Online Appendix for "Internal Credit and External Blame: Self-Attribution in Operations and Supply Chain Performance"

Yimeng Niu
Tilburg School of Economics and Management, Tilburg University y.niu@tilburguniversity.edu

Xiao Qiao
College of Computing, City University of Hong Kong xiaoqiao@cityu.edu.hk

Jing Wu
Chinese University of Hong Kong Business School, Chinese University of Hong Kong jingwu@cuhk.edu.hk

Xingsheng Yang
College of Computing, City University of Hong Kong xingsheng.yang@my.cityu.edu.hk

Online Appendix A. Example of A Conference Call Setup

Figure OA1 shows Hewlett-Packard's fourth quarter of 2023 earnings call.

MANAGEMENT DISCUSSION SECTION

Enrique Lores
President, Chief Executive Officer & Director, HP, Inc.

Thank you, Orit, and thank you, everyone, for joining our final earnings call of 2023. It was great to host many of you for our Securities Analyst Meeting last month. As I said at the time, we have made significant progress against our strategic priorities, and we see attractive opportunities ahead.

Turning to Print, net revenue was $4.4 billion. That's down 3% year-over-year or 2% in constant currency. Print revenue grew 4% sequentially, while units were flat. Supplies revenue was up in constant currency in Q4 on an easier compare and finished the year down 1% in constant currency, in line with our long-term outlook. We drove strong Print operating margins of 18.9%, reflecting disciplined execution and cost management.

Marie E Myers
Chief Financial Officer, HP, Inc.

Thanks, Enrique, and good afternoon, everyone. It's a pleasure to be here with you all today.

In Print, our results reflect our continued focus on execution and disciplined cost management as we navigate a weak and competitive Print market. In Q4, total Print revenue was $4.4 billion, down 3% nominally or 2% in constant currency. The decline was driven by soft demand in both Consumer and Commercial, market share loss, and currency, offset partially by supplies revenue growth.

In addition, we made great progress reducing our commodity complexity by decreasing the number of client SKUs at our Personal Systems portfolio. We expect these initiatives will optimize our overall supplier ecosystem, increase our supply chain efficiency, and drive better forecast capabilities.

QUESTION AND ANSWER SECTION

Toni Sacconaghi Q
Analyst, Sanford C. Bernstein & Co. LLC

Yes. Thank you. I was wondering if you could comment on supplies growth in the quarter. It was up sequentially. That hasn't happened at any point in the last 10 years on an organic basis. So what was driving the non-seasonal growth in supplies? Did you change prices and people bought in before? Or was there some channel fill?

And then could you also just comment on – you talked a lot about Future Ready and the strength. I think you've taken out maybe run rate about $600 million worth of costs, but OpEx is essentially identical to the first quarter of this year. So, when we think about Future Ready, should we actually think about OpEx going down in fiscal 2024, or we won't see the impact of Future Ready on OpEx? Thank you.

Marie E Myers A
Chief Financial Officer, HP, Inc.

Yeah. Hi, Toni. Good afternoon. So let me just start out with the question on Future Ready. So, first of all, obviously, really pleased with the performance. I think, as I mentioned in my prepared remarks for 2023 and then going into 2024, as you know, we raised our target to $1.6 billion in SAM, so very much on track.

Enrique Lores A
President, Chief Executive Officer & Director, HP, Inc.

Yeah. So let me complement that. So, I think there are multiple factors that drove the good performance of supplies in Q4. As I have said many times though, when we look at year-on-year compares, it's not the best way to look at the performance in the quarter, because there are many adjustments that are done ever quarter, and then comparisons can distort the overall perspective.

Erik W. Woodring Q
Analyst, Morgan Stanley & Co. LLC

Great. Thank you very much for taking my question. I just wanted to circle back to Toni's question on Print supplies. You obviously outperformed Print supplies this quarter. So I just wanted to make sure I understood specifically what drove the upside in the quarter.

Enrique Lores A
President, Chief Executive Officer & Director, HP, Inc.

Perfect. Thank you, Erik. And thank you because when I was answering to Toni, I missed to mention one of the elements that also help, which is the easy compare, because Q4 last year was not a strong quarter for supplies. So this also has an impact on the year-on-year compare.

Figure OA1 An Example of Hewlett-Packard's Earnings Conference Call in the Fourth Quarter of 2023

Online Appendix B. Pipeline for Fine-tuning Large Language Models

This appendix introduces the pipeline, shown in Figure OA2, for fine-tuning our LLMs. Using the 324 keywords listed in Table OA1, we extract Q&A pairs related to the company's operations and supply chain topic. This lexicon is derived from existing literature (Wu 2024, Ersahin et al. 2024, Voleti et al. 2025), and expansions conducted via GPT-4o, with manual verification. To validate the effectiveness of the bag-of-words approach, we first sample 1,000 Q&A pairs from the extracted dataset. A manual review confirms that 93.30% of these samples are related to operations and supply chain. Additionally, we obtain another 1,000 samples that are classified as non-operations and supply chain-related, and find that 5.30% of the samples are misclassified as non-related.

We then refine this extracted subset by classifying questions as focusing on either *positive* or *negative* performance and managers' answers as attributing outcomes to either *internal* or *external* factors. The fine-tuning process includes five steps: *task definition*, *open-source selection*, *custom data preparation*, *training*, and *performance evaluation*.

We successfully construct two LLMs tailored for downstream business tasks:

- LLM_Q: This model categorizes questions into "positive," "negative," or "other" types. A "positive" (or "negative") classification indicates that the question pertains to reasons behind the company's currently positive (or negative) operations management or supply chain performance. The "other" category includes questions outside this scope, primarily those seeking clarification or prediction related to operations and supply chain facts.

- LLM_A: This model classifies managers' attributions into "internal operations management," "supply chain partners," or "environmental conditions." "Internal operations management" means that managers attribute operational performance to the company's internal operational activities, while "supply chain partners" suggests attribution to the activities or performance of upstream or downstream partners. "Environmental conditions" refers to environmental factors beyond the company's control that impact performance.

Taking the construction of the LLM_Q model as an illustrative example, we provide a detailed overview of this process.

Task definition. In the task definition phase, it is essential to delineate clear task objectives and design prompts that ensure effective performance. Notably, the question classification task exemplifies a classic supervised learning challenge, which calls for targeted prompt engineering to harness the in-context learning capabilities of LLMs. To structure this approach, we create a template with the following configuration(Niu et al. 2024):

Table OA1 Keywords Related to Operations and Supply Chains

supplier	suppliers	supply	supplied	manufacturer	manufacturers
raw material	raw materials	procurement	procure	procured	procuring
spare part	spare parts	sourcing	component	components	oem
original equipment manufacturer	material	materials	distributor	distributors	value chain
value chains	vendor	vendors	foundry	foundries	shipment
shipments	assembly	assemblies	ingredient	ingredients	subcontractor
subcontractors	inventories	inventory	backlogs	stockouts	demand forecasting
demand planning	bullwhip effects	lead times	logistics	distributions	deliveries
milk runs	manufacturing	productions	production	commodities	backlog
stockout	out of stock	demand forecasting	demand planning	bullwhip effect	lead time
transportation	freight	production planning	distribution	direct shipping	drop shipping
milk run	cross-docking	just-in-time	commodity	order fulfillment	warehousing
warehouse	resilience	lean	jit	agility	trade
batch	scheduling	bill of materials	bom	capacity planning	cycle time
newsvendor	little's law	material requirements planning	mrp	continuous improvement	demand chain
materials management	last mile delivery	route planning	intermodal	abc analysis	abc classification
able supply	acquisition cost	acquisition costs	active inventory	adoption supply	advanced shipping
affect supplier	affects supplier	aggregate inventory	allocated stock	along supply	alongside ship
already shipping	anticipation stock	application blockchain	article numbering	chains international	channel inventory
cleaner production	closed loop	coming soon	coming sooner	companies supply	component part
component parts	finished goods	first batch	first in	first inventory	first out
first pick	first picked	forward supply	free board	free carrier	products services
products shipping	proof delivery	proximity supplier	pull system	purchase price	purchase prices
purchasing lead	purchasing price	push system	quarantine stock	radio frequency	random sample
rapid acquisition	raw material	ready packaging	rebuild order	reduce inventory	strategic stock
management integration	management international	management supply	managing supply	manufactured part	manufactured parts
manufacturing logistics	manufacturing resource	manufacturing supply	material requirements	materials management	matrix bar
maximum order	maximum stock	minimum order	minimum stock	model supply	much inventory
network design	product category	product delivery	product group	product groupings	product groups
product lifecycle	product recovery	product shipping	product supply	production economics	production lead
production leading	production research	products category	round time	rounding order	safety stock
safety stocking	safety stocks	shipping lines	shipping new	shipping now	shipping product
shipping products	shipping rates	start shipping	started shipping	still shipping	stock rotation
stock site	stock turn	stock turned	stock turning	stock turnover	stock turns
stock types	stock valuation	stock valuations	availability	backflushing	backhaul
backorder	benchmarking	category	component	consolidation	consumable
fifo	incoterms	fob	inventory	logistics	offshoring
outsourcing	rotatable	slotting	warehouse	stocktaking	traceability
transaction	supply chain	strategic fit	of capacity	the supply	of safety
chain and	demand the	chain to	a supply	production cost	the retailer
to order	of demand	of scale	the peak	the manufacturer	supply and
variable cost	if demand	the cycle	lead time	demand is	the aggregate
the safety	of product	the lot	to supply	demand to	chain is
the demand	chain the	the supplier	customer order	spot market	transportation cost
fixed costs	from supplier	of supply	time is	transportation costs	the network
an order	the lead	million units	expected profit	demand from	lead times
a forecast	demand per	demand and	stages of	uncertainty and	third party
supply chains	and supply	of orders	the plant	fixed cost	the spot
a demand	distribution network	the season	the capacity	lower price	the quantity
demand in	demand forecast	demand across	and demand	when demand	chain for
of transportation	demand at	in supply	capacity to	capacity is	product to
revenue management	demand of	chain management	production capacity	and transportation	response time
customer demand	revenue sharing	supplier to	future demand	demand uncertainty	the forecast
high demand	service level	customer orders	in transportation	global supply	order is
product in	square feet	level is	economies of	to forecast	

- Define the role of the LLM: *As an expert in operations management, your task is to classify whether the question is asking about the attribution of the company's performance.*
- Specify the task: *Step 1. Determine the topic of the question: If the question is related to the company's current operations management or supply chain performance, proceed to Step 2. If it does not, classify the question as "Others". Step 2. Check if the question seeks causes for performance: If the question asks about the reason for the performance mentioned (positive or negative), proceed to Step 3. If it does not, classify it as "Others". Step 3. Classify based on performance assessment: If the performance mentioned is positive, classify the question as "Positive". If the performance mentioned is negative, classify the question as "Negative". For all other cases, classify the question as "Others".*
- Design the output format: *Please return the result directly in the following JSON format.*{*"thinking process": "Provide a detailed 'Chain of Thought' reasoning process", "label": "Insert final classification after thinking"*}

Base Model Selection. Next, we need to choose an open-source LLM as our base model. Various open-source LLMs—such as Mixtral, DeepSeek, ChatGLM, and Qwen—are accessible for both research and individual use. In this study, we select the "LLaMA-3.1-8B-Instruct" model (Dubey et al. 2024), released by Meta, to serve as our foundational model. The reasons for this choice are not only its top-tier performance but also the advantages of its active open-source community and a suite of user-friendly tools that enhance its utility. Although constrained by GPU resources, the smallest 8-billion-parameter version of this model is sufficient to deliver robust classification outcomes through fine-tuning. The "Instruction" suffix signifies that this model has undergone specific instruction-based fine-tuning, enabling it to more effectively interpret and execute user directions compared to the standard model.

Custom Data Preparation. The quality of training data determines the upper bound of model capabilities, and preparing custom datasets is crucial for model performance optimization. Recent studies have validated the efficacy of synthetic data. For instance, the Nemotron-4 model, trained on synthetic data, demonstrates performance in text generation and commonsense reasoning comparable to models like LLaMA and Mixtral and in enhancing the inductive reasoning capabilities of LLMs (Adler et al. 2024). Our methodology draws inspiration from synthetic data approaches and employs state-of-the-art language models to construct a comprehensive training dataset through a two-phase approach.

The initial phase involves developing a benchmark to evaluate the performance of various open- and closed-source models on downstream tasks. For this purpose, we manually label a randomly selected subset of 1,000 questions and test three distinct models: GPT-4o-2024-08-06, Claude-3.5-Sonnet and Gemini-1.5-Pro. Specifically, we use the bootstrap method to assess the stability of the performance metric. We create 1,000 bootstrap replicates by resampling the 1,000 test cases with replacement from the test set. The mean and the 95% confidence interval of the metric are calculated from this bootstrap distribution. The comparative analysis in Table 4 shows that GPT-4o outperforms all other models, achieving exceptional classification metrics with an accuracy of 89.94% and an F1-score of 90.04% in the question classification task. In the next phase, we employ the optimal model to annotate an additional 19,000 questions, and use these labels as the training targets.

Training. After selecting the base model and preparing the training dataset, we proceed to the model training phase. The development of LLMs introduces numerous fine-tuning methods. The most intuitive approach involves full parameter fine-tuning, which requires updating all parameters in the pre-trained model. However, this method exhibits sensitivity to data quality and demands substantial time to identify suitable hyperparameters for achieving optimal accuracy. In this experiment, we implement the LoRA method, a popular and lightweight training technique (Hu et al. 2022). LoRA introduces trainable rank decomposition matrices to the weight updates while keeping the original model parameters frozen, enabling faster convergence and enhanced deployment flexibility. Although LoRA has certain limitations, such as high dependency on the quality of the base model and fine-tuning data, and its difficulty in generalizing to other tasks after fine-tuning for downstream applications, it proves highly effective in our specific context. By focusing exclusively on classifying earnings conference call text and conducting validation of training samples and the base model, the final fine-tuning results are satisfactory.

Since no single hyperparameter configuration universally applies to all tasks, a significant portion of our efforts is dedicated to identifying optimal parameter settings. During model training with LoRA, we set both conventional hyperparameters—such as batch size, which specifies the number of samples processed before each parameter update; epochs, the number of complete passes through the dataset; learning rate, which controls the speed of parameter adjustments; and validation size, the portion of data reserved for performance assessment—and LoRA-specific parameters like rank, which defines the dimensionality of

LoRA's low-rank matrices, and alpha, a scaling factor modulating the impact of LoRA's parameter updates. To achieve this, we need to monitor the loss on both the training set and the validation set to select the most effective and stable set of hyperparameters. Primarily, we should prevent the model from overfitting on the training data or experiencing underfitting. Next, by comparing the classification performance of different hyperparameter combinations on the validation set, we can obtain the optimal hyperparameters.

Performance Evaluation. Table 4 shows that our fined-tuned LLM_Q achieved an accuracy of 88.47% and an F1 score of 88.16% on the three-class question classification task. To further understand the nature of the misclassification, we conduct an analysis of misclassified text from text length and readability measured by the Flesch-Kincaid indicator (Kincaid et al. 1975). This analysis reveals the following pattern: on average, incorrectly classified texts are longer than correctly classified texts (mean length of misclassified questions = 84.55 words vs. 69.58 words for correctly classified questions; Mann-Whitney U test $p\,value < 0.01$). This suggests that LLMs have more difficulty when processing longer text, possibly due to a dispersion of attention. However, no significant differences are observed in the readability metric between the two groups, indicating a potential divergence between the superficial quality of the text and the model's deeper semantic understanding. Figure 3(b) displays the confusion matrix, which provides a detailed breakdown of the classification performance for each category.

Performance superiority against classic NLP techniques. Several tests are conducted in this study to demonstrate why leveraging LLMs is necessary in our setting and to identify the conditions under which fine-tuned LLMs outperform traditional NLP approaches. First, following the methodology outlined in Wu (2024), we construct a Continuous Bag-of-Words (CBOW) model to vectorize the textual input. The hyperparameters are set with a window size of 5, a minimum word frequency of 5, 100 training iterations, a word vector dimension of 300, 5 for negative sampling, and a learning rate of 0.025. This model leverages the context of target words to generate dense vector representations, which capture the semantic relationship among words. Our training sample comprises the set of OM-related Q&A samples excluding those reserved for subsequent training and testing. To validate the CBOW model result, we compute the top five most similar words to the word "supplier," which include "suppliers," "manufacturer," "vendor," "vendors," and "manufacturers." Next, we refer to Siano (2025) and select commonly used classic classification algorithms, including Gradient Boosting, Random Forest, and K-Nearest Neighbors (KNN), for the Questions and Answers classification tasks. The performance of both LLM-based classifiers and classical NLP classifiers is evaluated against 1,000 human-coded samples. Figure OA6 illustrates the nuanced differences in classification pipelines between classical NLP methods and fine-tuned LLMs. Table 4 reports the performance metrics of all classifiers and shows that our fine-tuned LLM, together with the GPT-4o model, outperforms all classical NLP models across all evaluation metrics, including F1 score, accuracy, macro precision, and macro recall. The highest accuracy values achieved by the machine learning models are 72.28% and 74.90% for the two classification tasks, compared to 88.47% and 87.75% for the fine-tuned LLMs. To further assess the statistical significance of these performance differences, we employ the McNemar test. As reported in Table OA2, the results indicate no statistically significant difference between our fine-tuned model and GPT-4o, while both achieve statistically significantly higher performance than all classical NLP models. Next, we examine error rates and error distributions for both LLM-based and classical NLP-based classifiers by comparing the confusion matrices of our model and the best-performing classical NLP approach, Gradient Boosting. Figure 3(b) shows that only a limited number of samples are misclassified, meaning they appear off the diagonal, in our model, whereas the classical NLP model exhibits substantial difficulty in distinguishing positive and negative classes from other categories. We attribute part of the superior performance of the LLM-based model to its ability to process complex contextual information. To substantiate this claim, we compare model accuracy across different text length quartiles in order to assess whether LLMs exhibit greater advantages when handling longer texts and richer contexts. Figure 3(c) shows that the fine-tuned LLM consistently outperforms classical NLP models across all text length groups. Although accuracy declines for both approaches as text length increases, the performance degradation of the LLM-based model is noticeably less pronounced than that of the classical NLP models.

By categorizing questions and answers using a fine-tuned LLM, we visualize the temporal changes of different types of questions and attribution answers in Figure OA3. This Figure shows the absolute value changes in negative/positive questions and internal/external factors over time, indicating that the Q&A primarily spans the period from 2008 to 2023. Figure OA4 shows the distribution of primary factor types in responses to OM-related questions. Questioners exhibit a stronger interest in exploring the reasons behind under-performance than outperformance. In contrast, managers tend to attribute performance first to internal factors, followed by the external environment, and lastly to supply chain partners. Figure OA5 presents a model-free analysis of self-attribution tendency, plotting the unconditional proportions of different Q&A types averaged across firms and time.

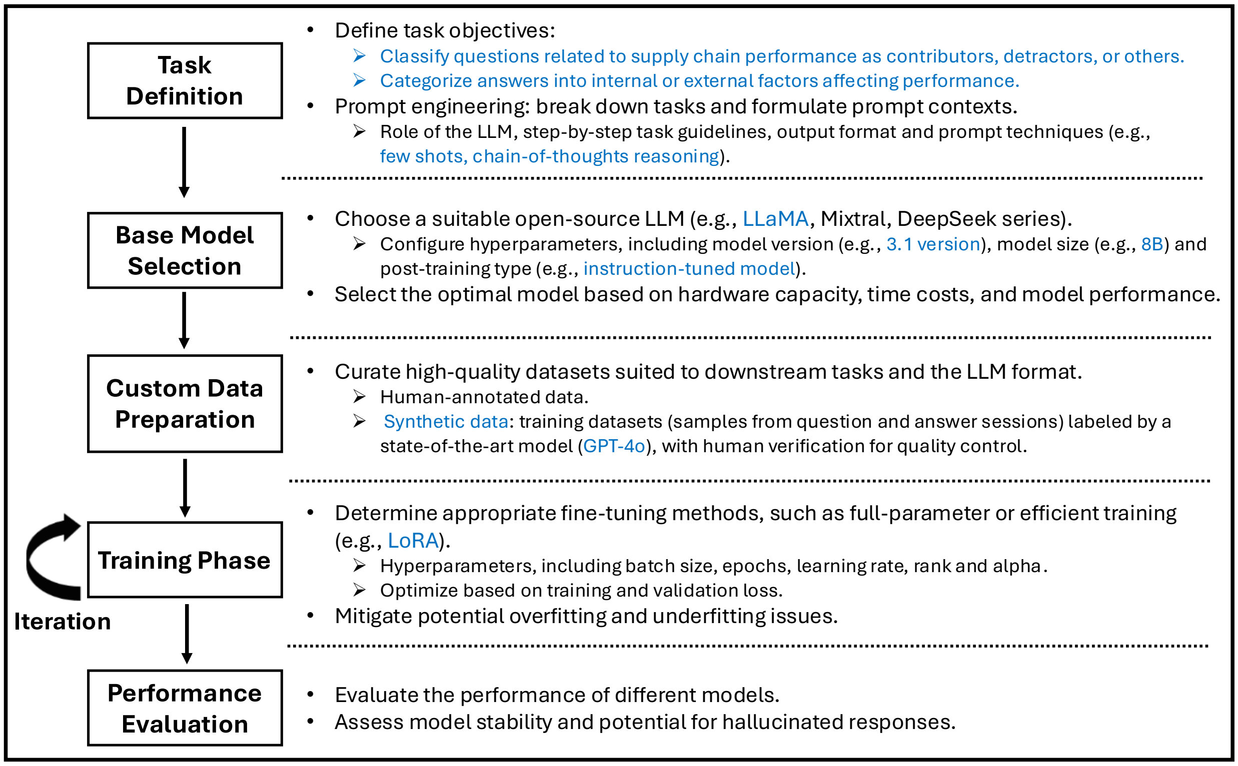
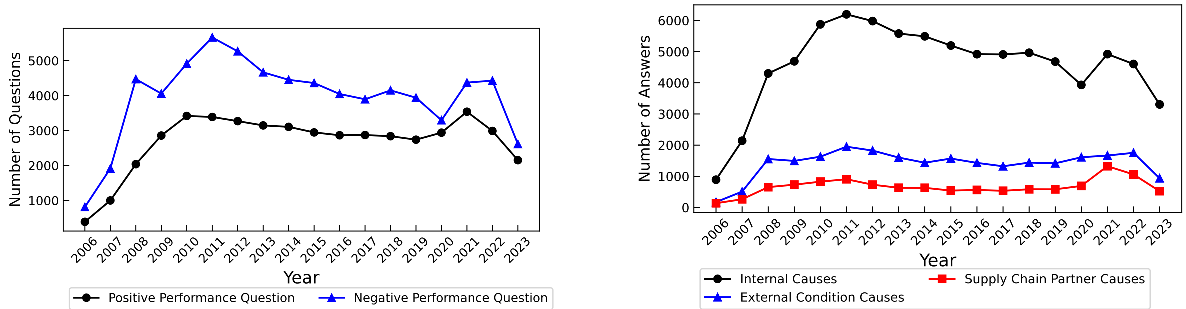


Figure OA2 **Large Language Model Fine-Tuning Pipeline**

Table OA2 **McNemar Test for Comparing Classifiers**

	GPT-4o-2024-08-06	Fine-tuned LLM	Claude-3-5-sonnet	Gemini-1.5-pro	Gradient Boosting	KNN	Random Forest
GPT-4o-2024-08-06	—	NS	NS	***	***	***	***
Fine-tuned LLM	NS	—	NS	***	***	***	***
Claude-3-5-sonnet	NS	NS	—	***	***	***	***
Gemini-1.5-pro	***	***	***	—	***	***	***
Gradient Boosting	***	***	***	***	—	***	***
KNN	***	***	***	***	NS	—	***
Random Forest	***	***	***	***	NS	NS	—

This table reports the pairwise comparison results using the McNemar test, employed to verify the statistical significance of the performance differentials between classification models. The upper-right triangle displays the statistical results for the question classification task, whereas the lower-left triangle presents the corresponding outcomes for the answer classification task. *Note.* $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$, 'NS' denotes no statistically significant difference.



(a) Number of Classified Questions Over Time

(b) Number of Classified Answers Over Time

Figure OA3 **Number of Classified Questions and Classified Answers Over Time**

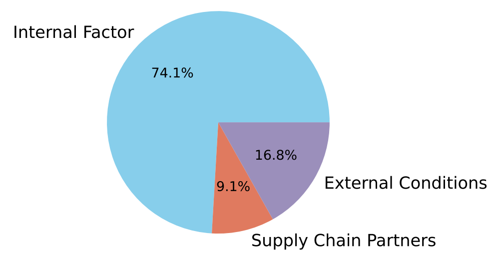


Figure OA4 **Distribution of Primary Factors in Responses to OM-Related Questions**

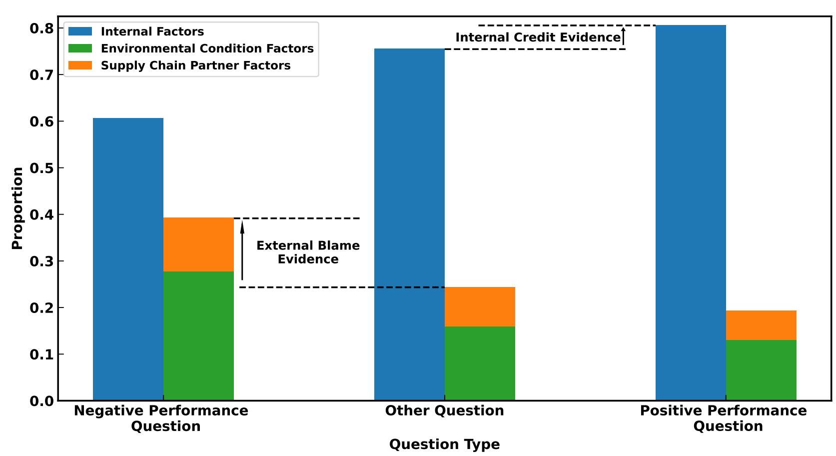


Figure OA5 **Proportion of Answers Surrounding Internal, Supply Chain, or Environmental Factors for Positive Performance, Negative Performance, and Other Questions**

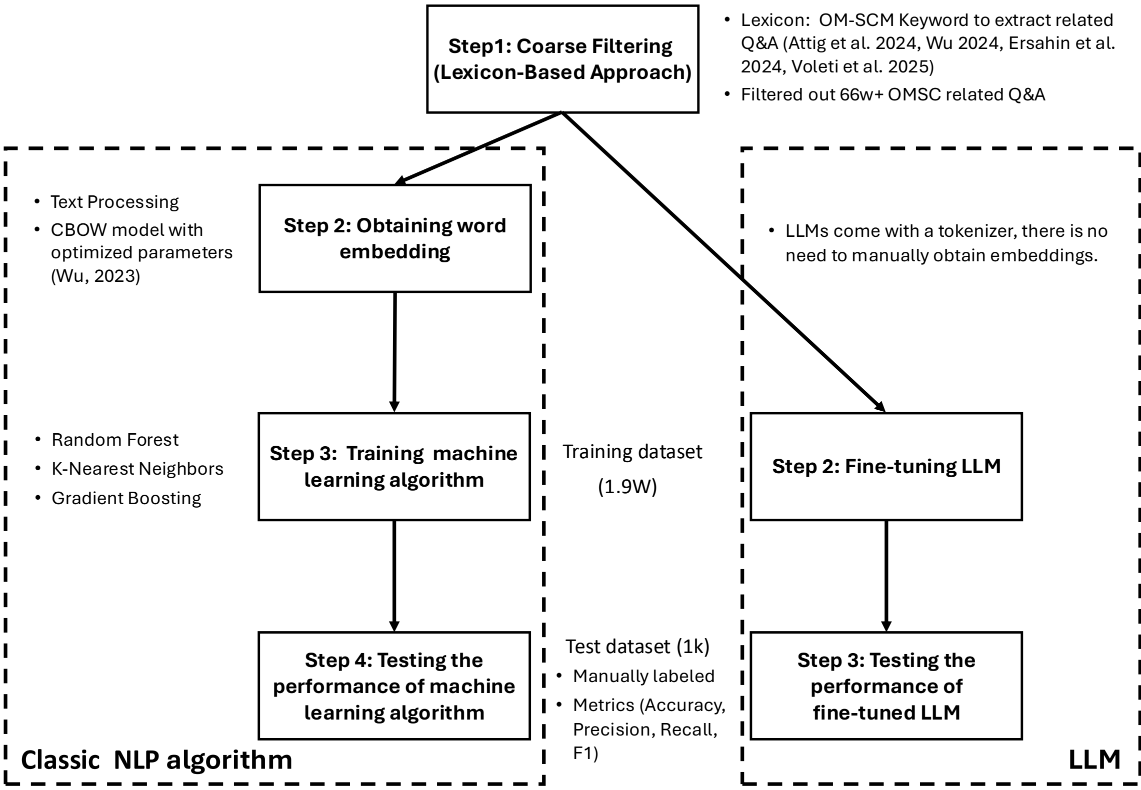


Figure OA6 Text Classification Workflow: Differences Between Classic NLP and Fine-tuning LLMs

Online Appendix C. Variable Measurement

Table OA3 presents the description, measurement, and data sources of the dependent variables, independent variables and controls.

Table OA3 Variable Descriptions

Variable	Measurement	Data Source
	Panel A: Question&Answer-Quarter Level	
$FactorType^{a}_{i,j,t}$	A categorical variable without order, taking values 0, 1, or 2 to indicate whether the answer a attributes the performance to internal factors, supply chain partners, or external conditions within the operations and supply chain-related Q&A pair j, respectively.	S&P Capital IQ Transcripts
$PositiveQuestion^{q}_{i,j,t}$	An indicator that equals 1 if the question q in the operations and supply chain-related Q&A pair j asks about the reasons for positive performance, and 0 otherwise.	S&P Capital IQ Transcripts
$NegativeQuestion^{q}_{i,j,t}$	An indicator that equals 1 if the question q in the operations and supply chain-related Q&A pair j asks about the reasons for negative performance, and 0 otherwise.	S&P Capital IQ Transcripts
$QuestionLength^{q}_{i,j,t}$	The total length of the question q in the Q&A pair j.	S&P Capital IQ Transcripts
$AnswerLength^{a}_{i,j,t}$	The total length of the answer a in the Q&A pair j.	S&P Capital IQ Transcripts
$PastPositiveQRate^{q}_{i,j,t}$	The proportion of positive questions prior to question q.	S&P Capital IQ Transcripts
$PastNegativeQRate^{q}_{i,j,t}$	The proportion of negative questions prior to question q.	S&P Capital IQ Transcripts
	Panel B: Firm-Quarter Level	
$SelfAttribution_{i,t}$	The proportion in firm i's earnings call during quarter t for operations and supply chain performance attribution Q&A, calculated as: $$AttributionTendency = \frac{(PI-OI)+(NS-OS)+(NE-OE)}{(PI+OI)+(NS+OS)+(NE+OE)},$$ $$SelfAttribution = AttributionTendency - \overline{AttributionTendency}$$ Section 3.2 provides details of the index.	S&P Capital IQ Transcripts
$PIRatio_{i,t}$	The proportion of Q&A pairs where management attributes positive performance to internal causes during firm i's earnings call in quarter t, compared to other questions and industry level. Section 3.2 provides details of the index.	S&P Capital IQ Transcripts
$NSRatio_{i,t}$	The proportion of Q&A pairs where management attributes negative performance to the actions of supply chain partners during firm i's earnings call in quarter t, compared to other questions and industry level. Section 3.2 provides details of the index.	S&P Capital IQ Transcripts
$NERatio_{i,t}$	The proportion of Q&A pairs where management attributes the negative performance to the environmental conditions during firm i's earnings call in quarter t, compared to other questions and industry level. Section 3.2 provides details of the index.	S&P Capital IQ Transcripts
$Inventory\ Turnover_{i,t}$	Cost of goods sold divided by the average inventory for firm i during quarter t, where the average inventory is calculated based on values from the prior quarter $t-1$ and the current quarter t.	Compustat Fundamentals Quarterly
$Size_{i,t}$	The natural logarithm of the total assets of the firm i during quarter t.	Compustat Fundamentals Quarterly
$Experienced\ Managers_{i,t}$	An indicator that equals 1 if at least one answer to operations and supply chain performance attribution questions in firm i's earnings call during quarter t comes from a manager familiar with the company's supply chain, and 0 otherwise.	S&P Capital IQ Transcripts
$SCRisk_{i,t}$	The frequency of co-occurrence of supply chain keywords and risk keywords in firm i's earnings call during quarter t.	S&P Capital IQ Transcripts
$QuestionLength_{i,t}$	The total length of the questions during firm i's earnings call in quarter t.	S&P Capital IQ Transcripts
$AnswerLength_{i,t}$	The total length of the answers during firm i's earnings call in quarter t.	S&P Capital IQ Transcripts
	Panel C: Implications and Mechanisms Variables	
$Gross\ Margin_{i,t+1}$	The gross profit divided by sales for firm i during quarter $t+1$	Compustat Fundamentals Quarterly
$ROA_{i,t+1}$	The ratio of net income to total assets for firm i during quarter $t+1$.	Compustat Fundamentals Quarterly
$EPS_{i,t+1}$	Earnings per share for firm i during quarter $t+1$.	Compustat Fundamentals Quarterly
$Tobin'sQ_{i,t+1}$	The ratio of market value to book value for firm i during quarter $t+1$	Compustat Fundamentals Quarterly
$Media\ Sentiment_{i,t}$	The average textual tone of the news for firm i during the six days following the earnings call in quarter t.	RavenPack
$CAR_{i,t}$	The three-day cumulative abnormal return for firm i centered on the earnings call date in quarter t.	Compustat Security Daily
$SGA_{i,t+1}$	The natural logarithm of selling, general, and administrative expenses over total assets for firm i during quarter $t+1$.	Compustat Fundamentals Quarterly
$Production\ Variability_{i,t+1}$	The standard deviation of the first-difference of the natural logarithm of the production series for firm i, calculated over a rolling window of 8 quarters ending in $t+1$.	Compustat Fundamentals Quarterly
$Demand\ Variability_{i,t+1}$	The standard deviation of the first-difference of the natural logarithm of the cost of goods sold series for the firm i, calculated over a rolling window of 8 quarters ending in $t+1$.	Compustat Fundamentals Quarterly
$BreakLinkage\ withSuppliers_{i,t+1}$	An indicator that equals 1 if firm i terminates any supplier relationships in quarter $t+1$, and 0 otherwise.	FactSet Revere
$BreakLinkage\ withGoodSuppliers_{i,t+1}$	An indicator equal to 1 if firm i terminates relationships with any "good" suppliers (inventory turnover $>$ median) in quarter $t+1$.	FactSet Revere & Compustat
$BreakLinkage\ withBadSuppliers_{i,t+1}$	An indicator equal to 1 if firm i terminates relationships with any "bad" suppliers (inventory turnover $<$ median) in quarter $t+1$.	FactSet Revere & Compustat
$Unassertiveness_{i,t+1}$	The fraction of co-occurrences between OM-SCM-related words and negative words within a 5-word range relative to the total number of words in the next 10-Q reports for firm i during quarter $t+1$.	SEC EDGAR

Online Appendix D. Robustness Checks.

This Appendix details robustness checks across various alternative empirical choices, including modifications to the construction of $SelfAttribution$, changes to the MNL specification, aggregation of Q&A-quarter data to the firm-quarter level, exclusion of leading questions and adjustments for systematic bias in $SelfAttribution$ introduced by LLMs.

Alternative Construction of *SelfAttribution*. We consider an alternative construction of the $SelfAttribution$ metric to remove potential computational artifacts associated with our main measure. Specifically, we recalculate the $SelfAttribution$ metric based on the length of Q&A pairs rather than their count. The distribution of the new $SelfAttribution$ is slightly changed compared to the one in the main text (Figure OA7 (a) and (b)). Panel A of Table OA4 shows that the drivers of self-attribution remain unchanged under this alternative variable construction. Table OA4 shows that the relationship between $SelfAttribution$ and future firm performance (Panel B) and the evidence for underlying mechanisms (Panel C) remain consistent with our main findings.

Alternative Approach to MNL Specification. To substantiate the existence of self-attribution, we re-evaluate the MNL regression setting using an alternative approach. While MNL offers a more theoretically consistent statistical framework for modeling choices among multiple mutually exclusive and unordered categories, the use of individual logistic regressions can provide complementary insights into specific attribution patterns. Specifically, we deconstruct the independent variable from MNL Model (3), which originally identified three unordered attribution categories, into three distinct indicator variables: $InternalFactor$, $SupplyChainPartnerFactor$, and $EnvironmentalConditionFactor$. Each indicator variable takes a value of 1 if performance is attributed to the corresponding category and 0 otherwise. Under this revised regression setting, the observed attribution pattern in Table OA5 remains nearly unchanged.

Aggregation to the Firm Level. While our main analysis centers around Q&A-quarter data, we also explore aggregating questions and answers to the firm-quarter level to validate the existence of self-attribution across different levels of granularity. Table OA6 illustrates that self-attribution also exists at the firm-quarter level.

Exclusion of Leading Questions. One challenge is the presence of *leading questions*, which steer managers toward predetermined answers via their phrasing. For instance, consider the following question: "*...Despite the geopolitics, Europe's demand remains strong.*

Given the 85.00% procurement from North America, should this not also be regarded as a favorable tailwind?" Here, the questioner is guiding the manager to draw a connection between Europe's demand and the market environment in North America. Our analysis of a random 1,000 questions reveals that 13.10% of them tend to be leading. Questions like this may bias responses by leading respondents to echo the factors suggested by the questioner, potentially skewing the measurement of self-attribution. We hypothesize that if the questioner successfully leads managers to attribute performance to a certain factor, the similarity between the question and answer embeddings should be high. Based on this assumption, we use an embedding-based approach, separately inputting each Q&A pair's question and answer into the LLM_A to extract the sentence embedding for capturing their attribution propensity. We split the 1,000 samples into validation and test sets at a 1:4 ratio. A 0.75 similarity threshold, determined via a validation set, distinguishes respondent influence from questioners. Applying this threshold to the test samples, we achieve a classification accuracy of 0.89, an F1 score of 0.71, a mean precision of 0.79, and a mean recall of 0.68. By applying a similarity threshold, we construct a subsample that excludes leading Q&As. Table OA7 provides robust evidence for the existence of self-attribution.

First Attribution and First Other Question Subsample. Another concern is that the questioner's attributional questions are influenced by the content of past communication. First, we summarize the proportion of attributional discussions within the 0-25%, 25-50%, 50-75%, and 75-100% quartiles of the conference call, and find that the proportion of attribution questions in each group decreases progressively, with percentages of 29.21%, 26.17%, 22.54%, and 22.09% observed in each respective quartile. Despite the decreasing trend, there remains a positive correlation between the attributional tendency of current questions and that of preceding inquiries. We find that positive (negative) questions tend to follow more previously asked positive (negative) questions. To mitigate the influence of prior conversational patterns on self-attribution, we construct a subsample that includes only the first instance of an attributional Q&A pair and a non-attributional Q&A pair. This approach helps remove the potential impact of preceding context. The results, presented in Table OA8, confirm the robustness of the existence of self-attribution.

The Impact of Multi-factor Responses. The classification of answers also contains some nuanced complexity. We observe that managers' responses sometimes enumerate multiple factors for a particular outcome. We test the occurrence of multiple factors and find that among 1,000 sample answers, 22.30% mention multiple factors. However, when multiple factors are given, they frequently exhibit a clear hierarchy, with a primary factor taking precedence over the secondary ones. 87.07% of these predominantly highlight a single, main factor. This finding guides our approach to designing prompts and assessing model performance by emphasizing the multiple factors. For the 30 answers without a clear primary factor, we observe that they consistently respond to "other" questions, which leads to an underestimation of self-attribution. To quantify this, we calculate the difference between self-attribution values derived from 1,000 manually-labeled samples (excluding those without a primary factor) and those from LLM-labeled samples (including those without a primary factor). The mean difference, across 100 simulations, is 0.00016. Therefore, to the extent there may be a bias in variable measurement introduced by multi-factor questions, its economic magnitude is small and in the direction of underestimation, thus the validity of subsequent analysis is unlikely to be affected.

The Impact of LLM Classification Bias. Although the LLM-based classifiers have high degrees of accuracy, they can sometimes misclassify Q&A types. To account for any potential bias introduced by LLMs in classification tasks, we incorporate manually annotated samples to conduct robustness checks. For each Q&A pair, we infer the potential true class based on the precision values of each category recorded in the confusion matrix. For instance, an observation classified as a positive question has an 83.18% likelihood of being correct, with probabilities 13.08% and 3.74% that it is in fact "other question" or "negative question". We sample from these probabilities 100 times for all samples, and for each iteration, we rerun the analysis presented in Table 5 on the simulated samples. We extract the coefficients of interest from each iteration and visualize their distribution in Figure OA8. The estimates in the adjusted samples consistently retain the same signs as those from our main analysis in Table 5.

Misclassification of Q&A categories may introduce bias in the measurement of $SelfAttribution$. To evaluate the impact of this error on operational and financial implications, we quantify its magnitude using out-of-sample data. Specifically, we prompt the LLM to perform inference on the test samples 100 times and measure the discrepancies between the actual $SelfAttribution$ and the predicted $SelfAttribution$ in each iteration. On average, the mean error between the actual and predicted $SelfAttribution$ is 0.0009, with a standard deviation of 0.2102. This error reflects both the bias introduced by the LLM's predictions and errors stemming from hallucinations. We assume that these observed errors follow a normal distribution and incorporate disturbances sampled from this distribution into the full sample estimates of $SelfAttribution$ values. We then re-run the analyses in Table 7 and Table 9, iterating this process 100 times. From each iteration, we extract the coefficients of $SelfAttribution$ and visualize their distribution in Figure OA8. The results show that the signs of the coefficients remain consistent, and their statistical significance mostly aligns with those reported in Table 7 and 9. These results indicate that our findings do not appear to be affected by LLM-induced classification biases.

Next, we explore more deeply the influence of LLMs' systematic bias on results by hypothesizing a strong, directional bias rather than random bias. Specifically, in cases of LLM misclassification, we consistently assign attributions to a fixed direction, such as entirely to internal factors, supply chain partners, or environmental conditions. This setup represents a deliberate extreme in potential classification bias. To investigate this, we conduct 100 simulations. In each simulation, for every sample in the dataset, we introduce an incorrect classification probability (calculated as 1 minus the precision rate of the corresponding category). If a classification is deemed incorrect, the corresponding attribution is assigned to a fixed factor. This 100-run simulation across the entire sample allows us to investigate the existence, implications, and mechanisms of LLM performance under strong systematic bias. As depicted in Figure OA9, which illustrates the coefficient distributions for various variables of interest under three distinct directional bias scenarios, the signs of the coefficients remain consistent. Theoretically, our main findings fall within the range of these simulated outcomes, thereby showing the robustness of our results.

Table OA4. Robustness Checks: Replace the Number of Q&A in Self-Attribution by Q&A Length

Panel A: Factors Influencing OM Self-Attribution

Variable	(1) $SelfAttribution$	(2) $PIRatio$	(3) $NSRatio$	(4) $NERatio$
$InventoryTurnover$	-0.002***	-0.000	-0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
$Size$	0.014**	0.002	0.007***	0.005***
	(0.005)	(0.005)	(0.002)	(0.002)
$ExperiencedManagers$	0.004***	0.003***	-0.000	0.001
	(0.001)	(0.001)	(0.000)	(0.000)
$SCRisk$	-0.001	0.001	-0.003***	-0.000
	(0.001)	(0.001)	(0.000)	(0.001)
$QuestionLength$	0.016***	0.010***	0.001	0.006***
	(0.002)	(0.002)	(0.001)	(0.001)
$AnswerLength$	-0.003	-0.006***	0.004***	0.000
	(0.002)	(0.002)	(0.001)	(0.001)
Constant	0.002***	0.002***	-0.002***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Year-Quarter FE	Y	Y	Y	Y
Company FE	Y	Y	Y	Y
Observations	91,473	91,473	91,473	91,473
R-squared	0.086	0.081	0.082	0.073

Panel B: Implications of OM-SCM Self-attribution

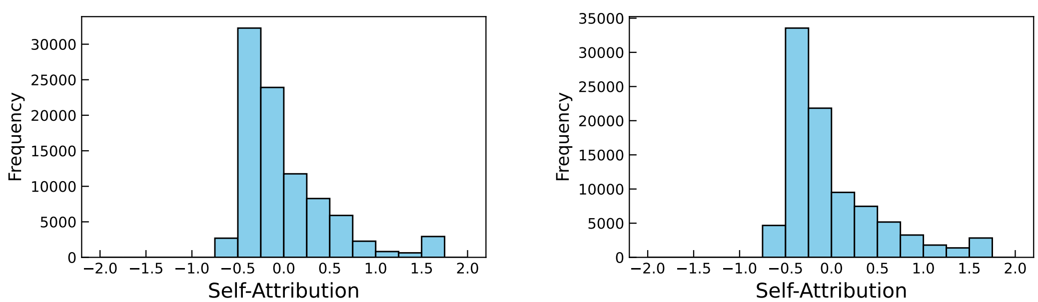
Variable	(1) $Gross\ Margin$	(2) ROA	(3) EPS	(4) $Tobin'sQ$
$SelfAttribution$	0.003***	0.002***	0.020*	0.070
	(0.001)	(0.000)	(0.011)	(0.051)
$InventoryTurnover$	-0.041***	-0.002	-0.004	-0.085
	(0.013)	(0.001)	(0.030)	(0.067)
$Size$	0.011	-0.004***	0.154**	-0.725**
	(0.011)	(0.001)	(0.060)	(0.290)
$ExperiencedManagers$	0.001	0.000	0.000	-0.023
	(0.001)	(0.000)	(0.003)	(0.015)
$SCRisk$	-0.004***	-0.000**	-0.004	-0.076***
	(0.001)	(0.000)	(0.005)	(0.023)
$QuestionLength$	0.003**	0.001***	0.017**	0.112***
	(0.001)	(0.000)	(0.007)	(0.041)
$AnswerLength$	-0.001	-0.000	0.011*	0.031
	(0.001)	(0.000)	(0.006)	(0.053)
Constant	0.397***	0.006***	0.452***	2.907***
	(0.000)	(0.000)	(0.000)	(0.015)
Year-Quarter FE	Y	Y	Y	Y
Company FE	Y	Y	Y	Y
Observations	90,628	90,832	90,420	85,622
R-squared	0.770	0.454	0.416	0.391

Table OA4. Robustness Checks: Replace the Number of Q&A in Self-Attribution by Q&A Length (Continued)

Panel C: Mechanisms of OM-SCM Self-attribution

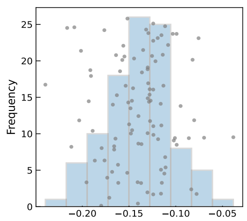
	Market Reactions		Internal Operational Activities						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	*Media Sentiment*	*CAR*	*SGA*	*Production Variability*	*Demand Variability*	*BreakLinkage withSuppliers*	*BreakLinkage withGoodSuppliers*	*BreakLinkage withBadSuppliers*	*Unassertiveness*
SelfAttribution	0.016*	0.246*	0.009***	-0.009***	-0.006***	0.010**	-0.003	0.015***	-0.002***
	(0.008)	(0.142)	(0.003)	(0.003)	(0.002)	(0.004)	(0.006)	(0.006)	(0.001)
InventoryTurnover	0.008**	0.122*	-0.002	-0.016	-0.003	-0.001	-0.001	-0.002	-0.000
	(0.003)	(0.066)	(0.007)	(0.011)	(0.006)	(0.002)	(0.004)	(0.005)	(0.001)
Size	-0.040***	-1.302***	-0.627***	0.012	0.009	0.011	0.014	-0.001	0.001
	(0.008)	(0.134)	(0.025)	(0.018)	(0.011)	(0.007)	(0.011)	(0.010)	(0.003)
ExperiencedManagers	0.000	-0.007	-0.000	-0.001	-0.000	-0.003**	0.001	-0.000	0.000
	(0.001)	(0.029)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.000)
SCRisk	-0.005***	-0.135***	-0.007***	0.002	-0.002	-0.002	-0.002	0.000	0.004***
	(0.002)	(0.026)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
QuestionLength	-0.001	-0.111**	0.002	-0.002	-0.000	0.001	-0.000	-0.005	0.001
	(0.002)	(0.043)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)
AnswerLength	-0.002	-0.040	-0.003	-0.000	-0.002	0.001	-0.002	0.002	0.000
	(0.003)	(0.039)	(0.003)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)
Constant	0.077***	-0.722***	-3.401***	0.262***	0.197***	0.088***	0.038***	0.063***	0.282***
	(0.004)	(0.007)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Company FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	39,819	81,774	83,569	90,325	90,855	69,435	24,223	24,223	43,267
R-squared	0.158	0.108	0.961	0.644	0.596	0.121	0.111	0.150	0.720

Note. This table provides robustness checks for the factors, implications and mechanisms of self-attribution. Panel A, Panel B and Panel C correspond to Table 6, Table 7 and Table 9, respectively. The results for the implications of *SelfAttribution* remain robust when disaggregating the *SelfAttribution* into *PIRatio*, *NSRatio* and *NERatio*. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests. All text-level and firm-level control variables adopt z-score standardization. OLS estimates with robust standard errors clustered by the industry are in parentheses.

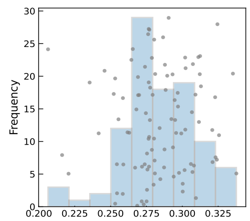


(a) Self-Attribution Measured by Count (b) Self-Attribution Measured by Length

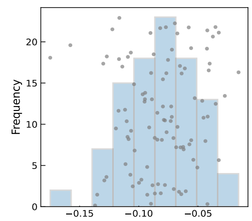
Figure OA7 Distribution of OM Self-Attribution



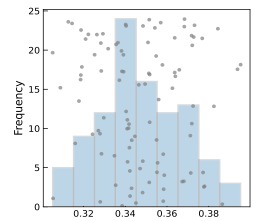
(a) Coefficient of $PositiveQuestion$ on $SupplyChainPartnerCause$



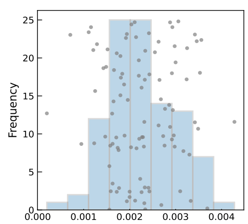
(b) Coefficient of $NegativeQuestion$ on $SupplyChainPartnerCause$



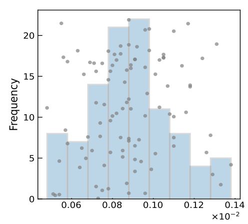
(c) Coefficient of $PositiveQuestion$ on $EnvironmentalConditionCause$



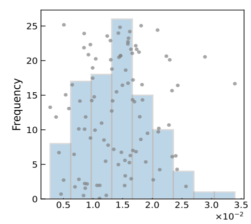
(d) Coefficient of $NegativeQuestion$ on $EnvironmentalConditionCause$



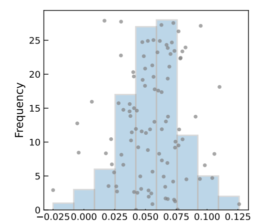
(e) Coefficient of $SelfAttribution$ on $GrossMargin$



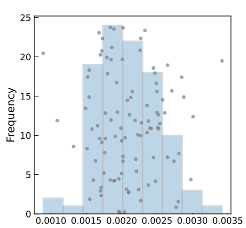
(f) Coefficient of $SelfAttribution$ on ROA



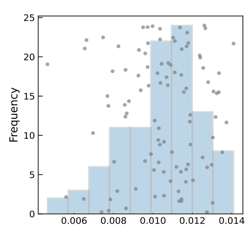
(g) Coefficient of $SelfAttribution$ on EPS



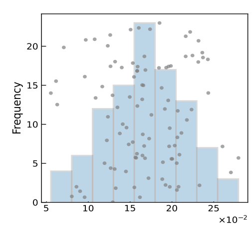
(h) Coefficient of $SelfAttribution$ on $Tobin'sQ$



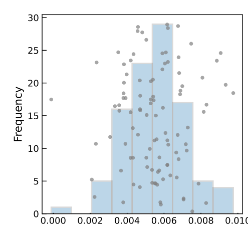
(i) Coefficient of $SelfAttribution$ on $SalesTurnover$



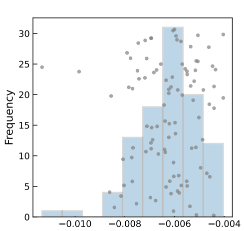
(j) Coefficient of $SelfAttribution$ on $MediaSent$



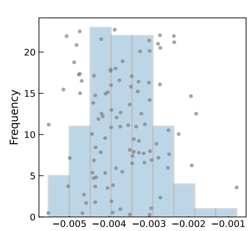
(k) Coefficient of $SelfAttribution$ on CAR



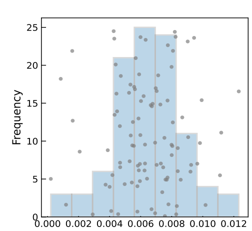
(l) Coefficient of $SelfAttribution$ on SGA



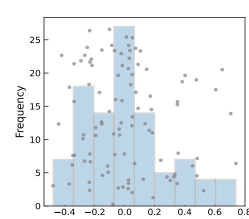
(m) Coefficient of $SelfAttribution$ on $ProductionVariability$



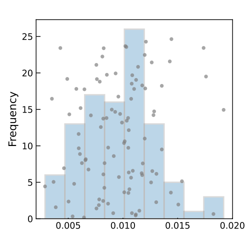
(n) Coefficient of $SelfAttribution$ on $DemandVariability$



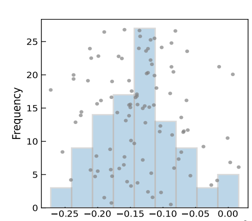
(o) Coefficient of $SelfAttribution$ on $BreakLinkagewithSuppliers$



(p) Coefficient of $SelfAttribution$ on $BreakLinkagewithGoodSuppliers$

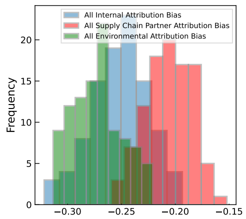


(q) Coefficient of $SelfAttribution$ on $BreakLinkagewithBadSuppliers$

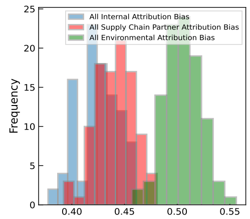


(r) Coefficient of $SelfAttribution$ on $Unassertiveness$

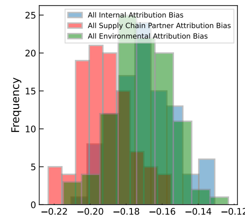
Figure OA8 Robustness Checks of OM Self-Attribution: Accounting for Random LLM Classification Bias



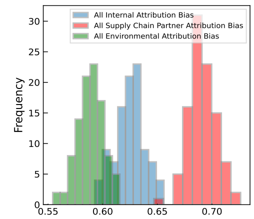
(a) Coefficient of $PositiveQuestion$ on $SupplyChainPartnerCause$



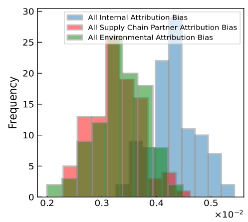
(b) Coefficient of $NegativeQuestion$ on $SupplyChainPartnerCause$



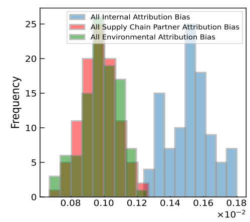
(c) Coefficient of $PositiveQuestion$ on $EnvironmentalConditionCause$



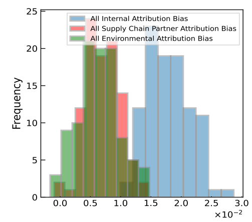
(d) Coefficient of $NegativeQuestion$ on $EnvironmentalConditionCause$



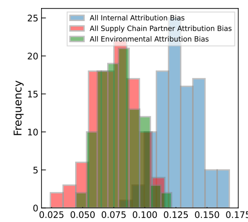
(e) Coefficient of $SelfAttribution$ on $GrossMargin$



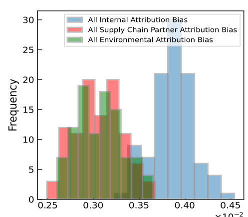
(f) Coefficient of $SelfAttribution$ on ROA



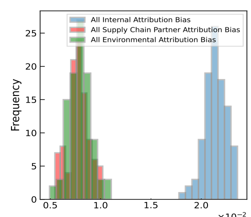
(g) Coefficient of $SelfAttribution$ on EPS



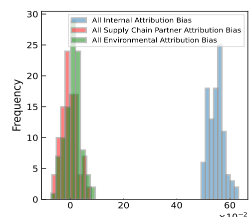
(h) Coefficient of $SelfAttribution$ on $Tobin'sQ$



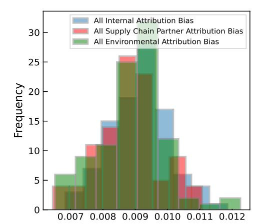
(i) Coefficient of $SelfAttribution$ on $SalesTurnover$



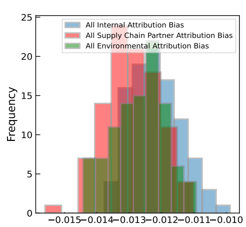
(j) Coefficient of $SelfAttribution$ on $MediaSent$



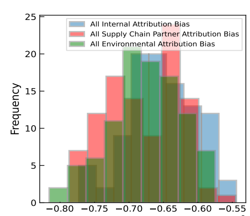
(k) Coefficient of $SelfAttribution$ on CAR



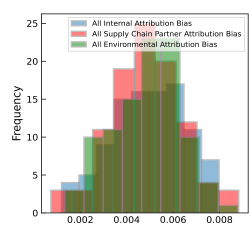
(l) Coefficient of $SelfAttribution$ on SGA



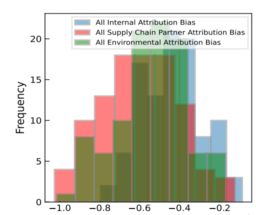
(m) Coefficient of $SelfAttribution$ on $Production$



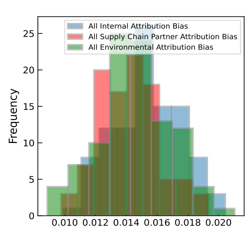
(n) Coefficient of $SelfAttribution$ on $DemandVariability$



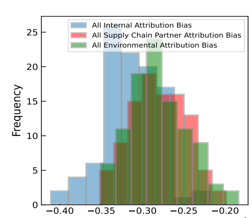
(o) Coefficient of $SelfAttribution$ on $BreakLinkagewithSuppliers$



(p) Coefficient of $SelfAttribution$ on $BreakLinkagewithGoodSuppliers$



(q) Coefficient of $SelfAttribution$ on $BreakLinkagewithBadSuppliers$



(r) Coefficient of $SelfAttribution$ on $Unassertiveness$

Figure OA9 Robustness Checks of OM Self-Attribution: Accounting for Strong LLM Classification Bias

Table OA5 Robustness Checks: Replace MNL with Single Logistic Regression

Variable	(1) $InternalFactor$	(2) $SupplyChainPartnerFactor$	(3) $EnvironmentalConditionFactor$
$PositiveQuestion$	0.182***	-0.240***	-0.054
	(0.046)	(0.058)	(0.054)
$NegativeQuestion$	-0.541***	0.247***	0.478***
	(0.033)	(0.039)	(0.037)
$QuestionLength$	0.012	-0.067***	0.048***
	(0.014)	(0.017)	(0.017)
$AnswerLength$	-0.042***	-0.035*	0.081***
	(0.015)	(0.019)	(0.017)
$PastPositiveQRate$	-0.009	-0.028	0.035
	(0.020)	(0.025)	(0.024)
$PastNegativeQRate$	0.064***	-0.061***	-0.009
	(0.017)	(0.019)	(0.019)
Company × Year-Quarter × Questioner FE	Y	Y	Y
Observations	36,821	25,728	27,301
Pseudo R-squared	0.018	0.007	0.016

Note. This table provides robustness checks for the existence of self-attribution in Table 5. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests. This table presents the results of the single logistic regression. Estimating coefficients necessitates the absence of exclusively positive or negative independent values within each Company#Year-Quarter#Questioner group.

Table OA6 Robustness Checks: Aggregate the Q&A-Quarter Level Samples to the Firm-Quarter Level

Variable	$\#InternalCause$	$\#SupplyChainPartnerCause$	$\#EnvironmentalConditionCause$
$PositiveQuestionRatio$	0.011***	-0.019***	-0.013***
	(0.001)	(0.003)	(0.003)
$NegativeQuestionRatio$	-0.047***	0.032***	0.084***
	(0.003)	(0.005)	(0.005)
Controls	Y	Y	Y
Year-Quarter FE	Y	Y	Y
Company FE	Y	Y	Y
Observations	93,295	93,295	93,295
R-squared	0.845	0.341	0.401

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests. We replace *InternalCause* at the Q&A-Quarter level in Table 5 with its absolute value at the Firm-Quarter level, denoted as *#InternalCause*. Similarly, the dependent variables *#SupplyChainPartnerCause* and *#EnvironmentalCause* follow the same adjustment. The independent variables *PositiveQuestion* and *NegativeQuestion* in Model (3) are expressed as their corresponding ratios for consistency. OLS estimates with robust standard errors clustered by the industry are in parentheses. We include the following control variables: *Inventory Turnover*, *Size*, *Experienced Executives*, *SCRisk*, *#Question*, and *#Answer*. Each of these variables is standardized using z-scores.

Table OA7 Robustness Checks: Excluding the Leading Questions

Variable	(1) *SupplyChainPartnerFactor*	(2) *EnvironmentalConditionFactor*
PositiveQuestion	-0.263***	-0.215***
	(0.049)	(0.040)
NegativeQuestion	0.484***	0.670***
	(0.035)	(0.030)
QuestionLength	-0.068***	0.017
	(0.016)	(0.013)
AnswerLength	-0.037**	0.065***
	(0.018)	(0.014)
PastPositiveQRate	0.002	0.014
	(0.018)	(0.015)
PastNegativeQRate	-0.064***	-0.033***
	(0.014)	(0.012)
Company × Year-Quarter × Questioner FE	Y	Y
Observations	77,794	77,794
Pseudo R-squared	0.0206	0.0206

This table provides robustness checks for the existence of self-attribution, corresponding to Table 5. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Table OA8 Robustness Checks: First Attribution and First Other Question Subsample

Variable	(1) *SupplyChainPartnerFactor*	(2) *EnvironmentalConditionFactor*
PositiveQuestion	-0.210*	-0.239***
	(0.117)	(0.091)
NegativeQuestion	0.419***	0.650***
	(0.082)	(0.063)
QuestionLength	-0.159***	0.002
	(0.052)	(0.040)
AnswerLength	0.024	0.081*
	(0.054)	(0.042)
PastPositiveQRate	-0.004	0.008
	(0.028)	(0.023)
PastNegativeQRate	-0.049**	-0.036**
	(0.021)	(0.018)
Company × Year-Quarter × Questioner FE	Y	Y
Observations	9,092	9,092
Pseudo R-squared	0.0531	0.0531

This table provides robustness checks for the existence of self-attribution, corresponding to Table 5. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Online Appendix E. Comparison Between Financial and OM Self-Attribution

In this section, we extend our analytical framework to the domain of financial performance to investigate managerial self-attribution behavior. As illustrated in Figure OA10, we replicate the data processing and analytical protocols established in the main text. This comparative approach allows us to study the existence, factors, implications, and underlying mechanisms of financial self-attribution.

To isolate relevant financial discussions, we leverage the subset of Q&A pairs previously classified as non-OM in our primary dataset. We employ a combination of a "bag-of-words" lexicon approach and fine-tuned LLMs to filter and categorize these texts. To isolate the financial discourse, we curate a specific lexicon of finance-related keywords (Table OA9), covering topics such as financing, investment, and market Performance. We synthesize this lexicon by integrating terminology from Schnatterly et al. (2021) and Astvansh and Simpson (2026), expanding it with GPT-4o, and then manually verifying its relevance. We apply stringent selection criteria to exclude terms overlapping between finance and OM, such as cash flow and operation leverage. Applying this filter yields a final sample of 301,618 finance-related Q&A pairs. Next, we employ a dual-LLM classification strategy similar to the methodology in Section B. Specifically, we fine-tune a distinct LLM to categorize non-OM questions into "positive," "negative," or "other" performance questions. This LLM model achieves an F1-score of 0.9006 and an accuracy of 0.9040. For the classification of managerial responses, we directly deploy the previously fine-tuned LLM_A model. Figure OA11(b) shows the distribution of financial self-attribution.

We re-estimate the whole empirical specifications from the main text within this financial context. To ensure the validity of the results, we recalibrate the control variables to reflect the specific characteristics of financial discussions. Particularly, we redefine $ExperiencedManagers$ to capture executives who demonstrate familiarity specifically with financial performance Q&A. Similarly, we substitute $SCRisk$ with $FinanceRisk$ to control for firm-specific financial uncertainties rather than supply chain exposure. Our empirical results provide a comprehensive view of financial self-attribution. Table OA10 reports the results of the existence of self-attribution tendencies in financial disclosures. Table OA11 details the factors driving these attribution behaviors. We further examine the consequences of these disclosures in Tables OA12 and OA13, which present the implications of financial self-attribution for future firm performance. Finally, Tables OA14 and OA15 elucidate the mechanisms through which financial self-attribution operates.

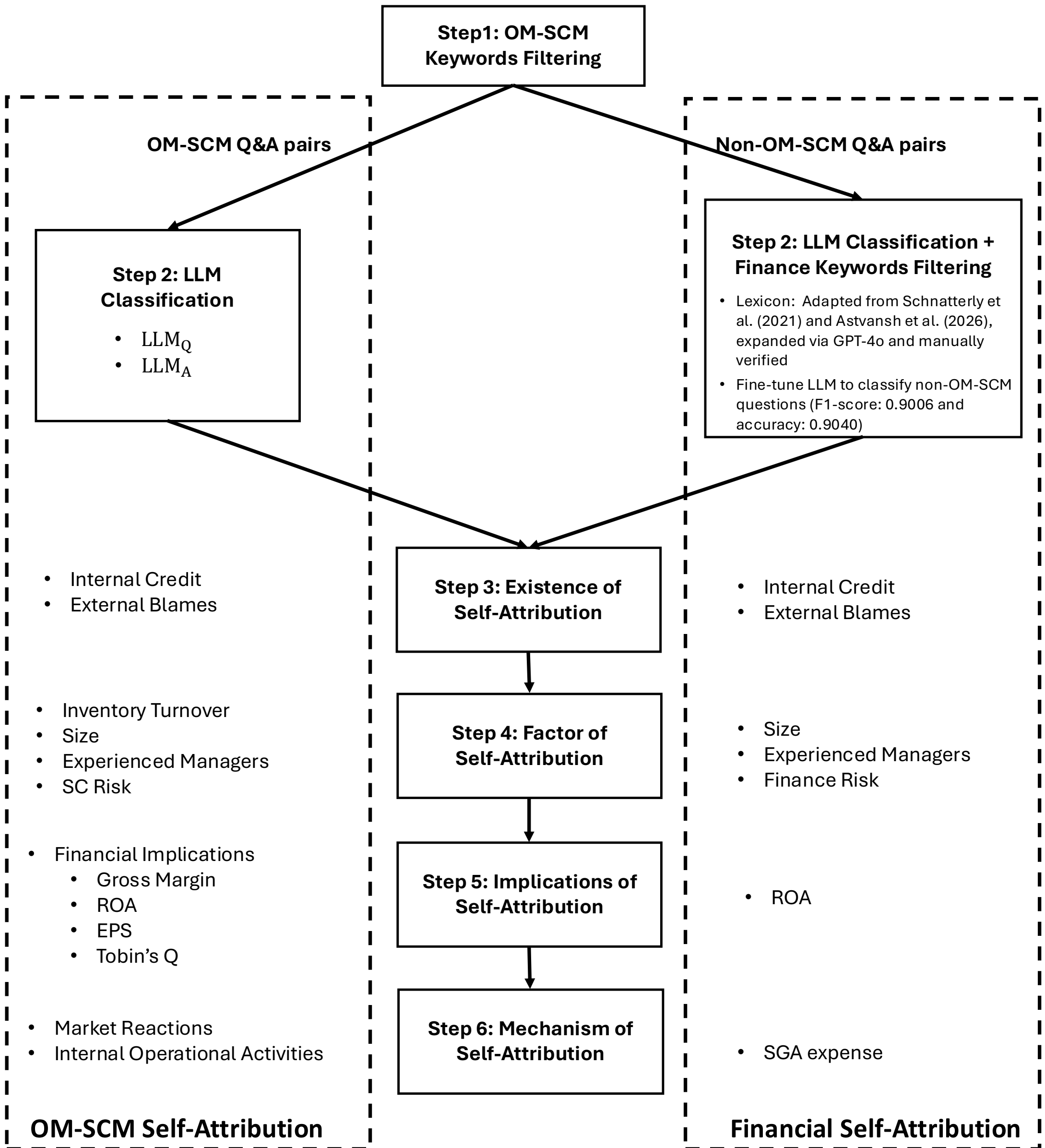
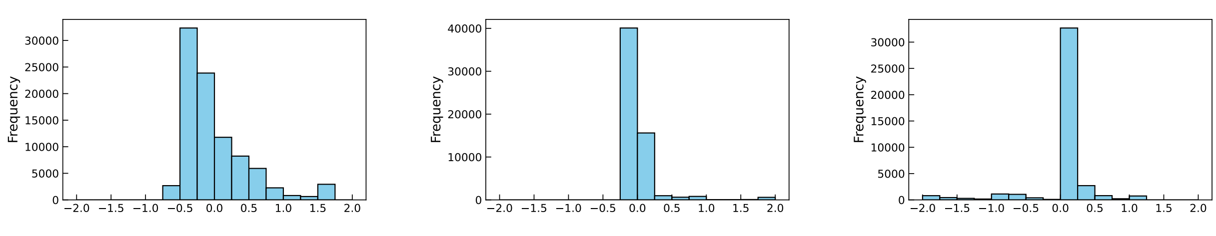


Figure OA10 **Comparison of Analytical Steps: OM Self-Attribution versus Financial Self-Attribution**



(a) OM Self-Attribution (b) Financial Self-Attribution (c) OM Self-Attributional Spread

Figure OA11 **Distribution of Self-Attribution**

Table OA9 Keywords Related to Finance

accretive	acquisition	asset sale	balance sheet strength	bond
capex	capital allocation	capital deployment	capital expenditure	capital structure
cost of debt	covenants	credit facility	credit rating	debt
debt burden	debt restructuring	deleverage	deleveraging	delisting
divestiture	drawdown	enterprise value	ev/ebitda	financing
financing cost	gross debt	indebtedness	inorganic	interest rate
issuance	listing	loan	loans	m&a
market cap	market capitalization	mergers and acquisitions	net debt	new financing
p/b ratio	p/e ratio	portfolio diversification	refinance	refinancing
reinvestment	return on invested capital	revolver	roe	roi
roic	share price	short interest	solvency	spin-off
stock price	underwriting	valuation	wacc	

Table OA10 Existence of Financial Self-attribution

Variable	(1) *SupplyChainPartnerFactor*	(2) *EnvironmentalConditionFactor*
PositiveQuestion	0.400	-0.433***
	(0.329)	(0.116)
NegativeQuestion	1.171***	0.665***
	(0.276)	(0.088)
QuestionLength	-0.014	0.017
	(0.044)	(0.014)
AnswerLength	-0.281***	-0.648***
	(0.051)	(0.018)
PastPositiveQRate	0.045	-0.009
	(0.068)	(0.022)
PastNegativeQRate	0.036	-0.038**
	(0.057)	(0.019)
Company × Year-Quarter × Analyst FE	Y	Y
Observations	48,867	48,867
Pseudo R-squared	0.0532	0.0532

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests. Estimating coefficients necessitates that the dependent variable doesn't exclusively belong to one category within each Firm × Year-Quarter × Questioner group.

Table OA11 Factors Influencing Financial Self-attribution

Variable	(1) $FinancialSelfAttribution$	(2) $PIRatio$	(3) $NSRatio$	(4) $NERatio$
$InventoryTurnover$	0.002	0.001	0.001**	-0.000
	(0.002)	(0.001)	(0.000)	(0.001)
$Size$	0.009***	0.004	0.005***	-0.004
	(0.002)	(0.004)	(0.001)	(0.004)
$ExperiencedManagers$	0.001*	-0.003***	0.000*	0.002***
	(0.001)	(0.000)	(0.000)	(0.001)
$FinanceRisk$	-0.004***	-0.000	0.000	-0.003***
	(0.001)	(0.001)	(0.000)	(0.001)
$QuestionLength$	0.006***	0.026***	-0.001***	-0.018***
	(0.001)	(0.002)	(0.000)	(0.001)
$AnswerLength$	0.002**	-0.027***	0.001***	0.029***
	(0.001)	(0.002)	(0.000)	(0.002)
Constant	-0.000***	-0.004***	-0.000***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Year-Quarter FE	Y	Y	Y	Y
Company FE	Y	Y	Y	Y
Observations	58,815	58,815	58,815	58,815
R-squared	0.075	0.102	0.092	0.108

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Table OA12 Implications of Financial Self-attribution

Variable	(1) $Gross$ $Margin$	(2) ROA	(3) EPS	(4) $Tobin'sQ$
$FinancialSelfAttribution$	0.001	0.001*	0.035	0.062
	(0.002)	(0.001)	(0.025)	(0.118)
$InventoryTurnover$	-0.025***	-0.001	-0.004	-0.091
	(0.007)	(0.001)	(0.037)	(0.060)
$Size$	0.005	-0.007***	0.131**	-0.783***
	(0.007)	(0.001)	(0.063)	(0.210)
$ExperiencedManagers$	-0.001	0.000	0.004	-0.008
	(0.001)	(0.000)	(0.004)	(0.012)
$FinanceRisk$	0.000	0.000	0.001	-0.065***
	(0.001)	(0.000)	(0.006)	(0.024)
$QuestionLength$	-0.001	-0.000	-0.016	0.011
	(0.001)	(0.000)	(0.011)	(0.026)
$AnswerLength$	0.003**	0.000	-0.000	-0.008
	(0.001)	(0.000)	(0.009)	(0.028)
Constant	0.402***	0.008***	0.482***	2.641***
	(0.000)	(0.000)	(0.000)	(0.011)
Year-Quarter FE	Y	Y	Y	Y
Company FE	Y	Y	Y	Y
Observations	58,342	58,489	58,170	54,934
R-squared	0.855	0.402	0.405	0.449

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Table OA13 Implications of Financial Self-attribution: Disentangling Effects of Internal Credit and External Blame

Variable	(1) *Gross Margin*	(2) *ROA*	(3) *EPS*	(4) *Tobin'sQ*
PIRatio	0.001	0.001	0.034	0.075
	(0.002)	(0.001)	(0.032)	(0.143)
NSRatio	0.012	0.001	-0.008	-0.198
	(0.008)	(0.002)	(0.073)	(0.280)
NERatio	0.004	0.004***	0.049	0.048
	(0.003)	(0.001)	(0.038)	(0.145)
InventoryTurnover	-0.025***	-0.001	-0.004	-0.090
	(0.007)	(0.001)	(0.037)	(0.060)
Size	0.005	-0.007***	0.132**	-0.782***
	(0.007)	(0.001)	(0.063)	(0.210)
ExperiencedManagers	-0.001	0.000	0.004	-0.008
	(0.001)	(0.000)	(0.004)	(0.012)
SCRisk	0.000	0.000	0.001	-0.065**
	(0.001)	(0.000)	(0.006)	(0.024)
QuestionLength	-0.001	-0.000	-0.016	0.010
	(0.001)	(0.000)	(0.011)	(0.025)
AnswerLength	0.003**	0.000	-0.000	-0.007
	(0.001)	(0.000)	(0.009)	(0.028)
Constant	0.402***	0.008***	0.482***	2.641***
	(0.000)	(0.000)	(0.000)	(0.011)
Year-Quarter FE	Y	Y	Y	Y
Company FE	Y	Y	Y	Y
Observations	58,342	58,489	58,170	54,934
R-squared	0.855	0.403	0.405	0.449

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Table OA14 Mechanisms of Financial Self-attribution

	Market Reactions		Internal Operational Activities						
Variable	(1) *Media Sentiment*	(2) *CAR*	(3) *SGA*	(4) *Production Variability*	(5) *Demand Variability*	(6) *BreakLinkage withSuppliers*	(7) *BreakLinkage withGoodSuppliers*	(8) *BreakLinkage withBadSuppliers*	(9) *Unassertiveness*
FinancialSelfAttribution	0.001	0.055	0.016**	0.017**	-0.000	-0.011	0.009	-0.007	-0.002
	(0.008)	(0.239)	(0.007)	(0.007)	(0.004)	(0.010)	(0.013)	(0.014)	(0.002)
InventoryTurnover	0.006	0.047	0.000	-0.021	0.000	0.001	-0.000	-0.002	-0.000
	(0.005)	(0.060)	(0.007)	(0.014)	(0.005)	(0.002)	(0.004)	(0.008)	(0.001)
Size	-0.051***	-1.338***	-0.577***	0.014	0.021	0.016**	0.020	0.001	0.003
	(0.010)	(0.184)	(0.028)	(0.023)	(0.013)	(0.008)	(0.012)	(0.014)	(0.004)
ExperiencedManagers	-0.001	-0.028	0.001	-0.002	-0.001	0.003**	0.001	0.002	0.001***
	(0.002)	(0.028)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.000)
SCRisk	0.001	0.077**	-0.008***	-0.000	0.002***	0.002	-0.001	0.004	-0.000
	(0.002)	(0.034)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)	(0.003)	(0.000)
QuestionLength	-0.011***	-0.080	0.002	0.002	0.000	0.000	0.005*	-0.004	0.001**
	(0.003)	(0.060)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.000)
AnswerLength	0.006**	-0.004	-0.003	0.001	0.001	0.000	-0.004*	-0.001	-0.001**
	(0.003)	(0.055)	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)	(0.001)
Constant	0.075***	-0.633***	-3.631***	0.278***	0.193***	0.085***	0.038***	0.064***	0.272***
	(0.006)	(0.009)	(0.002)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Company FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	22,665	53,001	53,010	58,035	58,485	43,890	14,895	14,895	28,235
R-squared	0.196	0.132	0.964	0.662	0.583	0.142	0.140	0.181	0.737

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Table OA15 Mechanisms of Financial Self-attribution: Disentangling Effects of Internal Credit and External Blame

	Market Reactions		Internal Operational Activities						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	*Media Sentiment*	*CAR*	*SGA*	*Production Variability*	*Demand Variability*	*BreakLinkage withSuppliers*	*BreakLinkage withGoodSuppliers*	*BreakLinkage withBadSuppliers*	*Unassertiveness*
PIRatio	-0.005	-0.101	0.022**	0.019**	-0.000	-0.017	0.005	-0.014	-0.002
	(0.012)	(0.295)	(0.009)	(0.009)	(0.005)	(0.012)	(0.017)	(0.018)	(0.003)
NSRatio	-0.002	0.810	0.036	0.004	-0.016	-0.039*	0.001	-0.013	-0.001
	(0.022)	(0.489)	(0.022)	(0.019)	(0.010)	(0.022)	(0.039)	(0.037)	(0.007)
NERatio	0.021*	0.485	0.018	0.021**	0.001	-0.009	0.011	-0.002	-0.004
	(0.012)	(0.334)	(0.011)	(0.010)	(0.005)	(0.014)	(0.018)	(0.024)	(0.004)
InventoryTurnover	0.006	0.046	0.000	-0.021	0.000	0.001	-0.000	-0.002	-0.000
	(0.005)	(0.060)	(0.007)	(0.014)	(0.005)	(0.002)	(0.004)	(0.008)	(0.001)
Size	-0.051***	-1.339***	-0.577***	0.014	0.022	0.016**	0.020	0.002	0.003
	(0.010)	(0.183)	(0.028)	(0.023)	(0.013)	(0.008)	(0.012)	(0.014)	(0.004)
ExperiencedManagers	-0.001	-0.029	0.001	-0.002	-0.001	0.003**	0.001	0.002	0.001***
	(0.002)	(0.028)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.000)
FinanceRisk	0.001	0.078**	-0.008***	-0.000	0.002***	0.002	-0.001	0.004	-0.000
	(0.002)	(0.034)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)	(0.003)	(0.000)
QuestionLength	-0.011***	-0.067	0.002	0.002	0.000	0.001	0.005*	-0.004	0.001**
	(0.003)	(0.060)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.004)	(0.000)
AnswerLength	0.005*	-0.021	-0.003	0.001	0.001	0.000	-0.005*	-0.001	-0.001**
	(0.003)	(0.054)	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)	(0.001)
Constant	0.075***	-0.634***	-3.631***	0.278***	0.193***	0.085***	0.038***	0.063***	0.272***
	(0.006)	(0.009)	(0.002)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Company FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	22,665	53,001	53,010	58,035	58,485	43,890	14,895	14,895	28,235
R-squared	0.197	0.132	0.964	0.662	0.583	0.143	0.140	0.181	0.737

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Online Appendix F. Comparison Between OM Self-Attributional Spread and OM Self-Attribution

In the main text, we construct OM self-attribution by benchmarking attributional OM questions against non-attributional OM questions to capture field-specific self-attribution, which is detailed in Equations (1)-(2). Here, we introduce the *OM Self-Attributional Spread* by changing the benchmark to financial performance self-attributional questions. This measure captures the extent to which a manager's self-attributional propensity in the OM diverges from their self-attributional style in financial discussions. Figure OA11(c) shows the distribution of *OM Self-Attributional Spread*. Using this alternative benchmark, we examine the existence, associated factors, implications, and mechanisms of this spread.

To formally test for the existence of OM Self-Attributional Spread, we modify the multinomial logit specification in Model (3):

$$FactorType^{a}_{i,j,t} = \alpha + \beta_1 PositiveQuestion^{q}_{i,j,t} + \beta_2 PositiveQuestion^{q}_{i,j,t} \times OMRelated^{q}_{i,j,t} + \beta_3 OMRelated^{q}_{i,j,t} + Z_{i,j,t}\gamma + \delta_i \times \delta_t \times \delta_k + \epsilon_{i,j,t,k}$$

We respecify the model by introducing a field indicator (*OMRelated*) and interacting it with the performance direction (*PositiveQuestion* $\times$ *OMRelated*) to capture the differential self-attribution tendency in the operational context.

The results in Table OA16 indicate a universal self-attribution tendency across both fields. The significantly negative coefficient for *PositiveQuestion* indicates that, across fields, managers are significantly less likely to attribute positive outcomes to external factors than to internal actions. Notably, the coefficient for the interaction term (*PositiveQuestion* $\times$ *OM_Related*) is statistically insignificant, suggesting that the self-attribution tendency in OM is statistically indistinguishable from that in financial Q&A.

Table OA17 reveals that the self-attribution tendency in OM can be divergent from the self-attribution tendency in finance for firms with lower inventory turnover and higher supply chain risks. This suggests that firms characterized by greater operational inefficiencies and external instability may face stronger pressure to explain complex operational failures.

In contrast to the broad performance implications and underlying mechanisms of OM self-attribution, the *OM Self-Attributional Spread* exhibits some correlations with firm performance and operational activities. As shown in Table OA18 and Table OA19, specific associations emerge: a wider *OM Self-Attributional Spread* correlates with higher Tobin's Q, increased CAR, more stable production variability, and active supplier termination.

Table OA16 Existence of OM Self-Attributional Spread

Variable	(1) *SupplyChainPartnerFactor*	(2) *EnvironmentalConditionFactor*
PositiveQuestion	-0.940**	-1.096***
	(0.479)	(0.238)
PositiveQuestion × *OMRelated*	0.179	0.136
	(0.483)	(0.242)
OMRelated	0.769**	-0.485***
	(0.312)	(0.162)
QuestionLength	-0.160***	0.017
	(0.047)	(0.035)
AnswerLength	0.160***	0.118***
	(0.052)	(0.036)
PastPositiveQRate	-0.031	-0.005
	(0.050)	(0.041)
PastNegativeQRate	-0.050	-0.001
	(0.038)	(0.030)
Company × Year-Quarter × Analyst FE	Y	Y
Observations	10,312	10,312
Pseudo R-squared	0.0516	0.0516

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests. Estimating coefficients necessitates that the dependent variable doesn't exclusively belong to one category within each Firm × Year-Quarter × Questioner group.

Table OA17 Factors Influencing OM Self-Attributional Spread

Variable	(1) *OMSelfAttributionalSpread*	(2) *PIRatio*	(3) *NSRatio*	(4) *NERatio*
InventoryTurnover	-0.003***	-0.002***	0.000	-0.001***
	(0.001)	(0.000)	(0.000)	(0.000)
Size	-0.002	-0.024***	0.002	0.019***
	(0.006)	(0.008)	(0.004)	(0.005)
ExperiencedManagers	-0.001	-0.000	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
SCRisk	0.013***	-0.007***	0.014***	0.006***
	(0.002)	(0.003)	(0.002)	(0.001)
QuestionLength	-0.002	-0.004*	0.000	0.002
	(0.002)	(0.002)	(0.001)	(0.002)
AnswerLength	-0.002	-0.002	-0.002**	0.002
	(0.002)	(0.002)	(0.001)	(0.002)
Constant	0.001***	-0.000***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Year-Quarter FE	Y	Y	Y	Y
Company FE	Y	Y	Y	Y
Observations	41,504	41,504	41,504	41,504
R-squared	0.148	0.119	0.127	0.117

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Table OA18 Implications of OM Self-Attributional Spread

Variable	(1) *Gross Margin*	(2) *ROA*	(3) *EPS*	(4) *Tobin'sQ*
OMSelfAttributionalSpread	-0.002	-0.000	-0.002	0.164**
	(0.001)	(0.000)	(0.014)	(0.081)
InventoryTurnover	-0.018***	-0.000	0.016	-0.127
	(0.005)	(0.001)	(0.027)	(0.109)
Size	0.005	-0.007***	0.255***	-0.789*
	(0.005)	(0.001)	(0.069)	(0.397)
ExperiencedManagers	0.000	0.000	0.002	-0.019
	(0.000)	(0.000)	(0.004)	(0.020)
SCRisk	-0.002**	-0.000	-0.001	-0.082**
	(0.001)	(0.000)	(0.007)	(0.033)
QuestionLength	0.000	0.000	0.007	0.059
	(0.001)	(0.000)	(0.007)	(0.049)
AnswerLength	-0.001	-0.000	-0.005	-0.019
	(0.001)	(0.000)	(0.009)	(0.051)
Constant	0.393***	0.011***	0.586***	3.141***
	(0.000)	(0.000)	(0.000)	(0.022)
Year-Quarter FE	Y	Y	Y	Y
Company FE	Y	Y	Y	Y
Observations	41,180	41,257	41,121	38,804
R-squared	0.908	0.444	0.464	0.439

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

Table OA19 Mechanisms of OM Self-Attributional Spread

	Market Reactions		Internal Operational Activities						
Variable	(1) *Media Sentiment*	(2) *CAR*	(3) *SGA*	(4) *Production Variability*	(5) *Demand Variability*	(6) *BreakLinkage withSuppliers*	(7) *BreakLinkage withGoodSuppliers*	(8) *BreakLinkage withBadSuppliers*	(9) *Unassertiveness*
OMSelfAttributionalSpread	0.003	0.305**	0.000	-0.016**	-0.001	0.011*	0.008	0.022	-0.002
	(0.007)	(0.152)	(0.005)	(0.006)	(0.002)	(0.006)	(0.008)	(0.013)	(0.002)
InventoryTurnover	0.012**	0.215**	0.002	-0.024	-0.005	0.002	-0.007**	-0.009	-0.000
	(0.005)	(0.106)	(0.006)	(0.016)	(0.004)	(0.004)	(0.003)	(0.008)	(0.001)
Size	-0.030***	-1.230***	-0.548***	0.020	0.019**	0.013	0.021*	-0.024*	0.005
	(0.010)	(0.189)	(0.023)	(0.016)	(0.008)	(0.011)	(0.012)	(0.012)	(0.004)
ExperiencedManagers	0.001	0.003	-0.000	0.001*	0.001**	-0.003*	-0.001	0.001	0.000
	(0.002)	(0.034)	(0.001)	(0.001)	(0.000)	(0.002)	(0.002)	(0.002)	(0.000)
SCRisk	-0.008***	-0.133***	-0.007***	0.003	-0.000	-0.001	-0.002	0.003	0.004***
	(0.002)	(0.039)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)
QuestionLength	-0.006*	-0.324***	0.004	-0.002	-0.001	-0.002	0.002	-0.003	-0.001
	(0.003)	(0.061)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.001)
AnswerLength	-0.006**	-0.056	-0.003	-0.001	-0.001	0.006**	-0.004*	0.001	0.001
	(0.002)	(0.056)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.001)
Constant	0.094***	-0.584***	-3.369***	0.234***	0.171***	0.091***	0.039***	0.068***	0.288***
	(0.006)	(0.007)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Company FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	16,714	37,268	38,882	41,027	41,249	32,033	11,302	11,302	20,870
R-squared	0.213	0.158	0.963	0.694	0.635	0.157	0.171	0.224	0.724

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ using two-tailed tests.

References

Adler, B., Agarwal, N., Aithal, A., Anh, D. H., Bhattacharya, P., Brundyn, A., Casper, J., Catanzaro, B., Clay, S., Cohen, J., et al. (2024). Nemotron-4 340b technical report. arXiv preprint arXiv:2406.11704.

Astvansh, V. and Simpson, J. J. (2026). A firm's operational risk: Data set and empirical evidence. Manufacturing & Service Operations Management, 28(1):326–341.

Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., et al. (2024). The Llama 3 herd of models. arXiv preprint arXiv:2407.21783.

Ersahin, N., Giannetti, M., and Huang, R. (2024). Supply chain risk: Changes in supplier composition and vertical integration. Journal of International Economics, 147:103854.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W., et al. (2022). Lora: Low-rank adaptation of large language models. ICLR, 1(2):3.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., and Chissom, B. S. (1975). Derivation of new readability formulas for navy enlisted personnel. Naval Technical Training Command Millington Research Branch.

Niu, Y., Wu, J., Jiang, S., and Jiang, Z. (2024). The bullwhip effect in servitized manufacturers. Management Science.

Schnatterly, K., Calvano, F., Berns, J. P., and Deng, C. (2021). The effects of board expertise-risk misalignment and subsequent strategic board reconfiguration on firm performance. Strategic Management Journal, 42(11):2162–2191.

Siano, F. (2025). The news in earnings announcement disclosures: Capturing word context using LLM methods. Management Science.

Voleti, S., Malladi, V., and Sohoni, M. G. (2025). Impact of OM content in earnings calls on a firm's stock performance. Management Science.

Wu, D. (2024). Text-based measure of supply chain risk exposure. Management Science, 70(7):4781–4801.