# The Moderating Effects of Semantic Alignment and Personalized Managerial Responses on the Relationship Between Response Characteristics and Review Helpfulness

***Altynsara Baktiyar[a], One-Ki Daniel Lee[b], Taeho Hong[a]***

[a]College of Business, Pusan National University
[b]School of Business, Virginia Commonwealth University

May 12, 2026

# CONTENT

# I. Introduction

## 1.1 Research Background

**Growing Importance of Online Reviews:**

- Online reviews have become a dominant form of electronic word-of-mouth (e-WOM) that strongly influences consumer decisions (Mudambi et al., 2010).

**Why Consumers Need Additional Signals**

- Consumers cannot directly experience products or services before purchase.
- They rely on available cues to judge whether review information is trustworthy and useful (Olson, 1972; Olson & Jacoby, 1972).

**Role of Managerial Responses as Extrinsic Cues**

- Review text alone may not be sufficient.
- Managerial responses function as extrinsic cues providing clarification, apology, and service recovery signals (Jin et al., 2023).
- Two types: Standard (generic/templated) vs. Personalized Managerial Response (PMR).
- Prior studies suggest **PMRs create more positive impressions and more information** than standard responses (Wei et al., 2013; Zhang et al., 2020).

**Role of Similarity between Manager Response and Review Texts**

- The similarity between managerial responses and review texts influences perceived helpfulness/company rating differently depending by variables(Yang et al., 2021; ; Zhang et al., 2020 ).

# I. Introduction

## 1.2 Research Gaps and Objectives

### *Research Gaps*

**1. Limited Focus on Response Characteristics**

Prior studies mainly examine **whether firms respond**, rather than **how response quality or characteristics influence helpfulness (**Jin et al., 2023; Wei et al., 2013; Hong et al., 2017)

**2. Inconsistent Findings on PMR Effectiveness**

Prior findings are **mixed**, and **PMR has been measured inconsistently**, often conflated with similarity/topic matching (Wang et al., 2020; Yang et al., 2021; Zhang et al., 2020)

**3. Limited Understanding of Semantic Alignment**

Existing studies assume **linear effects but many conclusions are mixed**, while the relationship between alignment and helpfulness may be **non-linear (**Yang et al. (2021); Wang & Chaudhry (2018)

### *Research Objectives*

- Examine the direct effects of response topic diversity and response length on review helpfulness.
- Introduce semantic alignment as a U-shaped moderator between response characteristics and review helpfulness.
- Investigate whether PMR strengthens this non-linear moderation as a second-order condition.

# II. Literature Review

## 2.1 Online Review Helpfulness

- Review helpfulness refers to the extent to which a review assists consumers in making better purchase decisions (Mudambi & Schuff, 2010).
- Helpful reviews reduce information overload, improve decision confidence, and shape purchase intentions.
- Known determinants include review length, readability, emotionality, rating extremity, reviewer credibility, and timeliness (Mudambi & Schuff, 2010; Hong et al., 2017).

## 2.2 Cue Utilization Theory

Grounded in Cue Utilization Theory (Olson, 1972; Olson & Jacoby, 1972): consumers evaluate quality based on available cues when direct assessment is difficult.

| Cue Type | Examples |
|---|---|
| Intrinsic (within the review) | Star rating, review length, sentiment, reviewer expertise |
| Extrinsic (outside the review) | Managerial response (PMR), response length, response speed |

# II. Literature Review

## 2.3. Review Helpfulness Determinants

### a) Review Characteristics

**Review length**: Longer reviews provide richer details and reduce consumer uncertainty (Hong et al., 2017; Mudambi & Schuff, 2010).
**Review topic diversity**: Balanced, multi-perspective information enhances helpfulness (Qazi et al., 2016; Stirling, 2007).
**Reviewer expertise**: Measured by total reviews written per reviewer (Lo & Yao, 2019).

### b) Response Characteristics

**Response speed**: Timely responses signal attentiveness and care (Li et al., 2017; Xie et al., 2017); responding within 7 days increases review helpfulness (Sparks et al., 2016).
**Response length**: Longer responses can improve usefulness through explanations and apologies (Liu & Ji, 2019), but the effect from a PMR perspective has received limited attention.
**Response topic diversity** *(this study)*: The breadth of distinct issues addressed in the response, measured by the Rao-Stirling index. Despite the established role of review topic diversity, its response-side counterpart has studied limitedly

# II. Literature Review

## 2.4. What are Standard Responses and Personalized Manager Response (PMR)?

***Standard Responses***

- Standard responses refer **to generic, template-based replies** posted by firms to customer reviews on online platforms. These responses usually contain broad expressions of apology, appreciation, or gratitude **without directly addressing the specific issues, experiences, or concerns mentioned in the review.**
- Because they are standardized and repetitive, they are often perceived as impersonal and less sincere, providing limited review-specific informational value (Wei et al., 2013; Wang & Chaudhry, 2018; Zhang et al., 2020; Wang et al., 2020; Xie et al., 2017).

***Personalized Manager Response (PMR)***

- Unlike standard responses, PMRs directly **address customer concerns, acknowledge detailed experiences, provide explanations or solutions, and demonstrate active engagement with the reviewer.**
- These responses are generally associated with stronger communication quality, higher credibility, reduced psychological distance, and greater consumer trust (Crijns et al., 2017; Sparks et al., 2016; Yang et al., 2021; Jin et al., 2023; Zhang et al., 2020).

# II. Literature Review

## 2.5. What is Semantic Alignment ?

***Semantic alignment*** refers to the degree to which a managerial response meaningfully matches the content of the original customer review — capturing meaning-level correspondence rather than surface-level keyword overlap.

- Based on Semantic Textual Similarity (STS): evaluates whether two texts express similar meaning even when different words are used (Guder et al., 2026).
- Operationalized using Sentence-BERT (SBERT) + cosine similarity (Devlin et al., 2018; Reimers & Gurevych, 2019).

Table 1. Difference of Semantic Alignment from Other Constructs

| Construct | What It Captures | Measurement / Method | Difference | Representative Study |
|---|---|---|---|---|
| **Topic Consistency** | Whether review and response discuss **similar topics** | LDA topic modeling + cosine similarity between topic distributions | Captures **topic overlap only**; cannot determine whether the response meaningfully addresses the concern | Yang et al. (2021) |
| **Topic Matching** | Degree of **shared themes/issues** between review and response | Machine-learning topic classification + topic overlap score | High overlap may reflect **repetition rather than quality or engagement** | Zhang et al. (2020) |
| **Linguistic Style Matching** | Similarity in **writing style and function words** | LIWC-based linguistic synchronization (pronouns, articles, conjunctions, etc.) | Style similarity ≠ **meaningful informational alignment** | Ren et al. (2024) |
| **Keyword / TF-IDF Similarity** | Similarity in **word frequency and lexical overlap** | TF-IDF / keyword matching | Cannot capture **synonyms, context, or implicit meaning** | Al-Anzi & AbuZeina (2017) |
| **Semantic Alignment (This Study)** | **Meaning-level contextual correspondence** between review and managerial response | **Sentence-BERT (SBERT) + cosine similarity** | Captures whether managers **genuinely address reviewer concerns beyond surface similarity** | **This Study** |

# II. Literature Review

| Table 2. Summary of Main Research Gaps | | | |
|---|---|---|---|
| **Research Gap** | **Prior Studies** | **This Study** | **Key Supporting Studies** |
| **Gap 1. Limited focus on response characteristics** | Prior studies mainly focus on **whether firms respond**, rather than **how response characteristics influence review helpfulness** | Examines **response topic diversity for the first time in terms of review helpfulness** and **response length** | Jin et al. (2023); Wei et al. (2013); Xie et al. (2017) |
| **Gap 2. Inconsistent understanding of review–response similarity** | Similarity constructs (topic consistency, topic matching, linguistic style matching) show **mixed findings** on outcomes (influence positively on rating with all types of reviews (not all types))and mostly assume **linear effects** | Introduces **semantic alignment** as a **meaning-level construct** and proposes a **U-shaped moderating role** | Yang et al. (2021); Wang & Chaudhry (2018); Zhang et al. (2020); Ren et al. (2024) |
| **Gap 3. Conceptual ambiguity of PMR** | PMR is often measured inconsistently or conflated with **topic similarity/matching** | Separates **PMR (binary personalization)** from **semantic alignment (continuous contextual similarity)** | Zhang et al. (2020); Jin et al. (2023); Wei et al. (2013) |
| **Gap 4. Limited understanding of boundary conditions** | Prior studies mainly examine **direct effects**, with limited explanation of **when response characteristics become effective** | Examines **semantic alignment and PMR as higher-order moderators through three-way interactions** | Guo & Zhou (2016); Jin et al. (2023) |

# III. Research Overview

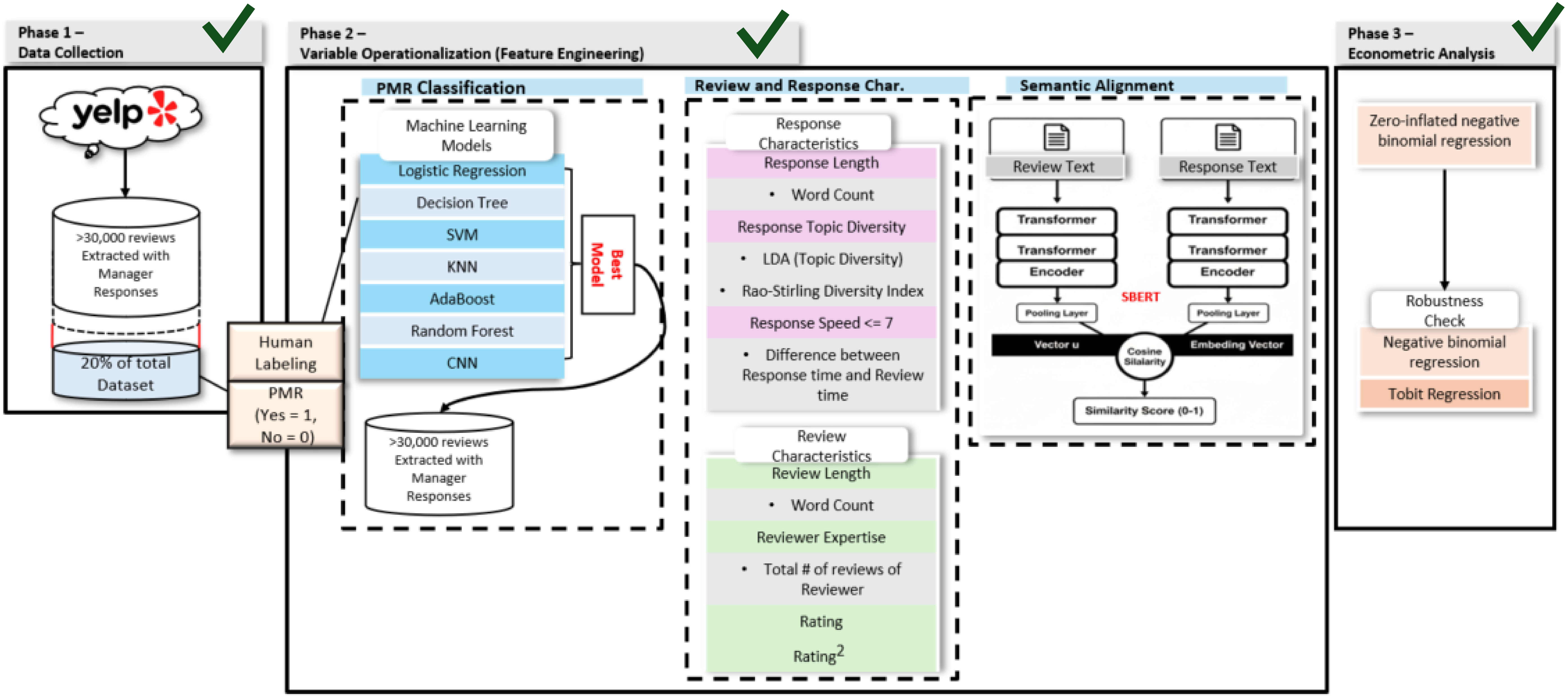


Figure 1. Research Framework

# IV. Hypotheses Development

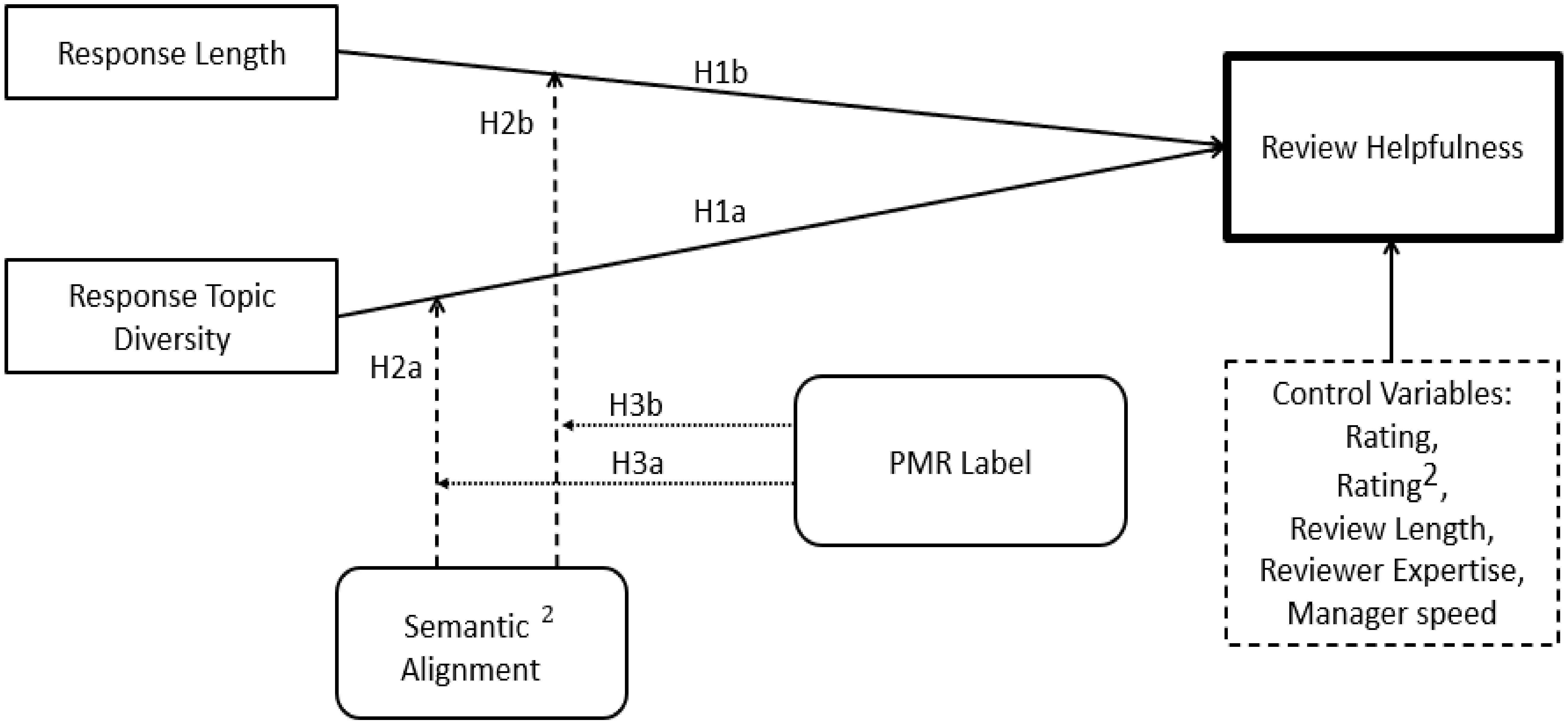


**Figure 2. Conceptual Framework**

# IV. Hypotheses Development

## 4.1. Response Characteristics

**Response Topic Diversity**

**Prior findings**

- Higher review topic diversity enhances diagnosticity and helps consumers evaluate products from multiple perspectives (Stirling, 2007; Qazi et al., 2016).
- Broader information improves **diagnostic value**

**Expected relationship**

- Since review topic diversity usually influences positively on review helpfulness, it is expected that Greater Response topic diversity reflects **more comprehensive issue coverage** and will have a positive effect on review helpfulness
- More diverse responses may provide **greater informational richness**

***H1(a): Response topic diversity positively affects review helpfulness.***

**Response Length**

**Prior findings**

- Response length is positively associated with review helpfulness and firm performance (Liu & Ji, 2019; Xie et al., 2017). However, excessively long or generic responses may reduce effectiveness if they do not add meaningful information.

***H1(b): Response length positively affects review helpfulness.***

# IV. Hypotheses Development

## 4.2. Why not a linear relationship?

**Positive side of alignment**

- High topic matching improves evaluations (**Zhang et al., 2020**)
- High stylistic similarity improves future ratings (**Ren et al., 2024**)

**Negative side of alignment**

- High topic consistency weakens the positive effect of **review length** (**Yang et al., 2021**)
- Similarity may create **redundancy**, especially when responses simply repeat review content(**Wang & Chaudhry, 2018**)

**Expected U-shaped moderation**

- **Moderate alignment** => repetitive / limited diagnostic value
- **Very high alignment** => genuine engagement and relevance

***H2(a): Semantic alignment (U-shape) weakens the positive effect of response topic diversity on review helpfulness.***
***H2(b): Semantic alignment (U-shape) weakens the positive effect of response length on review helpfulness.***

# IV. Hypotheses Development

## 4.3. Why PMR matters?

**Prior findings**

- Personalized managerial responses increase helpfulness, especially under ambiguity (**Jin et al., 2023**)
- Specific responses are perceived as **more sincere and diagnostic** than generic responses (**Wei et al., 2013**)

**Expected three-way interaction**

When semantic alignment is high:

**Without PMR**

=> Response may appear repetitive or generic

**With PMR**

=> Response appears sincere, tailored, and genuinely engaged

**Expected effect**

PMR weakens the negative moderating effect of semantic alignment

***H3(a): The negative moderating effect of semantic alignment (U-shape) on the relationship between response topic diversity and review helpfulness weakens when the response is personalized.***

***H3(b): The negative moderating effect of semantic alignment (U-shape) on the relationship between response length and review helpfulness weakens when the response is personalized.***

# V. Methodology

## 5.1. Phase 1 - Data Collection

- 117,416 reviews were collected from Yelp.com platform, which includes 141 restaurants from destinations such as San-Francisco, Hawaii, Washington, Las Vegas and New York. **However, all reviews did not have manager responses.**
- **30,692 reviews** were selected manually which include manager responses. Thus, there are 25 restaurants with solely manager responses.

Previously, TripAdvisor platform was used for manager response studies (Jin et al., 2023; Yang et al., 2021; Zhang et al., 2021; )

However, Yelp.com was also utilized as a trustworthy platform for review helpfulness studies

Table 3. Usage of Yelp.com in Prior Literature

| Study | Journal | Purpose of the Study |
|---|---|---|
| Zheng et al. (2023) | International Journal of Hospitality Management | Develop and test a multimodal deep learning framework to predict the number of helpfulness votes using text, images, metadata, etc. |
| Zhou et al. (2017) | Decision Support Systems | Examine the order effect on review helpfulness from a social influence perspective. |
| Xu et al. (2023) | Computers in Human Behavior | Analyze how negative emotions (anger, anxiety) in reviews affect helpfulness votes, moderated by product/business price. |
| Luo & Xu (2019) | Sustainability | Predict review helpfulness using machine learning algorithms (e.g., SVM with ontology) by extracting aspects and sentiments. |
| Ngo-Ye et al. (2014) | Decision Support Systems | Investigate the influence of reviewer engagement characteristics (RFM) on review helpfulness using text regression. |
| Kim et al. (2025) | Journal of Retailing and Consumer Services (or similar) | Explore unexpected consequences of helpful reviews and their impact on subsequent review behavior. |

# V. Methodology

## 5.2. Phase 2 – Variable Operationalization (Feature Engineering)

Table 4. Operational Definitions of Variables

| Category | Variable | Description |
|---|---|---|
| Dependent variable | Review Helpfulness | Total number of helpfulness votes. |
| Moderating Variable | Semantic Alignment between response and review text | Semantic similarity between the textual content of a review and its corresponding managerial response by using SBERT and cosine similarity |
| | Personalized manager response (PMR) | The type of manager response |
| Control Variable | Review Rating | The rating in Yelp from one to five |
| | Review Length (Huang et al, 2015) | The word count of a review |
| | Reviewer expertise (Liu et al., 2021) | The number of written reviews of each reviewer |
| | Manager Response Speed | The speed of the manager responds to a review, which is measured by the timeliness between the response and the review. |
| Independent Variable | Response length | The word count of a response |
| | Response topic diversity | Topic diversity measured by Rao-Stirling Diversity |

$$Semantic\ Alignment = \frac{R_i \times MR_i}{||R_i|| \times ||\text{MR\_i}||}$$

Nils Reimers and Iryna Gurevych (2019)

$$Rao - Stirling = \sum_{ij}^{k} p_i p_j \, d_{ij}$$

- $p_i$ - the proportion of the topic i
- $p_j$ - the proportion of the topic j
- $d_{ij}$ - cosine distance between topic i and topic j

(Zielinski et al., 2022; Nisar et al., 2020)

# V. Methodology

## 5.2. Phase 2 – Variable Operationalization (Feature Engineering)

## Binary Classification of Manager Response, creating a Moderating Variable

## (PMR = 1, Standard Response = 0)

**By adapting the previous study by Jin et al., 2023 for constructing PMR as 1 value and SR as 0**

This study also uses a **supervised machine learning approach** to classify Personalized Managerial Responses (PMRs).

Following prior research, **20% of the dataset** is manually labeled as:

- **PMR** = tailored, review-specific, issue-addressing responses
- **Standard Response** = generic or template-based replies

**Why Manual Labeling?**

PMR is a **qualitative construct** that cannot be directly observed from raw text.

Human labeling is required to create reliable training data.

**Classification Models**

Seven classifiers are compared:

- Logistic Regression, Decision Tree, K-Nearest Neighbors (KNN) , AdaBoost, Random Forest, Support Vector Machine (SVM) , Convolutional Neural Network (CNN)

**Evaluation Metrics**

(1)

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TN}{TP+FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

# V. Methodology

## 5.2. Phase 2 – Variable Operationalization (Feature Engineering)

### Binary Classification of Manager Response, creating a Moderating Variable (PMR = 1, Standard Response = 0)

Dataset was extracted randomly, considering 6142 rows as the training set.
PMR was labeled manually as a binary construct (1= PMR, 0= Standard Response).
Manual PMR distribution:
1 = 3696 and 0 = 2446.

Based on that, the prediction of PMR was applied
**CNN was selected as the best model for classifying PMR**

Following by> **18792 – PMR, 11,900 – Standard Response**

Table 5. Results of Classifiers

| Metric | Logistic Regression | Decision Tree | KNN | AdaBoost | Random Forest | SVM | CNN |
|---|---|---|---|---|---|---|---|
| Precision | 91.0% | 89.0% | 85.0% | 88.0% | 89.0% | 92.0% | **93.0%** |
| Recall | 95.0% | 87.0% | 91.0% | 90.0% | 94.0% | 93.0% | **95.0%** |
| F1-score | 93.0% | 87.7% | 88.1% | 89.3% | 91.3% | 92.6% | **94.2%** |
| Accuracy | 91.4% | 85.3% | 85.2% | 87.0% | 89.3% | 91.0% | **92.9%** |

# V. Methodology

## 5.2. Phase 2 – Variable Operationalization (Feature Engineering)

"Response speed" variable, where manager speed is no longer than 7 days.
Hence, **the final dataset has been established**

- **28554 rows**
- **PMR = 18015**
- **Standard Response = 10539.**

| Table 6. Statistical Description of the Variables | | | | | |
|---|---|---|---|---|---|
| Variable | Description | Mean | Std.dev | Min | Max |
| Review Helpfulness | Total number of helpfulness votes. | 0.533 | 2.810 | 0 | 117 |
| Personalized manager response (PMR) | The type of manager response | 0.631 | 0.483 | 0 | 1 |
| Review Rating | The rating in Yelp from one to five | 4.246 | 1.188 | 1 | 5 |
| Review Length (Huang et al, 2015) | The word count of a review | 82.396 | 83.677 | 6 | 952 |
| Reviewer expertise (Liu et al., 2021) | The number of written reviews of each reviewer | 110.895 | 289.342 | 1 | 9784 |
| Manager Response Speed | The speed of the manager responds to a review, which is measured by the timeliness between the response and the review. | 1.167 | 1.627 | 0 | 7 |
| Semantic Alignment between response and review text | Semantic similarity between the textual content of a review and its corresponding managerial response by using SBERT and cosine similarity | 0.452 | 0.175 | - 0.098*** | 0.917 |
| Semantic Alignment$^2$ between response and review text | Semantic similarity in square in order to see U-shape relationship | 0.234 | 0.155 | $3.13 \times 10^{-8}$ | 0.842 |
| Response length | The word count of a response | 50.308 | 32.413 | 1 | 417 |
| Response topic diversity | Topic diversity measured by Rao-Stirling Diversity | 0.256 | 0.159 | 0.021*** | 0.667 |
| Note: ***p < 0.01, **p<0.05, *p<0.1 | | | | | |

# V. Methodology

## 5.3. Phase 3 - Main Model: Zero-Inflated Negative Binomial (ZINB)

Table shows that dependent variable (**review helpful votes**) contains a **large number of zero values**.

Therefore, **ZINB regression** is appropriate because it handles:

- Excess zeros
- Overdispersed count data

Mafael (2019); Zhou & Guo (2017); Zhu et al. (2014)

| Table 7. Distribution of review helpfulness level | | |
|---|---|---|
| Reviews helpfulness level | Reviews number | Percentage |
| 0 | 21650 | 75.82% |
| 1 | 4550 | 15.93% |
| 2 | 1230 | 4.31% |
| 3 | 447 | 1.57% |
| 4 | 217 | 0.76% |
| 5 | 125 | 0.44% |
| 6 | 83 | 0.29% |
| >=7 | 252 | 0.88% |
| Note: ***p < 0.01, **p<0.05, *p<0.1 | | |

### Robustness Checks

To verify the reliability of results, additional models are applied:

**Negative Binomial Regression**

- Controls for **overdispersion** in count data
- Frequently used for helpfulness vote analysis

**Tobit Regression**

- Suitable when the dependent variable is **left-censored at zero**
- Used in prior review helpfulness studies

(Guo & Zhou, 2016).

$$\textit{Review Helpfulness}=\exp[\beta_0 + \beta_1 \textit{(Review Rating)} + \beta_2 \textit{(Review Rating2)} + \beta_3 \textit{(Review Length)} + \beta_4 \textit{(Reviewer Expertise)} + \beta_5 \textit{(Response Speed)} + \beta_6 \textit{(Response Length)} + \beta_7 \textit{(Response Topic Diversity)} + \beta_8 \textit{(Semantic Alignment)} + \beta_9 \textit{(Semantic Alignment2)} + \beta_{10}(\textit{Semantic Alignment}^2 \times \textit{Response Length}) + \beta_{11}(\textit{Semantic Alignment}^2 \times \textit{Response Topic Diversity}) + \beta_{12}(PMR) + \beta_{13}(\textit{Semantic Alignment}^2 \times \textit{Response Length} \times PMR) + \beta_{14}(\textit{Semantic Alignment}^2 \times \textit{Response Topic Diversity} \times PMR) + \varepsilon]$$

# VI. Results

## 6.1. Results of ZINB

- **H1a: Response Topic Diversity**
  Positive and significant (**β =** 0.535 **, p < 0.05**) - **Supported**
- **H1b: Response Length**
  Positive and significant (**β = 0.003, p < 0.1**) - **Supported**
- **H2a: Semantic Alignment$^2$ × Response Topic Diversity**
  Negative and significant (**β =** -3.054**, p < 0.05**) - **Supported**

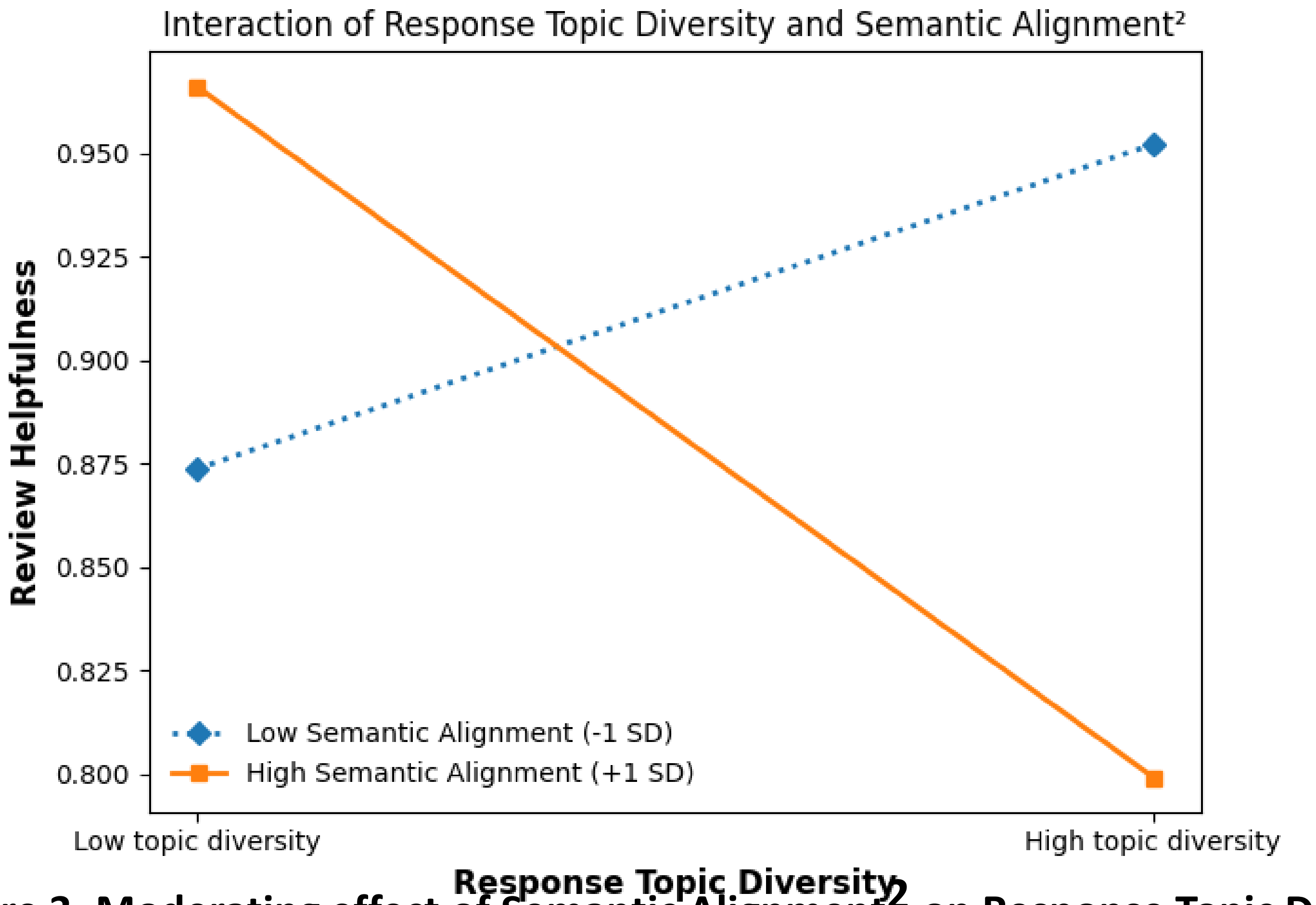


**Figure 3. Moderating effect of Semantic Alignment$^2$ on Response Topic Divesity**

Table 8. Results of ZINB

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Review Rating | -0.483 *** | -0.486*** | -0.482*** | -0.478*** | -0.477*** | -0.486*** |
| Review Rating$^2$ | 0.074*** | 0.075*** | 0.075*** | 0.074 *** | 0.073*** | 0.074*** |
| Review length | 0.652*** | 0.648*** | 0.649 | 0.650*** | 0.651*** | 0.650*** |
| Reviewer expertise | 0.471*** | 0.471*** | 0.471*** | 0.471 *** | 0.4701*** | 0.469*** |
| Response Speed | 0.017* | 0.017* | 0.017* | 0.018* | 0.016* | 0.015* |
| Response length | | 0.000 | 0.001 | 0.003** | 0.004** | 0.003* |
| Response Topic Diversity | | 0.210** | 0.208* | 0.322* | 0.382* | 0.535 ** |
| Semantic Alignment | | | -0.8620* | -1.304*** | -1.365*** | -1.120 |
| Semantic Alignment$^2$ | | | 0.838* | 1.788** | 2.061*** | 1.875** |
| Semantic Alignment$^2$ × Response Topic Diversity | | | | -0.380 | -0.480 | -3.054** |
| Semantic Alignment$^2$ × Response Length | | | | -0.008* | -0.009** | -0.001 |
| PMR Label | | | | | -0.0737 * | - 0.111 |
| Semantic Alignment$^2$ × Response Topic Diversity × PMR | | | | | | 2.351** |
| Semantic Alignment$^2$ × Response Length × PMR | | | | | | -0.007 |
| N | 28554 | 28554 | 28554 | 28554 | 28554 | 28554 |
| Log likelihood | -21768.88 | -21765.53 | -21762.08 | -21758.98 | -21757.01 | -21752.60 |
| Pseudo R2 | 0.142 | 0.142 | 0.142 | 0.143 | 0.143 | 0.143 |
| P value chi2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

Note: ***p < 0.01, **p<0.05, *p<0.1

# VI. Results

## 6.1. Results of ZINB

- **H2b: Semantic Alignment$^2$ × Response Length**
  Negative and Insignificant (**β =** -0.001, **p < 0.05**)
  BUT Model 4 and Model 5 (-0.008, 0.009, p < 0.1, p<0.05)

**- Not Supported**

- **H3a: Semantic Alignment$^2$ × Response Topic Diversity × PMR** (**β =** 2.351, **p < 0.05**)
  Positive and Significant **-Supported**
- **H3b: Semantic Alignment$^2$ × Response Length × PMR**
  **Negative and significant (β = -0.007)** - **Not Supported**

| Table 8. Results of ZINB | | | | | | |
|---|---|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| Review Rating | -0.483 *** | -0.486*** | -0.482*** | -0.478*** | -0.477*** | -0.486*** |
| Review Rating$^2$ | 0.074*** | 0.075*** | 0.075*** | 0.074 *** | 0.073*** | 0.074*** |
| Review length | 0.652*** | 0.648*** | 0.649 | 0.650*** | 0.651*** | 0.650*** |
| Reviewer expertise | 0.471*** | 0.471*** | 0.471*** | 0.471 *** | 0.4701*** | 0.469*** |
| Response Speed | 0.017* | 0.017* | 0.017* | 0.018* | 0.016* | 0.015* |
| Response length | | 0.000 | 0.001 | 0.003** | 0.004** | 0.003* |
| Response Topic Diversity | | 0.210** | 0.208* | 0.322* | 0.382* | 0.535 ** |
| Semantic Alignment | | | -0.8620* | -1.304*** | -1.365*** | -1.120 |
| Semantic Alignment$^2$ | | | 0.838* | 1.788** | 2.061*** | 1.875** |
| Semantic Alignment$^2$ × Response Topic Diversity | | | | -0.380 | -0.480 | -3.054** |
| Semantic Alignment$^2$ × Response Length | | | | -0.008* | -0.009** | -0.001 |
| PMR Label | | | | | -0.0737 * | - 0.111 |
| Semantic Alignment$^2$ × Response Topic Diversity × PMR | | | | | | 2.351** |
| Semantic Alignment$^2$ × Response Length × PMR | | | | | | -0.007 |
| N | 28554 | 28554 | 28554 | 28554 | 28554 | 28554 |
| Log likelihood | -21768.88 | -21765.53 | -21762.08 | -21758.98 | -21757.01 | -21752.60 |
| Pseudo R2 | 0.142 | 0.142 | 0.142 | 0.143 | 0.143 | 0.143 |
| P value chi2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Note: ***p < 0.01, **p<0.05, *p<0.1 | | | | | | |

# VI. Results

## 6.2. Robustness Check

- **Response topic diversity remained positively significant** (**β = 0.557–1.741**), confirming its positive effect on review helpfulness.
- **Semantic alignment maintained a stable U-shaped effect**, showing a **negative linear** and **positive quadratic relationship**.
- **SA² × Response Topic Diversity remained significantly negative** (**β = −3.188 to −9.035**), indicating that high alignment weakens the benefit of excessive topic breadth.
- **PMR moderation remained positive** (**β = 2.359–5.816**), suggesting that personalized responses improve effectiveness under high alignment.
- **Response length effects were less stable**, as interaction terms became insignificant in robustness models.

Table 9. Robustness Check

| | Model 7 Negative Binomial Regression | Model 8 Tobit Regression |
|---|---|---|
| Review Rating | -0.460*** | -1.757*** |
| Review Rating$^2$ | 0.071*** | 0.276*** |
| Review length | 0.656*** | 2.079*** |
| Reviewer expertise | 0.482*** | 1.396*** |
| Response Speed | 0.016** | 0.079* |
| Response length | 0.003*** | 0.009* |
| Response Topic Diversity | 0.557*** | 1.741** |
| Semantic Alignment | -1.223*** | -2.973* |
| Semantic Alignment$^2$ | 2.046*** | 4.822** |
| Semantic Alignment$^2$ × Response Topic Diversity | -3.188*** | -9.035** |
| Semantic Alignment$^2$ × Response Length | $-9.345\times 10^{-6}$ | -0.003 |
| PMR Label | - 0.111** | -0.270 |
| Semantic Alignment$^2$ × Response Topic Diversity × PMR Label | 2.359*** | 5.816* |
| Semantic Alignment$^2$ × Response Length × PMR Label | -0.009* | -0.021 |
| N | 28554 | 28554 |
| Log likelihood | -21901 | -28946.88 |
| Pseudo R2 | 0.361 | 0.149 |
| P value chi2 | 0.000 | 0.000 |

# VII. Results

## 7.1. Summary of Findings

**1. Response characteristics matter**
- **Response topic diversity** positively affects review helpfulness (**Response length** positively affects review helpfulness, although the effect is relatively weak

**2. Semantic alignment operates non-linearly**
- Semantic alignment showed a **U-shaped effect**
- Moderate alignment may reduce helpfulness due to **redundancy**
- Very high alignment restores helpfulness through **genuine engagement and contextual relevance**

**3. High semantic alignment weakens response characteristics**
- High semantic alignment weakens the positive effect of **response topic diversity** (**supported**)
- The moderating effect on **response length** was not fully supported in the final model, although earlier models showed significance

**4. PMR strengthens response effectiveness**
- PMR strengthened the relationship between **semantic alignment and response topic diversity**
- However, PMR did not significantly strengthen the effect involving **response length**

# VII. Results

## 7.2. Academic Contributions

**1. Introduces Semantic Alignment as a New Meaning-Level Construct**

- Extends prior similarity constructs (**topic consistency, topic matching, linguistic style matching**) by measuring **contextual meaning-level correspondence**
- Uses **Sentence-BERT + cosine similarity** to operationalize semantic alignment

**2. Explains Mixed Similarity Findings through a U-shaped Perspective**

- Resolves inconsistent findings in prior literature by showing that similarity is **not uniformly beneficial**
- Demonstrates that semantic alignment may become **redundant at moderate levels but valuable at very high levels**

**3. Separates PMR from Semantic Alignment**

- Distinguishes **PMR (binary personalization)** from **semantic alignment (continuous similarity construct)**
- Clarifies their **distinct yet complementary roles**

**4. Extends Review Helpfulness Literature**

- Introduces **response topic diversity** as an underexplored managerial response characteristic
- Examines **three-way interactions** to explain boundary conditions of response effectiveness

# VII. Results

## 7.3. Practical Contributions

**1. Firms should prioritize meaningful alignment, not simple repetition**

Responses should meaningfully address customer concerns rather than merely repeat review content

**2. More information is not always better**

High topic diversity may become ineffective when responses are already highly aligned

Managers should balance **breadth and relevance**

**3. Personalized responses improve effectiveness**

PMR increases perceived sincerity and usefulness, especially when responses contain diverse information

**4. AI-assisted customer service systems**

Firms may use **semantic alignment measures** to evaluate response quality and identify generic or redundant responses

**5. Prediction of Personalized Manager Response may help reader understand how this company will reply to further review text, and it can help in forecasting the reputation of a company**

- This can improve customer service consistency, monitor staff performance, support training, and strengthen online reputation.
- Readers may infer future service quality from response style, influencing trust and ratings.

# VII. Results

## 7.4. Future Research

### Future Directions

**1. Extend to other platforms and industries**

Future studies may examine **hotels, e-commerce, tourism, or healthcare services**

**2. Explore additional response characteristics**

Future research may investigate **response sentiment, emotional tone, response depth, or politeness**

**3. Examine different forms of non-linearity**

Future studies may explore whether semantic alignment follows **threshold effects or inverted U-shaped patterns** in different contexts

**4. Investigate AI-generated managerial responses**

Future research may compare **human vs. AI-generated responses** and their influence on review helpfulness

# THANK YOU