What Do Early Stage Investors Ask?
An LLM Analysis of Expert Calls

Victor Lyonnet Amin Shams Shaojun Zhang*

November 25, 2024

[Preliminary and Incomplete]

Abstract

We study how early-stage investors evaluate potential investments by using large language models (LLMs) to analyze 6,800 expert consultation calls. Not only do call volume and overall sentiment predict outcomes, but the specific content of discussions provides significant additional predictive power. Our topic-specific sentiment analysis shows that positive signals about technology integration and customer acquisition are associated with 15% and 16% higher deal likelihood, respectively. We find that the information content of the calls is particularly valuable for younger firms with limited track records, where information asymmetries are most severe. Our findings provide the first systematic evidence of how investors gather and process information in the absence of traditional financial metrics, and suggest some misalignment between topics that investors frequently discuss and those that best predict deal outcomes. Methodologically, we demonstrate the potential of LLMs to extract nuanced insights from complex qualitative data.

*Victor Lyonnet is with the University of Michigan (Email: vlyonnet@umich.edu), Amin Shams is with the Ohio State University (Email: shams.22@osu.edu), and Shaojun Zhang is with the Ohio State University (Email: zhang.7805@osu.edu).

1 Introduction

How do investors evaluate potential investments when traditional financial metrics are unavailable or uninformative? This is a key question in early-stage investing, where information asymmetries are severe and standard screening tools are often inadequate (Gompers and Lerner, 2001). Young ventures frequently develop novel technologies or business models in rapidly evolving sectors such as biotech, artificial intelligence, or renewable energy. The potential of these ventures is particularly difficult to evaluate due to limited historical data, emerging market dynamics, and technological uncertainty. While we know early-stage investors play a crucial role in funding innovation, the process through which they gather and evaluate information has remained largely a "black box" due to the complex, often qualitative nature of their due diligence.

We analyze the content and impact of expert consultations by developing a novel LLM-based methodology to study 6,800 calls between investors and industry experts. These calls are part of a growing expert network industry that connects investors with subject-matter experts. Our methodology identifies discussion topics and measures sentiment, revealing significant variations in how different topics predict investment decisions. We find different patterns in the predictive power of discussions on technology integration and customer adoption for investment decisions. We also find that investors probe different attributes of target companies in their conversations with industry experts, customers, and former executives. Our analysis suggests that consultations play a crucial role in investment decisions, with both the topics discussed and the sentiment of the discussions strongly predicting subsequent investment outcomes. Overall, our paper provides the first systematic evidence of what early-stage investors seek to learn and how they process and act on this information.

We use data from one key expert network used by over 2,500 institutions, including many of the top global VC firms. The dataset includes information on the date of the call, the name of the focal company, and the call transcript for 8,382 call transcripts covering 2,363 companies from January 2017 to January 2022. We merge this dataset with data on private deals from CBInsights and match 1,756 companies discussed in the call transcripts with deal events. In our preliminary results, we observe an 11 p.p. higher likelihood of a deal for a company in quarters following a call discussing this company. This relationship remains robust when accounting for quarter and company fixed effects, is stronger for calls with the companies' customers relative to other experts, and is particularly pronounced in sectors such as electronics, internet, mobile & telecom, and software.

Our analysis reveals several key patterns in how investors use expert networks. Series A and later-stage VC-backed companies are significantly more likely to be the subject of calls. This is also the case for companies in the digital sector, but not in hardware or retail & services. However, when controlling for funding history, younger companies at any given stage are more likely to receive expert consultation. Both the number and size of previous funding rounds positively predict expert calls, suggesting that investor interest and expert networks are particularly important for due diligence on younger but more mature startups that have achieved significant funding milestones.

We find a statistically and economically significant relationship between expert calls and subsequent deals. Companies that receive expert consultation in a given quarter are 11 percentage points more likely to complete a deal in the following quarter, a substantial effect given the baseline quarterly deal probability of 7.9 percent. This predictive relationship varies meaningfully across both expert types and sectors. Customer calls show the strongest association with future deals, followed by competitor and former executive consultations, while industry consultant calls have a weaker relationship. The link between expert calls and subsequent deals is particularly pronounced in technology-focused industries, with electronics, software, and internet companies showing the strongest effects.

Our LLM-based methodology allows us to study how information is conveyed through calls and how investors gather and process this information. Our approach uses ChatGPT4-Turbo to extract both discussion topics and their associated sentiment, along with a confidence score indicating the LLM's certainty about each extracted topic. We identify 40 topics ranging from technology integration to customer acquisition strategies, with sentiment scores ranging from -2 (strongly negative) to +2 (strongly positive).

Our topic modeling analysis reveals substantial variation in discussion topics and sentiments across experts. Competitors and partners focus heavily on technology integration, while former executives emphasize broader business strategy more than other experts. Industry consultants show the most balanced distribution of topics. Expert sentiment also varies systematically across both expert types and topics. Customers express the most consistently positive views, while other experts tend to be more measured in their assessments. Growth-related topics generally receive positive sentiment scores, while discussions of risk assessment and management evaluation generate more neutral or negative sentiment. The data suggest that investors strategically source different types of information from different experts, while potentially needing to adjust for systematic biases in their perspectives.

While these patterns reveal how investors gather information through expert networks, they do not tell us which aspects of these discussions most influence investment decisions. To analyze the predictive power of different discussion topics and sentiments for investment outcomes, we follow a machine learning approach using XGBoost to predict future deals. Our analysis incorporates call characteristics, topic sentiments, and firm-level controls. Using SHAP (SHapley Additive exPlanations) values to interpret feature importance, we find that technology integration and customer acquisition emerge as the strongest predictors of investment outcomes. Positive discussions of technology integration are associated with a 15% increase in deal likelihood, while positive customer-related signals predict a 16% increase. In contrast, topics that receive substantial discussion time such as market analysis, product development, and business strategy show values close to zero across all sentiment categories, indicating minimal predictive power for investment decisions.

The predictive power of expert discussions exhibits systematic variation across firm characteristics. For firms with 2-3 previous investment rounds, technology discussions show peak SHAP values corresponding to an 9.4% increase in deal odds. This predictive power declines monotonically with investment history and age, falling to 2% for firms with more than 20 previous rounds. This finding shows that a positive signal about technology integration having a significantly larger marginal impact on younger firms where information asymmetries about technological capabilities are most severe. Customer-focused discussions also show declining predictive power with firm age and funding history, but the effect is less pronounced than for technology discussions. SHAP values for customer-related signals decline from 8% to 5% across firm characteristics, compared to the steeper decline from 9% to 2% for technology discussions. The sharper decay in technology signals indicates that expert validation of technology is particularly more critical for younger firms with limited track records.

Our findings demonstrate how early-stage investors effectively resolve information asymmetries through systematic information gathering. The content and sentiment of discussions with industry experts significantly predict investment decisions, suggesting that investors extract valuable signals during their due diligence process. While investors appear to gather information systematically, we find some misalignments between information acquisition patterns and predictive power. What investors discuss most frequently is not necessarily what best predicts deal occurrence or amounts raised.

Our analysis contributes to the literature on information acquisition in venture capital by documenting the specific types of information investors seek. Gompers and Lerner (2001) emphasize that one of the primary functions of venture capitalists is to overcome information asymmetries, with Gompers et al. (2020) providing systematic evidence on how VCs approach this challenge through their due diligence processes and decision-making criteria. While the principal-agent problem in financial contracting has been extensively studied in the context of VC investments (e.g., Kaplan, Strömberg and Sensoy, 2002; Kaplan and Strömberg, 2004), the precise nature of the information asymmetry between investors and startups remains less explored. Kaplan and Strömberg (2001) provide a theoretical framework for understanding how VCs screen potential investments, but the specific mechanisms and types of information gathered during due diligence are still not well documented. We study this process by analyzing expert consultation calls, rich in what Kaplan and Strömberg (2004) term "soft information," offering a unique window into how investors gather and interpret crucial, often unquantifiable data that informs their decisions. The calls frequently delve into market dynamics and technological feasibility, which Kerr, Nanda and Rhodes-Kropf (2014) highlight as critical for entrepreneurial experimentation. By analyzing these discussions, our study provides insight into how investors evaluate market risk and technological uncertainty, thereby enriching our comprehension of the due diligence process in early-stage investing.

We contribute to the emerging literature on machine learning applications in finance (Giglio et al., 2021; Ke, Kelly and Xiu, 2020; Liu, Liu and Shahab, 2019; Eisfeldt and Schubert, 2024) by showing how LLMs can analyze complex, unstructured conversational data in investment contexts. Our LLM-based method connects the established literature on textual analysis in finance (Hoberg and Phillips, 2010; Loughran and McDonald, 2016) with contemporary developments in natural language processing and interpretable machine learning. Methodologically, we extend the Chain-of-Thought prompting framework of Wei et al. (2022) to financial text analysis, demonstrating its effectiveness in extracting nuanced insights from complex discussions. Our use of SHAP values for interpretation follows Lundberg and Lee (2017) but applies their framework to the novel context of LLM-based topic modeling, bridging the interpretability of traditional methods like Latent Dirichlet Allocation (LDA) and seeded LDA (Blei, Ng and Jordan, 2003; Watanabe and Zhou, 2022) and the power of modern language models. While traditional approaches like FinBERT

(Araci, 2019; Huang, Wang and Yang, 2023) primarily focus on information extraction and classification tasks, our approach leverages the semantic understanding and reasoning capabilities of large language models to discover more coherent and contextually relevant topics.

2 Data

2.1 Expert Network Data.

The expert network industry emerged in response to several regulatory changes in the early 2000s, including Regulation Fair Disclosure and the Global Analyst Research Settlement, which restricted traditional information channels and prompted investors to seek alternative sources of insight. The industry has grown to over 100 firms with estimated revenues of $1.9 Billion in 2021.

Expert calls are client-initiated consultations where investors engage subject matter experts for in-depth research on companies or market segments. The typical format is a 45-60 minute discussion, with expert compensation ranging from $100-$250 for junior professionals to over $1,000 for senior executives. Experts generally include competitors, customers, suppliers, industry consultants, or former employees of target companies. Following several insider trading cases, the industry has implemented strict compliance procedures, including mandatory call recordings and transcriptions.

Our sample comprises 8,382 call transcripts from 2017 to 2022, covering 2,363 companies, obtained from a major expert network's content library. Using company names, we match 1,756 of these companies with private deal data. The matched sample characteristics are presented in Panel A of Table 1, which shows the distribution of expert consultation calls across different expert types. The sample contains 6,783 calls, with customers representing the largest share at 43.18%, followed by industry consultants (25.28%) and former executives (18.28%).

Expert networks operate under strict legal and ethical guidelines.[1] Experts must not

[1] Section 204A of the Investment Advisers Act of 1940 requires advisers to implement written policies and procedures to prevent the misuse of material non-public information (MNPI). Experts must not be current employees of discussed companies and are prohibited from sharing material non-public information, trade secrets, or confidential data, as established in SEC Release No. 2011-38 and reinforced in the SEC's 2022 Risk Alert on MNPI compliance issues.

be current employees of discussed companies and are prohibited from sharing material non-public information, trade secrets, or confidential data. They may only discuss publicly available information, industry expertise, and general market insights. Topics typically include market trends, competitive dynamics, product assessment, and industry challenges. Experts must disclose any potential conflicts of interest and sign compliance agreements restricting discussion of privileged information. The networks actively monitor calls and maintain compliance databases to prevent unauthorized information sharing.

2.2 Private Deals Data.

We use CBInsights data on private deals from 2017q1 to 2023q3. It details 113,768 deals, including information on the funding round, amount, sector, industry, geography, and investors. Panel B of Table 1 reports the distribution of 9,614 deals by type, with venture capital deals across different stages (Seed through Series C) accounting for approximately 65% of all deals.

3 Calls and Deals Occurrence

3.1 Determinants of Expert Network Calls

To investigate the characteristics of firms that are the subject of expert network calls, we estimate:

$$Call_{it} = \alpha + \beta X_{it} + \tau_t + \epsilon_{it} \tag{1}$$

where $ExpertCall_{it}$ is a dummy equal to one if firm i received any expert calls in quarter t, X_{it} is a vector of firm characteristics including investment stage, industry, age, and funding history, and τ_t represents quarter fixed effects. All specifications cluster standard errors at the firm and quarter levels.

[Insert Table 2 here]

Table 2 reports the results. Column (1) explores how companies' funding round relates to expert call activity, using all other funding stages as the omitted baseline. Compared to companies at other funding stages, later-stage VC-backed companies are significantly more likely to receive expert calls, with a 0.79 percentage point higher probability. Series A companies are also the subject of more calls, with a 0.13 percentage point increase. In contrast, early-stage VC-backed firms are less likely to receive calls than companies at other stages, with a 0.083 percentage point lower probability. Growth/PE backed companies show no significant difference in call probability.

Column (2) investigates industry patterns, using all other sectors as the omitted baseline. Digital companies demonstrate the strongest relationship with expert calls, being 0.48 percentage points more likely to be the subject of a call. Hardware and retail & services firms show marginally significant positive relationships of 0.067 and 0.054 percentage points, respectively. Finally, healthcare companies show no significant difference in call activity compared to companies in all other sectors, despite the sector's general prominence in VC investing.

Column (3) examines firm characteristics. Company age exhibits a significant negative relationship with expert calls, with each additional year reducing call probability by 0.031 percentage points. Both the number and dollar value of previous funding rounds positively predict expert calls, with each additional round increasing call probability by 0.075 percentage points and each log dollar of previous funding increasing it by 0.22 percentage points.

These patterns suggest that expert networks are valuable for due diligence on more mature startups that have achieved significant milestones, especially in the digital sector. The negative age coefficient, controlling for funding history, indicates that younger firms at any given funding stage generate more expert calls, perhaps reflecting greater information asymmetry. The relatively low R^2 values (ranging from 0.008 to 0.011) suggest that while these characteristics help predict expert call activity, substantial variation remains unexplained by observable firm attributes.

3.2 Are there more Calls Around Deals?

To investigate whether experts are consulted more often around deals, we use the following baseline specification:

$$\mathbb{1}(Deal)_{i,t+s} = \alpha + \beta Call_{i,t-1} + \mu_i + \tau_t + \varepsilon_{it}, \tag{2}$$

where $\mathbb{1}(Deal)_{i,t+s}$ is a dummy equal to one if company i goes through a deal at $t+s$ and zero otherwise ($s \in [-8, 8]$). $Call_{i,t-1}$ is a dummy equal to one if a call occurred at $t-1$ and zero otherwise. The specification includes both firm (μ_i) and quarter (τ_t) fixed effects to control for time-invariant firm characteristics and common temporal patterns.

Baseline Results. Column (1) of Table 3 shows that having an expert call in the previous quarter is associated with an 11 p.p. increase in the probability of a deal in the current quarter (significant at the 1% level). The contemporaneous relationship is still significant but much smaller, with calls associated with a 2.6 p.p. higher probability of a deal in the same quarter.

[Insert Table 3 here]

The stronger predictive power of lagged calls than contemporaneous ones suggests that expert consultations typically precede deals rather than occurring simultaneously or after them. This timing pattern is consistent with investors using expert networks as part of their due diligence process before finalizing investment decisions. The magnitude of the effects is economically meaningful, given that the unconditional probability of a deal in any quarter is 7.9% in our sample.

Figure 1 shows how expert consultation calls relate to deal timing at a quarterly frequency. The coefficients represent estimates from equation (2) regressing a call indicator on quarters relative to deal events, controlling for firm and quarter fixed effects. The error bars represent 95% confidence intervals using standard errors clustered at the firm level. The results indicate minimal pre-deal consultation activity until two quarters before the deal, a sharp increase with a high and significant coefficient in the quarter immediately preceding the deal, a slightly lower but still significant coefficient during the deal quarter, and smaller yet significant levels of consultation activity during and after the deal quarter.

[Insert Figure 1 here]

Expert Types. To better understand which types of expert consultations are most informative for future deals, we modify equation (2) by replacing the single call indicator with separate indicators for calls with different categories of experts: competitors, customers, former executives, industry consultants, and partners.

[Insert Table 4 here]

The results in Table 4 reveal substantial heterogeneity in the predictive power of different expert types. Customer calls have the strongest relationship with future deals, associated with a 16 p.p. increase in deal probability. Competitor calls show the second strongest effect at 11 p.p., followed by former executives at 9.1 p.p. and industry consultants at 7 p.p.. Partner calls show a smaller and statistically insignificant effect.

This pattern reveals a strong predictive relationship between customer calls and subsequent deals, which could arise through three distinct channels. First, customers may provide valuable information about product-market fit and revenue potential. Second, customers might tend to convey more positive signals compared to other experts like industry consultants. Third, investors may strategically choose to conduct customer calls primarily when they are already leaning toward completing a deal, while preferring industry consultants for earlier-stage screening. The weaker relationship for industry consultant calls could similarly reflect any of these mechanisms. We analyze call content and sentiment in later sections to distinguish between these explanations.

Sector Analysis. To examine whether the relationship between expert calls and deals varies across industries, we augment equation (2) by interacting the lagged call indicator with industry fixed effects. This allows us to estimate industry-specific effects while continuing to control for firm and time fixed effects.

[Insert Table 5 here]

Table 5 shows significant heterogeneity across sectors. The strongest relationship between calls and subsequent deals are in technology-focused industries, with electronics showing a 16 p.p. increase (significant at 5%), software showing a 14 p.p. increase (significant at 1%), and internet companies showing a 10 p.p. increase (significant at 1%). Mobile and telecommunications companies also show a significant 11 p.p. increase (significant at 10%). In contrast, some traditional sectors like finance and industrials show small or even negative coefficients, though these are not statistically significant.

This sectoral pattern suggests that expert networks play a particularly important role in due diligence for technology investments. This could reflect greater information asymmetries in these sectors due to the technical complexity of products, rapid pace of innovation, and importance of intangible assets. The weaker relationships in traditional sectors may indicate that investors have access to other information sources or that information asymmetries are less severe in these industries.

4 LLM-based Topic Modeling

4.1 Methodology

Our approach builds on Pham et al. (2023) and uses Chain-of-Thought prompting for LLM-based topic modeling. It leverages the improved reasoning capabilities that Wei et al. (2022) demonstrated with this technique and identified by Meincke, Mollick and Terwiesch (2024) in innovation contexts. We are actively developing this methodology and plan to make the code publicly available, facilitating its dissemination and enabling other researchers to adopt and refine the tool. Following Pham et al. (2023), our LLM-based topic modeling approach relies on GPT-4-Turbo and consists of three steps.

Step 1: Topic Generation. In the first stage, we prompt ChatGPT to generate a range of high-level topics from a sample of consultation transcripts (see Appendix Prompt 1). This involves carefully engineered prompts that guide GPT in identifying generalizable themes relevant to early-stage investment decisions. This process results in a preliminary set of topics that capture the main discussion topics in the calls while trying to avoid topics that are too company- or industry-specific.

Step 2: Topic Refinement. The topic refinement stage involves further processing of the initially generated topics to ensure coherence, relevance, and non-redundancy (see Appendix Prompt 2). This stage includes (i) consolidating overlapping or redundant topics, (ii) adjusting topic labels for clarity and consistency, and (iii) eliminating overly specific or infrequent topics.

This stage is also executed using carefully engineered prompts for the LLM, guiding it to assess topic similarity, adjust specificity, and ensure overall coherence of the topic set. The refinement process helps to create a more concise, focused, and meaningful set of topics that accurately represent the key themes in the consultation calls.

Step 3: Topic Assignment. In the final stage, we use the LLM to assign topics to each consultation call transcript in our dataset (see Appendix Prompt 3). This process involves (i) analyzing the content of each transcript, (ii) assigning the most relevant topics from the refined topic list and identifying the main topic in the conversation, and (iii) assigning the sentiment score relevant to each topic. We define our topic sentiment scale as in Table A.2.

This three-step approach to topic modeling leverages new capabilities of LLMs. It provides a rich analysis to examine factors influencing early-stage investment decisions by testing how positive or negative signals about a specific factor lead to following investment decisions.

4.2 Advantages over Traditional Topic Modeling Methods

Our method builds upon recent advancements in natural language processing and aims to extract meaningful insights from complex, nuanced conversations in the investment domain. LLM-based topic modeling offers several key advantages over traditional methods like Latent Dirichlet Allocation (LDA), particularly for our research context:

Enhanced Interpretability. Traditional topic modeling methods often yield outputs that are difficult to interpret, requiring subjective manual analysis. Our LLM-based approach, however, generates topics in clear, natural language. This results in more readily understandable insights and a less subjective set of topics. For example, a topic like "Technology Management and Strategy" generated by GPT captures the broader theme of how technology impacts business decisions, rather than narrowly focusing on specific technological keywords and terms.

Contextual Understanding. Unlike LDA, which is based on statistical distributions of words, LLMs capture the contextual relationships between words and phrases. This is particularly important in investor consultations where conversations often involve complex and multi-faceted discussions, and where industry-specific and technology-specific terminology play a significant role. LLMs are capable of understanding and generating topics that capture the nuanced reasoning and strategic evaluations embedded in these conversations, leading to more accurate and coherent topic clusters.

Incorporating Sentiment Analysis. A major advantage of our LLM-based approach is the ability to incorporate sentiment analysis within the topic assignment stage. Our assignment prompt asks ChatGPT to assign topics in the finalized list to each transcript, along with a sentiment for each topic and for the overall conversation. Table A.2 describes the scale we ask ChatGPT to use.

LLM-based topic modeling allows us to capture not just what topics are discussed, but also how positively or negatively they are perceived in the context of the potential investment. Traditional methods lack the capacity to handle sentiment analysis in this integrated manner. For instance, LDA focuses purely on word frequency and co-occurrence, which means it cannot provide insights into the tone or sentiment surrounding the topics, thus missing a critical dimension in understanding investor decision-making.

4.3 Topic Distribution Across Expert Types

Table 6 presents the 10 most discussed topics and their definition obtained from the LLM analysis by ChatGPT. Figure 2 shows the relative frequency of these ten most discussed topics, broken down by expert type. While "Technology Integration Strategy" and "Competitive Analysis" are consistently the two most discussed topics across all expert types (approximately 50% of the topics discussed), we observe substantial heterogeneity in the emphasis and broader topic distribution across different categories of experts.

[Insert Table 6 here]

Competitors discuss technology integration the most (40%), while maintaining substantial coverage of competitive analysis (20%). This suggests that competitors provide valuable insights into technological feasibility and implementation challenges.

Customers also heavily emphasize technology integration (45%), though they uniquely show a higher propensity to discuss customer acquisition and retention issues (20%), reflecting their perspective as users of the product or service.

Former executives discuss technology integration relatively less than other expert types (15% vs. 30% on average for other experts). However, they discuss competitive analysis almost as much as competitors (30%), and have the strongest emphasis on business strategy across all expert types (15%). This shift in focus likely reflects their broader management perspective and a focus on strategy.

Industry consultants have the most balanced distribution of topics across all experts. Their higher propensity to discuss healthcare market analysis reflects both the importance of the healthcare sector in our sample and the consultants' focus on broader industry challenges. Similarly, their relatively greater emphasis on risk assessment compared to other experts suggests they bring a more comprehensive analytical perspective that spans both firm-specific and industry-wide considerations.

Similar to competitors, partners (while having the smallest sample size) show the highest concentration on technology integration and competitive analysis (jointly almost 60% of content). This suggests that partners are primarily consulted for targeted strategic insights rather than broad market understanding.

Overall, the systematic differences in the topic distribution in Figure 2 suggest that investors strategically source different types of information from different experts. This allows them to build a more comprehensive understanding of their investment targets. The pronounced emphasis on technology and competition across all expert types points to the centrality of these factors in early-stage investment decision-making.

4.4 Sentiment Across Expert Types

Figure 3 presents the average sentiment scores across all topics broken down by expert type, with 95% confidence intervals. Recall that the sentiment scores range from -2 to +2, with higher values indicating more positive sentiment. The first thing to note is that the average sentiment is positive across all expert types.

[Insert Figure 3 here]

We find that customers express the most positive sentiment in their consultations, with an average score of approximately 0.70, significantly higher than all other expert types. This positive bias might reflect selection effects if the companies being discussed tend to attract investors by being successful with satisfied customers.

The other experts express slightly less positive views, which are not statistically different from each other. While partners have a somewhat higher average sentiment score of around 0.57 compared to approximately 0.5 for competitors, former executives, and industry consultants, the confidence interval for partners is notably wider due to a smaller sample size.

These systematic differences in sentiment across expert types suggest that investors may need to adjust for potential biases when interpreting expert consultations, or that investors choose to talk to different experts when considering more or less promising companies.

4.5 Sentiment Across Topics

Figure 4 shows the average sentiment scores across topics. We observe substantial variation in sentiment across topics, ranging from approximately 0.2 for Risk Assessment and Management to nearly 0.8 for Growth and Scaling Strategy.

[Insert Figure 4 here]

Topics related to growth and operational capabilities tend to be discussed most positively, with the top four most positively discussed topics being Growth and Scaling Strategy, Security Strategy and Implementation, Customer Adoption Strategy, and Product Development and Market Fit. All these topics have sentiment scores of at least 0.7 on average. This suggests that experts are particularly optimistic when discussing companies' expansion potential and ability to execute core operational functions.

In contrast, topics related to evaluation and assessment generate much more neutral sentiment. Risk Assessment and Management shows the lowest sentiment score (around 0.2), while Management and Founder Assessment and Competitive Analysis also rank among the lowest, with scores below 0.5. This pattern suggests that experts adopt a more critical stance when evaluating potential challenges or assessing leadership capabilities.

Overall, one interpretation of Figure 4 is that experts tend to be more critical when discussing potential challenges versus opportunities.

4.6 Topic Distribution Across Industries

Figure 5 presents a heatmap of topic distribution across different industries. The intensity of blue represents the proportion of conversations dedicated to each topic, with darker shades indicating higher proportions. Several patterns emerge from Figure 5. First, the most discussed topics, Technology Integration Strategy (T1) and Competitive Analysis (T2), tend to dominate discussions in all industries. Beyond these common topics, discussions tend to concentrate in specific industries. For example, Data Management and Analytics (T7) appears frequently in the Application & Data Integration Software industry, while Product Development and Market Fit (T5) is more prominent in the Database Management Software industry.

[Insert Figure 5 here]

Some industries exhibit notably even distributions across topics - for instance, Biotechnology and Consulting & Outsourcing show relatively balanced distributions, with no single topic exceeding 21% of discussions. This even distribution suggest that multiple aspects of business and technology require equal attention during due diligence. Healthcare-related sectors (Medical Devices & Equipment, Medical Facilities & Services, Healthcare Software) show distinct topic distributions with greater emphasis on Healthcare Market Analysis (T10), an extracted topic that is industry-specific by nature.

4.7 Sentiment Variation Across Topics and Expert Types

Figure 6 displays a heatmap of sentiment scores across different topics and expert types, with colors ranging from red (slightly negative or small) through yellow (neutral to slightly positive) to blue (very positive). The most striking pattern is the consistently negative or only slightly positive sentiment around Risk Assessment and Management across all expert types, with scores ranging from -0.07 to 0.26. In contrast, Growth and Scaling Strategy generates the most positive sentiment across all expert types, with scores consistently above 0.75.

[Insert Figure 6 here]

Consistent with Figure 3, customers have the most consistently positive sentiment across almost all topics, particularly in Business Strategy (0.76), Customer Adoption Strategy (0.77), and Growth and Scaling Strategy (0.78). Industry consultants tend to show less positive sentiment, except for notably high sentiment in Cloud Computing Strategy (0.71). Former executives show particular variance in their sentiment, ranging from strongly positive on Growth and Scaling Strategy (0.77) to notably negative on Risk Assessment (-0.03) and Organizational Development (0.17).

Overall, Figure 6 shows rich variation in sentiment both across topics and expert types. While some topics like Risk Assessment and Growth Strategy elicit consistent sentiment from all experts (positive and negative, respectively), others such as Cloud Computing Strategy and Organizational Development show substantial variation across expert types. This heterogeneity in sentiment suggests that different experts bring distinct perspectives to the due diligence process. The systematic differences in sentiment levels across experts - from consistently positive customers to more varied assessments from industry consultants - highlight the importance of consulting diverse expert types to obtain a balanced view of investment opportunities.

4.8 What Drives Sentiment Variation?

To examine what drives positive sentiment in expert calls, we estimate:

$$\mathbb{1}(PositiveSentiment)_{it} = \alpha + \beta X_{it} + \tau_t + \epsilon_{it} \tag{3}$$

where $PositiveSentiment_{it}$ is a dummy equal to one if the expert expressed positive sentiment about firm i in quarter t, and zero if the sentiment was neutral or negative. As before, X_{it} represents firm characteristics and τ_t are quarter fixed effects. Standard errors are clustered by firm and quarter.

[Insert Table 7 here]

Table 7 presents the results. Column (1) shows that, relative to other financing stages which is the omitted baseline, early-stage VC-backed companies and Series A companies are 16-17 percentage points more likely to receive positive expert assessments. Later-stage VC and Growth/PE backed companies show smaller and statistically insignificant positive coefficients compared to the omitted category of other rounds, suggesting that experts' optimism is concentrated in earlier-stage ventures.

Column (2) reveals that, compared to other sectors which is the omitted baseline, digital companies receive more positive assessments, with a 14 percentage point higher likelihood of positive sentiment. Hardware firms show a similar magnitude effect of 15 percentage points relative to other sectors, though the estimate is not statistically significant at conventional levels. Healthcare and retail & services companies show smaller and statistically insignificant effects compared to other sectors, suggesting more mixed or neutral expert assessments in these industries.

Column (3) demonstrates that firm age is strongly negatively associated with positive sentiment, with each additional year reducing the probability of positive assessment by 0.83 percentage points. Previous funding history, whether measured in rounds or dollars, shows no significant relationship with expert sentiment once age is controlled for.

While earlier results showed that expert calls are more common for later-stage companies, the call sentiment tends to be more positive for younger and earlier-stage firms. This could reflect selection effects - experts may only be consulted on particularly promising early-stage companies - or could indicate genuine optimism about growth potential in younger ventures. The relatively low R^2 values (ranging from 0.010 to 0.018) suggest that observable firm characteristics explain only a small portion of the variation in expert sentiment.

5 Expert Network Calls and Investment Decisions

In this section, we examine how expert consultation calls and their content predict investment outcomes. We first examine this relationship using linear models and then extend the analysis to machine learning methods that capture non-linear effects and complex interactions.

5.1 Linear Analysis of Calls and Deal Prediction

We analyze how expert consultation calls predict subsequent investment deals using panel regressions at the firm-quarter level. Our specification builds on equation (2) by adding our LLM-based sentiment measures:

$$Deal_{i,t} = \alpha + \beta Call_{i,t-1} + \gamma Sentiment_{i,t} + \delta X_{i,t} + \mu_i + \tau_t + \epsilon_{i,t} \tag{4}$$

where $Deal_{i,t}$ indicates whether firm i receives investment in quarter t, $Call_{i,t-1}$ indicates if an expert consultation occurred at $t-1$, $Sentiment_{i,t}$ captures whether the call had positive sentiment, $X_{i,t}$ contains firm characteristics, and μ_i and τ_t are firm and quarter fixed effects.

[Insert Table 8 here]

Table 8 presents the results. Column 1 shows the baseline relationship - having a call is associated with an 6.8 percentage point higher probability of a subsequent deal. Adding firm characteristics in column 2 strengthens this effect - calls predict a 16.7 percentage point increase in deal probability. The controls reveal that VC-backed and older firms are more likely to receive deals, while firms with more previous funding rounds show lower deal probability.

Column 3 shows that call sentiment significantly predicts deals - positive sentiment is associated with a 6.2 percentage point higher likelihood of investment, even after controlling for the occurrence of calls and firm characteristics. This suggests that not just calls but also the overall conversation sentiment relates to investment decisions.

Column 4 examines how the predictive power of calls varies with firm age. The negative coefficient on the Call × Age interaction indicates that calls are more predictive for younger firms. For each additional year of age, the predictive effect of calls decreases by 73 basis points, suggesting that expert consultation may be particularly valuable for reducing information asymmetries in younger companies.

Finally, column 5 analyzes which discussion topics best predict deals. Technology Integration (Topic 1), Customer Acquisition (Topic 3), and Risk Assessment (Topic 6) emerge as the strongest predictors, each associated with 5.6-6.6 percentage points higher deal probability. Data Management (Topic 7) also shows significant predictive power but to a lower extent. This heterogeneity suggests that certain topics may contain information that is particularly valuable for investment decisions.

Overall, these results suggest that beyond just the occurrence of calls predicting deals, the sentiment expressed, topics covered, and firm characteristics all might play important roles in determining investment outcomes. However, these linear models may miss complex interactions and non-linear effects. For example, the impact of discussing technology integration might depend on firm age, previous funding, or discussion of other topics in ways not captured by simple interaction terms. Additionally, the sentiment around specific topics may matter differently across firm characteristics. To capture these rich patterns, we turn to machine learning methods in the next section, specifically employing gradient boosting models that can identify complex predictive relationships in the data.

5.2 Machine Learning Model Specification

To analyze the relationship between expert calls and investment outcomes, we use the XGBoost gradient boosting method. This approach offers several advantages for our setting. First, unlike linear models, XGBoost can identify non-linear relationships and complex interactions between features without requiring explicit specification. Second, compared to other tree-based methods like random forests, XGBoost offers superior handling of imbalanced data, which is important given the relative rarity of deals in our sample. Third, XGBoost provides built-in support for categorical variables and missing values, allowing efficient incorporation of industry, location, and other categorical firm characteristics.

Our feature set combines call characteristics, topic sentiments, and firm-level controls. Call characteristics include the occurrence of calls and overall sentiment. For topics, we incorporate both discussion indicators and topic-specific sentiment measures that range from -1 (negative) to 0 (neutral) to 1 (positive). Firm-level features include age, VC backing, previous funding rounds, total previous funding amount, geography (country and state), and industry classifications (primary sector, industry, and sub-industry). To address potential overfitting, we employ cross-validation and regularization through XGBoost's built-in L1 and L2 penalties. The model parameters are selected through grid search with 5-fold cross-validation across tree depth, learning rate, minimum child weight, and regularization parameters, fitting over 500 different model combinations to identify the parameter set that yields the highest out-of-sample predictive accuracy.

This specification allows us to systematically examine three key aspects of expert calls: their predictive power for deals, the relative importance of different discussion topics, and how these effects vary with firm characteristics. To interpret these complex relationships, we employ SHAP (SHapley Additive exPlanations) values, which provide a unified framework for understanding feature importance and interactions.

5.3 Key Predictors and Feature Importance

Our machine learning analysis reveals that expert consultation calls are strongly associated with investment outcomes. Figure 7 shows the distribution of the SHAP values for call occurrence. The average SHAP value for call occurrence for firms that receive a call is 0.64, which is associated with an 89.6% increase in the odds of a deal (exp(0.64) = 1.896). This strong predictive power likely reflects selection effects, as investors tend to consult experts more frequently just before investment decisions. However, when aggregating the SHAP values for the call occurrence and all the topic-specific sentiments and discussion content, the total SHAP values sum to 0.83, corresponding to a 129.5% increase in deal odds (exp(0.83) = 2.295). This substantial increase in predictive power when incorporating call content suggests that the substance of these consultations contains valuable information beyond the simple fact that a call took place.

[Insert Figure 7 here]

The analysis of topic-specific effects reveals clear patterns in how different topics of expert discussions predict investment outcomes. Figure 8 shows the SHAP values for each topic across negative, balanced, and positive sentiment categories. Technology Integration (Topic 1) and Customer Acquisition (Topic 3) emerge as the strongest predictors of deal occurrence. Positive discussions of technology integration are associated with a 15% increase in deal odds (exp(0.14) = 1.15), while positive customer-related signals predict a 16.1% increase (exp(0.15) = 1.161). Even balanced discussions of these topics show meaningful effects but to a lesser extent, with SHAP values of 0.05 and 0.14, respectively. These results suggest that substantive discussion of technology or customers helps resolve information asymmetries between investors and entrepreneurs, particularly regarding technical capabilities and market traction.

Risk Assessment (Topic 6), Data Management (Topic 7), and Financial Strategy (Topic 8) form a second tier of predictive importance. Risk Assessment shows consistently moderate positive SHAP values, while Data Management and Financial Strategy display more variation, where negative sentiment predicts lower deal probability and positive sentiment has a positive effect. In contrast, topics like Competitive Analysis (Topic 2), Market Analysis (Topic 4), Product Development (Topic 5), and Business Strategy (Topic 9) show relatively weak predictive power, with SHAP values close to zero across all sentiment categories.

[Insert Figure 8 here]

The relationship between sentiment and predictive power reveals how expert consultation calls inform investment decisions in early-stage ventures. The magnitude of effects and substantial variation across topics - ranging from a 16% and 15% increase in deal odds for positive signals about customer adoption and technology to effectively zero for competitive analysis and market and growth analysis - suggests highlights that the specific content of discussions matters significantly beyond overall call and sentiment. Indeed, once we account for topic-specific sentiment through our LLM-based approach, the predictive power of aggregate call sentiment becomes close to zero. This finding highlights the importance of distinguishing between different types of information evaluated during investment decisions and demonstrates the value of our granular approach to analyzing these contents.

5.4 Heterogeneous Effects Through SHAP Analysis

We next examine how the importance of the most relevant topics varies across firm characteristics. Figure 9 shows that the predictive power of technology-related discussions (Topic 1) exhibits systematic variation based on firms' lifecycle stage and funding history. For earlier-stage companies with fewer previous investments, technology discussions have substantially higher SHAP values, peaking at around 0.09 (corresponding to a 9.4% increase in deal odds) for firms with 2-3 previous rounds. This predictive power declines monotonically with investment history, falling to about 0.02 for firms with more than 20 previous rounds. Similarly, both firm age and total previous funding show consistent negative relationships with the predictive power of technology discussions. SHAP values decline from 0.09 for firms with minimal previous funding to 0.03 for those with over $700 million raised. The age pattern mirrors this trend, with the strongest effects observed for young firms.

[Insert Figure 9 here]

These patterns collectively suggest that expert consultation about technology plays a particularly crucial role in resolving information asymmetries for younger, less-funded firms for which traditional metrics provide limited guidance. The variation in predictive power across firm characteristics indicates that investors rationally adjust how much weight they place on expert technical validation relative to their own prior based on the availability of other information sources. This finding aligns with theories of information acquisition where the value of expert validation is highest when alternative sources of information are scarce.

Although Figure 10 shows that customer-related discussions (Topic 3) also contain meaningful variation across firm characteristics, the patterns differ somewhat from those observed for technology discussions. For firms with fewer previous investments, customer-focused discussions maintain relatively high SHAP values around 0.07, declining more gradually to 0.05 for firms with more investment history. Similarly, the relationship with total previous funding shows a more modest decline from SHAP values of 0.07 to 0.4. The age pattern also exhibits a gentler slope, suggesting that customer validation remains informative even as firms mature.

[Insert Figure 10 here]

The more persistent predictive power of customer discussions across firm characteristics contrasts with the sharp decline observed for technology validation. This pattern suggests that while technology assessment is crucial for young firms with limited track records, customer validation provides valuable information throughout a firm's lifecycle. This finding aligns with the intuition that customer relationships and market traction remain fundamental concerns even as technological uncertainty diminishes. The relatively stable importance of customer discussions may also reflect that even established firms face ongoing uncertainty about customer adoption and market evolution that expert consultation can help resolve.

6 Conclusion

We develop a novel LLM-based methodology to analyze the content and sentiment of 6,800 consultation calls between investors and industry experts. Our approach combines advanced language models with interpretable machine learning techniques to extract discussion topics and measure their predictive power for investment outcomes, offering unique insights into how investors gather and process information when evaluating potential investments.

Our analysis reveals several key findings. First, while expert calls cluster around Series A and later-stage companies, particularly in the digital sector, younger firms at any given stage generate more expert consultation. Second, different types of experts provide systematically different perspectives - customers express consistently positive views while industry consultants and former executives tend to be more measured in their assessments. Third, and most importantly, we find that the predictive power of expert discussions varies substantially across topics and firm characteristics. Technology integration and customer acquisition emerge as the strongest predictors of investment outcomes, with their signals being particularly valuable for younger firms where information asymmetries are most severe.

These findings have important implications for our understanding of early-stage investing and information production in financial markets more broadly. The systematic variation in how different expert types discuss and evaluate companies suggests that investors strategically source complementary information to build a comprehensive view of potential investments. However, the misalignment between frequently discussed topics and those that best predict outcomes raises questions about the efficiency of information gathering in due diligence processes. Future research could explore whether this pattern reflects institutional constraints, behavioral biases, or rational responses to objectives beyond deal completion. More broadly, our methodology demonstrates how advances in natural language processing can help illuminate previously opaque aspects of financial decision-making, potentially opening new avenues for research on information production in other contexts where traditional metrics are limited.

References

Araci, D. 2019. "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models." *arXiv preprint arXiv:1908.10063*.

Blei, David M, Andrew Y Ng, and Michael I Jordan. 2003. "Latent dirichlet allocation." *Journal of machine Learning research*, 3(Jan): 993–1022.

Eisfeldt, Andrea L, and Gregor Schubert. 2024. "AI and Finance." National Bureau of Economic Research.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. 2021. "Five facts about beliefs and portfolios." *American Economic Review*, 111(5): 1481–1522.

Gompers, Paul, and Josh Lerner. 2001. "The venture capital revolution." *Journal of Economic Perspectives*, 15(2): 145–168.

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

Hoberg, Gerard, and Gordon Phillips. 2010. "Product market synergies and competition in mergers and acquisitions: A text-based analysis." *The Review of Financial Studies*, 23(10): 3773–3811.

Huang, Allen H, Hui Wang, and Yi Yang. 2023. "FinBERT: A large language model for extracting information from financial text." *Contemporary Accounting Research*, 40(2): 806–841.

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

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

Kaplan, Steven N, Per Strömberg, and Berk A Sensoy. 2002. "How well do venture capital databases reflect actual investments?" *Working paper*.

Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as Experimentation." *Journal of Economic Perspectives*, 28(3): 25–48.

Ke, Zheng Tracy, Bryan T Kelly, and Dacheng Xiu. 2020. "Predicting returns with text data." *The Review of Financial Studies*, 33(8): 3381–3420.

Liu, Chen, Chao Liu, and Soran Shahab. 2019. "Machine learning from cases: A data-driven theory of judicial decisions." *The Review of Financial Studies*, 32(6): 2017–2066.

Loughran, Tim, and Bill McDonald. 2016. "Textual analysis in accounting and finance: A survey." *Journal of Accounting Research*, 54(4): 1187–1230.

Lundberg, Scott M, and Su-In Lee. 2017. "A unified approach to interpreting model predictions." *Advances in Neural Information Processing Systems*, 30: 4765–4774.

Meincke, Lennart, Ethan R Mollick, and Christian Terwiesch. 2024. "Prompting Diverse Ideas: Increasing AI Idea Variance." *arXiv preprint arXiv:2402.01727*.

Pham, Chau Minh, Alexander Hoyle, Simeng Sun, and Mohit Iyyer. 2023. "TopicGPT: A prompt-based topic modeling framework." *arXiv preprint arXiv:2311.01449*.

Watanabe, Kohei, and Yuan Zhou. 2022. "Theory-driven analysis of large corpora: Semisupervised topic classification of the UN speeches." *Social Science Computer Review*, 40(2): 346–366.

Wei, Jason, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. "Chain-of-thought prompting elicits reasoning in large language models." *Advances in neural information processing systems*, 35: 24824–24837.

Figures and Tables

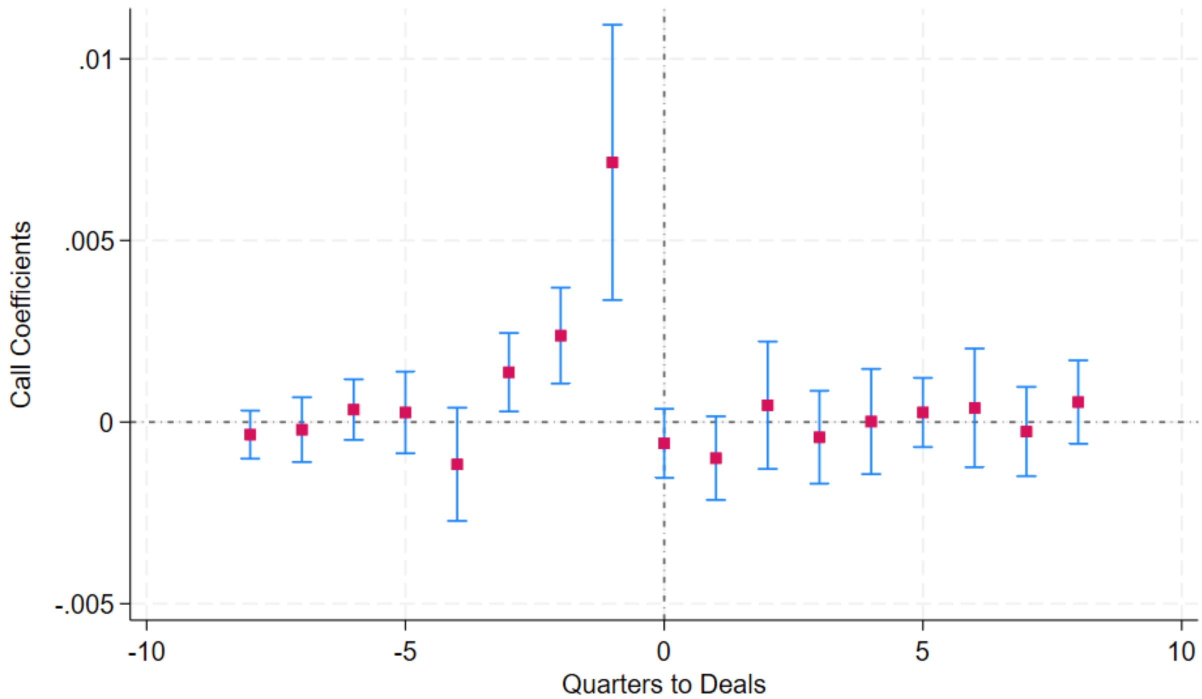


Figure 1: Event study around deal quarters. The figure plots quarterly regression coefficients and 95% confidence intervals from regressing an indicator for expert consultation calls on deal event time dummies, controlling for firm age, VC backing status, number and amount of previous funding rounds, and years since previous roundwith firm and quarter fixed effects. Time 0 represents the quarter in which a deal occurs for the company discussed in the call. The confidence intervals adjust for clustering by firm and quarter.

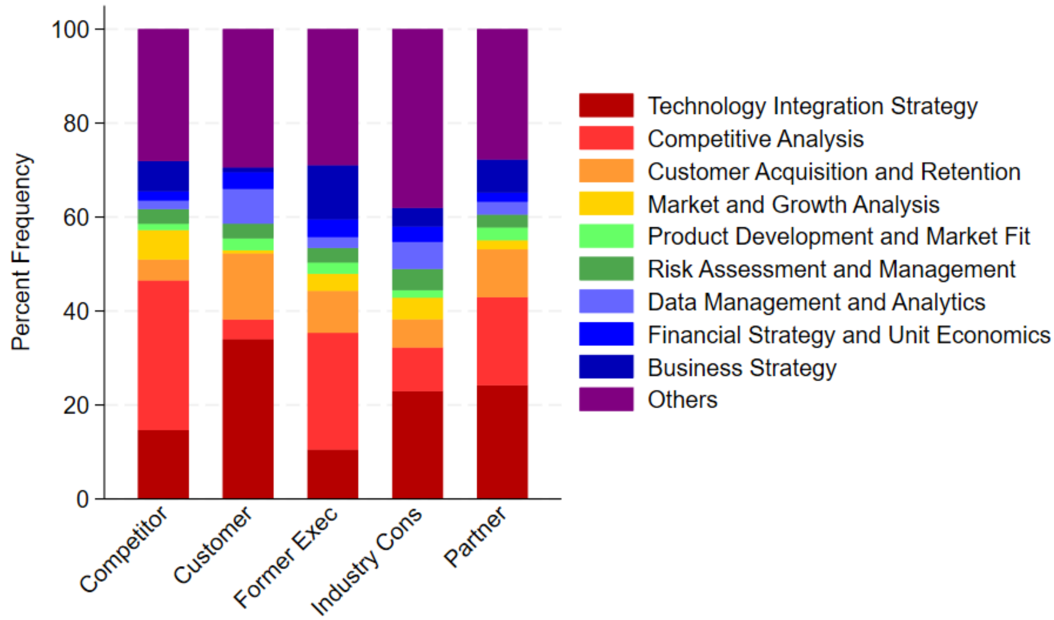


Figure 2: Topic Distribution by Expert Type. This figure shows the percentage frequency of different topics discussed during expert consultation calls across expert types: competitor, customer, former executive, industry consultant, and partner.

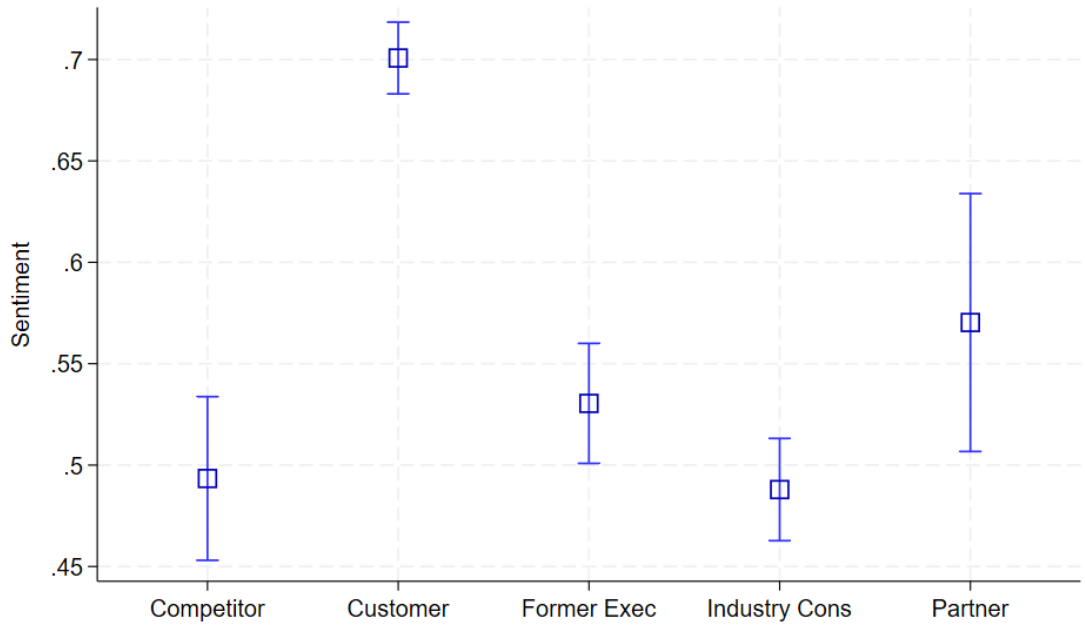


Figure 3: Average Sentiment by Expert Type. This figure plots the mean sentiment scores for each expert type, with 95% confidence intervals shown as vertical bars. Sentiment scores range from -2 to +2, with higher values indicating more positive sentiment.

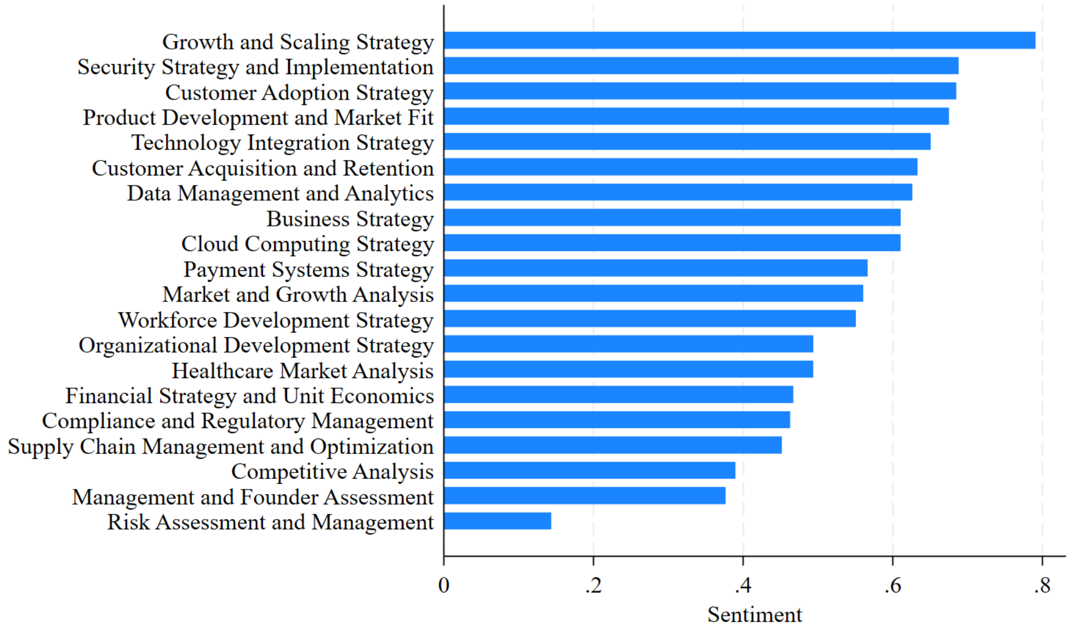
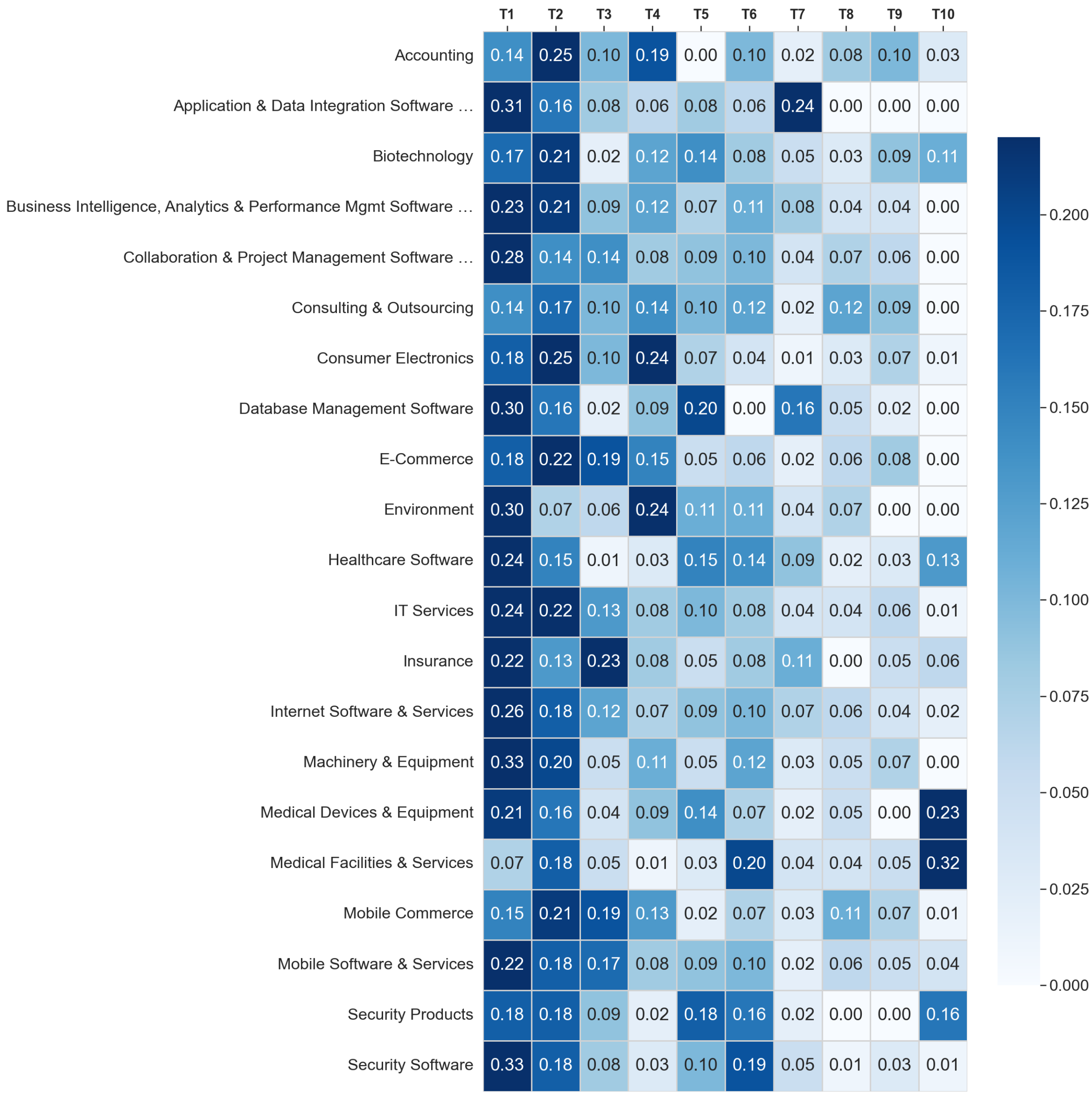


Figure 4: Average Sentiment by Topic. This figure shows the mean sentiment scores across different discussion topics in expert consultation calls. Topics are ordered by sentiment score, and sentiment ranges from -2 to +2, with higher values indicating more positive sentiment.



	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
Accounting	0.14	0.25	0.10	0.19	0.00	0.10	0.02	0.08	0.10	0.03
Application & Data Integration Software …	0.31	0.16	0.08	0.06	0.08	0.06	0.24	0.00	0.00	0.00
Biotechnology	0.17	0.21	0.02	0.12	0.14	0.08	0.05	0.03	0.09	0.11
Business Intelligence, Analytics & Performance Mgmt Software …	0.23	0.21	0.09	0.12	0.07	0.11	0.08	0.04	0.04	0.00
Collaboration & Project Management Software …	0.28	0.14	0.14	0.08	0.09	0.10	0.04	0.07	0.06	0.00
Consulting & Outsourcing	0.14	0.17	0.10	0.14	0.10	0.12	0.02	0.12	0.09	0.00
Consumer Electronics	0.18	0.25	0.10	0.24	0.07	0.04	0.01	0.03	0.07	0.01
Database Management Software	0.30	0.16	0.02	0.09	0.20	0.00	0.16	0.05	0.02	0.00
E-Commerce	0.18	0.22	0.19	0.15	0.05	0.06	0.02	0.06	0.08	0.00
Environment	0.30	0.07	0.06	0.24	0.11	0.11	0.04	0.07	0.00	0.00
Healthcare Software	0.24	0.15	0.01	0.03	0.15	0.14	0.09	0.02	0.03	0.13
IT Services	0.24	0.22	0.13	0.08	0.10	0.08	0.04	0.04	0.06	0.01
Insurance	0.22	0.13	0.23	0.08	0.05	0.08	0.11	0.00	0.05	0.06
Internet Software & Services	0.26	0.18	0.12	0.07	0.09	0.10	0.07	0.06	0.04	0.02
Machinery & Equipment	0.33	0.20	0.05	0.11	0.05	0.12	0.03	0.05	0.07	0.00
Medical Devices & Equipment	0.21	0.16	0.04	0.09	0.14	0.07	0.02	0.05	0.00	0.23
Medical Facilities & Services	0.07	0.18	0.05	0.01	0.03	0.20	0.04	0.04	0.05	0.32
Mobile Commerce	0.15	0.21	0.19	0.13	0.02	0.07	0.03	0.11	0.07	0.01
Mobile Software & Services	0.22	0.18	0.17	0.08	0.09	0.10	0.02	0.06	0.05	0.04
Security Products	0.18	0.18	0.09	0.02	0.18	0.16	0.02	0.00	0.00	0.16
Security Software	0.33	0.18	0.08	0.03	0.10	0.19	0.05	0.01	0.03	0.01

Figure 5: Topic Distribution by Industry. This heatmap shows the proportion of each topic discussed across different industries. Darker blue shades indicate higher proportions of discussion.

Figure 6: Sentiment by Topic and Expert Type. This heatmap shows the average sentiment scores for each combination of topic and expert type. Colors range from red (negative) through yellow (neutral) to blue (positive) sentiment.

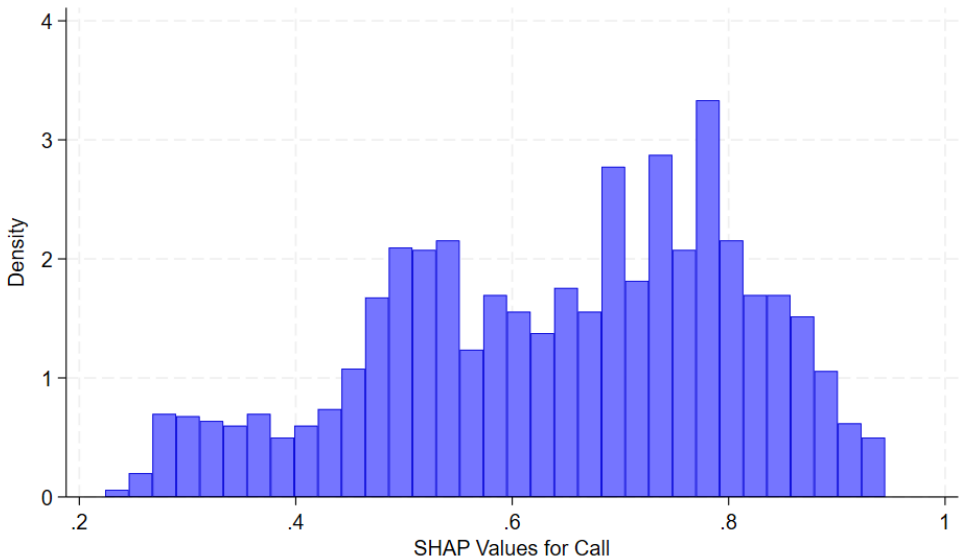


Figure 7: Distribution of SHAP Values for Expert Consultation Calls. The figure plots the histogram of SHAP values for firm-quarters with expert consultation calls. SHAP values measure contribution of call occurrence to the log-odds of deal probability in the XGBoost model, where a SHAP value of 0.1 corresponds to a 10.5% increase in deal odds relative to the baseline (exp(0.1) = 1.105).

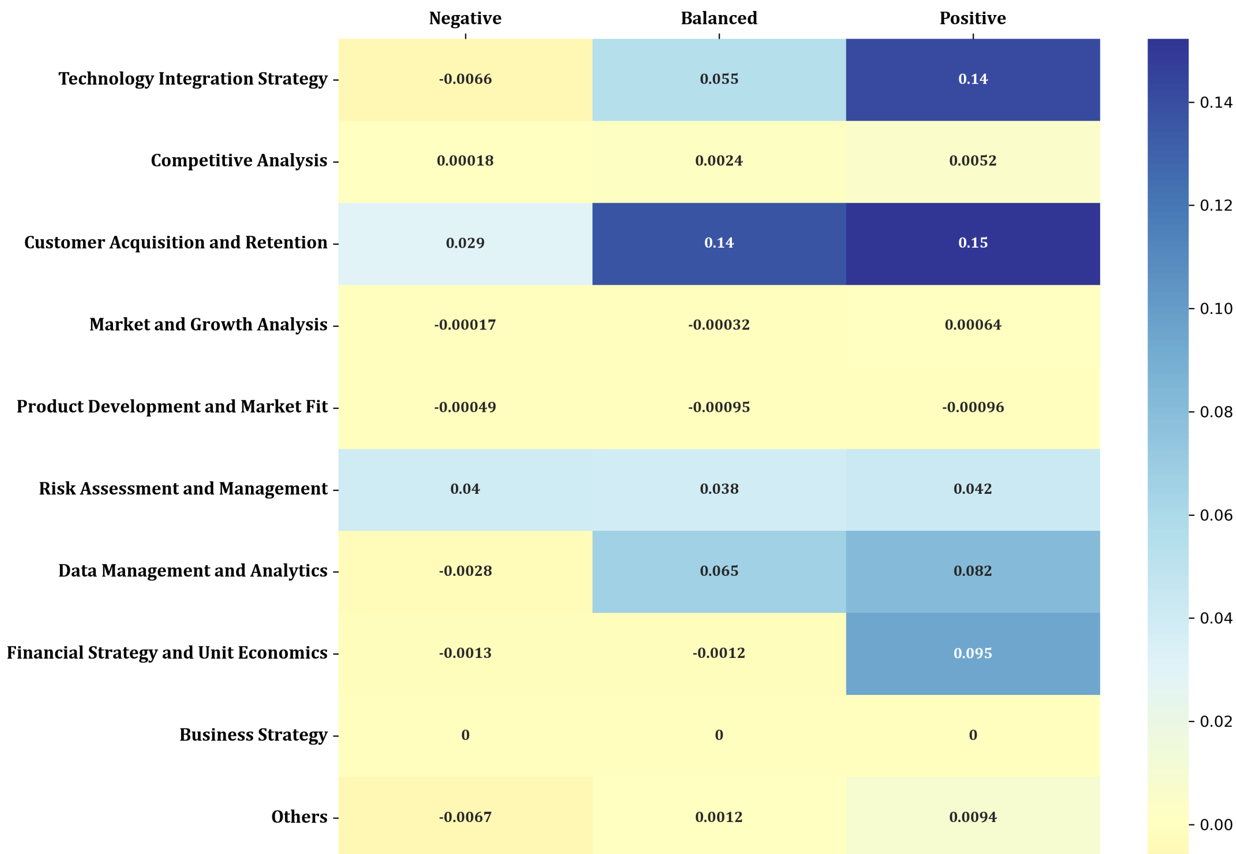


Figure 8: SHAP Values by Topic and Sentiment. The figure shows average SHAP values for different topics across negative, neutral, and positive sentiment categories in expert consultation calls. SHAP values measure each topic-sentiment combination's contribution to the log-odds of deal probability in the XGBoost model. Colors indicate the magnitude and direction of effects, with darker blue representing stronger positive contributions to deal prediction. Topic numbers correspond to: (1) Technology Integration, (2) Competitive Analysis, (3) Customer Acquisition, (4) Market Analysis, (5) Product Development, (6) Risk Assessment, (7) Data Management, (8) Financial Strategy, (9) Business Strategy, and (10) Others.

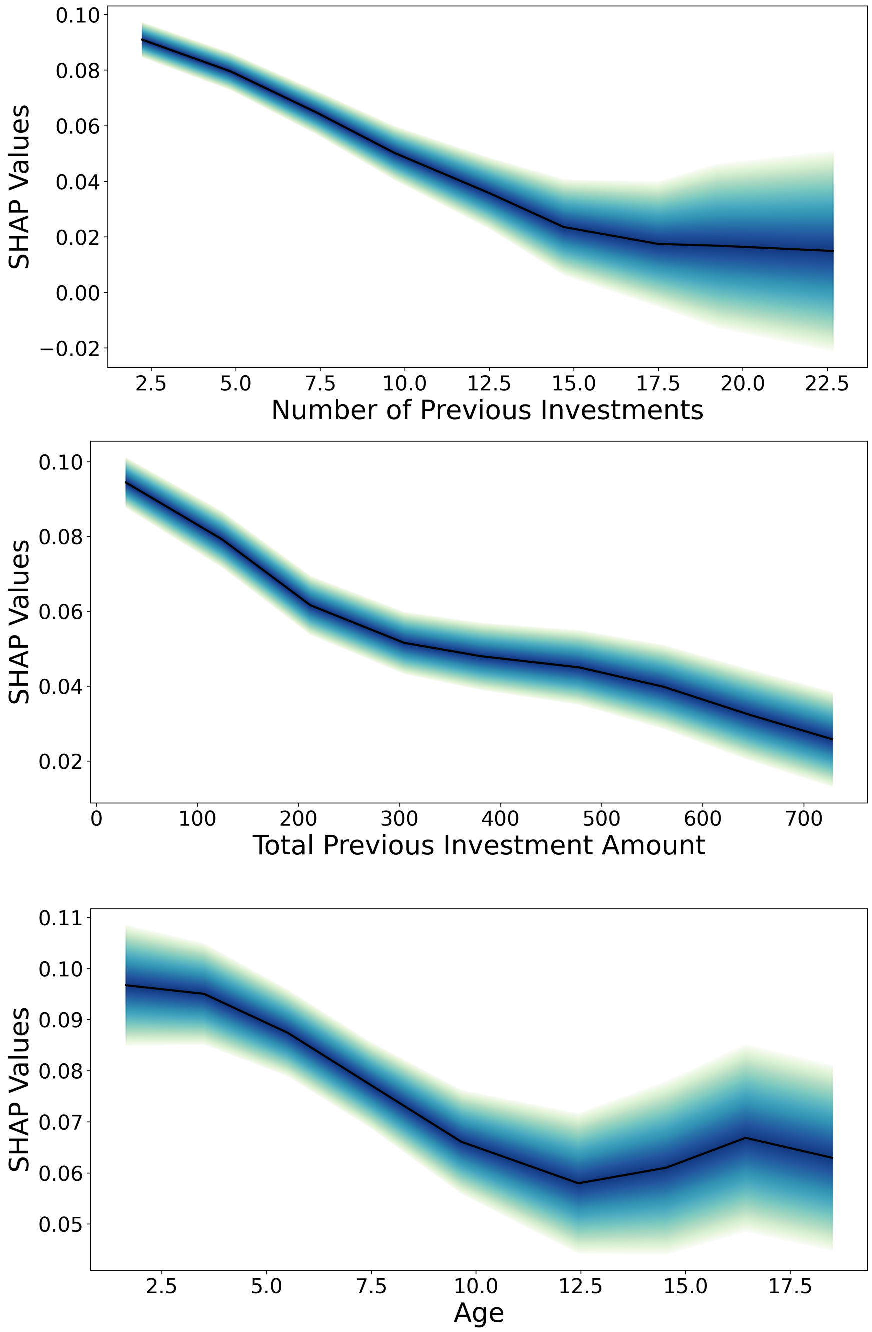


Figure 9: SHAP Values for Technology Integration (Topic 1) Across Firm Characteristics. The figure shows how the predictive power of technology discussions varies with firm characteristics. The top panel plots SHAP values against the number of previous investment rounds, the middle panel against total previous funding (in millions), and the bottom panel against firm age (in years). The black line shows the average SHAP value, with shaded regions representing 95% confidence intervals. SHAP values measure the topic's contribution to the log-odds of deal probability in the XGBoost model.

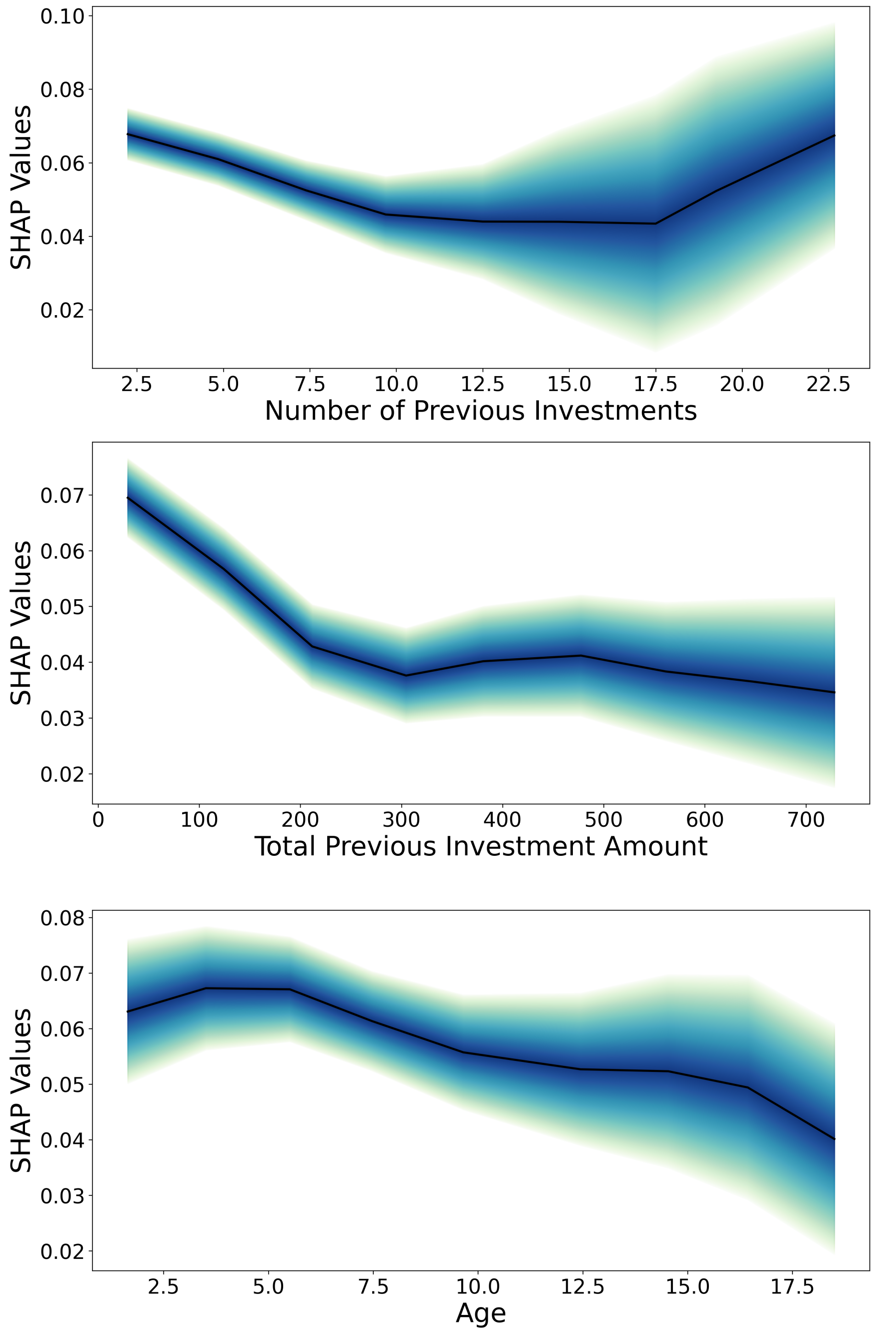


Figure 10: SHAP Values for Customer Acquisition (Topic 3) Across Firm Characteristics. The figure shows how the predictive power of technology discussions varies with firm characteristics. The top panel plots SHAP values against the number of previous investment rounds, the middle panel against total previous funding (in millions), and the bottom panel against firm age (in years). The black line shows the average SHAP value, with shaded regions representing 95% confidence intervals. SHAP values measure the topic's contribution to the log-odds of deal probability in the XGBoost model.

Table 1: Summary Statistics in Analysis Sample. This table reports the distribution of expert consultation calls and deals in our analysis sample from 2017 to 2021.

Panel A: Calls		
Expert Type	N	%
Competitor	629	9.22
Customer	2926	42.88
Former Executive	1255	18.39
Industry Consultant	1720	25.21
Panel	24	0.35
Partner	1	0.01
Missing	268	3.93
Total	6823	100.00
Panel B: Deals		
Deal Type	N	%
VC (Series A)	13908	22.23
VC (Seed)	10597	16.94
VC (Series B)	9906	15.84
Pre-VC Funding	8508	13.60
Acquisition/IPO	5262	8.41
VC (Series C)	4914	7.86
VC (Late-stage)	3877	6.20
Private Equity	1961	3.14
Miscellaneous	1836	2.94
VC (undisclosed stage)	1702	2.72
Debt	63	0.10
Undisclosed	17	0.03
Total	62551	100.00

Table 2: Determinants of Expert Calls. Panel regressions at the firm-quarter level for 41,686 firms over 2017q1–2023q2. The dependent variable in columns (1)–(3) is a dummy equal to one if at least one expert call occurred for a firm in a quarter, and in columns (4)–(6) is the count of expert calls for a firm in a quarter. Other round types is the omitted category in column (1) and other sectors is the omitted category in column (2). Standard errors (in parentheses) clustered by firm and quarter. *, **, *** denote significance at 10%, 5%, 1%.

		$\mathbb{1}$(Call)	
	(1)	(2)	(3)
Early VC	-.00083**		
	(.00031)		
Growth/PE	.0004		
	(.00032)		
Later VC	.0079***		
	(.0024)		
Series A	.0013**		
	(.00056)		
Digital		.0048***	
		(.0014)	
Hardware		.00067*	
		(.00039)	
Healthcare		4.9e-07	
		(.00017)	
Retail & Services		.00054*	
		(.00031)	
Age			-.00031***
			(.000098)
$ PrevRounds			.0022***
			(.00069)
# PrevRounds			.00075***
			(.00019)
Quarter FE	Yes	Yes	Yes
adj. R^2	.008	.0069	.011
Observations	549,474	618,601	563,716

Table 3: Calls and All Deals: Quarterly Frequency.. Panel regressions at the firm-quarter level for 41,686 firms over 2017q1–2023q2. Note: calls data only runs from 2017q1–2022q1. *dummy_call_quarter* is a dummy equal to one if at least one call occurred for a firm in a quarter. Standard errors clustered by year and firm are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively. Source: .

	Dependent Variable: $\mathbb{1}$(Deal)	
	(1)	(2)
L.dummy_call_quarter	.11***	
	(.012)	
dummy_call_quarter		.026***
		(.0077)
Quarter FE	✓	✓
Firm FE	✓	✓
adj. R^2	.057	.057
Observations	1271252	1338160

Table 4: Calls and Deals Across Expert Types.. Panel regressions at the firm-quarter level for 41,686 firms over 2017q1–2023q2. Note: calls data only runs from 2017q1–2022q1. *dummy_call_quarter* is a dummy equal to one if at least one call occurred for a firm in a quarter. Standard errors clustered by year and firm are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Deal)
	(1)
L.Competitor	.11***
	(.025)
L.Customer	.16***
	(.014)
L.Former Exec	.091***
	(.016)
L.Industry Consultant	.07***
	(.011)
L.Partner	.068
	(.077)
Quarter FE	✓
Firm FE	✓
adj. R^2	.057
Observations	1271252

Table 5: Calls and Deals Across Sectors.. Panel regressions at the firm-quarter level for 41,686 firms over 2017q1–2023q2. Note: calls data only runs from 2017q1–2022q1. *dummy_call_quarter* is a dummy equal to one if at least one call occurred for a firm in a quarter. Standard errors clustered by year and firm are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively. NOTE: We're only displaying interactions and the simple term of call occurrence. We're also dropping the base category consisting of all other sectors.

	$\mathbb{1}$(Deal)
	(1)
1L.$\mathbb{1}$(Call)	.01
	(.027)
1L.$\mathbb{1}$(Call) × Business Products and Services	.12
	(.11)
1L.$\mathbb{1}$(Call) × Computer Hardware and Services	.14
	(.086)
1L.$\mathbb{1}$(Call) × Consumer Products and Services	.1
	(.069)
1L.$\mathbb{1}$(Call) × Electronics	.16**
	(.062)
1L.$\mathbb{1}$(Call) × Energy and Utilities	.15
	(.1)
1L.$\mathbb{1}$(Call) × Finance	-.035
	(.072)
1L.$\mathbb{1}$(Call) × Healthcare	.066
	(.044)
1L.$\mathbb{1}$(Call) × Industrials	-.074
	(.085)
1L.$\mathbb{1}$(Call) × Internet	.1***
	(.032)
1L.$\mathbb{1}$(Call) × Mobile and Telecommunications	.11*
	(.054)
1L.$\mathbb{1}$(Call) × Software (non-internet/mobile)	.14***
	(.032)
Quarter FE	✓
Firm FE	✓
adj. R^2	.057
Observations	1270492

Table 6: Top 10 Refined Topics Assigned to Transcripts by ChatGPT, Ranked by Frequency.

Rank	Topic	Description
1	Technology Integration Strategy	Strategic analysis of technology implementation, digital transformation initiatives, and IT infrastructure decisions. Includes evaluation of technology investments, integration challenges, and alignment with business objectives.
2	Competitive Analysis	In-depth study of competitive landscape, market dynamics, and strategic positioning. Includes analysis of competitor strengths, market share dynamics, and competitive advantage development.
3	Customer Acquisition and Retention	Integrated strategies for acquiring and retaining customers through various channels, including analysis of customer lifetime value and engagement metrics.
4	Market and Growth Analysis	Comprehensive examination of current and anticipated market trends, including analysis of Total Addressable Market (TAM), Serviceable Addressable Market (SAM), and market growth trajectories. Includes analysis of consumption patterns, economic influences, and market expansion potential.
5	Product Development and Market Fit	Strategic approach to new product development and product-market fit validation. Includes feature planning, market alignment analysis, customer validation metrics, and iteration strategies. Also covers product roadmap development and competitive positioning.
6	Risk Assessment and Management	Comprehensive assessment of business risks, development of mitigation strategies, and analysis of risk-return trade-offs in strategic decision-making.
7	Data Management and Analytics	Comprehensive exploration of methods and technologies for data collection, analysis, and utilization in decision-making processes. Includes data quality management and analytical model development.
8	Financial Strategy and Unit Economics	Strategic planning for revenue generation, cost management, and financial optimization. Includes analysis of unit economics, customer acquisition costs, lifetime value metrics, and profitability analysis. Also covers pricing models, financial performance metrics, and scalability of the business model.
9	Business Strategy	Analysis of high-level strategic decisions including market positioning, competitive advantage, resource allocation, and long-term growth planning. Includes strategic partnerships, market expansion, and adaptation to industry trends.
10	Healthcare Market Analysis	Examination of healthcare market trends, regulatory environment, customer preferences, and industry-specific challenges. Includes analysis of market opportunities and barriers to entry.

Table 7: Determinants of Positive Expert Call Sentiment.. The dependent variable equals one if the sentiment of the expert call is positive, zero if neutral or negative. Panel regressions at the expert call level. Other round types and Other sectors are the omitted categories in columns (1) and (2) respectively. Standard errors (in parentheses) clustered by firm and quarter. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Positive Sentiment)		
	(1)	(2)	(3)
Early VC	.16*		
	(.079)		
Growth/PE	.024		
	(.13)		
Later VC	.054		
	(.052)		
Series A	.17*		
	(.079)		
Digital		.14*	
		(.072)	
Hardware		.15	
		(.096)	
Healthcare		.055	
		(.082)	
Retail & Services		.061	
		(.1)	
Age			-.0083***
			(.0016)
$ PrevRounds			.0033
			(.0089)
# PrevRounds			-.0046
			(.0028)
Quarter FE	✓	✓	✓
adj. R^2	.017	.0099	.018
Observations	771	2,290	2,262

Table 8: Relationship between Calls and Deals. Panel regressions at the firm-quarter level. The dependent variable is a dummy equal to one if a deal occurs in a firm-quarter. Call is a dummy for expert consultation calls, OverallSentiment captures call sentiment, and Topics 1-10 are dummies for specific topics. All specifications include firm and quarter fixed effects and standard controls. Standard errors (in parentheses) clustered by firm and quarter. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	(1) Deal	(2) Deal	(3) Deal	(4) Deal	(5) Deal
Call	0.0675**	0.1670***	0.1209***	0.1834***	
	(0.0276)	(0.0189)	(0.0262)	(0.0309)	
OverallSentiment	0.0570**		0.0616***	0.0529**	
	(0.0237)		(0.0211)	(0.0212)	
Age		0.0311***	0.0311***	0.0310***	0.0311***
		(0.0049)	(0.0049)	(0.0048)	(0.0049)
VC_Backed		0.0241***	0.0241***	0.0241***	0.0241***
		(0.0043)	(0.0043)	(0.0043)	(0.0043)
# PrevRounds		-0.1120***	-0.1120***	-0.1120***	-0.1121***
		(0.0112)	(0.0112)	(0.0112)	(0.0112)
Years_PrevRound		0.0007	0.0007	0.0007	0.0007
		(0.0054)	(0.0054)	(0.0053)	(0.0054)
$ PrevRounds		-0.0177***	-0.0177***	-0.0178***	-0.0177***
		(0.0027)	(0.0027)	(0.0027)	(0.0027)
$ PrevRounds_Oth		0.0077***	0.0077***	0.0077***	0.0077***
		(0.0012)	(0.0012)	(0.0012)	(0.0012)
Call $\times$ Age				-0.0073***	
				(0.0015)	
Topic1					0.0609***
					(0.0202)
Topic2					0.0300
					(0.0196)
Topic3					0.0662***
					(0.0179)
Topic4					0.0101
					(0.0131)
Topic5					0.0242
					(0.0199)
Topic6					0.0563***
					(0.0154)
Topic7					0.0391**
					(0.0138)
Topic8					0.0282
					(0.0186)
Topic9					-0.0097
					(0.0104)
Topic10					0.0231
					(0.0135)
Dep Var Mean	.103	.103	.103	.103	.103
Firm FE	YES	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES	YES
Observations	616947	616947	616947	616947	616947
Adjusted R^2	.0275	.0939	.0939	.0939	.094

A Additional Figures and Tables

Table A.1: Detailed Breakdown of Deal Types and Their Frequencies for Companies in the Call Sample (2017-2021)

Deal Type	N	%
Acq - P2P	63	0.10
Acquisition	3258	5.21
Acquisition (Financial)	413	0.66
Acquisition (Talent)	20	0.03
Angel	1389	2.22
Asset Sale	3	0.00
Bridge	287	0.46
Business Plan Competition	648	1.04
Convertible Note	8	0.01
Corporate Majority	197	0.31
Corporate Majority - P2P	9	0.01
Corporate Minority	1674	2.68
Corporate Minority - P2P	94	0.15
Corporate-Venture Partnership	1	0.00
Crowdfunding	103	0.16
Dead	11	0.02
Debt	14	0.02
Grant	2006	3.21
Growth Equity	222	0.35
IPO	1068	1.71
Incubator	4221	6.75
Line of Credit	3	0.00
Loan	38	0.06
Management Buyout	16	0.03
Merger	240	0.38
Option/Warrant	1	0.00
PIPE	7	0.01
Partnership	14	0.02
Pre-Seed	141	0.23
Private Equity	1199	1.92
Project Finance	21	0.03
Reverse Merger	2	0.00
Secondary Market	233	0.37
Seed	2915	4.66
Seed VC	7682	12.28
Series A	13908	22.23
Series B	9906	15.84
Series C	4914	7.86
Series D	2193	3.51
Series E	964	1.54
Series F	398	0.64
Series G	175	0.28
Series H	88	0.14
Series I	35	0.06
Series J	15	0.02
Series K	9	0.01

Continued next page

Continuation of Detailed Breakdown of Deal Types

Deal Type	N	%
Spinoff	1	0.00
Take Private	4	0.01
Undisclosed	17	0.03
Unit Acquisition	1	0.00
Venture Capital	1702	2.72
Total	62551	100.00

Table A.2: Sentiment Scale. This Table contains the sentiment scale we use in the assignment prompt in Stage 3 of our LLM-based topic modeling methodology.

Score	**Description**
+2	Strongly Positive: Clear positive discussion with multiple benefits; expert shows confidence in this aspect.
+1	Positive: Generally favorable discussion; more benefits than drawbacks mentioned.
0	Neutral: Balanced, neutral discussion without clear bias.
-1	Negative: More drawbacks than benefits; mild concerns are highlighted.
-2	Strongly Negative: Multiple clear problems discussed; expert expresses clear concerns.

Prompt 1: Topic Generation

```
Task Description
As part of an academic finance research project, we are analyzing the decision-making
    process of early-stage investors. We have compiled transcripts from approximately 8,000
    consultation calls between potential early-stage investors, referred to as "Clients" in
    the transcripts, and experts familiar with the companies in question. These experts may
    be former executives, industry specialists, or analysts. We will provide these
    transcripts one-by-one.

Your Task
Review each provided transcript and categorize the most prominent topics discussed. These
    topics should reflect the main focus of the consultation call and help in understanding
    the priorities and concerns of the clients.

Instructions
Identify ONLY the primary, top-level topics that are central to the consultation call.
Focus on topics that are broadly relevant and avoid overly technical details.
Topics should not be narrowly focused on specific lines of business or industries.
Prioritize clarity and relevance in your responses.
Output Format
Please output the name of the document first. Then, list each identified topic in the
    following format:

Document Name
- Topic Label: Brief definition of the topic

Document XYZ
- Market Trends: Overview of current market directions and investor sentiments
- Regulatory Impact: Description of how recent regulations affect investment decisions
```

Prompt 2: Topic Refinement

```
As part of an academic finance research project, we aim to investigate the primary factors influencing the decision-making process of early-stage investors. To accomplish this, we have collected transcripts from around 8,000 consultation calls between potential early-stage investors and experts familiar with the company. We will provide you a list of topics extracted from these transcripts one-by-one. Your task is to merge topics that are paraphrases, near duplicates, or broadly similar. Return "None" if no modification is needed. This analysis will serve as the topic refinement stage in a topic modeling analysis.

When merging topics:
- Create a title that captures the core process/concept, removing domain-specific details
- Merge topics that describe the same process even if applied to different domains/industries
- Merge more specific topics with their general versions when they describe the same core concept
- Create a general definition that describes the core concept or process, avoiding:
  * Company-specific examples
  * Industry-specific details
  * Platform-specific features
- NEVER use generic names like "Merged Topic"

Examples of GOOD merges (ordered by similarity):

1. Near-identical topics with slight variations:
Input topics:
Payment Processing System: Analysis of payment processing infrastructure including transaction flows, settlement times, and integration requirements.
Payment Processing Infrastructure: Discussion of payment processing systems covering transaction handling, settlement periods, and integration needs.
Response:
Payment Processing Analysis: Examination of transaction processing systems, including workflow management, settlement procedures, and integration requirements.

2. Specific and general versions of same concept:
Input topics:
SaaS Revenue Models in Healthcare: Analysis of revenue structures for healthcare software companies, including subscription tiers, usage-based pricing, and service fees.
Revenue Model Analysis: Examination of different revenue structures including subscription models, usage-based pricing, and service components.
Response:
Revenue Model Structure: Analysis of revenue generation frameworks, including subscription approaches, consumption-based pricing, and supplementary service components.

3. Same concept across different companies:
Input topics:
Salesforce Market Position: Analysis of Salesforce's competitive position in CRM, focusing on enterprise sales and pricing strategy.
HubSpot Competitive Analysis: Evaluation of HubSpot's market positioning against other marketing platforms.
Response:
Market Positioning Analysis: Assessment of competitive positioning in the market, including strategic differentiation, target segments, and value proposition across different market tiers.

4. Related concepts in different contexts:
Input topics:
Zoom Growth Strategy: Analysis of Zoom's expansion into enterprise markets and international regions.
TikTok Market Expansion: Examination of TikTok's strategy for entering new demographic segments.
Response:
Market Expansion Strategy: Analysis of approaches to market growth, including target segment identification, geographical expansion, and adaptation of offerings for new market opportunities.

Rules:
- Remove company names and specific product references from merged titles and definitions
- Focus on the underlying business concept or process being discussed
- Create definitions that could apply across any company or industry
- Keep examples and specifics only if they illustrate a broader pattern
- Return exactly "None" if topics shouldn't be merged
- Use format: [General Process Title]: [Comprehensive Definition]
```

```
Input topics:
{input_topics}
```

Prompt 3: Assignment Prompt

```
As part of an academic finance research project, we aim to investigate the primary factors influencing the decision-making process of early-stage investors. Your task is to assign the most relevant topics from the provided list that are heavily discussed and are the main focus of the conversation call provided. You must use the exact topic names from the list - do not modify them in any way.

[Available Topics]
{available_topics}

[Sentiment Scale]
For each assigned topic, evaluate the sentiment based on the conversation between the client and expert, using a scale of -2 to +2:
Strongly Positive (+2): Clear positive discussion with multiple benefits, expert shows confidence in this aspect
Positive (+1): Generally favorable discussion, more benefits than drawbacks mentioned
Neutral (0): Balanced, neutral discussion without clear bias
Negative (-1): More drawbacks than benefits, mild concerns are highlighted
Strongly Negative (-2): Multiple clear problems discussed, expert expresses clear concerns

[Format Requirements]
- Use "Overall Document Sentiment: [Sentiment Word] ([+/-]N)" format
- Number topics with square brackets: [1], [2], etc.
- Include confidence level in parentheses after the exact topic name
- Always include "Reasoning:" and "Quote:" labels
- Format topic sentiment identical to document sentiment format
- Only use exact topic names from the provided list - no modifications
- Main Topic of Concern must be one of the exact topics assigned above

[Examples]
Example 1 - Market opportunity excerpt:
"The market opportunity is absolutely massive. They're solving a critical problem that every enterprise faces, and their solution is years ahead of competitors. The potential ROI for customers is incredible, and they're already seeing strong adoption across multiple sectors."

[1] Market Opportunity Assessment (High Confidence)
Reasoning: Directly addresses market size and solution value proposition
Quote: "The market opportunity is absolutely massive. They're solving a critical problem that every enterprise faces"
Topic Sentiment: Strongly Positive (+2)
- Emphasizes massive market opportunity
- Highlights strong competitive position
- Notes proven customer adoption

Example 2 - Regulatory landscape excerpt:
"The regulatory framework is complex and varies by jurisdiction. Companies need to maintain compliance across multiple frameworks while balancing operational efficiency."

[1] Regulatory Compliance (High Confidence)
Reasoning: Addresses regulatory complexity and compliance requirements
Quote: "The regulatory framework is complex and varies by jurisdiction"
Topic Sentiment: Neutral (0)
- Objective description of regulatory landscape
- Neither emphasizes problems nor benefits
- Focuses on factual information

Example 3 - Risk assessment excerpt:
"The regulatory investigation poses an existential threat to their business model. The pending litigation could completely halt operations, and there's significant risk of substantial penalties. Customer trust has been severely damaged."

[1] Risk Assessment and Mitigation (High Confidence)
Reasoning: Addresses critical business risks and threats
Quote: "The regulatory investigation poses an existential threat to their business model"
Topic Sentiment: Strongly Negative (-2)
- Highlights existential business threats
- Emphasizes severe regulatory risks
- Notes significant damage to customer trust

[Instructions]
1. Begin with overall document sentiment using this exact format:
   Overall Document Sentiment: [Sentiment Word] ([+/-]N)
```

```
    [Brief explanation of why this sentiment was chosen]

2. List topic assignments using this exact format:
    [N] Topic Name (High/Medium/Low Confidence)
    Reasoning: [Explanation]
    Quote: "[Exact quote from document]"
    Topic Sentiment: [Sentiment Word] ([+/-]N)
    - [Supporting point 1]
    - [Supporting point 2]
    - [Supporting point 3]

3. Follow these rules:
    - Use only exact topics from the provided list - no modifications
    - Include exact quotes from the document
    - Choose most specific topic when multiple apply
    - Always include confidence level
    - Format all sentiments as [Word] ([+/-]N)
    - Use square brackets [N] for topic numbering
    - Include 2-3 supporting points for each topic
    - Label reasoning and quotes explicitly

4. Main Topic Identification: After listing all relevant topics, identify the single most
     critical topic that would materially impact an investment decision. This topic must be
     selected from the exact topics you have already assigned above. Use this exact format:
    Main Topic of Concern: [Topic Name from above assignments]

[Document]
{Document}

Your response:
```