Understanding the Relationship between Response Standardization and Sentiment in Customer-Company Conversations on Twitter

Group 11
Alessandro Bellini 3169373
Giulio Caputi 3153584
Giacomo Cirò 3160499
Tancredi Dorsi 3161375
Alessandro Morosini 3149076
Francesco Vacca 3140929

Abstract

In today's social media-dominated era, effective customer support plays a pivotal role in shaping a positive brand image and fostering customer loyalty. The primary objective of this research is to study the complex relationship between standardized responses and feedback sentiment in customer-company interactions on Twitter.

Our investigation focuses on how standardization in responses influences customer satisfaction and sentiment outcomes while considering different factors which can mitigate or magnify this effect. Our analytical approach includes OLS regression as a baseline model, further extended to account for confounding variables, moderator and mediator effects, industry-specific trends and language patterns in responses. This comprehensive analysis provides valuable marketing insights which can be used to shape efficient business strategies.

Contents

1 Introduction

In an era dominated by social media, effective customer support has become fundamental for businesses seeking to maintain a positive brand image and foster customer loyalty. The variable we use to measure the effectiveness of customer support is the feedback sentiment measured using RoBERTa (NLP model by Facebook AI) to assign a score to the last tweet posted by the user in the conversation. Satisfaction is in fact among the most critical determinants of a company's overall success, influencing everything from customer retention and brand reputation, to product development and market competitiveness.

In economics, Business Process Standardization (BPS) refers to creating company-wide standard procedures to improve operational performance, facilitate communication, reduce errors and costs. BPS is also applied to customer interactions, leading to faster and more effective responses. However, it can also be detrimental: answering a complex complaint with a vague and predetermined response can lead to a dissatisfied customer. Moreover, in the era of AI, the degree of standardization is even a more relevant topic. Indeed, it is increasingly more common to have interactions with conversational agents, and many researchers spend time in questioning to what extent chat-bots should be human-like and give personalised answers [6].

The study we are about to illustrate delves into the complex link between response standardization and sentiment within Twitter interactions between customers and companies' support teams, investigating how standardized responses influence customer satisfaction and how many other parameters and variables can tweak and modify the magnitude and direction of this effect.

We start from a dataset in which each row represents a distinct conversation between a customer and a firm, initiated when a user mentions the firm on Twitter and followed by the firm's response. We then apply data science methods to extract valuable marketing insights. First, we run a basic OLS regression between 'feedback sentiment' and 'standardization' to be used as a base model to compare with further analysis. Then, we expand the OLS model taking into account confounds, i.e., variables correlated with both the dependent and independent variable, that could absorb some part of the variation in feedback sentiment that was initially attributed to standardization alone. Afterwards, we account for moderators, namely those variables that can influence and potentially change the relationship between the independent and the dependent variable under different conditions or levels, finding out that the impact of standardized responses widely differs based on the degree of uniqueness of the message and its sentiment. Successively, we introduce possible mediators. In particular, we develop a theory on how standardization interacts with the sentiment influencing feedback through the mediation of time. We employed the causal steps approach proposed by Baron and Kenny to test this mediation theory. Then, we elaborate on the fact that we are dealing with data containing firms from various industries and we have not yet considered individual industry effects which might be responsible of latent heterogeneity. To solve this issue, we introduce a fixed effects model, finding out that there is a downward sloping trend shared among the majority of industries. For the sake of completeness, we run an Hausman test and reject the need for random effects estimator. Eventually, since we have the partial text of the first tweet posted by the company available and haven't analysed it yet, we decide to use it to understand whether or not particular ways of phrasing the initial message have different impact on user feedback. In particular, we study whether using the first name of the customer when addressing an inquiry is positive for the final feedback, and shall thus be a recommended practice to support teams.

2 Theoretical Framework

2.1 Feedback Sentiment and Standardization

The importance of the customer's sentiment measurement at the conclusion of their interaction with the company cannot be understated. It ranks among the most crucial factors affecting a company's overall success, exerting its influence across various aspects, from customer retention to brand reputation, from product development to market competitiveness. In our dataset, this variable is measured using "A Robustly Optimized BERT Pretraining Approach" a.k.a. RoBERTa, a highly advanced natural language processing model developed by Facebook AI, with exceptional language understanding capabilities. It measures and predicts contextual relationships among words in text, enabling tasks like sentiment analysis, language translation, and text summarization with remarkable accuracy. The variable takes values in the range $[-1, 1]$, from the most negative (-1) to the most positive (1) sentiment. What we seek to understand and quantify in this research is the link and influence that the variable "standardization" has on "feedback_sentiment".

The reason why this variable is of particular interest is that for a firm willing to optimize its online presence, is essential to understand how customer support procedures influence customer sentiment, and consequently brand perception and customer loyalty. It also helps identify strategies for enhancing operational efficiency and competitive advantage while informing data-driven decision-making in a customer-centric approach.

To state the importance of standardization in business, let's think for example at the assembly line, a pivotal invention of the Industrial Revolution that exponentially developed the automobile industry starting in the early 1900's with its impact still present today, whose underlying core idea is to standardize and automate a series of actions and practices. Moreover, BPS in general is an acknowledged source of corporate performance which companies must take into consideration in many activities, from manufacturing to supply chain management, and also when dealing with customer support (e.g. canned responses, FAQs etc.).

However, an optimal balance shall be found. In fact, adhering strictly to standardized responses in customer support may sometimes prove counterproductive: dealing with complex inquiries using generic and predetermined replies is likely to escalate the customer's frustration and may not contribute positively to customer retention.

Therefore, it is important for companies to understand how standardization in customer support influences customers satisfaction, enabling them to choose the best practices and maximize feedback sentiment.

2.2 Expectations

Overall, when dealing with customer relations, we expect standardization to have a negative effect on feedback sentiment, for a series of reasons. First of all, highly standardized responses may come across as impersonal and robotic, potentially leading to negative sentiment. Secondly, excessive standardization can limit the flexibility of support agents to address unique or complex issues, causing frustration and negative sentiment. Thirdly, encountering repetitive standardized responses may give customers the impression that their individual concerns are not valued, contributing to adverse feedback. And most importantly, standardized responses may lack the emotional connection and empathy often conveyed in personalized interactions, potentially leading to negative sentiment.

This is in line with theories from several published reasearch papers, according to which personalization has a positive effect on customer satisfaction [4], even though overpersonalization may have downsides[5].

2.3 Confounds

Beyond measuring the direct association between standardization and feedback sentiment with the OLS, we are going to add some confounds which are correlated with both the dependent and independent variable, and that could absorb some part of the variation in feedback sentiment that was attributed to standardization. The ones we deemed most relevant are:

- **Sentiment Score Company Mean**: the overall sentiment of the company's tweets could positively influence customer feedback sentiment. If companies generally have positive tones, this might require a higher degree of personalization, i.e., less standardization.
- **Response Time Company Mean**: faster response times may be associated with higher customer satisfaction, i.e., higher feedback. Companies with standardized responses might have quicker response times.
- **Word Count Company Mean**: the length of the company's messages may be positively correlated with the degree of standardization, assuming standardized responses tend to be longer and negatively affect feedback, because it could be that lengthy or verbose responses may be perceived as less engaging, or customers might prefer concise and to-the-point interactions.

2.4 Moderators

We now consider the hypothesis that some moderators could affect the effect of standardization on the dependent variable, such as:

- **Message Uniqueness**: in general, it is intuitive to think that a more unique request needs a more personalised and therefore less standardized response. Thus, a standardized response to a unique message will have a worst impact on feedback than if the message was not unique, which could have been addressed with a pre-drafted response already proven to be successful. In a nutshell, as message uniqueness increases we expect more standardized responses to have a worse effect on feedback.
- **Message Sentiment**: we expect standardization to have different impact on final feedback depending on the initial sentiment of the customer. Specifically, we hypothesise that the effect of message sentiment is mediated by the time needed for the company to respond:
  - A happy customer with a high message sentiment score probably does not need help. Instead, said customer is looking for a human acknowledgment of his comment. Then, it is reasonable to conclude high standardization will have a negative impact on feedback. We deem this type of satisfied customer more standardization than time sensitive.
  - An upset customer on the other hand, has the necessity of a timely response capable of solving the issue or answering the complaint. Non standardized responses usually require more time, thus we conclude that an untimely non standardized response has a worse effect compared to a timely standardized response. We deem this type of satisfied customer more time than standardization sensitive.

2.5 Fixed Effects

OLS models have a fundamental issue with the type of data we have at hands, which exhibits a cross dimensional structure where tweets of multiple firms across different industries are captured. Given the low number of observations per firm, we decide to exploit the panel data structure to account for possible latent heterogeneity among different industries. There could be various sources that determine industry-specific latent heterogeneity, for example:

- *Customer demographics*, different industries may attract distinct customer demographics, which could influence the relationship between standardization and feedback sentiment.
- *Technological maturity*, industries may differ in their technological maturity and adoption rates and potentially influence customer reaction to the firm behaviour on twitter.
- *Social media engagement*, industries may have different approaches to social media engagement. Younger industries make of social engagement their strength, while older industries care less.
- *Product* or *service characteristics*, each industry might offer products or services with unique characteristics, some of which require more personalized responses.

This kind of model can be developed introducing a dummy variable for each industry, hence the name LSDV (least squares dummy variables) for such type of model. Alternatively, a regression on the variables in group mean deviations can be implied. The LSDV formula is thus defined as follows:

$$Y_{it} = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \ldots + \beta_k D_k + X_{it}'\gamma + u_{it}$$

Where D_i is the dummy variable related to a specific industry. In this way, we make the module robust to individual specific effects and we enable the analysis of unique group dynamics.

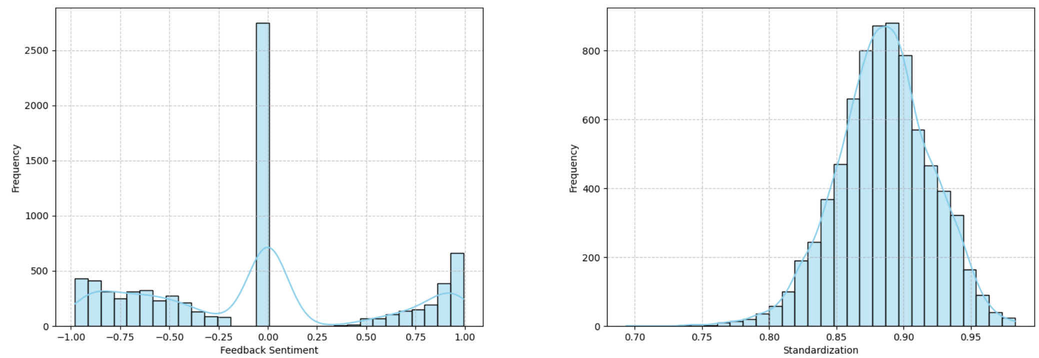
2.6 Addressing Customers using their Names

Eventually, we decide to exploit the partial text we have available of the first tweet posted by the company to analyze word patterns and how different ways of phrasing a message can lead to different feedback sentiment from the user.

In particular, we theorize that addressing customers using their first name gives a human-like tone to the conversation, channeling the subsequent exchange of messages in a more intimate and colloquial fashion, creating a temporary bond with the user which eventually leads to better feedback.

3 Data Overview

The data contains 7,975 observations, pertaining to 141 distinct companies across 26 distinct industries. First of all, we notice that there are 376 duplicates, and we decide to drop them, as they could possibly bias our findings. The dependent variable, "feedback_sentiment" and all the other variables measuring sentiment range from -1 to 1. As can be evicted by the graph (See Figure 1(a)), the majority of comments have neutral sentiment, while those not neutral tend to be polarized at the edges of the range, being either very positive or very negative. The independent variable "standardization" ranges from 0 to 1. From the graph (Figure 1(b)), we can see that standardization is normally distributed, with a mean score of 0.88. All the variables measuring anger or joy range from 0 to 1. Some other features such as offensive_company_count, with almost zero variability across the dataset, have been dropped from start, while some categorical columns such as MarketCap have been encoded with an ordinal encoder, so that the ordinal nature of the data was not lost. Finally, as good practice requires we centered all variables we plan to use as moderators and our main independent variable 'standardization', so to improve interpretability and reduce multicollinearity, enhancing the stability and reliability of the regression model.



(a) Distribution of 'feedback_sentiment' (b) Distribution of 'standardization'

Figure 1

4 Models and Analysis

4.1 Base Model

We use the centered variables to run a basic OLS regression and use it as the base model to compare with our further analysis. The coefficient of standardization is negative, which can also be inferred from the graph (Figure 2) as the fitting line slopes downward. Specifically, an increase in one unit of standardization is associated to a 0.75 decrease in feedback sentiment.The regressor is significant at any conventional level of significance, as the p-value is virtually zero.

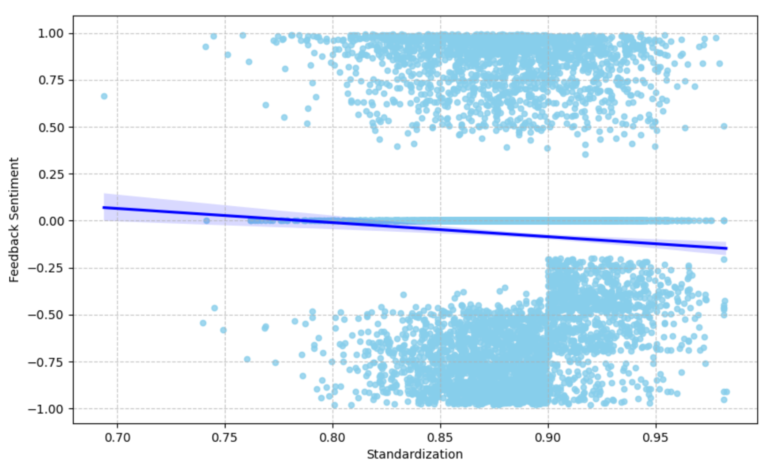


Figure 2: Base Model - OLS Regresion

	(1)
VARIABLES	c_feedback_sentiment
c_standardization	**-0.751*****
	(0.197)
Constant	2.44e-08
	(0.00694)
Observations	7,599
R-squared	0.002

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The analysis cannot be considered conclusive. First, we are assuming that the OLS assumptions hold. Second, we cannot infer causality. Third, we are not exploiting the panel data structure of our dataset. Moreover, we are not accounting for potential confounds or mediators in the relationship between independent and dependent variable. We proceed to expand the model.

4.2 Confounds

As previously introduced, we are now going to expand our model with some confounds that we expect to absorb part of the effect.

The estimated coefficient of standardization is still negative, but has notably decreased in absolute value, evidencing the fact that the confounds have absorbed some part of the variation in feedback sentiment previously attributed to standardization. Some of the regressors, such as the average response time per company or message uniqueness, are statistically insignificant. However, the overall model is significant as evidenced by the F-test of joint significance. Moreover, the R-squared is remarkably higher with respect to our base OLS model, meaning that the model explains better the variation in feedback sentiment.

	(B-path)
VARIABLES	**c_feedback_sentiment**
c_standardization	**-0.374***
	(0.197)
word_count_company_mean	-0.00446***
	(0.000579)
response_time_company_mean_s	**0.0149**
	(0.00980)
sentiment_score_company_mean	0.294***
	(0.0167)
Constant	0.0759***
	(0.0213)
Observations	7,599
R-squared	0.050

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.3 Moderators

Once we account for possible moderators too, the overall regression is still significant. Some considerations shall be done on the effect of standardization. In fact, the derivative with respect to standardization is now as follows:

$$0.93 - 35.37 \cdot message\_uniqueness - 2.63 \cdot message\_sentiment$$

In order to report our considerations, we distinguish different scenarios:

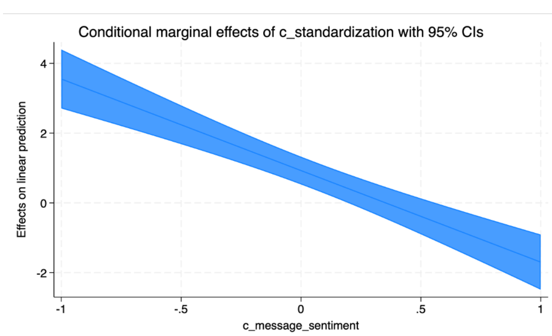
- Uniqueness and Sentiment at their mean: then the effect of standardization is positive, suggesting that when a message is not particularly unique nor emotionally polarized, a standardized answer is actually positive, probably because comes faster and straight to the point compared to a more elaborated and articulated one.
- Uniqueness < 0: then the effect is positive suggesting that when a message has already been received multiple time and thus is not unique, the standardized response provided by the company is useful and appreciated by the user
- Uniqueness > 0: then the effect is negative, which proves that the more unique a message is the more the company shall answer in a personalized manner
- Sentiment < 0: then the effect is positive, which hints at the fact that an animated customer prefers standardized answers. This can be explained by the fact that the user is seeking for a timely feedback on his complaint, thus preferring a standardized answer, which is provided faster. We test this theory of the mediation of time using the causal steps approach proposed by Baron and Kenny.
- Sentiment > 0: then the effect is negative. This is probably due to the fact that a customer who leaves a positive comment is actually seeking for an acknowledgement from a human-like counterpart, and despises a cold standardized reply.

In general, notice the absolute value of the coefficient on the interaction with uniqueness is greater than both standardization alone and its interaction with message sentiment. This suggests that no matter the sentiment, a unique and new complaint must be answered with a personalized response, otherwise causing detrimental effects on feedback.

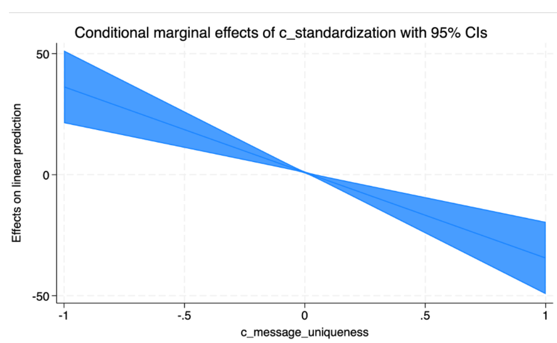
VARIABLES	(1) feedback_sentiment
c_standardization	**0.927***** (0.203)
word_count_company_mean	-0.00308*** (0.000567)
response_time_company_mean_s	0.0100 (0.00951)
sentiment_score_company_mean	0.201*** (0.0168)
o.c_standardization	-
c_message_uniqueness	0.813*** (0.304)
c.c_standardization#c.c_message_uniqueness	**-35.37***** (7.634)
c_message_sentiment	0.245*** (0.0152)
c.c_standardization#c.c_message_sentiment	**-2.625***** (0.365)
Constant	-0.0408** (0.0208)
Observations	7,599
R-squared	0.108

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In order to better understand the effects, we plot the marginal effect of standardization at different level of the moderator.



(a) Distribution of 'feedback_sentiment'



(b) Distribution of 'standardization'

Figure 3

We now proceed to test our theory on time mediation.

4.4 Causal Steps Mediator Test

In order to test our theory that standardization interacts with message sentiment and influences feedback through the mediation of time, we use the causal steps approach proposed by Baron and Kenny [7]. In particular, we test the significance of the regressions of time on standardization (A-path, with relevant confounds), of feedback on time and standardization (B-path, with relevant confounds) and of feedback sentiment on standardization (C-path, with relevant confounds).

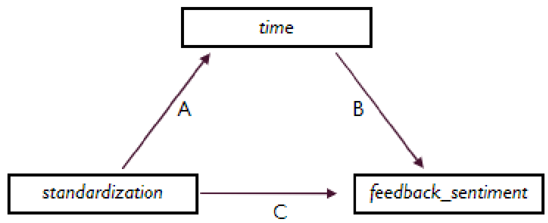


Figure 4: Causal Steps Schema

If our hypothesis is correct, at least the first regression should be relevant, proving that the time taken to answer a message is actually determined by the standardization of the message itself. We can then look at the other two regressions to determine partial or full mediation.

	(C-path)
VARIABLES	**c_feedback_sentiment**
c_standardization	**-0.386****
	(0.196)
word_count_company_mean	-0.00431***
	(0.000574)
sentiment_score_company_mean	0.294***
	(0.0167)
Constant	0.0728***
	(0.0212)
Observations	7,599
R-squared	0.050

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(A-path) **response_time_company_mean_s**
c_standardization	**-0.313**
	(0.231)
c_message_uniqueness	-0.471
	(0.354)
focal_user_followers_count_s	-0.000539
	(0.000420)
like_count_focal_user_first_s	0.384
	(0.393)
Constant	0.155***
	(0.00789)
Observations	7,599
R-squared	0.001

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(B-path) **c_feedback_sentiment**
c_standardization	**-0.374***
	(0.197)
word_count_company_mean	-0.00446***
	(0.000579)
response_time_company_mean_s	**0.0149**
	(0.00980)
sentiment_score_company_mean	0.294***
	(0.0167)
Constant	0.0759***
	(0.0213)
Observations	7,599
R-squared	0.050

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results contradict our theory on mediation, resulting in a non significant regression of time on standardization. Probably, this is also caused by the fact that the time variable we used is actually the average time taken to answer all messages in the conversation, and not just the first one. Alternatively, the observed positive effect of standardization when message sentiment is negative could be due to other factors. For example, an animated customer might not see the point in answering a standardized message as it's perceived as a waste of time, thus leaving the conversation without worsening the feedback.

4.5 Fixed Effects

In the previous models, we have not considered individual industry effects, which might be responsible for some latent heterogeneity. We now exploit panel data structure to account for such phenomenon, and hence implement a fixed effects regression model.

Figure 5: Industry-wise overview

A visual inspection shows that a downward sloping trend is shared among the majority of industries, with only a few such as Telecom Services and Travel Services that seem to be associated to a positive relation between standardization and customer feedback. Looking at our aggregated data we explore the between and the within variability, trying to understand if the fixed effect model is needed. Some variables show a high degree of within variability, such as user tweet count or the mean word count of the company. Few show a higher between variation, such as the presence of a dedicated support. We then implement FE and make use of statistical tests for significance, finding out that our previous results remain consistent under the FE model.

VARIABLES	(1) feedback_sentiment
c_standardization	1.013***
	(0.210)
word_count_company_mean	-0.00399***
	(0.000645)
response_time_company_mean_s	0.00496
	(0.00971)
sentiment_score_company_mean	0.193***
	(0.0172)
c_message_uniqueness	0.652**
	(0.307)
c.c_standardization#c.c_message_uniqueness	-36.10***
	(7.680)
o.c_standardization	-
c_message_sentiment	0.222***
	(0.0154)
c.c_standardization#c.c_message_sentiment	-2.623***
	(0.368)
Constant	-0.00702
	(0.0237)
Observations	7,599
Number of industry_id	**26**
R-squared	0.098

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

For the sake of completeness, we run a Hausman test to confirm the correctness of the model specification, testing the Fixed Effects model against the Random Effects. Since the p-value is virtually zero, we reject the null under which both the estimators are consistent, in favor of the alternative hypothesis where only the FE is.

4.6 Addressing Customers using their Names

To understand whether including or not the first name in the message has different effect of feedback, we first have to prepare the data. The goal is to come up with a binary column 'addressed_with_name' which shows whether or not the user in the conversation was addressed using the first name, namely whether or not the first tweet posted by the company in the conversation contained a first name (1) or not (0).

First, we take care of the user handle (@username) present in every tweet, locating it with a RegEx match and dropping it. Afterwards, we tokenize the words and use the standard NER (Named Entitiy Recognition) model from the Python NLTK package to tag each token, which however turns out to perform badly: it simply considers as first names all the capitalized words, hence classifying tokens such as "Hey", "Please" as names and "barbara", "jim" as not names. We proceed to do some research and decide to use the NER model developed and open-sourced by Stanford University. Since the base model only distinguishes between three classes, each token is tagged as either "PERSON", "LOCATION" or "OTHER". Thus, we encode the tweets containing a "PERSON" tag as tweets in which the user was addressed using the first name (1). Finally, we perform the regression with this additional feature and obtain the following results.

VARIABLES	(1) feedback_sentiment
c_standardization	1.041***
	(0.210)
word_count_company_mean	-0.00410***
	(0.000650)
response_time_company_mean_s	0.00439
	(0.00971)
sentiment_score_company_mean	0.191***
	(0.0173)
o.c_standardization	-
c_message_uniqueness	0.659**
	(0.307)
c.c_standardization#c.c_message_uniqueness	-36.29***
	(7.681)
c_message_sentiment	0.222***
	(0.0154)
c.c_standardization#c.c_message_sentiment	-2.635***
	(0.368)
addressed_with_name	**0.0210**
	(0.0146)
Constant	-0.00928
	(0.0238)
Observations	7,599
Number of industry_id	26
R-squared	0.098

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As expected, addressing customers with their name has a positive impact on the intercept, resulting in a 0.02 upshift of the curve. However, the confidence interval contains 0, suggesting non significance of this feature.

5 Further Work

Overall, our model has achieved good explanatory power and it is in line with our expectations, as well as past theory results. However, there are still several aspects under which it could be improved. One limitation of our approach is that we assumed throughout the research that the data respects the conditions of Ordinary Least Squares (OLS). Nevertheless, it is very likely that standardization is an endogenous variable. This is what motivated us to introduce confounds into the main regression, in order to control the correlation with other regressors. However, given that our independent variable standardization was extracted through NLP, there is a high chance that it was measured with error. If this is the case, introducing confounds might not be sufficient to solve the problems, and we should try more complex solutions.

One approach would be to use a Two Stage Least Squares Regression (TSLS). In this model, we use instrumental variables that are correlated with the endogenous variable, but are not directly related to the error term, to compute the estimated value of our independent variable. Then, we use the prediction rather than the real value in the linear regression. This helps overcome endogeneity issues and and control for omitted variable bias, as well as measurement errors, by capturing the variation in the independent variable that is unrelated to the error term. Our approach was to use the first user message uniqueness as an instrument. It should meet the relevance assumption, because the more unique a message, the higher the company's urge to respond in a personalized manner. It also satisfies exogeneity, as message uniqueness is less likely to be directly correlated with unobserved factors influencing the user sentiment. However, when running TSLS, the obtained results were statistically insignificant, and the instrument was proved to be invalid by a Hansen-Sargan test. In further work, we propose to try with some different instruments, for example the number of tweets received by the firm in the past days, or a flag representing whether the firm uses automated responses.

Additionally, we want to highlight that we understood this research as having the goal of explaining the relationship between standardization and feedback, aiming at understanding rather than predicting. In fact, different kinds of considerations shall be made for works that instead tackle this same issue from a predictive perspective. For example, all variables reporting the mean of measured parameters, such as anger, can be used to interpolate the data and obtain more complete features which can be useful to predict the final feedback. For example, taking the difference betweeen initial anger and its mean results in a indicator of the overall trend of the anger in the conversation. Namely, a positive delta reflects increasing anger, which is likely to lead to worse feedback compared to negative delta situation, where the anger decreased throughout the conversation. Anyway, given the purpose of this problem being understanding rather than predicting, we decided to discard these considerations and limit our work to reporting them in this section.

6 Conclusions

The project we just illustrated delves into the intricate relationship between standardized responses and feedback sentiment in customer-company interactions on Twitter. Thanks to different data analysis tools and techniques, such as OLS regression, confounds analysis, fixed effects modeling and moderator/mediator analysis, we have gained valuable insights into how these variables influence customer satisfaction and sentiment outcomes.

The findings from our OLS regression models revealed a significant negative association between standardization and feedback sentiment, indicating that as the degree of standardization increases, customer sentiment tends to become less positive, confirming our base theory that customers are more inclined to interact with human-like counterparts. However, this initial analysis provided a baseline understanding, but we recognized the need to go deeper.

Firstly, we introduced confounding variables, including message uniqueness, response time, sentiment score, and word count. This expanded model demonstrated that while standardization continues to have a negative impact on feedback sentiment, a portion of the effect has been absorbed by other significant variables. This highlights the importance of considering a comprehensive set of variables when studying this relationship.

The consideration of moderators provided additional perspectives to our analysis. We discovered that the impact of standardization varies depending on the uniqueness of the message and the initial sentiment. In particular, we found out that the attribute that most influences the effect of standardization on feedback is the uniqueness of a tweet, which thus must be taken into considerations. Additionally, we tested the hypothesis of time mediating the interaction between standardization and feedback sentiment with an approach proposed by Baron and Kenny in 1986, which rejected our theory.

Furthermore, we employed fixed effects modeling to account for latent heterogeneity among different industries. This allowed us to capture unique dynamics within specific sectors, highlighting how industry context shapes the interaction between standardization and feedback sentiment.

Finally, our exploration of addressing customers by their first names introduced an intriguing dimension to the study. Preliminary findings suggest that this personal touch can positively influence customer feedback, emphasizing the importance of humanizing interactions in customer support.

In conclusion, this research offers a comprehensive understanding of the relationship between response standardization and sentiment in Twitter conversations between customers and companies' support teams and how different factors can mitigate or magnify such effect. By examining various dimensions, we have provided valuable marketing insights that can help managers make strategic and efficient support-related choices, with the goal of finding an optimal balance between standardizing processes to optimize efficiency and maintaining human interaction with customers, which is recognized as of paramount importance for a company's success.

References

[1] P. Suchánek, M. Králová, *Effect of customer satisfaction on company performance*, 2015

[2] M. C. H. Yeung, C. T. Ennew, *Measuring the impact of customer satisfaction on profitability: A sectoral analysis*, 2001

[3] C. Tong, K. Lui, S. Wong, *The Influences of Service Personalization, Customer Satisfaction and Switching Costs on E-Loyalty*, 2020.

[4] Alan L. Montgomery, Michael D. Smith, *Prospects for Personalization on the Internet*, 2009.

[5] G. Miceli, F. Ricotta, M. Costabile, *Customizing customization: A conceptual framework for interactive personalization*, 2007.

[6] T. Nguyen, L. Waizenegger, A. Techatassanasoontorn, *"Don't Neglect the User!" – Identifying Types of Human-Chatbot Interactions and their Associated Characteristics*, 2021.

[7] Baron, R. M., & Kenny, *The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations.* Journal of Personality and Social Psychology, 1986