year × Country × Fund Type FE. Standard errors reported in parenthesis are clustered at an Investor level * p<.10; ** p<.05; *** p<.01.

	(1) Fraction of IPO Exits	(2) Fraction of Failed Exits	(3) Fraction of IPO Exits	(4) Fraction of Failed Exits
Average Fund Level Breadth Index	-0.016*	-0.020	-0.006***	-0.002
	(0.010)	(0.027)	(0.002)	(0.007)
Novelty (Distance to Closest Firm)	0.664***	0.602***		
	(0.050)	(0.143)		
Average Fund Level Breadth Index × Novelty (Distance to Closest Firm)	0.072*	0.082		
	(0.037)	(0.106)		
Fraction of Top Quartile Novelty Firms			0.110***	0.100***
			(0.009)	(0.025)
Average Fund Level Breadth Index × Fraction of Top Quartile Novelty Firms			0.033***	0.010
			(0.008)	(0.023)
log_exp	0.002	-0.032***	0.002	-0.033***
	(0.002)	(0.004)	(0.002)	(0.004)
Industry Controls	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓
Vintage Year × Country × FundType FE	✓	✓	✓	✓
Observations	2932.00	2932.00	2932.00	2932.00
R^2	0.34	0.53	0.34	0.53

Table 16: **Human Capital Breadth, Firm Novelty and Investor Performance in the yearly years of investor's existence** This table reports the results of a regression of investor performance measures on the interaction between the average novelty of a financed firm and the average human capital breadth in the first six years of investor's existence. The dependent variable in columns (1) and (3) Fraction of IPO Exits is the fraction of IPO exited deals relative to the total number of deals made by the investor. The dependent variable in columns (2) and (4) Fraction of Failed Exits of deals who have failed (have not received follow up financing or have not exited via an IPO or an Acquisition) relative to the total number of deals made by the investor. In columns (1) and (2) Novelty (Distance to Closest) firm is the average novelty across all deals financed by the investor in the first six years of the investor's existence. In columns (3) and (4) Fraction of Top Quartile Novelty Firms the fraction of deals that are in top quartile of novelty for deals made in the first six years of investor's existence. In columns (1)-(4) Average Fund Level Breadth Index is the average breadth index, averaged across partners in the investor who have made deals in the first six years of the investor's existence. In all columns we include Industry Controls, which are separate controls for the investor's deal industry composition, Stage Controls which are separate controls for the investor's stage of investment composition. In all columns we include Investor Founding Year × Country FE. Standard errors reported in parenthesis are clustered at a country × Founding Year level * p<.10; ** p<.05; *** p<.01.

References

Abuzov, Rustam, "Busy venture capitalists and investment performance," *Available at SSRN 3353343*, 2019.

Bartik, Timothy J, "Who benefits from state and local economic development policies?," 1991.

Bergemann, Dirk and Juuso Valimaki, "Bandit problems," 2006.

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

Berry, Donald A and Bert Fristedt, "Bandit problems: sequential allocation of experiments (Monographs on statistics and applied probability)," *London: Chapman and Hall*, 1985, *5* (71-87), 7–7.

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.

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.

Duevski, Teodor and Viacheslav Bazaliy, "Capital Allocation Concentration Measurement in Venture Capital," 2025. Working paper.

Ewens, Michael 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.

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.

González-Uribe, Juanita, "Exchanges of innovation resources inside venture capital portfolios," *Journal of Financial Economics*, 2020, *135* (1), 144–168.

Gornall, Will and Ilya A Strebulaev, "The economic impact of venture capital: Evidence from public companies," *Available at SSRN 2681841*, 2021.

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.

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.

Kaplan, Steven N and Josh Lerner, "It ain't broke: The past, present, and future of venture capital," *Journal of Applied Corporate Finance*, 2010, *22* (2), 36–47.

— **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.

Kempf, Alexander and Stefan Ruenzi, "Tournaments in mutual-fund families," *The Review of Financial Studies*, 2008, *21* (2), 1013–1036.

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, "Assessing the contribution of venture capital," *the RAND Journal of Economics*, 2000, *31* (4), 674–692.

_ **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.

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.

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

Samila, Sampsa and Olav Sorenson, "Venture capital, entrepreneurship, and economic growth," *The Review of Economics and Statistics*, 2011, *93* (1), 338–349.

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.

Appendix

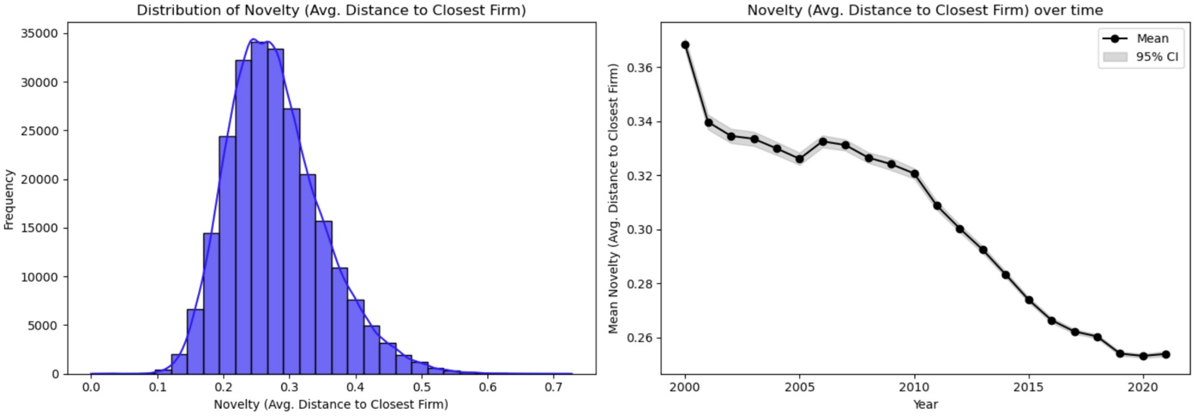

Figure A1: **Distribution of Novelty (Avg. Distance to Closest Firm) and time trend of Novelty (Avg. Distance to Closest Firm)** Left panel: This figure plots the distribution of the Novelty (Avg. Distance to Closest Firm) measure. Right Panel: Time trend of the mean of Novelty (Avg. Distance to Closest Firm) measure.

	(1) Failure	(2) Major Success	(3) IPO Exit	(4) Failure	(5) Major Success	(6) IPO Exit
Novelty (Avg. Distance to Closest Firms)	0.358***	0.753***	0.824***	0.141***	0.794***	0.846***
	(0.021)	(0.026)	(0.027)	(0.026)	(0.040)	(0.042)
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00
R^2	0.18	0.13	0.16	0.53	0.47	0.52

Table A2: **Association between startup novelty and likelihood of Failure, Major Success and IPO Exit - robustness to an alternative novelty measure** This table reports the results of a deal - level regression of Failure, Major Success and IPO Exit on the deal's novelty. The independent variable Novelty (Avg. Distance to Closest Firm) is a deal level measure of novelty defined as one - average of the cosine similarity of the textual description of the startup financed in the deal and top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. IPO Exit is an indicator variable taking a value of 1 if the firm exits via an IPO. Major Success is an indicator variable taking a value of 1 if the firm exits via an IPO or is acquired at a valuation at least five times greater than the total financing received. Failure is an indicator taking a value of 1 if the firm does not exit via an IPO, is not Acquired or does not receive any follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

Startup	Novelty Measure	Closest Startup	Novelty Percentile
Tesla	0.522353	Community Energy	99.875748
Facebook	0.507300	Mode Media	99.790265
Square	0.484537	Softgate Systems	99.571380
SpaceX	0.482793	Zero-G	99.529481
Slack	0.476927	Octopz	99.437496
Airbnb	0.454158	Dopplr	99.029344
Skype	0.448180	Arrival Communications	98.861507
SpaceX	0.445580	Blastoff!	98.801790
Dropbox	0.373166	Omnidrive	94.215553
Napster	0.266652	Deezer	64.103052
Revolut	0.195093	Wise (Application)	22.929740
Instagram	0.187039	Pixable	18.819387
Stripe	0.132001	Moip	2.470587

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.

	(1) Failure	(2) Major Success	(3) IPO Exit	(4) Failure	(5) Major Success	(6) IPO Exit
Novelty Quartile (Avg. Distance to Closest Firms)=2	0.029***	-0.004***	-0.000	0.020***	-0.002	0.002
	(0.003)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)
Novelty Quartile (Avg. Distance to Closest Firms)=3	0.052***	0.001	0.007***	0.032***	0.006***	0.009***
	(0.003)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)
Novelty Quartile (Avg. Distance to Closest Firms)=4	0.074***	0.105***	0.116***	0.036***	0.113***	0.121***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00
R^2	0.18	0.12	0.15	0.53	0.47	0.51

Table A3: **Association between startup novelty quartile and likelihood of Failure, Major Success and IPO Exit - robustness to an alternative novelty measure** This table reports the results of a deal - level regression of Failure, Major Success and IPO Exit on the deal's novelty quartile. The independent variables are novelty quartile dummies indicating whether a startup belongs to the i the novelty quartile within a given year and deal stage based on the Novelty (Avg. Distance to Closest Firm) which is a deal level measure of novelty defined as one - average of the cosine similarity of the textual description of the startup financed in the deal and top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. The omitted category is Novelty Quartile (Distance to Closest Firm) = 1. IPO Exit is an indicator variable taking a value of 1 if the firm exits via an IPO. Major Success is an indicator variable taking a value of 1 if the firm exits via an IPO or is acquired at a valuation at least five times greater than the total financing received. Failure is an indicator taking a value of 1 if the firm does not exit via an IPO, is not Acquired or does not receive any follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Number of Patents	(2) Number Citations	(3) Number of Claims	(4) Number of Patents	(5) Number Citations	(6) Number of Claims
Novelty (Avg. Distance to Closest Firms)	1.667*** (0.439)	2.908*** (0.506)	2.060*** (0.530)	1.398 (0.877)	3.096*** (1.081)	1.664* (0.998)
Exit Type Controls	✓	✓	✓	✓	✓	✓
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00

Table A4: **Association between startup novelty and innovation outcomes - Poisson count model - robustness to an alternative novelty measure** This table reports the results of a deal - level regression of innovation outcomes on the deal's novelty. The independent variable Novelty (Avg. Distance to Closest Firm) is a deal level measure of novelty defined as one - average of the cosine similarity of the textual description of the startup financed in the deal and top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. Number of Patents is the total number of patents granted to the firm after the deal date. Number Citations is the total number of adjusted citations (by grant year and NBER subcategory) that patents of the financed firm have received after the deal date. Number of Claims is total claims in all granted patents after the deal date of the financed firm. All columns include Exit Type controls which are separate controls specifying whether the startup goes public is acquired or receives follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Log(1+Number of Patents)	(2) Log(1+Number of Citations)	(3) Log(1+ Number of Claims)	(4) Log(1+Number of Patents)	(5) Log(1+Number of Citations)	(6) Log(1+ Number of Claims)
Novelty (Distance to Closest Firm)	0.052*** (0.013)	0.098*** (0.024)	0.198*** (0.039)	0.072*** (0.023)	0.145*** (0.042)	0.262*** (0.065)
Exit Type Controls	✓	✓	✓	✓	✓	✓
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00
R^2	0.17	0.17	0.20	0.47	0.47	0.49

Table A5: **Association between startup novelty and innovation outcomes - Logarithmic model** This table reports the results of a deal - level regression of innovation outcomes on the deal's novelty. The independent variable Novelty (Distance to Closest Firm) is a deal level measure of novelty defined as one - 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. log(1+Number of Patents) is logarithm of one plus the total number of patents granted to the firm after the deal date. log(1+Number of Citations) is the logarithm of one plus the total number of adjusted citations (by grant year and NBER subcategory) that patents of the financed firm have received after the deal date. log(1+Number of Claims) is the logarithm of one plus the total claims in all granted patents after the deal date of the financed firm. All columns include Exit Type controls which are separate controls specifying whether the startup goes public is acquired or receives follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Log(1+Number of Patents)	(2) Log(1+Number of Citations)	(3) Log(1+ Number of Claims)	(4) Log(1+Number of Patents)	(5) Log(1+Number of Citations)	(6) Log(1+ Number of Claims)
Novelty (Avg. Distance to Closest Firms)	0.064*** (0.014)	0.121*** (0.026)	0.220*** (0.041)	0.078*** (0.024)	0.156*** (0.044)	0.269*** (0.069)
Exit Type Controls	✓	✓	✓	✓	✓	✓
Industry × Deal Year × Deal Type FE	✓	✓	✓			
Industry × Deal Year × Deal Type × Investor FE				✓	✓	✓
Observations	241916.00	241916.00	241916.00	241916.00	241916.00	241916.00
R^2	0.17	0.17	0.20	0.47	0.47	0.49

Table A6: **Association between startup novelty and innovation outcomes - Logarithmic model - robustness to an alternative novelty measure** This table reports the results of a deal - level regression of innovation outcomes on the deal's novelty. The independent variable Novelty (Avg. Distance to Closest Firm) is a deal level measure of novelty defined as one - average of the cosine similarity of the textual description of the startup financed in the deal and top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. log(1+Number of Patents) is logarithm of one plus the total number of patents granted to the firm after the deal date. log(1+Number of Citations) is the logarithm of one plus the total number of adjusted citations (by grant year and NBER subcategory) that patents of the financed firm have received after the deal date. log(1+Number of Claims) is the logarithm of one plus the total claims in all granted patents after the deal date of the financed firm. All columns include Exit Type controls which are separate controls specifying whether the startup goes public is acquired or receives follow-up financing. Columns (1) - (3) include Industry × Deal Year × Deal Type FE. Columns (4) - (6) include Industry × Deal Year × Deal Type × Investor FE. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Novelty (Avg. Distance to Closest Firm)	(2) Novelty (Avg. Distance to Closest Firm)	(3) Novelty (Avg. Distance to Closest Firm)	(4) Novelty (Avg. Distance to Closest Firm)
Breadth Index	0.002*** (0.000)	0.002*** (0.001)	0.002** (0.001)	0.003** (0.001)
VC Experience	-0.001* (0.000)	0.000 (0.000)	0.000 (0.001)	0.000 (0.001)
Controls			✓	✓
Industry × Deal Year × Deal Type FE	✓		✓	
Industry × Deal Year × Deal Type × Investor FE		✓		✓
Observations	42317.00	42317.00	12372.00	12372.00
R^2	0.24	0.35	0.24	0.34

Table A7: **Association between lead partner's human capital breadth index and startup novelty - robustness to an alternative novelty measure** This table reports the results of an OLS regression of deal novelty and lead partner's human capital breadth. The dependent variable Novelty (Avg. Distance to Closest Firm) is a deal level measure of novelty defined as one - average of the cosine similarity of the textual description of the startup financed in the deal and top five closest 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 defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. 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) and (2) include Industry × Deal Year × Deal Type FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (4) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Novelty (Avg. Distance to Closest Firm)	(2) Novelty (Avg. Distance to Closest Firm)	(3) Novelty (Avg. Distance to Closest Firm)	(4) Novelty (Avg. Distance to Closest Firm)
Breadth Index	0.002***	0.002***	0.002*	0.003**
	(0.000)	(0.001)	(0.001)	(0.002)
VC Experience	-0.001	0.000	0.000	0.001
	(0.000)	(0.000)	(0.001)	(0.001)
Controls			✓	✓
Industry × Deal Year × Deal Type FE	✓		✓	
Industry × Deal Year × Deal Type × Investor FE		✓		✓
Observations	37896.00	37896.00	11270.00	11270.00
R^2	0.25	0.35	0.24	0.35

Table A8: **Association between lead partner's human capital breadth index and startup novelty - robustness to an alternative novelty measure & robustness to removing bottom decile of textual description lengths** This table reports the results of an OLS regression of deal novelty and lead partner's human capital breadth in the estimation we drop the bottom decile of length partner's background textual descriptions. The dependent variable Novelty (Avg. Distance to Closest Firm) is a deal level measure of novelty defined as one - average of the cosine similarity of the textual description of the startup financed in the deal and top five closest 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 defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. 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) and (2) include Industry × Deal Year × Deal Type FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (4) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) IPO Exit	(2) IPO Exit	(3) IPO Exit	(4) Failure	(5) Failure	(6) Failure
Breadth Index	-0.067***	-0.066***	-0.054**	0.043***	0.026**	0.016
	(0.009)	(0.009)	(0.021)	(0.010)	(0.010)	(0.022)
Novelty (Avg. Distance to Closest Firm)	0.973***	1.013***	1.072***	0.260***	0.168***	0.194***
	(0.037)	(0.039)	(0.080)	(0.037)	(0.038)	(0.070)
Breadth Index × Novelty (Avg. Distance to Closest Firm)	0.262***	0.247***	0.193**	-0.118***	-0.072**	-0.026
	(0.034)	(0.036)	(0.080)	(0.033)	(0.033)	(0.075)
VC Experience	0.005***	0.002	0.005*	-0.043***	-0.019***	-0.023***
	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.006)
Controls			✓			✓
Industry × Deal Year × Deal Type FE	✓			✓		
Industry × Deal Year × Deal Type × Investor FE		✓	✓		✓	✓
Observations	42317.00	42317.00	12372.00	42317.00	42317.00	12372.00
R^2	0.20	0.30	0.35	0.21	0.34	0.34

Table A9: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup outcome - robustness to an alternative novelty measure.** This table reports the results of an OLS regression of startup outcome on the interaction between deal's novelty and lead partner's human capital breadth. The independent variable Avg. Novelty (Distance to Closest Firm) is a deal level measure of novelty defined as one minus the average of the cosine similarity of the textual description of the startup financed in the deal and top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. The independent variable Breadth Index is a measure of human capital breadth defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. Breadth Index × Novelty (Distance to Closest Firm) is the interaction between the deal's novelty and the lead partner's human capital breadth index. In columns (1)-(3) the dependent variable IPO Exit is an indicator that equals 1 if the startup goes public. 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) and (2) include Industry × Deal Year × Deal Type FE. In columns (1)-(3) the dependent variable IPO Exit is an indicator that equals 1 if the startup goes public. In columns (4)-(6) the dependent variable Failure is an indicator taking a value 1 if the startup does not go public is not acquired or does not receive follow - up - financing. Columns (1) and (4) include Industry × Deal Year × Deal Type FE. Columns (2), (3), (5) and (6) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (6) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) IPO Exit	(2) IPO Exit	(3) IPO Exit	(4) Failure	(5) Failure	(6) Failure
Breadth Index	-0.067*** (0.008)	-0.065*** (0.009)	-0.064*** (0.021)	0.038*** (0.009)	0.028*** (0.010)	0.014 (0.022)
Novelty (Distance to Closest Firm)	0.956*** (0.038)	0.980*** (0.040)	1.029*** (0.081)	0.207*** (0.036)	0.131*** (0.038)	0.137* (0.071)
Breadth Index × Novelty (Distance to Closest Firm)	0.293*** (0.037)	0.275*** (0.038)	0.283*** (0.087)	-0.114*** (0.033)	-0.076** (0.034)	-0.024 (0.081)
VC Experience	0.004*** (0.001)	0.002 (0.001)	0.005* (0.003)	-0.040*** (0.003)	-0.017*** (0.003)	-0.024*** (0.007)
Controls			✓			✓
Industry × Deal Year × Deal Type FE	✓			✓		
Industry × Deal Year × Deal Type × Investor FE		✓	✓		✓	✓
Observations	37896.00	37896.00	11270.00	37896.00	37896.00	11270.00
R^2	0.21	0.31	0.36	0.20	0.33	0.35

Table A10: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup outcome - split by novelty quartile.** This table reports the results of an OLS regression of startup outcome on the interaction between deal's novelty and lead partner's human capital breadth. The independent variables are novelty quartile dummies indicating whether a startup belongs to the i the novelty quartile within a given year and deal stage based on the Novelty (Distance to Closest Firm) measure which is a deal level measure of novelty defined as one - 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 Breadth Index is a measure of human capital breadth defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. Breadth Index × Novelty Quartile (Distance to Closest Firm) is the interaction between the deal's novelty quartile and the lead partner's human capital breadth index. In columns (1)-(3) the dependent variable IPO Exit is an indicator that equals 1 if the startup goes public. 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) and (2) include Industry × Deal Year × Deal Type FE. In columns (1)-(3) the dependent variable IPO Exit is an indicator that equals 1 if the startup goes public. In columns (4)-(6) the dependent variable Failure is an indicator taking a value 1 if the startup does not go public is not acquired or does not receive follow - up - financing. Columns (1) and (4) include Industry × Deal Year × Deal Type FE. Columns (2), (3), (5) and (6) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (6) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) IPO Exit	(2) IPO Exit	(3) IPO Exit	(4) Failure	(5) Failure	(6) Failure
Breadth Index	-0.005***	-0.008***	-0.009	0.012**	0.005	0.005
	(0.002)	(0.002)	(0.006)	(0.005)	(0.005)	(0.010)
Novelty Quartile (Distance to Closest Firm)=2	-0.000	0.001	-0.005	0.027***	0.021***	0.004
	(0.001)	(0.002)	(0.004)	(0.006)	(0.006)	(0.010)
Novelty Quartile (Distance to Closest Firm)=3	0.006***	0.010***	0.007	0.037***	0.029***	0.029***
	(0.002)	(0.002)	(0.004)	(0.006)	(0.006)	(0.011)
Novelty Quartile (Distance to Closest Firm)=4	0.136***	0.140***	0.151***	0.047***	0.031***	0.026**
	(0.006)	(0.006)	(0.013)	(0.007)	(0.007)	(0.013)
Novelty Quartile (Distance to Closest Firm)=2 × Breadth Index	0.001	0.003	0.004	0.003	0.003	-0.001
	(0.002)	(0.002)	(0.005)	(0.006)	(0.006)	(0.011)
Novelty Quartile (Distance to Closest Firm)=3 × Breadth Index	0.006***	0.007***	0.004	0.001	0.006	0.018
	(0.002)	(0.003)	(0.005)	(0.006)	(0.006)	(0.013)
Novelty Quartile (Distance to Closest Firm)=4 × Breadth Index	0.044***	0.042***	0.036***	-0.012*	-0.006	-0.002
	(0.006)	(0.006)	(0.014)	(0.006)	(0.007)	(0.014)
VC Experience	0.005***	0.002*	0.006**	-0.043***	-0.019***	-0.023***
	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.006)
Controls			✓			✓
Industry × Deal Year × Deal Type FE	✓			✓		
Industry × Deal Year × Deal Type × Investor FE		✓	✓		✓	✓
Observations	42317.00	42317.00	12372.00	42317.00	42317.00	12372.00
R^2	0.18	0.28	0.33	0.21	0.34	0.35

Table A11: **Association between the interaction of lead partner's human capital breadth and deal novelty and startup outcome - split by novelty quartile.** This table reports the results of an OLS regression of deal novelty and lead partner's human capital breadth. The dependent variable Novelty (Distance to Closest Firm) 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 defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. 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) and (2) include Industry × Deal Year × Deal Type FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (4) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) IPO Exit	(2) IPO Exit	(3) IPO Exit	(4) Failure	(5) Failure	(6) Failure
Breadth Index	-0.005*** (0.002)	-0.008*** (0.002)	-0.009 (0.006)	0.012** (0.005)	0.005 (0.005)	0.005 (0.010)
Novelty Quartile (Distance to Closest Firm)=2	-0.000 (0.001)	0.001 (0.002)	-0.005 (0.004)	0.027*** (0.006)	0.021*** (0.006)	0.004 (0.010)
Novelty Quartile (Distance to Closest Firm)=3	0.006*** (0.002)	0.010*** (0.002)	0.007 (0.004)	0.037*** (0.006)	0.029*** (0.006)	0.029*** (0.011)
Novelty Quartile (Distance to Closest Firm)=4	0.136*** (0.006)	0.140*** (0.006)	0.151*** (0.013)	0.047*** (0.007)	0.031*** (0.007)	0.026** (0.013)
Novelty Quartile (Distance to Closest Firm)=2 × Breadth Index	0.001 (0.002)	0.003 (0.002)	0.004 (0.005)	0.003 (0.006)	0.003 (0.006)	-0.001 (0.011)
Novelty Quartile (Distance to Closest Firm)=3 × Breadth Index	0.006*** (0.002)	0.007*** (0.003)	0.004 (0.005)	0.001 (0.006)	0.006 (0.006)	0.018 (0.013)
Novelty Quartile (Distance to Closest Firm)=4 × Breadth Index	0.044*** (0.006)	0.042*** (0.006)	0.036*** (0.014)	-0.012* (0.006)	-0.006 (0.007)	-0.002 (0.014)
VC Experience	0.005*** (0.001)	0.002* (0.001)	0.006** (0.003)	-0.043*** (0.003)	-0.019*** (0.003)	-0.023*** (0.006)
Controls			✓			✓
Industry × Deal Year × Deal Type FE	✓			✓		
Industry × Deal Year × Deal Type × Investor FE		✓	✓		✓	✓
Observations	42317.00	42317.00	12372.00	42317.00	42317.00	12372.00
R^2	0.18	0.28	0.33	0.21	0.34	0.35

Table A12: **Association between lead partner's human capital breadth index and startup novelty** This table reports the results of an OLS regression of deal novelty and lead partner's human capital breadth. The dependent variable Novelty (Distance to Closest Firm) 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 defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. 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) and (2) include Industry × Deal Year × Deal Type FE. Columns (2) and (4) include Industry × Deal Year × Deal Type × Investor FE. Columns (3) and (4) include additional partner level controls. Standard errors reported in parenthesis are clustered at an Investor level. * p<.10; ** p<.05; *** p<.01.

	(1) Partner Leads a Deal	(2) Partner Leads a Deal	(3) Partner Leads a Deal	(4) Partner Leads a Deal
Breadth Index	-0.010** (0.004)	-0.012*** (0.004)	-0.019** (0.009)	-0.025*** (0.009)
Breadth Index × Novelty (Avg. Distance to Closest Firm)	0.053*** (0.014)	0.055*** (0.014)	0.073** (0.032)	0.084*** (0.031)
Partner Age			-0.002*** (0.000)	-0.003*** (0.000)
Female=1			0.010 (0.006)	0.011* (0.006)
VC Experience			0.058*** (0.002)	0.061*** (0.002)
Busy Partner			-0.016** (0.007)	-0.018*** (0.006)
Deal FE	✓	✓	✓	✓
Investor FE		✓		✓
Observations	207768.00	207768.00	64221.00	64221.00
R^2	0.14	0.17	0.27	0.31

Table A13: **Association between the interaction of lead partner's human capital breadth and deal novelty and the likelihood of a partner leading a deal - robustness to alternative novelty measure.** This table reports the results of a regression of the likelihood of leading a deal on partner's human capital breadth and the interaction between deal novelty and partner's human capital breadth. Each observation is a partner-deal-investor. The dependent variable Partner Leads a Deal is an indicator equal to 1 if a partner leads a deal and 0 otherwise. The independent variable Novelty (Avg. Distance to Closest Firm) is a deal level measure of novelty defined as one minus the average of the cosine similarity of the textual description of the startup financed in the deal and top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. The independent variable Breadth Index is a measure of human capital breadth defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. Breadth Index × Novelty (Avg. Distance to Closest Firm) is the interaction between the deal's novelty and the lead partner's human capital breadth index. In columns (3) and (4) Partner Age is the age of the partner at the time of the deal. Female is an indicator variable equal to 1 if the partner is female and 0 otherwise. 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. Busy Partner is an indicator equal to 1 if the partner leads a different deal that is at the time of the current deal undergoing a public listing (IPO) or acquisition. In columns (1) and (3) we include a Deal FE, in columns (2) and (3) an Deal FE and an Investor FE. Standard errors reported in parenthesis are clustered at a Deal level. * p<.10; ** p<.05; *** p<.01.

	(1) Partner Leads a Deal	(2) Partner Leads a Deal	(3) Partner Leads a Deal	(4) Partner Leads a Deal
Breadth Index	-0.009** (0.004)	-0.012*** (0.004)	-0.011 (0.009)	-0.020** (0.009)
Breadth Index × Novelty (Distance to Closest Firm)	0.050*** (0.017)	0.054*** (0.016)	0.054 (0.036)	0.077** (0.036)
Partner Age			-0.003*** (0.000)	-0.003*** (0.000)
Female=1			0.004 (0.008)	0.003 (0.008)
VC Experience			0.047*** (0.002)	0.052*** (0.002)
Busy Partner			-0.012* (0.007)	-0.015** (0.007)
Deal FE	✓	✓	✓	✓
Investor FE		✓		✓
Observations	179910.00	179910.00	56924.00	56924.00
R^2	0.15	0.18	0.27	0.32

Table A14: **Association between the interaction of lead partner's human capital breadth and deal novelty and the likelihood of a partner leading a deal - robustness to alternative horizon around a deal (-3, 1)** This table reports the results of a regression of the likelihood of leading a deal on partner's human capital breadth and the interaction between deal novelty and partner's human capital breadth. Each observation is a partner-deal-investor. The dependent variable Partner Leads a Deal is an indicator equal to 1 if a partner leads a deal and 0 otherwise. The independent variable Novelty (Avg. Distance to Closest Firm) is a deal level measure of novelty defined as one minus the average of the cosine similarity of the textual description of the startup financed in the deal and top five closest startups receiving venture capital financing within five years before the deal in the same deal stage. The independent variable Breadth Index is a measure of human capital breadth defined as the $\frac{S_b - S_n}{S_b + S_n}$ where S_b is the lead partner's textual description similarity to a broad human capital partner prototype and S_n is the lead partner's textual similarity to a narrow human capital partner prototype. Breadth Index × Novelty (Avg. Distance to Closest Firm) is the interaction between the deal's novelty and the lead partner's human capital breadth index. In columns (3) and (4) Partner Age is the age of the partner at the time of the deal. Female is an indicator variable equal to 1 if the partner is female and 0 otherwise. 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. Busy Partner is an indicator equal to 1 if the partner leads a different deal that is at the time of the current deal undergoing a public listing (IPO) or acquisition. In columns (1) and (3) we include a Deal FE, in columns (2) and (3) an Deal FE and an Investor FE. Standard errors reported in parenthesis are clustered at a Deal level. * p<.10; ** p<.05; *** p<.01.

	(1) Breadth Index	(2) Breadth Index × Novelty (Avg. Distance to Closest Firm)	(3) IPO Exit	(4) IPO Exit	(5) Failure	(6) Failure
Avg. Available Breadth	0.313*** (0.061)	-0.157*** (0.020)				
Avg. Available Breadth × Novelty (Avg. Distance to Closest Firm)	-0.216 (0.185)	0.836*** (0.069)				
Breadth Index			-0.196*** (0.039)	-0.068*** (0.024)	0.090 (0.083)	-0.006 (0.028)
Breadth Index × Novelty (Avg. Distance to Closest Firm)			0.582*** (0.124)	0.248*** (0.092)	-0.253 (0.159)	0.074 (0.094)
Novelty (Avg. Distance to Closest Firm)			0.978*** (0.088)	0.951*** (0.086)	0.202** (0.092)	0.220** (0.087)
Controls FE	✓	✓	✓	✓	✓	✓
Investor FE	✓	✓	✓	✓	✓	✓
Deal Stage × Industry × Year FE	✓	✓	✓	✓	✓	✓
Observations	8371.00	8371.00	8371.00	8371.00	8371.00	8371.00
R^2	0.71	0.70	0.09	0.33	0.00	0.37
F-statistic of Instrument			61.32		61.32	

Table A15: **Effect of the interaction of lead partner's human capital breadth and deal novelty on startup outcomes. - Robustness to an Alternative Novelty Measure** This table presents the results of a instrumental variable regression of deal performance on the interaction between lead partner breadth and deal novelty. Column (1) presents the first stage regressions of the first 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 at a VC firm level across the firm's available partners. Column (2) presents the first stage regression of the second instrumented variable Breadth Index × Novelty on the two instruments Avg. Available Breadth and Avg. Available Breadth × Novelty. Columns (3) and (5) present the IV regression where the dependent variables are IPO Exit and Failure Respectively. Columns (4) and (6) present the OLS estimates of an equivalent model.

	(1) Net Multiple	(2) Net Multiple	(3) Net Multiple	(4) Net Multiple	(5) Net Multiple	(6) Net Multiple
Fraction of IPO Exits	3.404* (1.907)			4.468** (2.129)		
Fraction of Failed Exits		-0.818** (0.367)			-0.776 (0.491)	
Fraction of Acquisition Exits			-0.481 (1.036)			-1.193 (1.448)
Fund Size	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
First Fund=1	0.302 (0.246)	0.228 (0.244)	0.246 (0.243)	-0.273 (0.427)	-0.313 (0.426)	-0.415 (0.475)
Industry Controls	✓	✓	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓	✓	✓
Vintage Year FE	✓	✓	✓			
Country FE	✓	✓	✓			
Fund Type FE	✓	✓	✓			
Vintage Year × Country × FundType FE				✓	✓	✓
Observations	690.00	690.00	690.00	690.00	690.00	690.00
R^2	0.26	0.24	0.24	0.48	0.46	0.46

Table A16: **Validation of Exit Measure Used: Pitchbook VC Funds with Performance Measure in Preqin** This table reports the results of a regression of realized fund performance on the Fraction of IPO, Failed and Acquisition Exits. The dependent variable in columns (1)-(6) is the Net Multiple of a fund defined as the total value of distributions and unrealized gains to investors relative to the total value invested. In columns (1) and (4) Fraction of IPO Exits is the fraction of IPO exited deals relative to the total number of deals made by the fund. In columns (2) and (5) Fraction of Failed Exits of deals who have failed (have not received follow up financing or have not exited via an IPO or an Acquisition) relative to the total number of deals made by the fund. In columns (3) and (6) Fraction of Acquisition Exits is the fraction of Acquisition exited deals relative to the total number of deals made by the fund. Fund Size is the size of the fund. First Fund is an indicator taking a value of 1 if the Fund is a First Fund raised by a given VC firm. In all columns we include Industry Controls, which are separate controls for the fund's industry composition, Stage Controls which are separate controls for the fund's stage of investment composition. In columns (1) - (3) we include Vintage Year FE, Country FE and Fund Type FE. In columns (4)-(6) we include Vintage Year × Country × Fund Type FE. Standard errors reported in parenthesis are clustered at an Investor level * p<.10; ** p<.05; *** p<.01.

	(1) Fraction of IPO Exits	(2) Fraction of Failed Exits	(3) Fraction of IPO Exits	(4) Fraction of Failed Exits
Average Fund Level Breadth Index	-0.100*** (0.023)	0.101*** (0.031)	-0.010** (0.005)	0.015** (0.007)
Novelty (Avg. Distance to Closest Firm)	1.175*** (0.123)	0.044 (0.162)		
Average Fund Level Breadth Index × Novelty (Avg. Distance to Closest Firm)	0.402*** (0.088)	-0.366*** (0.104)		
Novelty3_emb1_Quartile_Q4_Fraction			0.174*** (0.018)	0.047* (0.027)
Average Fund Level Breadth Index × Novelty3_emb1_Quartile_Q4_Fraction			0.111*** (0.025)	-0.083*** (0.027)
Fund Size	0.000*** (0.000)	-0.000 (0.000)	0.000*** (0.000)	-0.000 (0.000)
First Fund=1	-0.015*** (0.004)	0.008 (0.009)	-0.015*** (0.004)	0.009 (0.009)
Industry Controls	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓
Vintage Year × Country × FundType FE	✓	✓	✓	✓
Observations	7364.00	7364.00	7364.00	7364.00
R^2	0.53	0.61	0.52	0.61

Table A17: **Human Capital Breadth, Firm Novelty and Fund Performance - Robustness to Alternative Novelty Measure**This table reports the results of a regression of fund performance measures on the interaction between the average novelty of a financed firm and the average human capital breadth in a fund partnership. The dependent variable in columns (1) and (3) Fraction of IPO Exits is the fraction of IPO exited deals relative to the total number of deals made by the fund. The dependent variable in columns (2) and (4) Fraction of Failed Exits of deals who have failed (have not received follow up financing or have not exited via an IPO or an Acquisition) relative to the total number of deals made by the fund. In columns (1) and (2) Novelty (Distance to Closest) firm is the average novelty across all deals financed by the fund. In columns (3) and (4) Fraction of Top Quartile Novelty Firms the fraction of deals that are in top quartile of novelty. In columns (1)-(4) Average Fund Level Breadth Index is the average breadth index, averaged across partners who have made deals in the fund. Fund Size is the size of the fund. First Fund is an indicator taking a value of 1 if the Fund is a First Fund raised by a given VC firm. In all columns we include Industry Controls, which are separate controls for the fund's industry composition, Stage Controls which are separate controls for the fund's stage of investment composition. In all columns we include Vintage Year × Country × Fund Type FE. Standard errors reported in parenthesis are clustered at an Investor level * p<.10; ** p<.05; *** p<.01.

	(1) Fraction of IPO Exits	(2) Fraction of Failed Exits	(3) Fraction of IPO Exits	(4) Fraction of Failed Exits
Average Fund Level Breadth Index	-0.017*	-0.019	-0.005**	-0.004
	(0.010)	(0.028)	(0.002)	(0.007)
Novelty (Avg. Distance to Closest Firm)	0.636***	0.569***		
	(0.050)	(0.142)		
Average Fund Level Breadth Index × Novelty (Avg. Distance to Closest Firm)	0.065*	0.068		
	(0.035)	(0.099)		
Fraction of Top Quartile Novelty Firms			0.100***	0.092***
			(0.009)	(0.024)
Average Fund Level Breadth Index × Fraction of Top Quartile Novelty Firms			0.028***	0.019
			(0.008)	(0.022)
log_exp	0.002	-0.033***	0.002	-0.033***
	(0.002)	(0.004)	(0.002)	(0.004)
Industry Controls	✓	✓	✓	✓
Stage Controls	✓	✓	✓	✓
Vintage Year × Country × FundType FE	✓	✓	✓	✓
Observations	2932.00	2932.00	2932.00	2932.00
R^2	0.33	0.53	0.33	0.53

Table A18: **Human Capital Breadth, Firm Novelty and Investor Performance - Robustness to Alternative Novelty Measure** This table reports the results of a regression of investor performance measures on the interaction between the average novelty of a financed firm and the average human capital breadth in the first six years of investor's existence. The dependent variable in columns (1) and (3) Fraction of IPO Exits is the fraction of IPO exited deals relative to the total number of deals made by the investor. The dependent variable in columns (2) and (4) Fraction of Failed Exits of deals who have failed (have not received follow up financing or have not exited via an IPO or an Acquisition) relative to the total number of deals made by the investor. In columns (1) and (2) Novelty (Distance to Closest) firm is the average novelty across all deals financed by the investor in the first six years of the investor's existence. In columns (3) and (4) Fraction of Top Quartile Novelty Firms the fraction of deals that are in top quartile of novelty for deals made in the first six years of investor's existence. In columns (1)-(4) Average Fund Level Breadth Index is the average breadth index, averaged across partners in the investor who have made deals in the first six years of the investor's existence. In all columns we include Industry Controls, which are separate controls for the investor's deal industry composition, Stage Controls which are separate controls for the investor's stage of investment composition. In all columns we include Investor Founding Year × Country FE. Standard errors reported in parenthesis are clustered at a country × Founding Year level * p<.10; ** p<.05; *** p<.01.

A Optimal Strategy Benchmark Case Example

We study a two-sector decision model where an agent aims to maximize the probability of reaching a threshold T over d draws (steps). At each step, the agent can choose between:

- **Sector 1 (Safe)**: Provides a higher probability of moderate payoffs (x_{int}), but a lower probability of large payoffs (x_{good}).
- **Sector 2 (Risky)**: Provides a higher probability of failure (x_{fail}), but a higher probability of large payoffs (x_{good}).

The agent's decision at each step depends on the accumulated reward R and the number of steps left d. The goal is to maximize the probability of reaching or exceeding the threshold T.

We assume the following payoffs:

- Failure payoff: $x_{fail} = 0$.
- Intermediate payoff: x_{int}, where $x_{int} > 0$.
- Good payoff: x_{good}, where $x_{good} > x_{int}$.

The probabilities for the two urns are assumed to satisfy the following conditions:

Safe sector

- $P_1(x_{fail})$: Lower probability of failure.
- $P_1(x_{int})$: Higher probability of intermediate payoff x_{int}.
- $P_1(x_{good})$: Lower probability of the good payoff x_{good}.

Risky sector

- $P_2(x_{fail})$: Higher probability of failure, i.e., $P_2(x_{fail}) > P_1(x_{fail})$.
- $P_2(x_{int})$: Lower probability of intermediate payoff x_{int}, i.e., $P_2(x_{int}) < P_1(x_{int})$.
- $P_2(x_{good})$: Higher probability of good payoff x_{good}, i.e., $P_2(x_{good}) > P_1(x_{good})$.

The value function $V(d, R)$ represents the maximum probability of reaching or exceeding the threshold T given d steps remaining and an accumulated reward R. The Bellman equation is:

$$V(d, R) = \max \{V_1(d, R), V_2(d, R)\},$$

where:

$$V_1(d,R) = P_1(x_{fail}) \cdot V(d-1,R) + P_1(x_{int}) \cdot V(d-1,R+x_{int}) + P_1(x_{good}) \cdot V(d-1,R+x_{good}),$$

and

$$V_2(d,R) = P_2(x_{fail}) \cdot V(d-1,R) + P_2(x_{int}) \cdot V(d-1,R+x_{int}) + P_2(x_{good}) \cdot V(d-1,R+x_{good}).$$

B Some background on Dirichet priors

In this section, I provide a brief overview of Dirichlet updating rules. Suppose there is one sector of the economy and the true quality distribution of startups in this sector is given by a vector $P = (P(x_{low}), P(x_{int}), P(x_{high}))$. Suppose an agent starts with a Dirichlet prior over the outcomes given by a vector $\alpha_0 = (\alpha_0(x_{low}), \alpha_0(x_{int}), \alpha_0(x_{high}))$. let $\alpha_0 = \alpha_0(x_{low}) + \alpha_0(x_{int}) + \alpha_0(x_{high})$ Since the prior distribution is a Dirichlet with a prior described by α_0 the prior distribution satisfies:

$$p(P) = Dirichlet(P|\alpha_0) = \frac{1}{B(\alpha_0)} P(x_{low})^{\alpha_0(x_{low})-1} P(x_{int})^{\alpha_0(x_{int})-1} P(x_{high})^{\alpha_0(x_{high})-1}, \quad (22)$$

where $B(\alpha_0) = \frac{\Gamma(\alpha_0(x_{low}))\Gamma(\alpha_0(x_{int}))\Gamma(\alpha_0(x_{high}))}{\Gamma(\alpha_0)}$ is a normalization constant. Suppose we observe data, D from this sector generated by N draws and we observe outcomes the following number of times $n = (n_{low}, n_{int}, n_{high})$.

Then the likelihood function of the multinomial distribution is given by:

$$p(D|N,P) = \frac{N!}{n_{low}!n_{int}!n_{high}!} P(x_{low})^{n_{low}} P(x_{int})^{n_{int}} P(x_{high})^{n_{high}}, \quad (23)$$

Now we apply Bayes' formula $\left(p(P|D) = \frac{p(D|N,P)p(P)}{P(D)} \right)$ to calculate the posterior and obtain:

$$p(P|D) = \frac{N!}{B(\alpha_0)n_{low}!n_{int}!n_{high}!} P(x_{low})^{\alpha_0(x_{low})+n_{low}-1} P(x_{int})^{\alpha_0(x_{int})+n_{int}-1} P(x_{high})^{\alpha_0(x_{high})+n_{high}-1}. \quad (24)$$

From (24) it is obvious that the posterior is a Dirichlet distribution with parameters $\alpha_N = (\alpha_0(x_{low}) + n_{low}, \alpha_0(x_{int}) + n_{int}, \alpha_0(x_{high}) + n_{high})$