[Inner endpaper]

HEC MONTRÉAL
École affiliée à l'Université de Montréal

From Machine Learning to Theoretical Concepts and Wellbeing: Two Different Applications for Branding and Service Failures

par
Chau Minh Nguyen

Thèse présentée en vue de l'obtention du grade Ph. D. en administration (option Marketing)

Août 2024

HEC MONTRÉAL
École affiliée à l'Université de Montréal

Cette thèse intitulée :

From Machine Learning to Theoretical Concepts and Wellbeing: Two Different Applications for Branding and Service Failures

Présentée par :

Chau Minh Nguyen

a été évaluée par un jury composé des personnes suivantes :

Bruno Lussier
HEC Montréal
Président-rapporteur

Yany Grégoire
HEC Montréal
Directeur de recherche

Marcelo Vinhal Nepomuceno
HEC Montréal
Codirecteur de recherche

Renaud Legoux
HEC Montréal
Membre du jury

Matthijs Meire
IÉSEG
Membre du jury

Paolo Antonetti
EDHEC
Examinateur externe

Claudia Rebolledo
HEC Montréal
Représentant du directeur de HEC Montréal

Résumé

À mesure que le Big Data et l'apprentissage automatique gagnent en importance dans la recherche en marketing, ils dépassent progressivement la littérature basée sur les méthodes pour devenir un outil important dans l'amélioration des théories.

Notre premier essai applique un modèle d’apprentissage profond (« deeplearning ») avancée, « Bidirectional Encoder Representations from Transformers » (BERT), pour identifier différentes dimensions dans les messages des musiciens. En utilisant des données d'entraînement préalablement classées par des humains, nous avons développé plusieurs modèles BERT, chacun pour une variable, et nous avons appliqué ces modèles pour prédire les variables sur un grand ensemble de données.

Plus précisément, nous étudions l'impact du signal caritatif des musiciens (c'est-à-dire l'envoi de messages liés à la défense d'œuvres caritatives) et d'autres types de signaux sur leurs ventes de musique à court et à long terme par la médiation de l'engagement dans les médias sociaux, en utilisant la médiation multiniveau. Nous examinons également les différences entre les musiciens de différents degrés d'influence. Nous contribuons ainsi à la fois à l'éthique des affaires et à la littérature sur la signalisation.

Dans le deuxième essai, nous avons utilisé un modèle d'apprentissage profond déjà entraîné dans un autre travail pour prédire deux types de plaintes dans un ensemble de données privées d'une compagnie aérienne. L'apprentissage automatique nous permet de distinguer deux types de plaintes avec des niveaux de précision élevés, ce qui nous permet de les relier aux facteurs précédents et à la conséquence, qui est la valeur d'achat de ces

plaignants. Nous utilisons également des théories liées aux ressources, jusqu'à présent souvent utilisées dans la littérature sur les organisations et les relations sociales, pour expliquer les décisions des consommateurs à différents stades.

En d'autres termes, en appliquant l'apprentissage automatique, nous pouvons identifier de nouveaux concepts, les mesurer et établir leurs diverses relations, et à partir de là, développer de nouveaux concepts théoriques dans deux domaines matures : l'action caritative et les défaillances de service. Dans les deux cas, les expériences et les méthodes quantitatives traditionnelles sont difficiles à mettre en œuvre pour répondre à nos questions de recherche, en raison de la nature complexe de ces dernières.

En outre, le bien-être est un élément important des deux essais. Le premier essai encourage le co-marquage à long terme entre les musiciens et les organisations à but non lucratif, contribuant ainsi au bien-être social, car les deux parties vont au-delà des relations transactionnelles à court terme pour apporter des changements plus permanents à la communauté. Dans le deuxième essai, nous aidons les entreprises à examiner les décisions des consommateurs à chaque étape du parcours de l'échec et de la récupération de service, en tant que mécanismes de protection du bien-être de ces derniers, en utilisant le prisme des ressources. Cela permet aux entreprises de conserver leurs clients à long terme et de contribuer à la viabilité de la communauté.

Mots clés : l'apprentissage automatique, le bien-être, l'action caritative, les défaillances de service, concepts théoriques

Méthodes de recherche : exploitation de données, économétrie, recherche longitudinale, recherche quantitative

Abstract

Big Data and machine learning become more prominent in marketing research, they are also gradually moving beyond methods-based literature to become an important tool in enhancing theories.

Our first essay applies an advanced deep-learning approach, Bidirectional Encoder Representations from Transformers (BERT), to identify different dimensions in musicians' messages. Using previously human-classified training data, we developed multiple BERT models, each for a variable, and applied those models to predict the variables on a large dataset.

Specifically, we study the impacts of musicians' charity signaling (i.e., sending messages related to charity advocacy) and other types of signaling on their music sales in the short and long terms through the mediation of social media engagement, using multilevel mediation. We also examine how this differs among musicians of different degrees of influence. We thus make contributions to both business ethics and signaling literature.

In the second essay, we used a deep learning model already trained in another work to predict two types of complaints in a private dataset of an airline company. Machine learning allows us to distinguish two types of complaints with high levels of accuracy, allowing us to link them to the preceding factors and the consequence, which is the purchase value of these complainers. We also use theories related to resources, until now often used in organization and social relationship literature, to explain consumer decisions at different stages.

In other words, by applying machine learning, we can identify new constructs, measure them, and establish their diverse relationships, and from these develop new theoretical concepts in two mature fields: charity advocacy and service failures. In both cases, experiments and traditional quantitative methods are challenging to answer our research questions, due to the latter's complex nature.

Furthermore, well-being is an important part of both essays. The first essay encourages long-term co-branding between musicians and non-profits, thereby contributing to social well-being, as both parties go beyond transactional, short-term relationships to make more permanent changes for the community. In the second essay, by using the lenses of resources, we help firms to examine consumer decisions in each stage of the failure and recovery journey as mechanisms to protect the latter's wellbeing. This helps firms to retain their customers in the long term and to contribute to community viability.

Keywords : machine learning, wellbeing, charity advocacy, service failures, theoretical concepts

Research methods : data mining, econometrics, longitudinal research, quantitative research

Table of contents

List of tables and figures

List of abbreviations

BERT : Bidirectional Encoder Representations from Transformers

CSR : Corporate Social Responsibility

LIWC: Linguistic Inquiry and Word Count

Dedicated to my parents, who went and still go through so much trouble for me, my dear daughter Bo, Dr Ram, who helped me to start the journey on marketing research, and Dr Hair, Dr Flight and my qualitative professor Sean Jordan, who supported my PhD adventure.

Acknowledgements

I would like to thank my supervisors, Yany Grégoire and Marcelo Vinhal Nepomuceno, who have guided me so committedly, Renaud Legoux, whose knowledge of quantitave methods and pedagogical skills gave me a much-needed confidence, Danilo Dantas and François Colbert, who helped me in the difficult early steps of my research, and Matthijs Meire, who taught me how to do machine learning properly.

I would also like to thank Professor Michel Laroche who gave me precious advice from early on of the research and Hyunhwan "Aiden" Lee, whose works and passion inspired me to study machine learning.

I also would like to thank Jonathan Deschênes who is a great human being and was always available when I needed some advice, Marie-Agnès Parmentier, who opened me up to the wonderful world of CCT, Aya Aboelenien, who taught me how to think like a qualitative researcher and Shang Lin Chen and Davood Wadi, who helped me to survive the machine learning classes.

To all my other professors, colleagues and friends, many thanks for all your help throughout this journey. I have learned so much from each of you and without you, I would not have made it this far.

Preface

Qualitative researchers have long analyzed text, to explore new phenomena and concepts, such as consumer engagement (Brodie et al. 2013), branding through social media firestorms (Scholz and Smith 2019), racialized brands (Veresiu 2023). Quantitative approaches to text, referred to as text-mining (Berger et al. 2020), often focus on extracting text-features, e.g. sentiments, style variance (Herhausen et al. 2019), subjects (Humphreys 2010a) and dimensions (Nepomuceno et al. 2020). The advantage of qualitative approaches refers to exploring complex constructs and relationships, while that of quantitative approaches is measuring them.

Lexicon-based approach Linguistic Inquiry and Word Count (LIWC) has established itself as an important text-mining tool in marketing research, e.g., to examine legitimization through linguistic changes (Humphreys 2010b) and complaint de-esclation (Herhausen et al. 2023). With the growth of Artificial Intelligence and Natural Language Processing, the complexity of the methods also increased over time, moving from lexicon-based approaches (e.g, LIWC and VADER), linear classification, (e.g. Naïve Bayes and Support Vector Machine), non-linear classification (e.g., Random Forest), to deep-learning (Hartmann et al. 2019). As Big Data and machine learning become more prominent in marketing research, they are also gradually moving beyond methods-based literature (e.g., Moon et al. 2010; Wei et al. 2022 ; Nepomuceno et al. 2020) to become an important tool in enhancing theories.

Machine learning is using computational algorithms by teaching the machine to do different repeated tasks (El Naqa and Murphy 2015). In supervised learning, the machine learns to do these tasks by working with a large amount of input data and adapts its "architecture through repetition" (p. 4, El Naqa and Murphy 2015) to achieve better results. Various mechanisms are in place to help the algorithms improve the results during training, e.g., adjustable parameters, using probability distributions from input data and different computational approaches (El Naqa and Murphy 2015). Another form of machine learning is unsupervised learning. In this case, no training data is necessary. The machine figures out features of the text using its own knowledge of existing features. They are thus less costly than supervised learning.

The machine learning tasks for our two essays are mainly supervised textual classification. In the first essay, we trained the machine to classify different types of social media messages, using human-classified text. In the second essay, we used a model already trained to interpret different types of complaints. The machine imitates the ways human beings classify text through repeated exposure to the same similarities and differences (El Naqa and Murphy 2015). An ideal situation would have been the entire text being human-coded. But for large datasets, this was not feasible and typically could only be done on a sample (e.g., Johnen and Schnittka 2019). By using advanced machine-learning models, however, we could reach levels of precision comparable to human classifications.

The first essay applies an advanced deep-learning approach, Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al. 2019), to identify different dimensions in musicians' messages. Using previously human-classified training data, we developed multiple BERT models, each for a variable, and applied those models to predict the variables on a large dataset.

Specifically, we study the impacts of musicians' charity signaling (i.e., sending messages related to charity advocacy) and other types of signaling on their music sales in the short and long terms through the mediation of social media engagement, using multilevel mediation. We also examine how this differs among musicians of different degrees of influence. We thus make contributions to both business ethics and signaling literature.

In the second essay, a deep learning model already trained in another related work was used to predict two types of complaints in a private dataset of an airline company. Thus, two types of complaints were identified with high levels of accuracy, making it possible to analyze their preceding factors and their consequence, which is the purchase value of these complainers.

In other words, machine learning allows for the identification of new constructs and their measurements. This helps to establish their diverse relationships and to develop new theoretical concepts in two mature fields: charity advocacy and service failures. In both cases, experiments and traditional quantitative methods are challenging to answer the relevant research questions, due to the latter's complex nature.

Thus, machine learning could play a fundamental role in contributing to theoretical development. In current marketing literature, many researchers focus on introducing new advanced methods, while many others use mainstream approaches (e.g, experiments, interviews) with a focus on theoretical contributions. The two essays aim to fill in this gap, by using advanced approaches to make theoretical contributions. This is important, as the combined growth and influence of artificial intelligence (AI) and social media provide enormous opportunities for breakthroughs in research.

In addition, long-term wellbeing is an important part of both essays. In the first essay, our research shows the how advocating for charity consistently helps musicians' financial performance. This finding also encourages long-term co-branding between musicians and non-profits, thereby contributing to social well-being, as both parties go beyond transactional, short-term relationships to make more permanent changes. We also compared charity advocacy with mission signals (i.e., commercial messages) and non-mission signals (i.e self-disclosing messages) and showed how in the long term, charity advocacy could benefit financial performance more than other types of signals. This unintuitive finding showcases the importance of being committed to social causes for brands in a complex world.

In the second essay, service failures and recoveries were studied from a well-being perspective. Management and organisation research has long examined employee wellbeing issues (e.g. stress and burnouts), as they are considered important for the survival of organisations. However, in service literature, consumer well-being itself is a relatively new concept often linked to transformative consumer research. By examining consumer decisions through the lenses of resources, we can explain consumer decisions in each stage of the failure and recovery journey as ways to protect their own well-being. Understanding this is important for firms, as it allows them to approach their consumers from the perspective of preserving the consumers' resources and this not only helps them to retain their customers in the long term, but also to contribute to community viability.

References

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the Tribes: Using Text for Marketing Insight. *Journal of Marketing*, *84*(1), 1–25. https://doi.org/10.1177/0022242919873106

Brodie, R. J., Ilic, A., Juric, B., & Hollebeek, L. (2013). Consumer engagement in a virtual brand community: An exploratory analysis. *Journal of Business Research*, *66*(1), 105–114. https://doi.org/10.1016/j.jbusres.2011.07.029

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*. http://arxiv.org/abs/1810.04805. Accessed 5 July 2021

El Naqa, I., & Murphy, M. J. (2015). What Is Machine Learning? In I. El Naqa, R. Li, & M. J. Murphy (Eds.), *Machine Learning in Radiation Oncology* (pp. 3–11). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-18305-3_1

Hartmann, J., Huppertz, J., Schamp, C., & Heitmann, M. (2019). Comparing automated text classification methods. *International Journal of Research in Marketing*, *36*(1), 20–38. https://doi.org/10.1016/j.ijresmar.2018.09.009

Herhausen, D., Grewal, L., Cummings, K. H., Roggeveen, A. L., Villarroel Ordenes, F., & Grewal, D. (2023). Complaint De-Escalation Strategies on Social Media. *Journal of Marketing*, *87*(2), 210–231. https://doi.org/10.1177/00222429221119977

Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, Preventing, and Mitigating Online Firestorms in Brand Communities. *Journal of Marketing*, *83*(3), 1–21. https://doi.org/10.1177/0022242918822300

Chapter 1:

Striking the right notes: Long- and short-term financial impacts of musicians' charity advocacy versus other signaling types[1][2]

Abstract

By using multilevel mediation involving 322,589 posts made by 384 musicians over 104 weeks, we simultaneously analyze the short-term and long-term effects of charity-related signaling on sales, with social media engagement as the mediator. Specifically, we compare the effects of charity-related signals with those of two other types of signals: mission-related (i.e., promoting music and commercial products) and non-mission-related (i.e., other posts that do not relate to the other two categories). In the short term, the indirect effect of using charity signaling on sales (through engagement) is positive, though smaller than the effects of mission-related and non-mission related signals. However, in the long term, the indirect effect of regularly using charity-related signaling on sales (through long-term engagement) is greater than for the effects involving the other types of signals. We derive from these findings three main implications for the business ethics literature. First, in the long term, the mutual economic benefits of charity signaling should encourage both entities (i.e., musicians and charities) to go beyond short-term, transactional philanthropy. Second, because it is profitable for musicians to partner with charities in the long-

[1] _This article has been published in Journal of Business Ethics (2024) 193, 217–233.

[2] This article is co-authored with Marcelo Vinhal Nepomuceno, Yany Grégoire and Renaud Legoux.

term, our research argues that charities have extensive bargaining power in such co-branding decisions. Third, our research highlights the importance of studying the longitudinal aspects of co-branding decisions involving non-profit organizations; the financial outlook of such decisions could greatly vary depending on the timeframe (i.e., short vs. long).

Key words: human brands, musicians, charity, CSR, social media, engagement, financial performance, multilevel analyses

1.1 Introduction

For-profit companies enter partnerships with non-profit organizations or create their own charity entities for many different reasons. For instance, such associations can allow firms to promote their own products while engaging in social causes, which are close to their heart and endorsed by their target audiences (Vanhamme et al. 2012). Such partnerships can also be a way for firms to show that "they care" about their environment and stakeholders (Adkins 1999). Although signaling support for charity has been shown to have a positive impact on financial performance (Hasan et al. 2018), organizational legitimacy (Liston-Heyes and Liu 2010) and stakeholder involvement (Liu et al. 2010), it remains to be seen how charity-related signals fare compared to other signals, which are more closely related to firms' core business and mission (Connelly et al. 2011; Guo and Saxton 2014).

To better understand this issue, here is an example with the LEGO Group using different types of signals. When LEGO showcases children's creativeness through the usage of its products, it uses a *mission-related signal* that is directly linked to its core business (Guo and Saxton 2014). In turn, LEGO also employs *charity-related signals* when it publicizes its collaboration with sight-loss organisations; in these partnerships, the LEGO Foundation freely provides blocks with Braille numbers and letters (Dixon 2019). In this LEGO case, which signal—between the charity-related or the mission-related one—would be the most effective at generating engagement and sales? To the best of our knowledge, it is unknown whether the impact of charity signals on financial performance would be greater or smaller compared to the effects of other signals (Connelly et al. 2011). Given this general gap, the broad purpose of our research is to examine the different effects of charity signals versus other signals—related to a brand's mission (Guo and Saxton 2014), for instance—on the sales made by human brands, such as musicians.

Compared with corporate brands, research on human brands is emerging (Osorio et al. 2020). This context is of special interest because musicians have been known to support numerous charities on social media. Although some research has examined the phenomenon of celebrity philanthropy on social media (e.g., Bennett 2014; Dieter and Kumar 2008), it is unclear if musicians and charities mutually benefit from their partnerships in financial terms (Santos et al. 2019). Documenting this issue is important for both entities. Indeed, musicians need to have a better understanding of the effects of their public advocacy and the potential frictions that could exist between their economic interests and their social responsibilities (Harlow and Benbrook 2019). In addition, if such associations are profitable for musicians, this situation could position charities as valuable "business partners," which may have more leverage than is often assumed.

In light of these two gaps—that is, 1) contrasting charity signaling with other signals and 2) understanding the financial effects of charity signaling for musicians—our research answers the following three questions. First, we wonder if the impact of charity signalling on musicians' financial performance is larger or smaller than other types of signalling? Second, what would be the long-term and short-term effects of using charity signals on financial performance in this context? Third, does social media engagement play a mediation role to explain the different effects (short-long vs. long-term) of musicians' charity signaling on financial performance?

To answer these questions, we conducted a study with the posts (322,589 posts) and sales of 384 musicians; these data were collected over 104 weeks. In this research, we identified three main types of posts: charity-related (i.e., when artists directly discussed a given charity cause), mission-related (i.e., when artists discussed their music, shows, events or other commercial products and brands) and non-mission-related (i.e., when posts mentioned any other subjects than those belonging to the first two categories). To be able to examine simultaneously the effects of short-term versus long-term postings of different signals, we use a novel multilevel mediation recently advanced by Hayes and Rockwood (2020).

Before discussing the relevance of this method, we first define our two multilevel effects, which we label *short-term* and *long-term*. These two effects can be understood differently if we define them conceptually or analytically. To reconcile these views, we define these two effects by using three attributes: unit level ("week" versus "artist"), temporality of effect ("intermittent" versus "cumulative") and analytical level ("within-person" versus "between-person"). Please see Table 1 for a summary of these definitions. The *short-term* effect (level 1) captures the average changes (in engagement or sales) from one week to another for a given artist (Wang and Maxwell 2015). Because this effect captures the average longitudinal change between two weeks, we qualify it by using the adjectives "intermittent" and "short-term". The *long-term* effect (level 2) aggregates all the information over 104 weeks for each of the 384 artists, who become the unit of analysis (Wang and Maxwell 2015). This analysis is not longitudinal, and it uses all the information cumulated for each artist. Given these characteristics, we use the adjectives "long-term" and "cumulative" to qualify this effect (Tasca and Gallop 2009).

In other words, the long-term effect refers to the impact of doing an activity on a consistent basis. The short-term effects indicates the impact of doing an activity when it deviates from the regular pattern. For simplicity of exposition, we mainly use the labels "short term" and "long term" in the rest of our manuscript; the only exception is for the methods section, which requires the usage of technical terms.[3]

[3] In the methods section, we refer to the technical terms—within-person and between-person—that are used by methodologists and software (Zhang et al. 2009; Curran and Bauer 2011; Hayes and Rockwood 2020).

Table 1.1: Description of the notions "short term" and "long term" in this research

Concept	Short term	Long term
Unit (level)	39,936 observations (104 weeks * 384 artists) (level 1)	384 artists (level 2)
Temporality of effect	Intermittent: average change between two weeks for a given artist	Cumulative: an aggregation of all the weeks for each artist
Analytical level	Within-person	Between-person

The multilevel, longitudinal analyses used in this research (Preacher et al. 2010) allow estimating these two effects simultaneously. Differentiating these effects is important because an inability to do so could lead to biased results and inaccurate interpretations. Here is an illustration of such potential biases. For instance, people are more likely to have heart attacks while exercising (short-term effect), but those who regularly exercise over years are less likely to suffer from heart attacks (long-term effect) (Curran and Bauer 2011). So, according to this example, people need to account for the coexistence of both effects; it would be unreasonable to focus only on the short-term effect and to recommend people to stop exercising. Applying a similar logic to our context, our multilevel analyses provide insightful responses to our three questions, and they enable us to make three specific contributions.

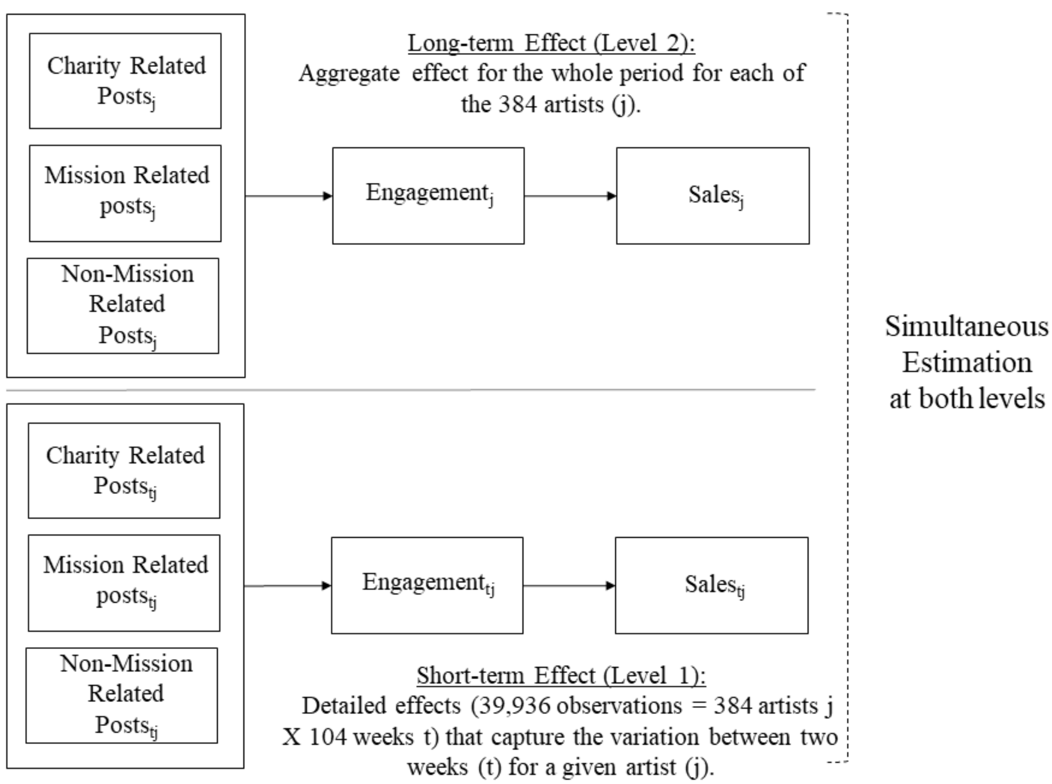
First, we find that posting about charity is beneficial to sales in *the short term*, even if this effect is relatively modest and weaker than those of other signals (i.e., mission-related or non-mission related). In other words, there is a positive effect of using a given charity signal in a given week (i.e., short-term effect), although this effect is somewhat limited. Here, we argue that artists who occasionally support charities could still benefit a little from mentioning them from time to time. This finding might persuade hesitant artists to experiment with charity advocacy.

Second and importantly, we show that the *long-term effects* of regularly using charity-related signals on sales is not only positive but also greater than the long-term effects of the other types of signals (i.e., mission-related and others). This set of results is encouraging for both music artists and charities; it shows that long-term partnerships could be mutually beneficial in financial terms and possibly many other ways (e.g., societal, reputational). These results are important because they demonstrate the important role that charity advocacy can play over time for brand building. By developing well-crafted strategic partnerships, both musicians and non-profits could enhance their financial situation and societal impact (Austin and Seitanidi 2012; Macdonald and Chrisp 2005).

Third, we pay special attention to understand the process explaining the effects of signaling on sales. To do so, we argue that engagement plays an important role of mediation in explaining the short-term and long-term effects of signaling (all three types) on sales. Here, engagement on social media is conceptualized as being composed of three core indicators: shares, reactions and comments (Kumar and Pansari 2016). Please see our conceptual framework (Figure 1), which displays the mediation role of engagement in the short and long terms. Although the concept of engagement has been very influential in marketing strategy (Kumar et al. 2010; Soares et al. 2022; Ji et al. 2017; Li et al. 2021), this notion has been rarely discussed in the non-profit and signaling literatures, to the best of our knowledge. Addressing this gap, we find that the relative effects of charity-related signaling on sales (short- and long-term) are mediated by an engagement mechanism, which becomes especially strong when charity signaling is regularly used by musicians. With this finding, we extend prior research that has focused mainly on the direct impact of charity-related initiatives on engagement (Kucukusta et al. 2019; Chu et al. 2020) or financial performance (van Beurden and Gössling 2008; Clacher and Hagendorff 2012; Kang et al. 2016).

Building on this research, we integrate both streams by arguing for a sequence "charity-related posts → engagement → sales", which is tested at two levels (short- and long-term).

Figure 1.1: Short-term and long-term impacts of different types of posts on sales through engagement



By making these contributions, we derive three key implications for the business ethics literature and the management of co-branding with non-profit organizations (e.g., charities). First, the long-term, mutual economic benefits of musicians' charity signaling should encourage both organizations to go beyond short-term, transactional philanthropy. Both musicians and non-profits are encouraged to build long-term partnerships that aim to co-create durable value for societies (Austin and Seitanidi 2012); doing so would generate financial and societal benefits for both parties (Knoll and Matthes 2017). Second, because it is profitable for musicians to partner with charities in the long term, our research argues for a change in the relational dynamics between musicians and non-profits, with charities or non-profits having extensive bargaining power in strategic co-branding decisions. Third, our research highlights the importance of studying the longitudinal aspect of co-branding decisions involving non-profit organizations. The financial outlook of such decisions could greatly vary depending on the timeframe (i.e., short vs. long term). Indeed, we find that the financial benefits of such partnership are more advantageous when considered over a long (vs. a short) period. Importantly, this long-term beneficial effect is mainly explained by a long-term engagement mechanism. To the best of our knowledge, engagement-based processes have rarely been discussed in the business ethics literature.

1.2 Theoretical background and hypotheses

1.2.1. Signals

Through their music, activities, public statements and social media messages, artists send out different types of signals (Higgins and Gulati 2006; Connelly et al. 2011), such as their implicit and explicit emotions (Waterman 1996), personal and social identity (MacDonald et al. 2002), and even political resistance (Street et al. 2008). Charity-related posts and other types of posts are therefore just a small part of these signals. As they wear away over time, the information asymmetry between signal senders and signal receivers also increases (Janney and Folta 2003), especially in the presence of other conflicting signals from the musicians themselves or from other signalers (e.g., the media and social influencers). Thus, repeating the same kind of signals is important to reduce the uncertainty of the branding interpretation. Repetition also leads to the cumulative impact of consistent signals over time (Heil and Robertson 1991), which increases their credibility in the long term (Connelly et al. 2011).

However, while the signaling literature often stresses the importance of repetition, in some cases, intermittent, short-term signals can also have their own value. Irregular signals could be more effective when they give the impression of rarity, which increases their worth in the eyes of the signal receiver (e.g., an annual instead of a monthly prize) (Phau and Prendergast 2000). Surprise is another potential advantage of irregular signals. It garners them more attention so that they become more memorable, helping the signalers to achieve more influence (Loewenstein 2019).

Musicians' charity-related posts share the characteristics of non-profit social media messages—that is, they provide information related to the causes, show attempts at community building and promote calls for action (Lovejoy and Saxton 2012). Charity-related posts are thus associated with warmth (Bernritter et al. 2016), belonging to a community (Chwialkowska 2019), and caring for others (Bernritter et al. 2016). Communities (e.g., from a book club or a church to civic engagement) ideally help members fight isolation and look after each other's well-being (Block 2009). The sense of belonging to a social group increases meaning in an individual's life (Lambert et al. 2013).

These messages are distinct from mission-related posts (Guo and Saxton 2014), which imply the intrinsic quality of the products and brands, such as competence (Bernritter et al. 2016; Nepomuceno et al. 2020), credibility (Erdem et al. 2006) and symbolic values (Schembri et al. 2010). They are also different from non-mission-related posts, which are self-revealing and personal (Chung and Cho 2017; Nepomuceno et al. 2020). In the context of this research, the popularity of a musician, or their social capital (Bourdieu et al. 2003), is also a signal. A message is likely to reach more people from a well-known artist than a new singer and is also considered more credible (Guo and Saxton 2014). Because of these distinct features, the impact of a long- or short-term posting of each type of signal can be different, even contradictory. In the next section, we explain how different signals can influence viewers' engagement on social media in the short or long term.

1.2.2. Signals and engagement

Consumers use reactions, shares and comments on social media to show their *engagement* with brands (Kumar et al. 2010); these actions are instrumental to sway other users and have persuasive effects beyond social networks (Kumar and Pansari 2016; Geng et al. 2020). Here, engagement can be viewed as playing the role of "feedback", or countersignals, to any type of message (Connelly et al. 2011; Saxton et al. 2019). As we argue more comprehensively below, the engagement mechanism is different for mission-related posts and non-mission-related posts in comparison to charity-related posts in both the short and long terms.

Belonging to a parasocial relationship—that is, a relationship that a person builds with a musician who does not personally know him or her (Gong and Li 2017)— "true" fans closely follows musicians' careers and personal lives, and they intensively react to any of their posts, regardless their type. Compared to "casual" followers—who occasionally follow an artist—the "true" fans engage more intensively with mission-related and non-mission-related signals in the short and long terms. For these two types of signals, we expect casual followers to show some engagement with such posts, but to a much lesser extent than the fanbase. Research has found similar effects for external commercial brands supported by musicians (Aw and Labrecque 2020). When celebrities endorse a brand, the engagement increases among followers (i.e., fanbase), while there is no effect for casual or non-followers (Song and Kim 2020).

We argue that the engagement with charity-related posts follows a different pattern compared to the other two types of signals. We expect this because charity-related posts belong to the category "community-centric content" (Chwialkowska 2019) and not to the "parasocial relationship" category (Gong and Li 2017). Community-centric messages encourage interactions among consumers, thus giving them social benefits and making them feel part of a social movement (Wirtz et al. 2013). We predict that because they are different in kind from the other two signals of interest, charity-related signals create engagement in a different manner in the *short term* versus the *long term*.

In the *short term*, a given charity-related post should produce high engagement from the "true" fanbase, as these individuals always tend to support their favorite artists. In contrast, we posit that the casual followers would show little engagement with charity-related posts in the short term. These latter followers could seriously doubt the sincerity of artists occasionally supporting a charity, and they could make negative attributions about the artists' true intentions. In this case, casual followers could infer a lack of *authenticity* in the posts (Park and Cho 2015). Here, casual followers could discount musicians' actions and believe that they advocate a cause to gain political capital (Kane et al. 2009), public image (Babiak et al. 2012) or for tax purposes (Dieter and Kumar 2008). Because of this ambiguity, when musicians endorse charities on an intermittent basis, casual followers are cautious in their engagement. By combining the responses of the true "fanbase" and "casual followers", we propose the following:

H1: *In the short term*, the positive effect of musicians' charity-related posts on user engagement is *weaker* than with mission-related and non-mission-related posts.

However, in the *long term*, artists' repeated charity signals and their continued advocacy of a given cause could create strong support and engagement from casual followers. By regularly displaying their support to a given cause, celebrities build up the stability of their signals and establish the authenticity of their advocacy (Moulard et al. 2015). In this case, casual followers will see these signals as being authentic, credible and truthful, and they will infer strong, positive motives from such repeated signals (Frey and Meier 2004). In the long term, musicians' charity-related posts should attract engagement not only from their "true" fans but also from casual followers and new followers coming from different social network circles. Indeed, given the community-centric nature of charity signaling, new followers could be encouraged to be part of the movement and to engage in pro-social behaviors (Bernritter et al. 2016, Herzog and Yang 2018). Such signals could lead "friends" of the fanbase to support a given cause, especially when pro-social actions become a frame of reference (Frey and Meier 2004). Formally:

H2: *In the long term*, the positive effect of musicians' charity-related posts on user engagement is *stronger* than with mission-related and non-mission-related posts.

1.2.3. Engagement and sales

Engagement, as a measure of the "social media capital" of musicians, contributes to their financial income or "returns" (Saxton and Guo 2020, p. 1). For instance, the engagement on Facebook with automobile makers was associated with an increase in offline car sales (Wang et al. 2021). Similarly, the rate of social media interactions per user had a positive effect on sales in the food and beverage industry (Yost et al. 2021). For non-profit organizations, the number of Facebook shares was a key predictor of the success of a fundraising campaign (Bhati and McDonnell 2020), and organisations with more Facebook fans also received more donations through this channel (Saxton and Wang 2014). Building on H1 and H2, we explain in this section how engagement plays a different mediation role in the short vs. the long term (see Figure 1).

In the *short term,* an increase in engagement helps information to reach more people in a network, which activates "latent ties" (Ellison et al. 2007, p. 1162), leading to the conversion of potential consumers. Even if this engagement takes place only intermittently, it signals a momentary rise in an artist's influence and encourages people to look for and buy his or her music (Ellison and Vitak 2015; Lin 2019). Given this logic and the predicted positive effects between short-term charity-related posts and engagement, we argue that this latter construct (i.e., engagement) mediates the linkage between charity-related posts and sales. However, in the short-term, we expect that this indirect effect (i.e., "short-term charity-related posts → engagement → sales") has less amplitude than the indirect effects involving the other two short-term signals of interest (see H1 for explanations). For precision, these two comparative indirect effects are: "short-term mission-related posts → engagement → sales" and "short-term non-mission-related posts → engagement → sales". Formally:

H3: *In the short term*, the indirect effect "charity-related posts → engagement → sales" is *weaker* than similar indirect effects with mission-related posts and non-mission-related posts.

In the *long term*, when engagement is sustained over time, the growth in perceived influence is steadier, which enhances artists' competitive advantages and sales over time. The fact that strong customer engagement leads to strong sales is a key premise explaining the success of this research stream in marketing strategy (e.g., Kumar and Pansari 2016). Since the long-term effect of charity-related posts on engagement is hypothesized to be greater than that of mission-related or non-mission-related posts over time (see H2), we argue for a similar logic for the indirect effects involving these three signals. For precision, we expect that the long-term indirect effect (i.e., "long-term number of charity-related posts → engagement → sales") has greater amplitude than the indirect effects involving the other two signals of interest (i.e., mission- and non-mission-related). Formally:

H4: *In the long term*, the indirect effect "charity-related posts → engagement → sales" is *stronger* than similar indirect effects with mission-related posts and non-mission-related posts.

1.3. Methodology

1.3.1. Multilevel modeling for longitudinal data

The assumption of independence of observations often does not hold with data in multilevel, nested structures (Moerbeek 2004). The first type of nested structure has members nested within groups. For example, for research on smoking with students grouped (i.e., nested) in schools, the smoking patterns of the students in one school could have correlations with each other because of peer influence, teacher influence and school policies (Moerbeek 2004).

The second type of nested structure relates to different observations nested within individuals—such as, the physical activity of a given person measured at different times (Burton et al. 2009). Our dataset is similarly structured, and it includes 104 weekly observations nested within 384 artists. In this case, multilevel analyses evaluate how individuals change "within themselves" on average between two weeks (i.e., within-person effects), and how individuals differ from one another on average over the whole period (i.e., between-person effects) (Hair Jr. and Fávero 2019). In technical terms, the *within-person effects* represent the short-term effects discussed in our theory, and the *between-person effects* capture the long-term effects previously presented (see Table 1).

1.3.2. Operationalization of constructs

1.3.2.1. Building a weekly dataset

Our dataset includes 322,589 posts made over 104 weeks by 384 artists. In terms of organisation, our databank includes 39,936 observations—which represent the number of weeks by artist (104 weeks * 384 artists)—in which the variable "artist" is nested with the variable week. This form of nested, multilevel databank allows simultaneous testing for "within-person" effects and "between-person" effects, which respectively correspond to the "short-term" and "long-term" effects in our theory.

Our posts originated from three platforms (Facebook, Twitter, and Instagram) and were collected for 384 artists/music groups of different nationalities over two years (2016-2017). Firstly, we applied machine learning to identify the constructs we needed. The first construct demonstrated whether a post invited social media users to play an active role in social causes, in other words, whether a post was related to charity causes (Nepomuceno et al. 2020). The other three constructs are explicit selling (indicating whether a post explicitly promoted a music product or merchandise), show-related (whether a post explicitly or implicitly promoted a show) and merchandise-related (whether a post explicitly or implicitly promoted merchandise, a brand or a company for a commercial purpose).

To achieve this, we used another much smaller dataset (of 5,413 posts) already human-coded on the four variables (with three raters and an inter-rater agreement ranging from 81.8% to 97.6%; see Web Appendix 1.A, Table 1.A1) to train four classifying models, one for each variable. We used Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al. 2019), an advanced Natural Language Processing method. The results indicate 98%, 83%, 83% and 78% accuracy in predicting the presence of charity-related, explicit selling, show-related and merchandise-related posts, respectively (See Web Appendix 1.A, Table 1.A2).

We then applied the four models to predict the constructs on the original dataset (n = 322,589). They identified 5,168 charity-related, 53,305 explicit selling, 109,508 show-related and 21,917 merchandise-related posts. We then created a new variable called mission-related post (Guo and Saxton 2014), namely whether a post is related to shows, merchandise, or explicit selling.

In total, 141,919 posts are mission related. Finally, non-mission-related posts capture the rest of the social media posts when they are neither mission-related nor charity-related.

Next, we converted this dataset into a weekly one to merge it with Nielsen's weekly sales data, which covered album sales, digital song sales and streaming of these artists on the Canadian market in 2016 and 2017. To achieve this, we summed up the number of posts of each type for each week. Similarly, for the engagement information for different types of posts, instead of using the number of reactions (i.e., likes, love, anger, laughter, sadness, surprise, and thankfulness), comments (i.e., comments and replies) and shares (i.e., shares and retweets) for each post of each artist, we summed up the number of reactions, comments and shares for each type of post (i.e., charity-related posts, mission-related posts and non-mission-related posts) of each artist each week. The new dataset includes 39,936 weekly observations (104 weeks * 384 artists).

1.3.2.2. Control variables

We adopted different approaches to find the control variables. First, we used a scraping engine on Python to find news articles mentioning each artist on Google News in the titles or leads. Second, through the *application programming* interfaces (API's) of Twitter and Spotify, we accessed the tweet volume of each artist for each day, the information of their tracks and the release dates of these tracks. We then built the weekly variables for news volume (i.e., the number of articles mentioning an artist each week in the titles or leads), tweet volume (i.e., the number of tweets mentioning an artist each week) and track volume (i.e., the number of tracks featuring an artist each week). We calculated the total number of tracks over all the artists' careers until the end of 2017. Third, Facebook fanbase size information, artists' age, and their experience (i.e., the number of years from the beginning of their career until 2017) were manually collected from Facebook, MusicBrainz, Wikipedia, or past press articles. Facebook fanbase size was coded from

0 to 4: less than 10,000 Facebook fans (0), 10,000 to 100,000 fans (1), 100,000 to one million fans (2), one million to 5 million fans (3) and more than 5 million fans (4).

Apart from the fanbase size, all other variables were ln-transformed to remedy skewed distributions and ensure construct consistency. Collinearity is not an issue for the ln-transformed variables of the number of charity-related posts (M = .07, SD = .26), the number of mission-related posts (M = 1, SD = .93), and the number of non-mission-related posts (M = 1.19, SD = 1), for which correlations range from .13 to .49.

Table 1.2: Definitions of the variables

Variables	**Definition**	**Data source**
Charity-related posts$_{tj}$	The ln value of the sum of the number of charity-related posts in week t of artist j	FB, Twitter, Instagram
Charity-related posts$_{j}$	The mean of charity-related posts$_{tj}$ over 104 weeks for artist j	FB, Twitter, Instagram
Charity-related posts$_{tj(w)}$	Deviation of charity-related posts$_{tj}$ from charity-related posts$_{j}$ per week (i.e., charity-related posts$_{tj(w)}$ = charity posts$_{tj}$ - charity-related posts$_{j}$)	FB, Twitter, Instagram
Engagement$_{tj}$	The ln value of the sum of the volume of all reaction (i.e., likes, loves, anger, laughter, sadness, surprise, thankfulness), retweets/shares, and comments/replies for artist j in week t	FB, Twitter, Instagram
Engagement$_{j}$	The mean of engagement$_{tj}$ (i.e., reactions$_{tj}$, shares$_{tj}$, and comments$_{tj}$) over 104 weeks for artist j	FB, Twitter, Instagram
Engagement$_{tj(w)}$	Deviation of engagement$_{tj}$ from engagement$_{j}$ per week (i.e., engagement$_{tj(w)}$ = engagement$_{tj}$ - engagement$_{j}$)	FB, Twitter, Instagram
Sales$_{tj}$	The ln value of the sum of volume for album sales$_{tj}$, digital songs$_{tj}$, and streaming$_{tj}$ for artist j in week t	Nielsen
Sales$_{j}$	The mean of album sales$_{tj}$, digital songs$_{tj}$, and streaming$_{tj}$ over 104 weeks for artist j	Nielsen
Sales$_{tj(w)}$	Deviation of sales$_{tj}$ from sales$_{j}$ per week (i.e., sales$_{tj(w)}$ = sales$_{tj}$ - sales$_{j}$)	Nielsen
Mission-related posts$_{tj}$	The ln value of the sum of the number of posts related to merchandise, shows or explicit selling of artist j in week t and 1	FB, Twitter, Instagram
Non-mission-related posts$_{tj}$	The ln value of the sum of the number of posts not related to mission or charity of artist j in week t and 1	FB, Twitter, Instagram
Fanbase size$_{j}$	Artist j's fanbase size: < 10K Facebook fans (coded as 0), 10K – 100K fans (1), 100K – 1M fans (2), 1M – 5M fans (3), >5M fans (4).	Facebook
Track volume$_{tj}$	The ln value of the sum of artist j's volume of tracks released in week t	Spotify
Tweet volume$_{tj}$	The ln value of the sum of the number of tweets mentioning artist j on Twitter in week t	Twitter
News volume$_{tj}$	The ln value of the sum of the number of news articles on an artist on Google News that week	Google News

$Experience_j$	The ln value of the sum of the number of years from when artist j's first musical product was publicly released until 2017	MusicBrainz
Age_j	The ln value of the sum of artist j's age until 2017	Wikipedia and press
Total track $volume_j$	The ln value of the total volume of tracks released by artist j in their career until the end of 2017	Spotify
$Week_t$	Week t (ranging from 1 to 104) of the relevant data	

1.3.2.3. Measuring engagement and sales

Consumers use reactions, shares and comments on social media to show their engagement with brands and firms (Kumar et al. 2010). This conceptualization of engagement is consistent with Kumar et al.'s (2013) construct "customer influence"—that is, a key dimension of engagement (Kumar 2018). Our conceptualization also aligns with previous research that finds strong correlations between reactions, shares and comments (Soares et al. 2022; Ji et al. 2017; Li et al. 2021).[4]

To measure music sales, we combined sales of digital songs, sales of albums (both digital and physical) and streaming. Recent literature has confirmed that streaming reflects music industry revenues (Wlömert and Papies 2016), music consumption (Datta et al. 2018), digital music sales (Aguiar and Martens 2016) and even physical album sales (Lee et al. 2020). While research on streaming has also found a cannibalizing effect of streaming adoption on other sales channels (e.g., Wlömert and Papies 2016; Aguiar and Waldfogel 2018) at the industry level, it still has a positive relationship with other types of sales at the artist and song levels (Aguiar and Waldfogel 2018).

[4] In addition to the conceptual reasons for aggregating reactions, comments and shares, there is also a conciseness reason. In our case, a concise measure is important, so we combine different sources of data in a multilevel mediation model. Cole and Preacher (2014) argue against the use of manifest variable paths because of measurement errors, especially in complex models. Accordingly, the use of latent variables formed by multiple measures is usually recommended.

After the ln-transformations, the three social media engagement measures: reactions (*M* = 7.68, *SD* = 4.59), comments (*M* = 4.7, *SD* = 3.2) and shares (*M* = 3.99, *SD* = 3.4) achieved high correlations (between .81 and .95, $p < .001$), high alpha, composite reliability (CR) and average variance extracted (AVE) (between .89 and .96) (See Appendix B). Thus, we created the engagement construct (*M* = 5.46, *SD* = 3.58) from the mean of these three ln-transformed variables.

Similarly, because of the high correlations (between .73 and .88, $p < .001$) and high alpha, CR and AVE (between .81 and .93) of the three sales-related constructs (See Appendix B)—that is, album sales (*M* = 2.69, *SD* = 2.2), digital songs sales (*M* = 4.28, *SD* = 2.68) and streaming (*M* = 10.77, *SD* = 3.38)—we created the sales construct (*M* = 5.91, *SD* = 2.58) from the mean of these three ln-transformed measures.

Finally, we tested the two-factor structure (with sales and engagement) for longitudinal measurement invariance. The high number of waves (104), the large number of parameters, and the modest sample size for each wave (384) made it technically challenging to conduct the tests for all the waves (Kyriazos 2018). Thus, we decided to test invariance at five equally-spaced times (weeks 20, 40, 60, 80, and 100), using Laavan in R (Mackinnon et al. 2022). Cross-sectional CFA's for each of these periods provide good fit indexes, with high alpha, CR and AVE for each latent construct (between .81 and .96) (See Web Appendix C) (Radanielina Hita et al. 2022). The longitudinal measurement invariance test confirms that our repeated constructs attain configural and metric invariance (Chen 2007; Putnick and Bornstein 2016) over these five periods (See Web Appendix E).

Thus, all our main variables (i.e., number of charity-related posts, number of mission-related posts, number of non-mission-related posts, engagement and sales), tweet volume, news volume and track volume are measured for each week for each artist (level 1 variables). The others (i.e., fan-based size, age, total track volume and experience) are assumed to remain the same over the whole period for a given artist (level 2 variables). Explanations of each construct are available in Table 2. Because of the collinearity issue between tweet volume and fanbase size (r = .77, p < .001) and between age and experience (r = .78, p < .001), tweet volume and experience were removed from the analyses (See Web Appendix F).

1.3.3. Multilevel mediation in MLmed

Central to our research is the multilevel mediation method proposed by Zhang et al. (2009). Building on this initial work, Hayes and Rockwood (2020) renamed it "multilevel conditional process analyses" and developed a macro in SPSS to help in the usage of these relatively complex analyses. The name of this macro is MLmed, for multilevel mediation, and we use it in this research. Given the multilevel structure of our data, MLmed—which relies on mixed linear modeling—is appropriate for the following reasons. MLmed accounts for the hierarchical, nested nature of our data, which is organized in two levels (i.e., week and artist). Importantly, MLmed simultaneously analyzes these two types of effects by separating each observation into two parts: the average effect for each artist (i.e., between-person or long-term effect) and the average difference between two weeks within each artist (i.e., within-person or short-term effect). Please see the work of Hayes and Rockwood (2020) for an effective summary of this analysis.

Our weekly constructs are of level 1 and are nested in different artists (level 2). Since our data is longitudinal (Tasca and Gallop 2009), MLmed allows us to simultaneously analyze the effects within each artist over time (i.e., within-person effects) and the variations from artist to artist (i.e. between-subject effects). Specifically, MLmed separates the cluster mean (i.e., the average effect for each artist: the between-person effect) from the deviation to the cluster mean for each observation (i.e., the difference between an observation and the average effect for each artist: the within-person effect). At the end, this procedure allows us to analyze both effects simultaneously. In addition, MLmed calculates multilevel indirect effects between the independent variable of interest (e.g., charity-related posts), the mediator (e.g., engagement) and the dependent variable (e.g., sales); and it tests their significance by conducting Monte-Carlo simulations. To the best of our knowledge, MLmed is one of the rare solutions that can test the significance of multilevel indirect effects.

In a classical longitudinal model, an observation in week t belongs to an artist j (i.e., cluster j). Charity-related posts_{tj} and engagement_{tj} are the respective observed values of the number of charity-related posts and the level engagement in week t for artist j. Charity-related posts_{j} and engagement_{j} refer to the cluster averages of charity-related posts and engagement for artist j. In MLmed, the within-person effects are the differences between the observed and mean values for an artist j (Zigler and Ye 2019). Charity-related $\text{posts}_{tj(w)}$ is the difference between charity-related posts_{tj} and charity-related posts_{j}, whereas engagement $_{tj(w)}$ is equal to engagement_{tj} minus engagement_{j}. Thus, charity-related $\text{posts}_{tj(w)}$ and $\text{engagement}_{tj(w)}$ capture the within-person variations effects, while charity-related posts_{j} and engagement_{j} show the between-person effects (Song 2018; Zigler and Ye 2019). Similarly, the dependent variable sales_{tj} is composed of sales_{j}, which is the cluster mean, and the $\text{sales}_{tj(w)}$, which is the deviation from the cluster mean. See Table 2 for formal definitions.

1.3.4. Test of hypotheses by using MLmed

We used MLmed to test our four hypotheses (see Table 1.3 for an overview of the results). In terms of within-person effects (i.e., short-term effects in our theory), charity-related posts ($\beta = .73$, $p < .001$) have a positive but weaker effect on engagement compared to mission-related posts ($\beta = 1.21$, $p < .001$) and non-mission related posts ($\beta = 1.59$, $p < .001$). These effects occur at level 1. According to these results, H1 is supported.

Table 1.3: Fixed effects and indirect effects of the mediation model

Independent variables	**β**	**t**	**B**	**t**
Within effects (short-term)	**On Engagement$_{tj(w)}$**		**On Sales$_{tj(w)}$**	
Constant	2.79	3.64***	5.2	5.42***
Charity-related posts$_{tj(w)}$	.73	11.1***	-.007	-.45
Mission-related posts$_{tj(w)}$	1.21	113.58***	.16	26.08***
Non-mission-related posts$_{tj(w)}$	1.59	136.73***	.06	7.99***
Engagement$_{tj(w)}$	---	---	.06	22.64***
Track volume$_{tj(w)}$	-.05	-1.94 .	.26	20.27***
Between effects (long-term)	**On Engagement$_j$**		**On Sales$_j$**	
Fanbase size$_j$	1.04	20.27***	1.3	14.3***
Charity-related posts$_j$	2.13	3.97***	-.15	-.22
Mission-related posts$_j$	1.11	8.59***	.02	1.47
Non-mission-related posts$_j$	1.65	15.55***	-.61	-3.64***
Engagement$_j$	---	---	.23	3.66***
Track volume$_j$	2.49	2.42*	-.85	-.67
Age$_j$	-.78	-3.22**	.02	.14
Total track volume$_j$	-.04	-.82	.16	2.54*
Charity-related posts -> Engagement -> Sales	**Coefficient**	**Z**	**p value**	**95% CI**
Within indirect effects (short-term)	.04	9.96	<.001	[.03, .05]
Between indirect effects (long-term)	.49	2.65	.008	[.18, .9]

Note: *** $p < 0.001$, ** $p< 0.01$, * $p< 0.05$, **.** $p < 0.1$; BIC = 240,729.9

In terms of between-person effects (i.e., the long-term effects in our theory), the effect of charity-related posts on engagement is positive and stronger (β = 2.13, p < .001) than that of mission-related posts (β = 1.11, p < .001) or non-mission related posts (β = 1.65, p < .001); these results occur at level 2. Accordingly, H2 is confirmed.

We test H3 and H4 by reporting the indirect effects of interest, as calculated by MLmed. The significance of the indirect effects is determined by using Monte-Carlo simulations (i.e., 10,000 samples) that produce 95% confidence intervals (CI).

In terms of within-person effects (i.e., short-term effects), the indirect effect "charity-related posts → engagement → sales" is positive and significant ($\beta = 0.04$, $p < .001$); in addition, the CI do not contain zero (*95% CI* [.03, .05]). We also test the within-person indirect effects involving the two other signals—that is, "mission-related posts → engagement → sales" ($\beta = 0.08$, $p < .001$; *95% CI* [.07, .09]) and "non-mission-related posts → engagement → sales" ($\beta = 0.1$, $p < .001$; *95% CI* [.08, .11].). Consistent with H3, in the short term, the indirect effect involving charity-related posts is weaker than the indirect effects involving the other two signals.

In terms of between-person effects (i.e., long-term effects), the indirect effect "charity-related posts → engagement → sales" is positive, significant and of large amplitude ($\beta = 0.49$, $p < .01$; *95% CI* [.18, .9]). The between-person indirect effects involving the other two signals—that is, "mission-related posts → engagement → sales" ($\beta = 0.26$, $p < .001$; *95% CI* [.11, .41] and "non-mission-related posts → engagement → sales" ($\beta = 0.38$, $p < .001$; *CI* [.18, .60]—are also positive and significant, although of lesser amplitude than the indirect effect involving charity-related signal. These results support H4.

1.3.5. Additional Analyses

1.3.5.1. Direct effects on sales

We also test the direct effects of the different signals on the level of sales—that is, the final outcomes. First, we ran a cross-sectional regression on $sales_j$, with the cluster mean of $sales_{tj}$ for each artist being the dependent variable (Table 1.4, Model 1), and the cluster means of other variables for each artist being the predictive variables. This model is the equivalent of testing between-subject (long-term) effects. Here, we note that charity-related $posts_j$ are not significant on $sales_j$; neither are mission-related $posts_j$. In turn, non-mission related $posts_j$ have a negative relationship with $sales_j$ ($\beta = -.55$, $p < .001$).

Table 1.4: Cross-sectional and longitudinal models on sales

	Cross-sectional (Model 1)		Longitudinal (Model 2)	
Independent variables	**β**	**t**	**β**	**t**
Within effects (intermittent, short-term)	**On $Sales_j$**		**On $Sales_{tj}$**	
Constant	5.27	5.61***	5.27	5.61***
Charity-related $posts_{tj}$	---	---	-.02	-1.32
Mission-related $posts_{tj}$	---	---	.14	22.90***
Non-mission-related $posts_{tj}$	---	---	.05	7.04***
$Engagement_{tj}$	---	---	.05	21.16***
News $volume_{tj}$	---	---	.20	21.21***
Track $volume_{tj}$	---	---	.23	18.23***
Between effects (long-term)	**On $Sales_j$**		**On $Sales_{tj}$**	
Charity-related $posts_j$	.14	.22	.16	0.25
Mission-related $posts_j$	.08	.45	-.07	-.39
Non-mission-related $posts_j$	-.55	-3.34***	-.60	-3.64***
$Engagement_j$	.16	2.55*	.11	1.71 .
News $volume_j$	1.51	3.86***	1.30	3.34***
Track $volume_j$	-1.18	-.94	-1.41	-1.12
Fanbase $size_j$	1.20	12.92***	1.20	12.92***
Age_j	-.96	-3.23**	-.96	-3.23**
Total track $volume_j$	1.20	12.92***	.16	2.59*

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, **.** $p < 0.1$
R^2 and Adjusted R^2 of Model 1 = .77; BIC of Model 2 = 93,579.05

Second, we conducted a mixed linear regression[5] on $sales_{tj}$, including both the cluster means of other variables (for each artist) and their observed values as predictive variables (Table 4, Model 2). Both charity-related $posts_j$ (within-person, or short-term effects) and charity-related $posts_{tj}$ (between-person, or long-term effects) are not significant. For this signal, it seems that there are only indirect effects on sales through engagement in the short and long terms (see results for H3 and H4).

[5] It should be noted that the mixed linear regression model (Table 4, Model 2) used the observed values of the independent variables (e.g., charity-related $posts_{tj}$) for short-term effects. In contrast, MLmed (Table 3) applied the differences between their observed and cluster mean values (e.g., charity-related $posts_{tj(w)}$) for short-term effects.

The effect of mission-related $posts_{tj}$ is positive in the short term ($\beta = 0.14$, $p < .001$), while mission-related $posts_j$ are not significant in the long term. This signal only has a direct effect on sales when it is done in the short term. In the long term, like charity-related signals, mission-related posts indirectly generate sales through engagement.

Finally, non-mission-related $posts_{tj}$ have a positive effect ($\beta = 0.05$, $p < .001$) in the short term, but non-mission-related $posts_j$ have a negative effect on $sales_{tj}$ ($\beta = -0.6$, $p <.001$) in the long term. At first sight, this result is surprising, and it will be discussed in detail in the general discussion.

1.3.5.2. Endogeneity test

To investigate possible omitted bias linked to our charity-related $posts_{tj}$ variable, we applied Kim and Frees' (2007) generalized method of moments. This technique applies to multilevel models without requiring external instrumental variables and is implemented through the R package REndo (Gui et al. 2021). It provides a reference random effects model (REF) along with two more robust models: the generalized method of moments (GMM) and a fixed effects (FE) model. Both tests comparing between REF and FE: $x^2(11, N = 39{,}936) = 333.2$, $p <.001$ and between GMM and FE: $x^2(10, N = 39{,}936) = 333.1$, $p < .001$ are significant. This indicates that there could be omitted variable bias. However, the parameters for the three models are analogous to the third decimal. This demonstrates that while there might be an omitted variable bias, our results are not affected and can be deemed robust.

1.3.5.3. The moderation effect of fanbase size

Additionally, we ran a *post hoc* analysis, with fanbase $size_j$ as a moderator in the link between charity-related $posts_j$ and $engagement_j$ in the sequence "charity-related posts – engagement – sales" (i.e., between-subject, or long-term effects). The interaction between fanbase $size_j$ and charity-related $posts_j$ on $engagement_j$ is positive and significant ($\beta = 1.24$, $p < .05$). The 95% confidence interval for the index of moderated mediation does not contain 0 (β = 1.1, *95% CI* [.17, 2.06]) (Hayes and Rockwood 2020)*;* this result indicates that the value of the indirect effect is different for different values of the moderators. Accordingly, we probe the indirect effects at different values of the moderating variable (Preacher et al. 2007). The long-term indirect effect of charity-related $posts_j$ on $sales_j$ is not significant when fanbase $size_j$ is less than 3, or when artists have less than one million Facebook fans. The results are significant when fanbase $size_j = 3$ (i.e., when artists have 1 – 5 million fans, $\beta = 1.67$, $p < .001$), and when fanbase $size_j = 4$, (i.e., when artists have more than 5 million fans, $\beta = 2.77$, $p < .001$). In other words, only when musicians have more than 1 million fans do their charity-related posts have a positive long-term indirect effect on sales through engagement. This effect increases as their fanbase increases.

1.3.5.4. Robustness check

To confirm the validity of our measures for engagement and sales, we ran a confirmatory factor analysis (CFA) on their respective component factors in the entire dataset. Two composite variables were created from their component factors, with each factor being multiplied by its respective loading (Lefcheck 2016). We used these new composite variables for sales and engagement and replicated our previous analyses with these new variables. All the earlier findings were replicated, confirming again H1-H4 and the moderation role of fanbase size (Web Appendix G).

1.4. General Discussion

1.4.1. Theoretical Implications

Despite the rich literature on the financial impact of charity initiatives and corporate social responsibility (CSR) programs, previous studies have often focused on the *direct relationship* between charity or CSR and performance indicators (e.g., Wang et al. 2008; van Beurden and

Gössling 2008; Clacher and Hagendorff 2012; Kang et al. 2016) by accounting for the influence of factors such as size, industry (van Beurden and Gössling 2008), marketing capability (Mishra and Modi 2016) and geographical differences (Lu et al. 2020). A separate literature has also examined charity or CSR *engagement* on social media (e.g. Kucukusta et al. 2019; Chu et al. 2020). This second literature considers engagement as an important goal to achieve so that firms can communicate their ethical practices and enhance their reputation. While Saxton and Guo (2020) have already discussed the mediating role of "social media capital" in the linkage between "social media presence" and "organizational outcomes" (Saxton and Guo 2020, p. 2), the current research takes an extra step by confirming the sequence *charity-related posts → engagement → sales* with field data.

Also, previous research on the impact of charity or CSR on financial performance has rarely separated short-term from long-term effects, resulting in possible biases. Some studies find a positive impact (Lev et al. 2009; Cai et al. 2012), while others report insignificant (Clacher and Hagendorff 2012) or even negative impacts (Sipilä et al. 2021). By analyzing short-term and long-term effects simultaneously through multilevel mediation (Wang and Maxwell 2015), we provide new insights into reconciling previous conflicting results. Indeed, we find that among musicians, charity signaling on social media does not have a significant direct effect on sales in either the short or the long term. This finding is also confirmed through cross-sectional and longitudinal models. Instead, charity signaling has a *positive indirect effect* on sales through social media engagement in the short and long terms. The indirect effect of charity signaling on sales (through engagement) is lower than the impacts of other types of signaling in the short term. Yet, in the long term, it is higher than other types of signaling, including mission signaling, despite the latter's focus on artists' core business.

While engagement drives sales, the motivations for engagement and sales are not always the same. Previous literature examined the impacts of content signals on engagement (Schreiner et al. 2021) and on sales (Babić Rosario et al. 2016; Yost et al. 2021), as well as the effect of engagement on sales (Yoon et al. 2018). The current research takes the extra step by studying three types of constructs (i.e., signals, engagement and sales) in tandem and by comparing their short-term versus long-term impacts.

For charity signaling, authenticity is important for engagement (Wymer and Akbar 2019), which impacts sales in the long term. In this case, engagement is linked to self-identification (Chapman et al. 2020). In other words, users are engaged with causes supported by musicians in order to express their own prosocial identities (Hitlin 2007). Signal authenticity, though related to, goes beyond signal reliability (Connelly et al. 2011). It considers the signaler's honesty and the fit between the signaler and the signals (Connelly et al. 2011), and it distinguishes the signaler from the others (Moulard et al. 2015). Charity signaling could thus be a powerful channel to distinguish brands. Previous literature has already examined philanthropy and CSR for brands' legitimacy building (Sánchez 2000; Werther and Chandler 2005). By comparing the long-term impact of charity signaling versus other signaling types on sales, we showcase its strategic potential.

We note that non-mission signaling has significant direct effects on sales in the short and long terms, though in opposite directions. That is, the short-term direct effect of non-mission signaling on sales is positive, but its long-term direct effect is negative. We explain this surprising result by arguing that the authenticity of non-mission related signals may decrease with repetition, which is the opposite for charity signaling. Non-mission-related signals can be the positions that artists take on different issues, such as their thoughts on current events, politics, sports, or society, or simply moments in their own lives. The rarity (Moulard et al. 2015) and spontaneity (Kreling et al. 2022) of such signals, rather than their regularity, may make them more authentic. However, the abundance of artists' messages on their personal lives and experiences over time could make the messages come across as planned and framed, which would result in less perceived authenticity and lower sales (Kreling et al. 2022).

Our research also contributes to the non-profit literature. Research on charity donations often focuses on the direct connections between the non-profits and their donors (Kumar and Chakrabarti 2023). By studying charity-advocating artists, we contribute to the understanding of the actors in the non-profit-related network by examining the role of an additional intermediary (i.e., supporting artists). The exploration of a finer-grained conceptualization of the different actors involved in charity donation is important as non-profit research keeps growing as a field.

1.4.2. Managerial Implications

Given the importance of charity signaling on sales, musicians would benefit from making it a significant part of their long-term branding strategy. The public's scrutiny of celebrities (Dieter and Kumar 2008; Duvall 2015; Haynes 2014) could make them hesitate to voice their advocacy. Even in the short term, when its indirect impact is modest, charity signaling already contributes to sales. This contribution becomes much greater over time when the signals are consistent. It should be noted that only when artists achieve significant stardom will charity signaling help their engagement and sales in the long run. Future research could examine human brands in different areas—such as entertainment, fashion and sports—to see whether the findings are valid in other contexts and industries. Furthermore, scholars could also examine possible circumstances where charity signaling is an effective long-term branding channel for smaller brands.

Also, while previous literature has stressed the importance of bonding with consumers through content marketing (Geng et al. 2020), we show that self-revealing actions by artists should be done with care because of their negative impact on sales over time, as their authenticity may weaken in the long term. This is the case observed for non-mission-related signals. For corporate brands, the bonding mechanism and authenticity perception might not follow the same pattern (Napoli et al. 2014). Authenticity in corporate brands (Södergren 2021) refers to at least one of three aspects: brands' honesty (Morhart et al. 2015), consumers' emotional attachment to them (Beverland 2005) and their history as cultural icons (Holt 2004). Even for humanized corporate brands, consumers perceive them through brand stimuli (e.g., logos, slogans and interactions) and their own individual inferences (Sharma and Rahman 2022). Thus, future research could investigate the long-term impact of non-mission-related signaling among corporate brands. The choice of stimuli, decisions on content, interactions and consumers' expectations could all influence perceptions of authenticity over time.

Our research also has managerial implications for non-profits. Their partnerships with celebrities tend to stop at the philanthropic (e.g., Thrall et al. 2008) and transaction (e.g., volunteering or cause-related marketing, Thamaraiselvan et al. 2017) levels (Austin and Seitanidi 2012). Weak partnerships could lead to criticisms of inefficiency (Kane et al. 2009; Dieter and Kumar 2008; Duvall 2015). Even though artists have plenty of good will and connections, they may lack the necessary expertise on issues and the required magnitude of activism (Alexander 2013; Bennett 2014; Haynes 2014). However, with their long-term economic self-interest linked to the social good, artists will be more motivated to co-create values with non-profits (Schiller and Almog-Bar 2013). In other words, there is potential for higher-level partnerships to ensure sustainable results for both artists and charities. Future research could also study options for such partnerships. Examples are transformative partnerships (Austin and Seitanidi 2012) or partnerships with more than one non-profit or more than one celebrity to maximize each partner's strengths.

1.4.3. Future Research Avenues

The link between charity advocacy and artists' long-term economic interests also means that non-profits could be in the position to be more selective about co-branding celebrities. In addition, smaller non-profits with unique value offerings could also benefit, as artists will need to choose the causes they truly care about in long-term partnerships, rather than selecting well-known charities for short-term publicity. Such diverse dynamics could also be a promising area for research. As artists wish to distinguish themselves in terms of charity advocacy, and as non-profits have increasing bargaining powers, the partnerships between these two entities are likely to become much more complex in years to come.

A popular aspect in the co-branding literature is the fit between celebrities and non-profit causes because of its effect on perceived authenticity (e.g., Ilicic and Baxter 2014; Park and Cho 2015). This "fit" could change over time, as continuous active support by celebrities could improve perceived fit. Future research could compare the impact of such changes on non-profit donations when the fit stays static versus when it changes over time. Here, a strengthening fit, rather than the non-changing fit, could indicate the growing influence of a cause, which, in turn, could generate more donations. Similarly, the fit could weaken over time, when for example, the celebrities change their priorities. In co-branding relationships with more than one celebrity, how might the presence of both weakening and strengthening celebrity-cause fits influence donations? Answering these questions might help non-profits to better prepare their long-term strategies.

While we focus on charity-advocating artists, other scholars could investigate the multi-actor network which is built around non-profits. For instance, future research could examine the role of social influencers for or against a cause, corporate partners, celebrity partners, governments, experts, consumers, and of course, non-profits (Van Royen et al. 2022). The examination of topics related to multilateral exchanges, benefits and conflicts, network hierarchy, and simultaneous and sequential actions (Bruijn and Heuvelhof 2018) could produce rich findings insights for non-profits. Future research could also examine how artists can influence the norms linked to a non-profit (Kumar and Chakrabarti 2023) in the context of diverse and conflicting discourses. The impact of artists could become more prominent if they strengthen their partnerships with non-profits in a way that goes beyond simple awareness-raising. Building such partnerships would require that non-profits effectively coordinate the roles of their partners at different stages of their relationship over time.

References

Adkins, S. (1999). *Cause related marketing: Who cares wins*. Oxford; Boston: Butterworth-Heinemann.

Aguiar, L., & Martens, B. (2016). Digital music consumption on the Internet: Evidence from clickstream data. *Information Economics and Policy*, *34*, 27–43. https://doi.org/10.1016/j.infoecopol.2016.01.003

Aguiar, L., & Waldfogel, J. (2018). As streaming reaches flood stage, does it stimulate or depress music sales? *International Journal of Industrial Organization*, *57*, 278–307. https://doi.org/10.1016/j.ijindorg.2017.06.004

Alexander, J. (2013). The case of the green vampire: Eco-celebrity, Twitter and youth engagement. *Celebrity Studies*, *4*(3), 353–368. https://doi.org/10.1080/19392397.2013.831625

Anderson, A. (2011). Sources, media, and modes of climate change communication: The role of celebrities. *WIREs Climate Change*, *2*(4), 535–546. https://doi.org/10.1002/wcc.119

Austin, J. E., & Seitanidi, M. M. (2012). Collaborative Value Creation: A Review of Partnering Between Nonprofits and Businesses: Part I. Value Creation Spectrum and Collaboration Stages. *Nonprofit and Voluntary Sector Quarterly*, *41*(5), 726–758. https://doi.org/10.1177/0899764012450777

Aw, E. C.-X., & Labrecque, L. I. (2020). Celebrity endorsement in social media contexts: understanding the role of parasocial interactions and the need to belong. *Journal of Consumer Marketing*, *37*(7), 895–908. https://doi.org/10.1108/JCM-10-2019-3474

Babiak, K., Mills, B., Tainsky, S., & Juravich, M. (2012). An Investigation into Professional Athlete Philanthropy: Why Charity Is Part of the Game. *Journal of Sport Management*, *26*(2), 159–176. https://doi.org/10.1123/jsm.26.2.159

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. *Journal of Marketing Research*, *53*(3), 297–318. https://doi.org/10.1509/jmr.14.0380

Bennett, L. (2014). 'If we stick together we can do anything': Lady Gaga fandom, philanthropy and activism through social media. *Celebrity Studies*, *5*(1–2), 138–152. https://doi.org/10.1080/19392397.2013.813778

Bernritter, S. F., Verlegh, P. W. J., & Smit, E. G. (2016). Why Nonprofits Are Easier to Endorse on Social Media: The Roles of Warmth and Brand Symbolism. *Journal of Interactive Marketing*, *33*, 27–42. https://doi.org/10.1016/j.intmar.2015.10.002

Beverland, M. B. (2005). Crafting Brand Authenticity: The Case of Luxury Wines*. *Journal of Management Studies*, *42*(5), 1003–1029. https://doi.org/10.1111/j.1467-6486.2005.00530.x

Bhati, A., & McDonnell, D. (2020). Success in an Online Giving Day: The Role of Social Media in Fundraising. *Nonprofit and Voluntary Sector Quarterly*, *49*(1), 74–92. https://doi.org/10.1177/0899764019868849

Block, P. (2009). *Community: the structure of belonging*. San Francisco: Berrett-Koehler Publ.

Bourdieu, P., Thompson, J. B., & Raymond, G. (2003). *Language and symbolic power* (Reprinted.). Cambridge: Harvard Univ. Press.

Bruijn, J. A. de, & Heuvelhof, E. F. ten. (2018). *Management in networks: on multi-actor decision making* (Second edition.). New York: Routledge.

Burton, N. W., Haynes, M., Wilson, L.-A. M., Giles-Corti, B., Oldenburg, B. F., Brown, W. J., et al. (2009). HABITAT: A longitudinal multilevel study of physical activity change in mid-aged adults. *BMC Public Health*, *9*(1), 76. https://doi.org/10.1186/1471-2458-9-76

Cai, Y., Jo, H., & Pan, C. (2012). Doing Well While Doing Bad? CSR in Controversial Industry Sectors. *Journal of Business Ethics*, *108*(4), 467–480. https://doi.org/10.1007/s10551-011-1103-7

Chapman, C. M., Masser, B. M., & Louis, W. R. (2020). Identity motives in charitable giving: Explanations for charity preferences from a global donor survey. *Psychology & Marketing*, *37*(9), 1277–1291. https://doi.org/10.1002/mar.21362

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chu, S.-C., Chen, H.-T., & Gan, C. (2020). Consumers' engagement with corporate social responsibility (CSR) communication in social media: Evidence from China and the United States. *Journal of Business Research*, *110*, 260–271. https://doi.org/10.1016/j.jbusres.2020.01.036

Chung, S., & Cho, H. (2017). Fostering Parasocial Relationships with Celebrities on Social Media: Implications for Celebrity Endorsement: CELEBRITY PARASOCIAL RELATIONSHIPS ON SOCIAL MEDIA. *Psychology & Marketing*, *34*(4), 481–495. https://doi.org/10.1002/mar.21001

Chwialkowska, A. (2019). The Effectiveness of Brand- and Customer-Centric Content Strategies at Generating Shares, 'Likes', and Comments. *Journal of Promotion Management*, *25*(2), 270–300. https://doi.org/10.1080/10496491.2018.1443307

Clacher, I., & Hagendorff, J. (2012). Do Announcements About Corporate Social Responsibility Create or Destroy Shareholder Wealth? Evidence from the UK. *Journal of Business Ethics*, *106*(3), 253–266. https://doi.org/10.1007/s10551-011-1004-9

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300–315. https://doi.org/10.1037/a0033805

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling Theory: A Review and Assessment. *Journal of Management*, *37*(1), 39–67. https://doi.org/10.1177/0149206310388419

Curran, P. J., & Bauer, D. J. (2011). The Disaggregation of Within-Person and Between-Person Effects in Longitudinal Models of Change. *Annual Review of Psychology*, *62*(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356

Datta, H., Knox, G., & Bronnenberg, B. J. (2018). Changing Their Tune: How Consumers' Adoption of Online Streaming Affects Music Consumption and Discovery. *Marketing Science*, *37*(1), 5–21. https://doi.org/10.1287/mksc.2017.1051

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019a). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*. http://arxiv.org/abs/1810.04805. Accessed 5 July 2021

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019b, May 24). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv. http://arxiv.org/abs/1810.04805. Accessed 2 July 2022

Dieter, H., & Kumar, R. (2008). The Downside of Celebrity Diplomacy: The Neglected Complexity of Development. *Global Governance: A Review of Multilateralism and International Organizations*, *14*(3), 259–264. https://doi.org/10.1163/19426720-01403001

Dixon, E. (2019, April 24). Lego releases Braille bricks to teach blind and visually impaired children | CNN. *CNN*. https://www.cnn.com/2019/04/24/health/lego-braille-bricks-scli-intl/index.html. Accessed 19 January 2023

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Ellison, N. B., & Vitak, J. (2015). Social Network Site Affordances and Their Relationship to Social Capital Processes. In S. S. Sundar (Ed.), *The Handbook of the Psychology of*

Communication Technology (1st ed., pp. 203–227). Wiley. https://doi.org/10.1002/9781118426456.ch9

Erdem, T., Swait, J., & Valenzuela, A. (2006). Brands as Signals: A Cross-Country Validation Study. *Journal of Marketing*, *70*(1), 34–49. https://doi.org/10.1509/jmkg.70.1.034.qxd

Frey, B. S., & Meier, S. (2004). Pro-social behavior in a natural setting. *Journal of Economic Behavior & Organization*, *54*(1), 65–88. https://doi.org/10.1016/j.jebo.2003.10.001

Geng, R., Wang, S., Chen, X., Song, D., & Yu, J. (2020). Content marketing in e-commerce platforms in the internet celebrity economy. *Industrial Management & Data Systems*, *120*(3), 464–485. https://doi.org/10.1108/IMDS-05-2019-0270

Gong, W., & Li, X. (2017). Engaging fans on microblog: the synthetic influence of parasocial interaction and source characteristics on celebrity endorsement: GONG and LI. *Psychology & Marketing*, *34*(7), 720–732. https://doi.org/10.1002/mar.21018

Gui, R., Markus, M., Algesheimer, R., & Schilter, P. (2021). REndo: A Package to Address Endogeneity Without External Instrumental Variables.

Guo, C., & Saxton, G. D. (2014). Tweeting Social Change: How Social Media Are Changing Nonprofit Advocacy. *Nonprofit and Voluntary Sector Quarterly*, *43*(1), 57–79. https://doi.org/10.1177/0899764012471585

Hair Jr., J. F., & Fávero, L. P. (2019). Multilevel modeling for longitudinal data: concepts and applications. *RAUSP Management Journal*, *54*(4), 459–489. https://doi.org/10.1108/RAUSP-04-2019-0059

Harlow, S., & Benbrook, A. (2019). How #Blacklivesmatter: exploring the role of hip-hop celebrities in constructing racial identity on Black Twitter. *Information, Communication & Society*, *22*(3), 352–368. https://doi.org/10.1080/1369118X.2017.1386705

Hasan, I., Kobeissi, N., Liu, L., & Wang, H. (2018). Corporate Social Responsibility and Firm Financial Performance: The Mediating Role of Productivity. *Journal of Business Ethics*, *149*(3), 671–688. https://doi.org/10.1007/s10551-016-3066-1

Hayes, A. F., & Rockwood, N. J. (2020). Conditional Process Analysis: Concepts, Computation, and Advances in the Modeling of the Contingencies of Mechanisms. *American Behavioral Scientist*, *64*(1), 19–54. https://doi.org/10.1177/0002764219859633

Haynes, D. F. (2014). The Celebritization of Human Trafficking. *The ANNALS of the American Academy of Political and Social Science*, *653*(1), 25–45. https://doi.org/10.1177/0002716213515837

Heil, O., & Robertson, T. S. (1991). Toward a theory of competitive market signaling: A research agenda. *Strategic Management Journal*, *12*(6), 403–418. https://doi.org/10.1002/smj.4250120602

Herzog, P. S., & Yang, S. (2018). Social Networks and Charitable Giving: Trusting, Doing, Asking, and Alter Primacy. *Nonprofit and Voluntary Sector Quarterly*, *47*(2), 376–394. https://doi.org/10.1177/0899764017746021

Higgins, M. C., & Gulati, R. (2006). Stacking the deck: the effects of top management backgrounds on investor decisions. *Strategic Management Journal*, *27*(1), 1–25. https://doi.org/10.1002/smj.495

Hitlin, S. (2007). Doing good, feeling good: Values and the self 's moral center. *The Journal of Positive Psychology*, *2*(4), 249–259. https://doi.org/10.1080/17439760701552352

Holt, D. B. (2004). *How brands become icons: the principles of cultural branding*. Boston, Mass: Harvard Business School Press.

Ilicic, J., & Baxter, S. (2014). Fit in celebrity-charity alliances: when perceived celanthropy benefits nonprofit organisations: Influence on attitudes and donation intentions. *International Journal of Nonprofit and Voluntary Sector Marketing*, *19*(3), 200–208. https://doi.org/10.1002/nvsm.1497

Janney, J. J., & Folta, T. B. (2003). Signaling through private equity placements and its impact on the valuation of biotechnology firms. *Journal of Business Venturing*, *18*(3), 361–380. https://doi.org/10.1016/S0883-9026(02)00100-3

Ji, Y. G., Li, C., North, M., & Liu, J. (2017). Staking reputation on stakeholders: How does stakeholders' Facebook engagement help or ruin a company's reputation? *Public Relations Review*, *43*(1), 201–210. https://doi.org/10.1016/j.pubrev.2016.12.004

Kane, J., Patapan, H., & 't Hart, P. (2009). *Dispersed Democratic Leadership*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199562992.001.0001

Kang, C., Germann, F., & Grewal, R. (2016). Washing Away Your Sins? Corporate Social Responsibility, Corporate Social Irresponsibility, and Firm Performance. *Journal of Marketing*, *80*(2), 59–79. https://doi.org/10.1509/jm.15.0324

Kelly, S. P., Morgan, G. G., & Coule, T. M. (2014). Celebrity altruism: the good, the bad and the ugly in relationships with fundraising charities: Celebrity altruism. *International Journal of Nonprofit and Voluntary Sector Marketing*, *19*(2), 57–75. https://doi.org/10.1002/nvsm.1487

Kim, J.-S., & Frees, E. W. (2007). MULTILEVEL MODELING WITH CORRELATED EFFECTS. *Psychometrika*, *72*(4), 505–533. https://doi.org/10.1007/S11336-007-9008-1

Knoll, J., & Matthes, J. (2017). The effectiveness of celebrity endorsements: a meta-analysis. *Journal of the Academy of Marketing Science*, *45*(1), 55–75. https://doi.org/10.1007/s11747-016-0503-8

Kreling, R., Meier, A., & Reinecke, L. (2022). Feeling Authentic on Social Media: Subjective Authenticity Across Instagram Stories and Posts. *Social Media + Society*, *8*(1), 205630512210862. https://doi.org/10.1177/20563051221086235

Kucukusta, D., Perelygina, M., & Lam, W. S. (2019). CSR communication strategies and stakeholder engagement of upscale hotels in social media. *International Journal of Contemporary Hospitality Management*, *31*(5), 2129–2148. https://doi.org/10.1108/IJCHM-06-2018-0484

Kumar, A., & Chakrabarti, S. (2023). Charity Donor Behavior: A Systematic Literature Review and Research Agenda. *Journal of Nonprofit & Public Sector Marketing*, *35*(1), 1–46. https://doi.org/10.1080/10495142.2021.1905134

Kumar, V. (2018). A Theory of Customer Valuation: Concepts, Metrics, Strategy, and Implementation. *Journal of Marketing*, *82*(1), 1–19. https://doi.org/10.1509/jm.17.0208

Kumar, V., Aksoy, L., Donkers, B., Venkatesan, R., Wiesel, T., & Tillmanns, S. (2010). Undervalued or Overvalued Customers: Capturing Total Customer Engagement Value. *Journal of Service Research*, *13*(3), 297–310. https://doi.org/10.1177/1094670510375602

Kumar, V., Bhaskaran, V., Mirchandani, R., & Shah, M. (2013). **Practice Prize Winner** — Creating a Measurable Social Media Marketing Strategy: Increasing the Value and ROI of Intangibles and Tangibles for Hokey Pokey. *Marketing Science*, *32*(2), 194–212. https://doi.org/10.1287/mksc.1120.0768

Kumar, V., & Pansari, A. (2016). Competitive Advantage through Engagement. *Journal of Marketing Research*, *53*(4), 497–514. https://doi.org/10.1509/jmr.15.0044

Kyriazos, T. A. (2018). Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. *Psychology*, *09*(08), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Lambert, N. M., Stillman, T. F., Hicks, J. A., Kamble, S., Baumeister, R. F., & Fincham, F. D. (2013). To Belong Is to Matter: Sense of Belonging Enhances Meaning in Life. *Personality and Social Psychology Bulletin*, *39*(11), 1418–1427. https://doi.org/10.1177/0146167213499186

Lee, M., Choi, H. S., Cho, D., & Lee, H. (2020). Can digital consumption boost physical consumption? The effect of online music streaming on record sales. *Decision Support Systems*, *135*, 113337. https://doi.org/10.1016/j.dss.2020.113337

Lefcheck, J. S. (2016). PIECEWISESEM : Piecewise structural equation modelling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579. https://doi.org/10.1111/2041-210X.12512

Lev, B., Petrovits, C., & Radhakrishnan, S. (2009). Is doing good good for you? how corporate charitable contributions enhance revenue growth. *Strategic Management Journal*, n/a-n/a. https://doi.org/10.1002/smj.810

Li, J. (Justin), Kim, W. G., & Choi, H. M. (2021). Effectiveness of social media marketing on enhancing performance: Evidence from a casual-dining restaurant setting. *Tourism Economics*, *27*(1), 3–22. https://doi.org/10.1177/1354816619867807

Lin, N. (2019). Buidling a Network Theory of Social Capital. In *Social Capital, Social Support and Stratification* (pp. 50–76). Edward Elgar Publishing. https://doi.org/10.4337/9781789907285.00009

Liston-Heyes, C., & Liu, G. (2010). Cause-Related Marketing in the Retail and Finance Sectors: An Exploratory Study of the Determinants of Cause Selection and Nonprofit Alliances. *Nonprofit and Voluntary Sector Quarterly*, *39*(1), 77–101. https://doi.org/10.1177/0899764008326680

Liu, G., Liston-Heyes, C., & Ko, W.-W. (2010). Employee Participation in Cause-Related Marketing Strategies: A Study of Management Perceptions from British Consumer Service Industries. *Journal of Business Ethics*, *92*(2), 195–210. https://doi.org/10.1007/s10551-009-0148-3

Loewenstein, J. (2019). Surprise, Recipes for Surprise, and Social Influence. *Topics in Cognitive Science*, *11*(1), 178–193. https://doi.org/10.1111/tops.12312

Lovejoy, K., & Saxton, G. D. (2012). Information, Community, and Action: How Nonprofit Organizations Use Social Media*. *Journal of Computer-Mediated Communication*, *17*(3), 337–353. https://doi.org/10.1111/j.1083-6101.2012.01576.x

Lu, J., Liang, X., & Wang, H. (2020). Geographical influences on the relationship between corporate philanthropy and corporate financial performance. *Regional Studies*, *54*(5), 660–676. https://doi.org/10.1080/00343404.2019.1668551

MacDonald, R. A. R., Hargreaves, D. J., & Miell, D. (Eds.). (2002). *Musical identities*. Oxford ; New York: Oxford University Press.

Mackinnon, S., Curtis, R., & O’Connor, R. (2022). Tutorial in Longitudinal Measurement Invariance and Cross-lagged Panel Models Using Lavaan. *Meta-Psychology*, *6*. https://doi.org/10.15626/MP.2020.2595

McQuitty, S. (2004). Statistical power and structural equation models in business research. *Journal of Business Research*, *57*(2), 175–183. https://doi.org/10.1016/S0148-2963(01)00301-0

Mishra, S., & Modi, S. B. (2016). Corporate Social Responsibility and Shareholder Wealth: The Role of Marketing Capability. *Journal of Marketing*, *80*(1), 26–46. https://doi.org/10.1509/jm.15.0013

Moerbeek, M. (2004). The Consequence of Ignoring a Level of Nesting in Multilevel Analysis. *Multivariate Behavioral Research*, *39*(1), 129–149. https://doi.org/10.1207/s15327906mbr3901_5

Morhart, F., Malär, L., Guèvremont, A., Girardin, F., & Grohmann, B. (2015). Brand authenticity: An integrative framework and measurement scale. *Journal of Consumer Psychology*, *25*(2), 200–218. https://doi.org/10.1016/j.jcps.2014.11.006

Moulard, J. G., Garrity, C. P., & Rice, D. H. (2015). What Makes a Human Brand Authentic? Identifying the Antecedents of Celebrity Authenticity: ANTECEDENTS OF CELEBRITY AUTHENTICITY. *Psychology & Marketing*, *32*(2), 173–186. https://doi.org/10.1002/mar.20771

Napoli, J., Dickinson, S. J., Beverland, M. B., & Farrelly, F. (2014). Measuring consumer-based brand authenticity. *Journal of Business Research*, *67*(6), 1090–1098. https://doi.org/10.1016/j.jbusres.2013.06.001

Nepomuceno, M. V., Visconti, L. M., & Cenesizoglu, T. (2020). A model for investigating the impact of owned social media content on commercial performance and its application in large and mid-sized online communities. *Journal of Marketing Management*, *36*(17–18), 1762–1804. https://doi.org/10.1080/0267257X.2020.1825112

Osorio, M. L., Centeno, E., & Cambra-Fierro, J. (2020). A thematic exploration of human brands: literature review and agenda for future research. *Journal of Product & Brand Management*, *29*(6), 695–714. https://doi.org/10.1108/JPBM-02-2019-2274

Park, S., & Cho, M. (2015). Celebrity Endorsement for Nonprofit Organizations: The Role of Celebrity Motive Attribution and Spontaneous Judgment of Celebrity-Cause Incongruence. *Journal of Promotion Management*, *21*(2), 224–245. https://doi.org/10.1080/10496491.2014.996802

Phau, I., & Prendergast, G. (2000). Consuming luxury brands: The relevance of the 'Rarity Principle.' *Journal of Brand Management*, *8*(2), 122–138. https://doi.org/10.1057/palgrave.bm.2540013

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227. https://doi.org/10.1080/00273170701341316

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*(3), 209–233. https://doi.org/10.1037/a0020141

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Radanielina Hita, M. L. , Grégoire, Y., Lussier, B., Boissonneault, S., Vandenberghe, C., & Sénécal, S. (2022). An extended health belief model for COVID-19: understanding the media-based processes leading to social distancing and panic buying. *Journal of the Academy of Marketing Science*.

Sánchez, C. M. (2000). Motives for Corporate Philanthropy in El Salvador: Altruism and Political Legitimacy. *Journal of Business Ethics*, *27*(4), 363–375. https://doi.org/10.1023/A:1006169005234

Santos, A. L., Barros, F., & Azevedo, A. (2019). Matching-up celebrities' brands with products and social causes. *Journal of Product & Brand Management*, *28*(2), 242–255. https://doi.org/10.1108/JPBM-03-2017-1439

Saxton, G. D., Gomez, L., Ngoh, Z., Lin, Y.-P., & Dietrich, S. (2019). Do CSR Messages Resonate? Examining Public Reactions to Firms' CSR Efforts on Social Media. *Journal of Business Ethics*, *155*(2), 359–377. https://doi.org/10.1007/s10551-017-3464-z

Saxton, G. D., & Guo, C. (2020). Social media capital: Conceptualizing the nature, acquisition, and expenditure of social media-based organizational resources. *International Journal of Accounting Information Systems*, *36*, 100443. https://doi.org/10.1016/j.accinf.2019.100443

Saxton, G. D., & Wang, L. (2014). The Social Network Effect: The Determinants of Giving Through Social Media. *Nonprofit and Voluntary Sector Quarterly*, *43*(5), 850–868. https://doi.org/10.1177/0899764013485159

Schembri, S., Merrilees, B., & Kristiansen, S. (2010). Brand consumption and narrative of the self: Brand Consumption and Narrative of the Self. *Psychology & Marketing*, *27*(6), 623–637. https://doi.org/10.1002/mar.20348

Schiller, R. S., & Almog-Bar, M. (2013). Revisiting Collaborations Between Nonprofits and Businesses: An NPO-Centric View and Typology. *Nonprofit and Voluntary Sector Quarterly*, *42*(5), 942–962. https://doi.org/10.1177/0899764012471753

Schreiner, M., Fischer, T., & Riedl, R. (2021). Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda. *Electronic Commerce Research*, *21*(2), 329–345. https://doi.org/10.1007/s10660-019-09353-8

Sharma, M., & Rahman, Z. (2022). Anthropomorphic brand management: An integrated review and research agenda. *Journal of Business Research*, *149*, 463–475. https://doi.org/10.1016/j.jbusres.2022.05.039

Sipilä, J., Alavi, S., Edinger-Schons, L. M., Dörfer, S., & Schmitz, C. (2021). Corporate social responsibility in luxury contexts: potential pitfalls and how to overcome them. *Journal of the Academy of Marketing Science*, *49*(2), 280–303. https://doi.org/10.1007/s11747-020-00755-x

Soares, J. C., Limongi, R., & Cohen, E. D. (2022). Engagement in a social media: an analysis in higher education institutions. *Online Information Review*, *46*(2), 256–284. https://doi.org/10.1108/OIR-06-2020-0242

Södergren, J. (2021). Brand authenticity: 25 Years of research. *International Journal of Consumer Studies*, *45*(4), 645–663. https://doi.org/10.1111/ijcs.12651

Song, H. (2018). A Primer on Multilevel Mediation Models for Egocentric Social Network Data. *Communication Methods and Measures*, *12*(1), 1–24. https://doi.org/10.1080/19312458.2017.1416343

Song, S., & Kim, H.-Y. (2020). Celebrity endorsements for luxury brands: followers vs. non-followers on social media. *International Journal of Advertising*, *39*(6), 802–823. https://doi.org/10.1080/02650487.2020.1759345

Street, J., Hague, S., & Savigny, H. (2008). Playing to the Crowd: The Role of Music and Musicians in Political Participation. *The British Journal of Politics and International Relations*, *10*(2), 269–285. https://doi.org/10.1111/j.1467-856x.2007.00299.x

Tan, Y., Fan, Z., Wei, X., & Yang, T. (2022). School Belonging and Reading Literacy: A Multilevel Moderated Mediation Model. *Frontiers in Psychology*, *13,* 816128. https://doi.org/10.3389/fpsyg.2022.816128

Tasca, G. A., & Gallop, R. (2009). Multilevel modeling of longitudinal data for psychotherapy researchers: I. The basics. *Psychotherapy Research*, *19*(4–5), 429–437. https://doi.org/10.1080/10503300802641444

Thamaraiselvan, N., Arasu, B. S., & Inbaraj, J. D. (2017). Role of celebrity in cause related marketing. *International Review on Public and Nonprofit Marketing*, *14*(3), 341–357. https://doi.org/10.1007/s12208-017-0176-0

Thrall, A. T., Lollio-Fakhreddine, J., Berent, J., Donnelly, L., Herrin, W., Paquette, Z., et al. (2008). Star Power: Celebrity Advocacy and the Evolution of the Public Sphere. *The International Journal of Press/Politics*, *13*(4), 362–385. https://doi.org/10.1177/1940161208319098

Van Royen, K., Pabian, S., Poels, K., & De Backer, C. (2022). Around the same table: Uniting stakeholders of food-related communication. *Appetite, 173,* 105998. https://doi.org/10.1016/j.appet.2022.105998

Vanhamme, J., Lindgreen, A., Reast, J., & van Popering, N. (2012). To Do Well by Doing Good: Improving Corporate Image Through Cause-Related Marketing. *Journal of Business Ethics*, *109*(3), 259–274. https://doi.org/10.1007/s10551-011-1134-0

Wang, H., Choi, J., & Li, J. (2008). Too Little or Too Much? Untangling the Relationship Between Corporate Philanthropy and Firm Financial Performance. *Organization Science*, *19*(1), 143–159. https://doi.org/10.1287/orsc.1070.0271

Wang, L. (Peggy), & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63–83. https://doi.org/10.1037/met0000030

Wang, Y.-Y., Guo, C., Susarla, A., & Sambamurthy, V. (2021). Online to Offline: The Impact of Social Media on Offline Sales in the Automobile Industry. *Information Systems Research*, *32*(2), 582–604. https://doi.org/10.1287/isre.2020.0984

Waterman, M. (1996). Emotional Responses to Music: Implicit and Explicit Effects in Listeners and Performers. *Psychology of Music*, *24*(1), 53–67. https://doi.org/10.1177/0305735696241006

Werther, W. B., & Chandler, D. (2005). Strategic corporate social responsibility as global brand insurance. *Business Horizons*, *48*(4), 317–324. https://doi.org/10.1016/j.bushor.2004.11.009

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial Invariance Within Longitudinal Structural Equation Models: Measuring the Same Construct Across Time. *Child Development Perspectives*, *4*(1), 10–18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Wlömert, N., & Papies, D. (2016a). On-demand streaming services and music industry revenues — Insights from Spotify's market entry. *International Journal of Research in Marketing*, *33*(2), 314–327. https://doi.org/10.1016/j.ijresmar.2015.11.002

Wlömert, N., & Papies, D. (2016b). On-demand streaming services and music industry revenues — Insights from Spotify's market entry. *International Journal of Research in Marketing*, *33*(2), 314–327. https://doi.org/10.1016/j.ijresmar.2015.11.002

Wymer, W., & Akbar, M. M. (2019). Brand Authenticity's Influence on Charity Support Intentions. *Journal of Nonprofit & Public Sector Marketing*, *31*(5), 507–527. https://doi.org/10.1080/10495142.2018.1526754

Yoon, G., Li, C., Ji, Y. (Grace), North, M., Hong, C., & Liu, J. (2018). Attracting Comments: Digital Engagement Metrics on Facebook and Financial Performance. *Journal of Advertising*, *47*(1), 24–37. https://doi.org/10.1080/00913367.2017.1405753

Yost, E., Zhang, T., & Qi, R. (2021). The power of engagement: Understanding active social media engagement and the impact on sales in the hospitality industry. *Journal of Hospitality and Tourism Management*, *46*, 83–95. https://doi.org/10.1016/j.jhtm.2020.10.008

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing Multilevel Mediation Using Hierarchical Linear Models: Problems and Solutions. *Organizational Research Methods*, *12*(4), 695–719. https://doi.org/10.1177/1094428108327450

Zigler, C. K., & Ye, F. (2019). A Comparison of Multilevel Mediation Modeling Methods: Recommendations for Applied Researchers. *Multivariate Behavioral Research*, *54*(3), 338–359. https://doi.org/10.1080/00273171.2018.1527676

Web Appendices

A. Classification on the human-coded dataset (5,413 posts)

Table A1: Inter-coder agreement

Charity	Merchandise	Explicit selling	Show
97.62%	87.88%	82.30%	81.80%

Table A2: Classification quality of the four constructs using the BERT approach

Variable	Charity	Explicit Selling	Show	Merchandise
Precision	.98	.83	.83	.78
Recall	.98	.83	.83	.76
F1 Score	.98	.83	.82	.75

An issue of working with Big Data is that it is very time consuming to conduct content analysis on hundreds of thousands of social media messages to identify appropriate variables. Thus, we used a small dataset (5,413 posts) already classified on the variables of interest, i.e., whether a post is related to charity, explicit selling, merchandise, or shows (Nepomuceno et al. 2020), to train four respective models. For this, we applied Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al. 2019) in Python. We then employed the trained model to predict the same variables on the main dataset of 322,589 posts. Compared with many previous methods, BERT is superior in its classification quality, thanks to its pre-training model absorbing a very large English corpus and its ability to learn information from both the left and right sides of a word.

B. Correlations of engagement and sales items

Table 1.B: Correlations of engagement and sales items (after ln transformations)

	1	2	3	4	5	6
1. Reactions	1					
2. Shares	**.81****	1				
3. Comments	**.95****	**.87****	1			
4. Album sales	.42**	.45**	.47**	1		
5. Digital songs	.52**	.55**	.56**	**.81****	1	
6. Streaming	.48**	.49**	.52**	**.73****	**.88****	1

** $p < .01$

Engagement items: $\alpha = .94$, Composite Reliability (CR) = .96, Average Variance Extracted (AVE) = .89

Sales items: $\alpha = .91$, CR = .93, AVE = .81

C. Measurements and cross-sectional factor analyses

C1. Engagement (Reactions, shares, and comments after ln transformations)

Time 20: α = .95, Composite Reliability (CR) = .97, Average Variance Extracted (AVE) = .91; Time 40: α =.94, CR = .96, AVE = .89; Time 60: α =.95, CR = .96, AVE = .90; Time 80: α =.94, CR = .96, AVE = .89; Time 100: α = .94, CR = .96, AVE = .90.

C2. Sales (Total album sales, digital songs, and streaming after ln transformations)

Time 20: α = .91, CR = .94, AVE =.84; Time 40: α = .91, CR = .93, AVE = .82; Time 60: α = .92, CR = .93, AVE = .81; Time 80: α = .92, CR = .93, AVE = .82; Time 100: α = .92, CR = .93, AVE = .81.

C3. Five cross-sectional confirmatory factor analyses (CFA):

We conducted five confirmatory factor analyses for five equally spaced time points for engagement and sales, using Laavan in R, with maximum likelihood estimations. The five CFA models give good fit indices [Model of Time 20: $\chi2$ (9) = 29.73, p <.001, CFI = .99, and RMSEA = .077; Model of Time 40: $\chi2$ (9) = 7.47, p = .59, CFI = 1, and RMSEA = .00; Model of Time 60: $\chi2$ (9) = 31.86., p < .001, CFI = .99 , and RMSEA = .08; Model of Time 80: $\chi2$ (9) = 30.89, p < .001, CFI = .99 , and RMSEA = .08; Model of Time 100: $\chi2$ (9) = 29.02, p < 0.01, CFI = .99, and RMSEA = .076]. The results indicate that the models fit the data adequately (McQuitty 2004). All alphas, composite reliability, and AVE scores (See C1 and C2) are above .8.

D. Correlation sales and engagement over time

Table 1.D: Correlation matrix of sales and engagement over time

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. Engagement T20	1									
2. Engagement T40	.64**	1								
3. Engagement T60	.61**	.63**	1							
4. Engagement T80	.51**	.52**	.58**	1						
5. Engagement T100	.5**	.48**	.59**	.65**	1					
6. Sales T20	.58**	.49**	.48**	.46**	.44**	1				
7. Sales T40	.55**	.52**	.49**	.46**	.46**	.95**	1			
8. Sales T60	.52**	.49**	.54**	.5**	.5**	.91**	.94**	1		
9. Sales T80	.49**	.48**	.52**	.58**	.56**	.86**	.89**	.93**	1	
10. Sales T100	.49**	.47**	.52**	.57**	.59**	.83**	.86**	.91**	.96**	1

** $p < .01$

E. Longitudinal measurement invariance

Table 1.E. Longitudinal confirmatory factor analysis of engagement and sales (6 items) across T20, T40, T60, T80 and T100

Step	χ^2	**df**	**CFI**	**RMSEA**	**SRMR**	**ΔCFI**	**ΔRMSEA**	**ΔSRMR**
Configural invariance	1068.96	355	.965	.072	.045			
Metric invariance	1195.79	379	.961	.075	.072	≥ -.01	≤ .015	≤ .03

We conducted longitudinal CFA to confirm the equivalence of the repeated measures over time (Widaman et al. 2010). We tested configural and metric invariance by comparing nested models over five equally spaced time points of the two latent constructs: engagement and sales (Time 20, Time 40, Time 60, Time 80 and Time 100) (Mackinnon et al. 2022; Radanielina Hita et al. 2022). We used Laavan in R, with maximum likelihood estimations. In the first step, we estimated the factor loadings without constraints. We thus establish the same factor structure across five waves, i.e., the same number of latent constructs with the same number of items on each construct (configural invariance) (Chen 2007). In the second step, we constrained the factor loadings to be equal across five waves. We confirm that the items do not vary in how they represent the latent constructs over time (metric invariance) (Chen 2007; Putnick and Bornstein 2016) (See Table E).

F. Correlations of constructs in the three models

Table 1.F: Correlations of constructs considered for the models

	1	2	3	4	5	6	7	8	9	10	11	12
1. Engagement$_{ti}$	1											
2. Sales$_{ti}$	.55**	1										
3. Charity-related posts$_{ti}$	.2**	.12**	1									
4. Mission-related posts$_{ti}$	.6**	.22**	.13**	1								
5. Non-mission-related posts$_{tj}$	.73**	.28**	.14**	.49**	1							
6. Fanbase size$_{ti}$	.58**	.82**	.14**	.18**	.29**	1						
7. Track volume$_{ti}$	.10**	.08**	.02**	.1**	.09**	.05**	1					
8. Tweet volume$_{ti}$	.69**	.78**	.14**	.35**	.44**	.77**	.10**	1				
9. News volume$_{ti}$	.38**	.37**	.1**	.27**	.25**	.34**	.13**	.41**	1			
10. Experience$_{i}$	.24**	.40**	.12**	.09**	.11**	.54**	.03**	.32**	.16**	1		
11. Age$_{i}$	.11**	.18**	.1**	.06**	.06**	.30**	.01**	.08**	.08**	.78**	1	
12. Total track volume$_{i}$	.34**	.53**	.13**	.1**	.19**	.63**	.08**	.45**	.21**	.78**	.57**	1

** $p < .01$

G. Composite variables for engagement and sales in the robustness check

We ran a CFA on the respective component factors of engagement (reactions, shares, and comments after ln-transformations) and sales (total album sales, digital songs, and streaming after ln-transformations) using Laavan in R, with maximum likelihood estimations. The CFA model gives good fit indices ($\chi 2$ (9) = 1694.85, p <.001, CFI = .99, and RMSEA = .068). Composite variables for engagement and sales were created from the sum of the component factors, with each factor multiplied by its unstandardized loading. Thus, the composite variable for engagement = 2.98*shares + 3.195*comments + 4.376*reactions; and the composite variable for sales = 1.812*total album sales + 2.64*digital songs + 3.009 streaming (Lefcheck 2016).

In the multilevel mediation model using the newly created composite variables engagement and sales, charity-related posts again do not have a significant direct effect on sales in either the short or the long term. Mission-related posts have a positive direct relationship on sales in the short term (β =1.25, p <.001) but their direct long-term effect is not significant. The direct relationship between non-mission-related posts and sales in the short term is positive (β = .47, p <.001), but negative in the long term (β =-4.14, p <.01).

In the short term, charity-related posts have a positive relationship with engagement (β = 7.91, p < .001), but the effect is lower than that of mission-related posts (β =13.15, p < .001) and non-mission-related posts (β =17.43, p < .001). Also, their indirect effect on sales through engagement is positive (β = .33, p < .001, *95% CI* [.27, .4]), but lower than that of mission-related posts (β = .61, p < .001, *95% CI* [.54, .69]) and non-mission-related posts (β = .75, p < .001, *95% CI* [.67, .84]).

In the long term, charity signaling has a positive association with engagement (β = 23.17, p < .001). This effect is higher than that of mission signaling (β = 12.11, p < .001) and that of non-mission signaling (β = 18.13, p < .001). Long-term charity signaling also has a positive indirect effect on sales through engagement (β = 3.69, p < .01, *95% CI* [1.34, 6.84]). This effect is higher than that of mission-related posts (β = 1.93, p < .001, *95% CI* [.84, 3.13]) and non-mission-related posts (β = 2.89, p < .001, *95% CI* [1.27, 4.59]).

In the moderated mediation model, when fanbase size is the moderator between charity-related posts and engagement in the long-term sequence charity-related posts – engagement – sales, the interaction between fanbase size and charity-related posts on engagement is positive and significant (β = 13.32, p < .05). The 95% confidence interval for the index of moderated mediation does not contain 0 (β = 1.1, *95% CI* [.17, 2.06]) (Hayes and Rockwood 2020). The long-term indirect effect of charity-related posts on sales is not significant when fanbase size is less than 3, or when artists have fewer than one million Facebook fans. However, it is significant when fanbase $size_j$ = 3 (i.e., when artists have 1 – 5 million fans, β = 13.05, p < .001), and when fanbase $size_j$ = 4, (i.e., when artists have more than 5 million fans, β = 21.51, p < .001).

Chapter 2 :
Antecedents and consequences of complaint types – a field study on service failures in the airline industry

Abstract

We applied a deep learning model to interpret two types of complaints in service failures: conciliatory complaints, which sought solutions and were open to compromises, and vigilant complaints, which aimed to punish the firms. We then examined the real financial impact of vigilant complaints and analyzed the antecedents of the two types of complaints, thus validating them and helping to prevent the occurrence of vigilant complaints in the future. Furthermore, by applying resource literature in the context of service failures, we help firms to develop long-term approaches in the relationship with customers with the latter's well-being at the center.

2.1 Introduction

For the airline industry, causes of service failures are diverse, e.g., cancelled and late flights, lost luggage, food poisoning or airline staff rudeness. Some of them became so serious that they even ended up in the news headlines. ‘My daughter has a brain tumour but Ryanair refuses to refund our flights’ or ‘Help! My Lost Luggage Wasn’t Delivered for 12 Days. I Want Restitution’ are among the examples. Large volumes of daily passengers and flights unavoidably also mean more risk of service failures. Understanding the nature of the complaints is vital to protect the relationship of these complainers. This context allows us to examine multiple factors and different stages of the relationship between firms and customers.

From the perspective of the resources of these customers, due to different circumstances and personal characteristics (Hobfoll et al. 1990), service failures affect them differently. The amount of emotional and psychological resources customers lose during the experiences might be different; this situation influences how customers complain and make decisions to recover their lost resources or not to lose any more of their resources. In this light, vigilant complaints, which seek to punish the firm, and conciliatory complaints, which seek solutions with the firm, are two different mechanisms of using resources.

Furthermore, in terms of consequences of complaints, literature on service failures measured the harm of customer anger through perceived negative word-of-mouth (Grégoire et al. 2010) and purchase intention (Cai and Qu 2018). Consequences of service recovery include satisfaction with the recovery and customer churn probability (i.e., the probability that they will stop using the firm’s products or services) (Van Vaerenbergh & Orsingher, 2016; Stakhovych & Tamaddoni, 2020). However, vigilant complaints should drive “real" loyalty in the post-compensation period, and the extent of change in “real” behaviors is still an under-researched issue. This requires linking two stages—complaint and post complaint—with each other as well as using real behavior data. Here, recent editorial and systematic reviews note that the usage of quantitative metrics are rare and much needed in this literature (Khamitov et al., 2020).

Previous literature has also described certain characteristics of these two types of complaints (Grégoire et al. 2019) and factors driving customers' desire to retaliate against the firm, including violation of relationship norms (Grégoire and Fisher 2008), employee failures (Obeidat et al. 2018) and contract breaches (Obeidat et al. 2018). However, discrepancies exist among the findings, due to choice of models (Grégoire et al. 2010) and cultural elements (Obeidat et al., 2018), among others. It is also not known how these factors compare with each other, as it is possible that some could overshadow the others as they are closely related.

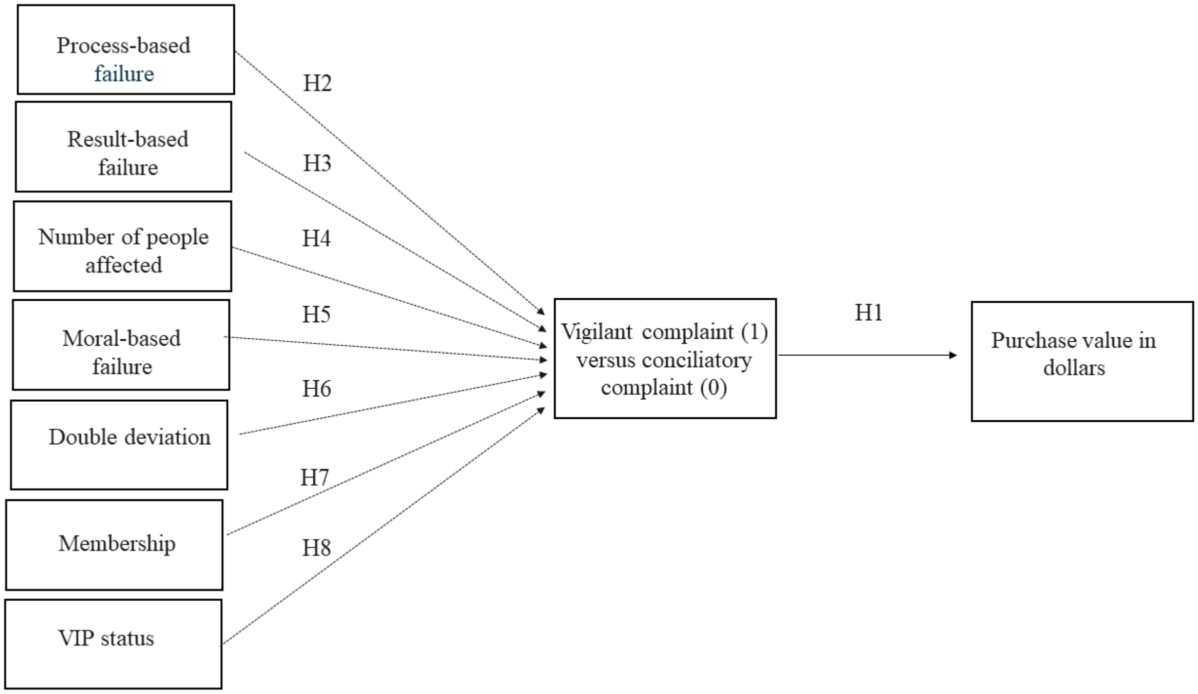
In this essay, we focus on two main research questions (1) What are the consequences of vigilant complaints on consumer loyalty, measured objectively? (2) What are the antecedents of the complaint types? By answering the first question, we will help researchers and firms to understand the long-term impacts of vigilant complaints after service recovery. The second question examines in details diverse factors that might influence how complainers decide to use their resources in the complaint process.

Thus, our research makes two main contributions. First, to the best of our knowledge, we believe that this is the first time two types of complaints are linked to an objective measure with real financial impacts. Our dependent variable is purchase value in dollars, which is the product of loyalty point redemption frequency (i.e. the number of redemption transactions) and the value in dollar per redemption transaction. Redemption regularity increases later purchase amounts and purchase incidences (Dorotic et al. 2014) and retains consumers in the loyalty program (Danaher et al. 2020). In other words, purchase value allows us to capture both the level of consumer loyalty and its financial magnitude. Confirming this link helps us to materialize the financial and relational impacts of vigilant complaints.

Second, we examine the antecedents leading customers to become vigilantes or conciliators. This allows us to validate the two complaint types. Specifically, we prove that the two types of complaints are distinct from each other, as factors associated with them are also distinct. Further, we measure different factors alongside each other, making it possible to contrast their importance in leading to vigilant or conciliatory complaints. Last, identifying key factors in the complaining process will also prevent the occurrence of vigilante complaints.

Our research also has important managerial implications. Measuring financial risks of vigilant complaints allows managers to estimate properly their importance in the firm's performance. Furthermore, by using a well-being perspective for service failures and recovery, we not only help firms to reduce the number of vigilant complaints, but also identify ways for firms to protect consumers' resources. Thus, firms could maintain relationships with these consumers in the long term and contribute to the community's well-being in the same process. Examples include paying attention to vulnerable consumers or different elements that make certain customers more vulnerable to the depletion of resources. Lastly, the complex dynamics of different factors in their relationships with vigilant complaints (i.e. the failure stage) and later purchase value (i.e. the post-recovery stage) give firms more knowledge to prioritize factors effectively at different stages of the customer experience.

Figure 2-1: The antecedents and consequence of vigilant complaints vs. conciliatory complaints



2.2. Theoretical background and hypotheses

2.2.1. Conciliatory and vigilant complaints

Methodologies on complaints mainly focused on distinguishing negative, neutral and positive text, with negative eWOM considered complaints (Vermeer et al. 2019; Herhausen et al. 2019). Some scholars also paid attention to certain aspects of complaints, e.g., high versus low-arousal words (Herhausen et al. 2019) and outcome-focused versus process-focused complaints (Gunarathne et al. 2018).

In this research, we focus on two different types of complaints: conciliatory and vigilante complaints. Conciliators refer to those who attempt to ameliorate the situation (Grégoire et al. 2019), are supportive (Weitzl and Hutzinger 2019) and are interested in maintaining relationships

(Ringberg et al. 2007). The vigilantes are unforgiving (Weitzl and Hutzinger 2019) and adversarial (Ringberg et al. 2007).

Conciliators are often amenable, try to resolve the problem with the relevant firm and gain tangible results (Ringberg et al. 2007; Beverland et al. 2010). Their main purpose is regaining justice for themselves and other consumers (Beverland et al. 2010). They are interested in service recovery and tend not to seek punishment on firms (Grégoire et al. 2019). Thus, the likelihood that they will reach a solution with the firm is also higher.

The complainers of the second type, vigilantes, however, seek personal and public revenge. They feel that the firm betrayed their trust and they need to punish that firm (Ward and Ostrom 2006) with their vocal resistance (Ringberg et al. 2007). The wish to hurt the brands could also come from ideological or symbolic conflicts (Aziz and Rahman 2022). More upfront than the conciliators, vigilantes tend to show strong negative feelings in their communications with the firm.

Furthermore, the desire to revenge (Grégoire et al. 2009) makes vigilant complainers willing to hurt the brands both in the short and long terms (Fetscherin 2019). Emotions linked to active hate (i.e., anger, disgust and contempt) (Aziz and Rahman 2022) drive them to aggressive actions against the brands, notably negative WOM (Aziz and Rahman 2022). Seeing their mission as punishing the wrong-doers for the public good (Grégoire et al. 2019), they could go beyond avoiding the brands (Grégoire et al. 2009) to invite the others join in the actions, thus potentially greatly harm a brand’s interests (Simon 2011).

While both see their courses of action as morally justified, conciliators consider themselves to be more moral, as they do not seek revenge (Grégoire et al. 2019). Those who adopt the conciliatory approach often use a more formal writing style than the vigilante, more often use the pronoun “I” and the past tense and are less likely to blame the firm for the failures (Grégoire et al. 2019). Vigilante complainers use a less formal writing style, are more likely to blame the firms and use the pronoun “you” and the present tense.

2.2.2. Vigilant and conciliatory complaints through the lenses of resources

Resources are tangible or intangible “units” (Hobfoll 2002) that hold value, and such resources could be material (e.g., money), social (e.g., social support or status) or psychological (e.g., self esteem) (Hobfoll 1989). Loss of resources leads to stress, distress, and anger (Hobfoll et al. 1990). Individuals are therefore motivated to obtain and protect their resources. Vigilant and conciliatory complainers adopt different strategies. Vigilant complainers stop devoting their resources to the firm altogether and encourage others to do the same; conciliatory complainers invest “resources” in the firm with the hope to regain their losses (Doane et al. 2012).

From this perspective, by retaliating against the firm, vigilant complaints try to regain equity in invested resources between them and the transgressor (Adams and Freedman 1976) and take away the resources from the firm. Perceived losses of resources often affect individuals more deeply than the gain of the equivalent amount of resources (Doane et al. 2012). Thus, vigilant complainers are willing to go for less ideal options (e.g., by switching to a more expensive competitor), as the satisfaction from the revenge offsets the material loss (Bechwati & Morrin, 2003).

Conciliatory complainers, instead, adopt a pragmatic approach, relying on the firm’s social support (Hobfoll et al. 1990). An example of the relationship between social support and personal resources is a research on patients with breathing disorders (Hobfoll et al. 1990). As the disorder symptoms became more serious, the patients felt the loss of resources and became hostile to their nurses, leading to a deterioration in the relationship with these careers. To regain the support of the nurses, these patients made efforts to be less angry, and as a result, received more help and social support.

Thus, conciliatory complainers are prepared for compromises to achieve solutions. In this case, they appeal to the support of the firm to protect their resources (Hobfoll et al. 1990). This investment is expected to help them to regain personal resources (Hobfoll et al. 1990). However, the other side of this investment is that if their efforts do not lead to expected results, they will lose even more resources (Doane et al. 2012) and could end up becoming vigilant.

2.2.3. The consequence of vigilant complaints

In this research, we use change in purchase value, which is the product of change in redemption frequency and the average value in dollar per redemption, as the consequence of vigilant complaints. Redemption frequency, i.e., how often one redeems loyalty points, is an important measurement of consumer loyalty programs. In the short term, decisions to redeem points lead to a later increase in the number of purchase occasions and purchase amounts (Dorotic et al. 2014). In the long term, redemption regularity encourages customers to maintain their active state in the loyalty program, (e.g., by expanding purchases to the firm's partners), thus helping firms to retain these customers (Danaher et al. 2020). We combine redemption frequency with the average value in dollars to measure both consumer loyalty and its financial impact.

Vigilant complainers are more likely to retaliate than conciliatory complainers are, due to the feeling of being betrayed (Grégoire and Fisher 2008). They will thus go through different routes to punish the firm and make the relationship even between the two parties (Adams and Freedman 1976), such as through negative words of mouth or complaining to the third party (Grégoire et al. 2018). In other words, they will take away the resources of the firm to make up for their lost resources. Apart from calling on other people, they themselves will reduce the engagement with the firm and their dependence on these services.

Furthermore, harboring anger puts stress on the vigilant complainers (Hobfoll et al. 1990). To conserve their resources (Hobfoll et al., 1990), they could limit their own engagement with the firm. From their perspective, the risk of being reminded of negative emotions and possibly repeating the same unpleasant experiences outweighs the consumption benefits. Being more affected by resource loss, vigilant complainers will also be more selective in their future investment of resources (Doane et al. 2012). However, being defensive could also make their decisions irrational (Doane et al. 2012).

H1: Vigilant complaints will have more negative effect than conciliatory complaints on future purchase value in dollars.

2.2.4. Antecedents of complaint types

2.2.4.1. Failure types

Failure types identify intrinsic and situational characteristics of service failures which influence how consumers perceive them. As we discuss below, previous literature has long studied separately process-based and result-based failures, double deviations and to some extent moral-based failures, although how they drive different complaint types and how they appear alongside each other is not yet known. Group actions received much less attention, as field data are needed to identify the real size of a group of affected consumers.

2.2.4.1.1. Process-based failures

Process-based failures are related to the experience(s) of a complainer—i.e., how they received a service (Smith et al. 1999). Rather than actual outcomes, interpersonal elements of experiences often lead to stronger revenge desires of customers (Bechwati and Morrin 2003). Perceptions of the experiences themselves are relational, contextual, and introspective (Katja Wiemer-Hastings and Xu 2005) and this tension drains people's resources (Curci et al. 2013).

Poor customer treatment is also perceived as being controllable (Choi and Mattila 2008), due to integrity or competence issues. When the failures are related to integrity, they are intentional, and will be considered major violation of trust (Kim et al. 2004). And even when they are perceived to be related to staff competence (e.g., impoliteness due to lack of training), the controllability of these types of failures makes customers more likely to resent the firm for the harm (Choi & Mattila, 2008), as they see the failures as evidences of the broken social contract between them and the firm (Weiner et al. 1987), which expects the commitment of adequate resources. Accordingly:

H2: *Process-based service failures are more positively associated with vigilant complaints than with conciliatory complaints.*

2.2.4.1.2. Outcome-based failures

Outcome-based failures are related to what the customers received (Smith et al. 1999). The loss is based on single transactions, and it reflects an exchange relationship (Aggarwal 2004). The failure is thus not viewed as psychological loss and consumers are more likely to try to get a concrete and tangible solution with the firm. Furthermore, reasons behind outcome-based failures can be diverse and not necessarily always attributed to the firm (e.g., due to the third party or other uncontrollable factors like bad weather or safety concerns). This ambiguousness makes it more difficult for customers to justify to themselves their own anger. Unless there are evidences that the firm or their employees were intentional, to save their own resources (Curci et al. 2013), complainers are likely to be more focused on the final results.

H3: Outcome-based service failures are more negatively associated with vigilant complaints than with conciliatory complaints.

2.2.4.1.3. Group action

A larger group involves more social state, relational factors, and complex psychological characteristics, in comparison with a smaller group. The composition of a larger group is often more complicated than a smaller group, e.g., including people of different age and issues. A larger group also means more diverse relationships among members. All these factors lead to more challenges if a service failure occurs. Emotion contagion leads to more anger in a group failure (Du et al. 2014), as individuals reinforce emotions of their companions. Furthermore, the more similar the emotions among the group members, the stronger the reinforcement will be. All these combined factors lead to a greater load on the resources of the complainer, making them more likely to be vigilant.

*H4**:** Number of people affected in a service failure is more positively associated with vigilant complaints than with conciliatory complaints.*

2.2.4.1.4. Moral-based failures

Moral-based transgressions, representing violations of social norms, are serious and also often behind calls for brand boycotts (Glazer et al. 2010). Even outsiders, who are economically unaffected by the transgressions, feel the need to punish norm violators and restore justice, sometimes even at their own expense (Leliveld et al. 2012), i.e. by sacrificing a little of their own resources. And for those who are directly affected, being vigilant also means protecting the others in future similar situations (Grégoire et al. 2019b).

In addition, moral transgressions hurt the victims' self-esteem and efforts to forgive the firm exhaust their self-regulatory resources (Baumeister et al. 1998). Being willing to compromise could lead to the same transgression happening to them again (Zhang et al. 2023). In other words, even when social relations between these consumers and the firm are important Hobfoll et al., 1990), these consumers face difficulties devoting cognitive resources to regulate their emotions and behavior (Bies et al. 2016). Instead, due to ego depletion (Baumeister et al. 1998), they could turn to more destructive courses (Bies et al. 2016).

H5: Moral-based service failures are more positively associated with vigilant complaints than with conciliatory complaints.

2.2.4.1.5. Double deviation

Double deviation refers to failed service recovery or multiple service failures (Khamitov et al. 2020). Failed recovery leads to eroded trust (Basso and Pizzutti 2016) between the firm and the consumers. The longer the process lasts (Baumeister et al. 1998), the more emotional and cognitive resources the complainer needs to cope with the growing stress (Suresh and Chawla 2022). The complainer, having invested resources in the hope of a satisfactory solution, become more and more disillusioned. Compared with single deviation, double deviation intensifies frustration and anger (Basso and Pizzutti 2016). As the recovery process prolongs, consumers' love could turn into hate and desire to revenge (Grégoire et al. 2009).

H6: Repetitive characteristics of double deviation are more positively associated with vigilant complaints than with conciliatory complaints.

2.2.4.2. Relational factors

Furthermore, depending on their current ties with the firm, customers are likely to assess their experiences against different benchmarks. Previous service failure literature often examined the relationship between the customers and the firm from the customers' subjective perspective (e.g., high or low level of relationship, Grégoire and Fisher 2008). Our data allowed us to analyze instead objective measures, i.e., whether a customer was currently in the loyalty program, and if that was the case, their rankings in this program. These allow the findings to go beyond subjective variations and be more feasible for managerial implications.

2.2.4.2.1. Loyalty program membership

Human beings have the need of belonging to a group or community, as it helps them feel that they are part of a larger body that goes beyond their own limitations and boundaries and gives meaning to their life (Lambert et al. 2013). People forgive more easily transgressions committed by ingroup members than those committed by people outside their group (Zourrig et al., 2015; Wohl & Branscombe, 2009). These social resources from belonging make complainers more likely to be cautious about attacking the firm. Furthermore, loyalty program members are likely to have more experience with the firm and have more knowledge of compensation policies than non-members. They are, thus, more likely than non-members to seek a beneficial solution, and this knowledge, also a type of resources, allows them to invest to gain new resources more effectively than non-members (Doane et al. 2012)..

*H7**:** Being loyalty members (vs. not being loyalty members) is more negatively associated with vigilant complaints than with conciliatory complaints.*

2.2.4.2.2. VIP status

The high ranks in the loyalty system confirm these members' social standings and their self-identity. As the relationship gives them social and economic benefits, higher ranked customers are more likely to avoid hurting the brand than those who receive fewer benefits. From the perspective of psychological contracts (Cullinane and Dundon 2006), high-ranked customers, with more ties with the firms, will feel that they have more unspoken obligations in the relationship than low-ranked customers.

As it also takes more investment of resources to be promoted to higher ranks, from the perspective of conservation of resources (Roskes et al. 2013), it will take more for higher-ranked members to go against the brand, as they invested more resources in the relationship than lower ranked members. Lower-ranked members have less to lose, as they decide to leave relationship. Also, people with more resources are less likely to suffer from resource loss and are also more capable of regaining resources (Doane et al. 2012). In other words, higher-ranked consumers should be less affected than low-ranked customers even when faced with similar service failures, as they are likely to have more coping options.

H8: Among loyalty members, higher-ranked members are less likely than lower-ranked members to send vigilant complaints.

2.3. Methodology

2.3.1. Study 1: The consequence of vigilant complaints

2.3.1.1. Propensity score matching and difference in difference

To identify vigilant and conciliatory complaints, we used a Transformer model, developed by Meire et al (2024). Thanks to the advanced architecture of the model, it could identify different relationships between words in the same sentence, making it a better performer in textual classification than its predecessors. This model was trained on 5,830 public complaints, which were collected from different platforms and manually coded into vigilant and conciliatory complaints. Appendix A includes a codebook for coders to identify vigilant and conciliatory complaints on Twitter, one of the above-mentioned platforms. The accuracy of the model on the training dataset is .82 and its the area under the receiver operator curve (AUC) is .9. The accuracy indicates the total percentage of correctly classified complaints (both vigilant and conciliatory complaints). AUC evaluates classification performance across vigilant as well as conciliatory complaints and is robust to imbalance issues of the data.

We used this model to classify 1,293,453 English complaints sent to a North American airline company in 2019, 2022 and the first five months of 2023 into vigilant and conciliatory complaints. Of these complaints, 34,139 are vigilant, or only 2.6% of all the sample.

Recovery types include mileage, promotion codes, giftcards, vouchers, ecoupons or credits and cash or refund, all being count variables. Mileage indicates the number of occasions loyalty points are offered as compensations, promotion codes (i.e., the number of codes given, giftcards, vouchers, ecoupons or credits), the number of times giftcards, vouchers, ecoupons or credits provided, cash or refund (i.e., the number of direct money transfer or refund actions).

The categories of service failures were identified based on the available categories that the customers chose when filing the complaints. These are all dummy variables. The variables related to process-based failures are service at airports, service onboard, delays and other experience issues. The variables related to result-based transgression include damaged bags, missing bags, shipped equipment issues, other bag issues, refund requests and call center and customer support. We also examine whether a complaint is related to a disability issue (e.g., accessibility or service animal problem) or a disability type (e.g., impaired vision), and consider such cases to be related to moral-based failures.

Apart from these categories, we also consider factors related to double deviation (i.e., enumeration of complaints over the same flight, count of complaints of the same flight from the same consumer and whether a complaint is a duplicate of a previous complaint). The number of people affected represents the group action concept. We also use dummy variables for loyalty membership (i.e. whether the complainer is a loyalty member) and different membership groups (i.e. whether a loyalty member is a bottom tier member or a lower mid-tier, upper mid-tier, top tier level member). All the levels above the bottom tier represent the VIP status. Other variables that we collected on the consumers and their relevant flights are control variables. Table 2-1 explains the variables we use for study 1 and study 2.

Table 2-1: Definitions of variables

Variables	Definition
Vigilant complaint	Complaint type: vigilant (1) vs conciliatory (0).
Change in redemption frequency	The difference between the total number of transactions of redemption of loyalty points during 12 months after a complaint was sent and that during 12 months before a complaint was sent
Purchase value change in dollars	The product of change in redemption frequency and the value in dollars per redemption[7].
Mileage (loyalty points)	Count of separate loyalty point offers for the complainer and their trip companions.
Promotion codes	Count of separate promotion codes for the complainer and their trip companions.
Giftcard, voucher, ecoupon or credit	Count of separate giftcards, vouchers, ecoupons or credits for the complainer and their trip companions.
Cash or refund	Count of separate direct money transfer or refund actions for the complainer and their trip companions).
Service at airport	Whether a complaint is related to service issues at the airport (1 or 0)
Service onboard	Whether a complaint is related to service issues onboard (1 or 0)
Delays	Whether a complaint is related to flight delay issues (1 or 0)
Other experience issues	Whether a complaint is related to other experience issues (1 or 0)
Missing bags	Whether a complaint is related to missing luggage issues (1 or 0)
Damaged bags	Whether a complaint is related to damaged bags (1 or 0).
Shipped equipment issues	Whether a complaint is related to shipped equipment issues (1 or 0).
Other bag issues	Whether a complaint is related to other luggage issues (1 or 0).
Refund requests	Whether the consumer requests refunds in the complaint (1 or 0).
Call center and customer support	Whether the consumer makes requests related to the call center or customer support section (1 or 0).
Number of people affected	The number of people affected in the incident
Disability-related	Whether a complaint is related to support for people with disability (1 or 0)

Variables	Definition
Enumeration of complaints over the same flight	Enumeration of complaints over the same flight until that time
Count of complaints of the same flight from the same consumer	The total number of all the complaints over the same flight from the same consumer
Duplicate complaints	Whether a complaint contains similar contents to another one (1 or 0).
Member (vs being non-member)	Whether a complainer is a member of the firm's loyalty program (1 or 0).
Top tier	Whether a complainer is a loyalty member belonging to top tier, i.e., the highest level (1 or 0).
Upper mid-tier	Whether a complainer is a loyalty member belonging to the upper mid-tier (1 or 0).
Lower mid-tier	Whether a complainer is a loyalty member belonging to the lower mid-tier (the level only higher than the bottom tier) (1 or 0).
Temperature of the destination location	The temperature of the destination location at the landing time of the flight in the complaint
Precipitation rate of the destination location	The precipitation rate of the destination location at the landing time of the flight related to the complaint
Temperature of the departure location	The temperature of the departure location at the take-off time of the flight related to the complaint.
Precipitation rate of the departure location	The precipitation rate of the departure location at the take-off time of the flight related to the complaint.
Month of the complaint	The month the complaint was sent
Year of the complaint	The year the complaint was sent
Female ratio of the traveling group	Ratio of female passengers in the traveling group under the same reservation
Ratio of unidentified gender of the traveling group	Ratio of passengers of unidentified gender in the traveling group under the same reservation
Complainer being of unidentified gender	Whether the complainer is of unidentified gender (vs other genders) (1 or 0)
Daily journey duration in the last three years	Total duration of journeys (including both flight and connection time) during the last three years, until the complaint was sent, in minutes divided by 365*3
Duration of the journey related to the complaint	Duration of the journey (in minutes) related to the complaint including both flight and connection time
Minimum age among passengers traveling in the group	The minimum age among passengers traveling under the same reservation
Ticket costs	The total costs of tickets bought under the same reservation
Number of tickets	The total number of tickets bought under the same reservation
Case duration	The length in days from when a complaint was sent until it was resolved

Variables	Definition
Number of days affected by COVID	Number of days during the period in consideration (i.e. from 12 months before until 12 months after the complaint) when COVID-related travel restrictions were in effect.

As clarified earlier, we use purchase value in dollars as the dependent variable. This value is the product of redemption frequency and the average value in dollar per redemption[6]. Redemption frequency is defined as the number of transactions a consumer conducted to redeem their loyalty points during a period of time. In this study, we included only observations of loyalty members, because only they could redeem loyalty points. Also, as we measured the purchase value change one year after the complaints versus one year before the complaints, we could only include complaints which were sent at least one year before the latest redemption date (i.e. November 30, 2023). The number of complaints meeting these conditions is 70,809 (see descriptive statistics data in Appendix 2.B).

As we did not include observations with missing information, we have the bias risk due to self-selection (De Haan et al. 2018). Thus, we used the propensity score matching approach, by matching observations from the vigilant group (i.e., the treated group) with those from the conciliatory group (i.e., the untreated group) with similar distributions of covariates (Austin 2009). We did this by running a multivariable logistic regression on the group of the vigilant complaints and that of conciliatory complaints using all the variables available before a complaint was launched, with the calipers of width of .2 of the standard deviation of the propensity score, as recommended by Austin (2011). The assessment of the balance of the distribution of covariates using standardized mean differences (Zhang et al., 2019) indicates that the covariates of the two

[6] We do not disclose the exact amount of value in dollars per redemption to protect the anonymity of the firm in our research. It is in the range of 25 to 100 dollars.

vigilant and conciliatory groups after propensity score matching have good balance, and this balance is also much better than that of the original dataset (see Appendix F for details).

We combined propensity score matching with difference in difference approaches to establish the causal impact of complaint types on purchase value change between one year before and one year after a complaint was sent (Goldfarb et al. 2022). The highest VIF score among the variables is 3.61, well below the threshold of 5, indicating no multicollinearity issue.

Table 2-2 shows the results of the mixed-linear regressions on the original sample (N= 70,809, with 1.31 observations per customer and 1,914 vigilant complaints) and the propensity score matched sample (N = 55,882, with 1.27 observations per customer and 1,914 vigilant complaints). We will refer to the original sample as the *before-correction sample*, and to the propensity score matched sample as the *after-correction sample*. We include in these two regressions all the variables previously used for the logistic regression, together with compensation types: mileage (*M* = .02, *SD* = .17), promotion codes (*M* = .33, *SD* = .51), gift cards, vouchers, ecoupons or credits (*M* = .23, *SD* = .46) and cash or refund (M = .3, SD = .65), the main independent variable - vigilant complaints (*M* = .03, *SD* = .18) and the dependent variable - change in purchase value (*M* = 146.67, *SD* = 836.34).

Table 2-2: Mixed linear regression on change in purchase value among members before and after propensity score matching

Concept		Before correction		After correction	
	Construct	**β**	**t value**	**β**	**t value**
	Vigilant complaint	-37.18	-2.03*	-35.49	-2.02*
Recovery types	**Mileage (loyalty points)**	-36.12	-2.03*	-37.8	-2.04*
	Promotion codes	-3.32	-0.53	2.18	0.32
	Giftcard, voucher, ecoupon or credit	-17.77	-2.42*	-14.94	-1.97*
	Cash or refund	-9.63	-1.96*	-9.67	-1.77 .
Process-related failures	**Service at airport**	-8.7	-0.63	-8.54	-0.63

Concept		Before correction		After correction	
	Service onboard	-12.83	-0.91	-14.39	-1.02
	Delays	-5.94	-0.49	-10.52	-0.86
	Other experience issues	-0.95	-0.06	0.67	0.04
Result-related failures	**Missing bags**	-60.1	-3.18**	-67.42	-2.9**
	Damaged bags	-42.18	-2.64**	-50.16	-2.01
	Shipped equipment issues	-29.57	-0.88	-46.35	-1.17
	Other bag issues	-22.42	-1.58	-19.89	-1.26
	Refund requests	-8.17	-0.61	-5.97	-0.39
	Call center and customer support	-15.3	-1.09	-29.51	-2.03*
Group actions	**Number of people affected**	-19.41	-4.14***	-21.05	-4.43***
Moral-based failures	**Disability-related**	3.53	0.08	-13.81	-0.33
Double deviation	**Enumeration of complaints over the same flight**	7.06	1.74	5.88	1.3
	Count of complaints of the same flight from the same consumer	-3.59	-0.94	-3.78	-0.93
	Duplicate complaints	-3.26	-0.21	0.5	0.03
VIP status	**Top tier**	415.64	38.34***	384.71	32.99***
	Upper mid-tier	175.11	19.79***	166.33	17.32***
	Lower mid-tier	39.83	4.05***	35.76	3.43***
Control variables	**Temperature of the destination location**	-0.54	-2.2*	-0.3	-1.13
	Precipitation rate of the destination location	1.1	0.34	1.97	0.6
	Temperature of the departure location	-1.13	-4.85***	-1.11	-4.41***
	Precipitation rate of the departure location	13.5	3.81***	1.06	0.26
	Month of the complaint	-23.11	-22.78***	-22.3	-20.32***
	Year of the complaint	81.22	31.52***	79.71	29.16***
	Female ratio of the traveling group	-32.89	-4.07***	-29.07	-3.36***
	Ratio of unidentified gender of the traveling group	-44.99	-0.26	-8.39	-0.05
	Complainer being of unidentified gender	-3.23	-0.14	0.01	0.0004
	Daily journey duration in the last three years	-4.47	-28.09***	-4.23	-26.02***
	Duration of the journey related to the complaint	0.01	6.21***	0.01	5.27***
	Minimum age among passengers traveling in the group	-1.2	-6.12***	-1.25	-5.92***
	Ticket costs	0.01	6.49***	0.01	6.6***
	Number of tickets	-7.8	-6.64***	-7.58	-5.68***
	Case duration	0.06	0.81	0.04	0.59
	Number of days affected by COVID	0.12	3.68***	0.1	2.85**
	Intercept	-163770	-31.47***	-160723.4	-29.11***

Concept		Before correction	After correction
	BIC	1145936	899613.9

Vigilant complaints in the before-correction sample ($\beta = -37.18$, $p < .05$) and that in the after-correction sample ($\beta = -35.49$, $p < .05$) both have a negative relationship with change in purchase value. In other words, vigilant complaints lead to a reduction in purchase value (H1 confirmed). We also found similar results when we ran linear regression on the before and after correction samples (See results of linear regression for the two samples in Appendix C). The following findings are based on the results of the after-correction sample.

Among the recovery types, mileage ($\beta = -37.8$, $p < .05$) and giftcards, vouchers, ecoupons or credits ($\beta = -14.94$, $p < .05$) have negative direct effects on change in purchase value, while promotion codes and cash or refund have no significant effect.

For the before-correction model, BIC is 1,145,936, AIC is 1,145,551, and log likelihood is -572,733.5. For the after correction model, BIC is 899,613.9, AIC is 899,238.8 and log likelihood is -449,577.4

2.3.1.2. Robustness checks

2.3.1.2.1. Endogeneity test

To test for omitted bias on a multilevel model without external instrumental variables, we use Kim & Frees' (2007) generalized method of moments, using the R package REndo (Gui et al., 2023 ; Nguyen et al., 2023). This approach produces three tests comparing among the reference random effects (REF), the generalized method of moments (GMM) and the fixed effects (FE) models, with vigilant complaints being the endogenous regressor. All the three tests are not significant. Specifically, for the test between REF and GMM, $x2(4, N=55,882) = 2.23$, p =.69. The result of the test between FE and REF is $x2(40,N =55,882) = 29.35$, p = .89 and the one between FE and GMM is $x2(39,N=55,882) = 29.35$, p = .87). This indicates no significant omitted variable bias.

2.3.1.2.2. Interactions

We tested for interactions between complaint types and compensation types on the after correction sample (Table 2-3). When we used mixed linear regression, we found no significant interaction. However, when we tested interactions using linear regression between vigilant complaint and giftcards, vouchers, ecoupons or credits, this interaction is positive ($\beta = 75.67$, $p <.05$). Nevertheless, as we choose mixed linear regression as our main approach, we cannot make interpretations from this finding. Mixed linear regression allows us to account for the hierarchical nature of the data. In this sample, with many customers complaining more than once (1.27 complaints per customer in the after-correction sample, as earlier indicated), this approach considers the customers as a grouping variable.

Table 2-3: Mixed linear regression vs linear regression on change in purchase value with interactions

		Mixed linear		Linear	
	Construct	**β**	**t value**	**β**	**t value**
	Vigilant complaint	-34.79	-1.35	-51.44	-1.84 .
Recovery types	**Mileage (loyalty points)**	-34.93	-1.84 .	-22.87	-1.1
	Promotion codes	4.06	0.59	2.64	0.35
	Giftcard, voucher, ecoupon or credit	-17.66	-2.28*	-18.10	-2.15*
	Cash or refund	-10.40	-1.88 .	-15.95	-2.69**
Process-related failures	**Service at airport**	-8.13	-0.6	-9.88	-0.67
	Service onboard	-14.22	-1	-19.14	-1.26
	Delays	-10.30	-0.84	-10.82	-0.81
	Other experience issues	0.94	0.06	5.15	0.3
Result-related failures	**Missing bags**	-67.71	-2.91**	-73.20	-2.91**
	Damaged bags	-50.47	-2.02*	-48.79	-1.82 .
	Shipped equipment issues	-46.32	-1.17	-55.83	-1.32

		Mixed linear		Linear	
	Other bag issues	-20.01	-1.27	-24.55	-1.43
	Refund requests	-6.07	-0.4	-2.23	-0.14
	Call center and customer support	-29.28	-2.01*	-25.85	-1.64
Group actions	**Number of people affected**	-21.02	-4.42***	-15.21	-2.99**
Moral-based failures	**Disability-related**	-15.36	-0.37	5.99	0.14
Double deviation	**Enumeration of complaints over the same flight**	5.86	1.3	1.38	0.27
	Count of complaints of the same flight from the same consumer	-3.83	-0.95	7.21	2.02*
	Duplicate complaints	0.39	0.02	-20.19	-1.13
VIP status	**Top tier**	384.85	33***	406.19	36.52***
	Upper mid-tier	166.24	17.31***	169.23	18.08***
	Lower mid-tier	35.96	3.45***	36.74	3.59***
	Temperature of the destination location	-0.30	-1.13	-0.48	-1.75 .
Control variables	**Precipitation rate of the destination location**	2.06	0.62	2.39	0.69
	Temperature of the departure location	-1.11	-4.41***	-1.06	-3.96***
	Precipitation rate of the departure location	1.06	0.26	-0.68	-0.16
	Month of the complaint	-22.27	-20.3***	-19.20	-16.68***
	Year of the complaint	79.79	29.19***	82.85	29.69***
	Female ratio of the traveling group	-29.06	-3.36***	-27.31	-3.19**
	Ratio of unidentified gender of the traveling group	-11.65	-0.07	-34.62	-0.19
	Complainer being of unidentified gender	0.10	0.004	4.99	0.2
	Daily journey duration in the last three years	-4.23	-26.02***	-3.93	-25.72***
	Duration of the journey related to the complaint	0.01	5.28***	0.01	4.67***
	Minimum age among passengers traveling in the group	-1.25	-5.93***	-1.50	-7.12***
	Ticket costs	0.01	6.61***	0.01	5.27***
	Number of tickets	-7.55	-5.66***	-7.28	-5.28***
	Case duration	0.04	0.59	0.00	-0.05
	Number of days affected by COVID	0.10	2.86**	0.04	1.09
Interactions	**Type*Mileage (loyalty points)**	-68.20	-0.79	-170.70	-1.76 .
	Type*Promotion codes	-59.32	-1.62	-28.56	-0.71
	Type*Giftcard, voucher, ecoupon or credit	55.60	1.6	75.67	1.99*
	Type*Cash or refund	20.85	0.68	17.19	0.51
	Intercept	-160882.5	-29.14***	-167066.00	-29.64***
		BIC	899612.2	Adjusted R squared	0.0681

2.3.1.2.3. Another approach to the data

We also tested another approach to the data, with a dummy variable process-related failure for all process-failure factors and another for all result-failure factors. However, these two dummy variables are strongly corelated (-.84), we kept only the variable for process-related failure. Interestingly enough, process-related failures have a positive relationship with purchase value ($\beta = 16.77, p < .05$), while as study 2 later shows, they also have a positive relationship with vigilant complaints. Also, in this approach, among the recovery types, mileage and gift cards, vouchers, ecoupons and credits no longer have a significant relationship with purchase value. Instead, cash and refund have a negative relationship with purchase value ($\beta = -10.96, p < .05$). To avoid possible confusion, we will discuss only the results of the main approach.

2.3.1.3. Discussion of study 1

Even in the presence of different types of recovery, vigilant complaints still have a negative effect on purchase value. However, some other important factors have no significant relationship with purchase value (to be discussed in discussion of study 2, e.g. process-related factors and double-deviation factors). Among result-related factors, missing bags and call center and customer support are associated with decrease in purchase value, while the others have no significant relationships.

Group actions, i.e., number of people affected, are also associated with a decrease in purchase value. However, VIP-related factors are associated with an increase in purchase value. We will discuss further after study 2.

All the recovery types do not help to increase purchase value. Mileage and gift cards, vouchers, credits or ecoupons both have negative relationship with change in purchase value. Promotion codes and cash or refund have no significant relationship with change in purchase value. Among the different forms of compensations, the fixed value of mileage makes it less attractive than promotion codes, which are in discount percentages and give consumers more leeway for large purchases. As for gift cards, vouchers, credits or ecoupons, they are preferrable to points, due to the psychological effect of saving money (Antonides and Ranyard 2017). However, as their amounts are also fixed, in general, they are less popular than promotion codes. There might be two reasons explaining the unintuitive effect of cash or refund (i.e. not leading to increase in purchase value). In the airline context, this recovery type is often related to compensations required by law, (e.g., delays for more than three hours or cancellations within the airline's control), while the other recovery types are more likely to be goodwill compensations. Customers paid compulsory compensations are unlikely to feel indebted toward the firm. Also, as cash or refund does not tie customers to the loyalty program, even though they might react most positively to it, this alone might not be sufficient to encourage customers to make more purchases, unless other binding elements exist (e.g., a promotion program or shortage of alternative options).

Among the control variables, minimum age has a negative relationship with increase in purchase value, which means that customers with younger children are more likely to consume in the future. We did not include the maximum age in the models (study 1 and study 2) due to its strong correlation with the minimum age. Ticket costs are associated with an increase in purchase value. We will discuss these factors further after study 2.

2.3.2. Study 2: Antecedents of complaint types

2.3.2.1. Logistic regression

To measure the relationship between different preceding factors with vigilant complaints vs. conciliatory complaints, we ran a binary logistic regression model, with all these variables as predictors. From the original dataset of 1,293,453 observations (see study 1), to be able to conduct logistic regression in study 2, however, we kept only the observations with all the needed variables available, i.e., 354,539 complaints. In this sample, we have 1.23 observations per customer (see descriptive data in Appendix 2) and 9,576 vigilant complaints, or 2.7% of all the sample.

The correlations of the included variables all fall far below .7 (Appendix D). The highest variance inflation factor (VIF among the variables is 3.02, much lower than the cut off value of 5, indicating no multicollinearity issue).

Table 2-4: Logistic regression and mixed effects logistic regression on vigilant complaints vs. conciliatory complaints (N = 354,539)

		Logistic regression		**Mixed effects logistic regression**	
Concepts	**Variables**	**β**	**z**	**β**	**z**
Process-based failures	Service at airport	1.05	27.31***	1.01	24.98***
	Service onboard	0.48	10.42***	0.47	9.82***
	Delays	0.06	1.6	0.06	1.55
	Other experience issues	-0.03	-0.56	-0.03	-0.47
Result-based failures	Missing bags	-0.68	-10.6***	-0.67	-10.23***
	Damaged bags	-1.2	-16.61***	-1.18	-16.04***
	Shipped equipment issues	-0.25	-1.94 .	-0.25	-1.85
	Other bag issues	-0.42	-8.08***	-0.42	-8.02***
	Refund requests	-0.54	-12.05***	-0.52	-11.14***
	Call center and customer support	0.03	0.73	0.04	0.78
Group actions	Number of people affected	0.06	6.09***	0.06	5.63***
Moral-based failures	Disability-related	0.41	5.00***	0.44	5.04***
Double deviation	Enumeration of complaints over the same flight	0.03	2.73**	0.03	3.06**
	Count of complaints of the same flight from the same consumer	0.05	7.49***	0.05	4.42***
	Duplicate complaints	0.29	6.38***	0.24	4.92***
Belonging	Member (vs being non-member)	-0.07	-2.89**	-0.07	-2.76**
VIP status	Top tier	-0.23	-3.85***	-0.23	-3.68***

		Logistic regression		Mixed effects logistic regression	
	Upper mid-tier	-0.18	-3.75***	-0.18	-3.61***
	Lower mid-tier	0.01	0.26	0.00	0.05
Control variables	Temperature of the destination location	0.004	4.21***	0.00	3.81***
	Precipitation rate of the destination location	0.02	1.49	0.01	1.35
	Temperature of the departure location	0.003	3.38**	0.003	3.19**
	Precipitation rate of the departure location	-0.02	-1.4	-0.02	-1.45
	Month of the complaint	-0.01	-3.57***	-0.01	-3.35***
	Year of the complaint	-0.04	-5.54***	-0.04	-5.5***
	Female ratio of the traveling group	-0.02	-0.81	-0.02	-0.72
	Ratio of unidentified gender of the traveling group	0.9	1.96 .	0.83	1.67
	Complainer being of unidentified gender (vs other genders)	-0.07	-0.71	-0.06	-0.58
	Daily journey duration in the last three years	0.001	0.73	0.00	0.32
	Duration of the journey related to the complaint	0.00002	3.91***	0.00002	3.68***
	Minimum age among passengers traveling in groups	-0.002	-3.18***	-0.002	-2.87**
	Ticket costs	-0.00002	-2.68**	-0.00002	-2.58**
	Number of tickets	-0.02	-4.73***	-0.02	-4.09***
	Intercept	79.47	5.3***	83.2	5.27***
	BIC	85168.79		84274.58	

Note: . p <.1, * p<.05, ** p< .01,

The results (Table 2-4) indicate that process-based factors often have positive relationships with the probability of a complaint being vigilant vs. conciliatory: service at airports (M = .06; SD = .24; β = 1.05, p < .001), and service onboard (M = .07, SD = .26; β = .48, p < .001), though delays and other experience issues have no significant relationship (H2 confirmed).

In turn, result-based factors often have negative relations with the odds of a complaint being vigilant—that is, missing bags (*M* = .07, *SD* = .25; β = -.68, $p < .001$; damaged bags (*M* = .09, *SD* = .28; β = -1.2, $p < .001$; other bag issues (*M* = .16, *SD* = .37; β = -0.42 , $p < .001$), and refund request (*M* = .17, *SD* = .37; β = -.54 , $p < .001$), although shipped equipment issues (*M* = .01, *SD* = .11) and call center and customer support (*M* = .16, *SD* = .36) have no significant relationship (H3 confirmed).

Also, confirming H4 on the role of the group actions, the number of affected people in a service failure has a positive relationship with vigilant complaints vs. conciliatory complaints *(M =1.28, SD = .84,* β = .06, $p < .001$). A complaint that mentions a disability issue (*M* = .01, *SD* = .09, β = .41, $p < .001$) has a positive relationship with the probability complaint being vigilant (H5 on moral-based failures confirmed). Similarly, variables which indicate the perception of the transgressions being repeated are also positively associated with the odds of vigilant complaints: enumeration of complaints over the same flight (*M* = 1.39, *SD* = 1.13; β = .03, $p < .01$), count of complaints of the same flight from the same consumer (*M* = 1.56, *SD* = 1.27; β = .05, $p < .001$) and duplicate complaints (*M* = .06, *SD* = .23; β = .29 , $p < .001$) (i.e., H6 on double deviation confirmed).

However, being a loyalty member has a negative relationship with this probability (*M* = .66, *SD* = .48; β = -.07, $p < .01$) (H7 confirmed). So does having the VIP status. For the most prestigious, being a top tier member (*M* = .05, *SD* = .22), this negative relationship is the strongest (β = -.23, $p < .001$), while for the middle level members, i.e., upper mid-tier (*M* = .06, *SD* = .25) members, this relationship is a little weaker (β = -.18, $p < .001$), though still being negative (H8 confirmed). However, being a lower mid-tier (*M* = .05, *SD* = .22) member has no significant difference than being a bottom tier member (*M* = .49, *SD* = .5) in the association with vigilant complaints. The model has pseudo R squared of .038, log likelihood of -42367.16, AIC of 84802.31 and BIC of 85168.79.

2.3.2.2. Robustness checks

2.3.2.2.1. Haussmann test for endogeneity

In this study, we do not have a unique variable that we can test as an endogenous regressor (in study 1, it is vigilant complaints). Therefore, we used the Durbin–Wu–Hausman test, which does not require us to specify an endogenous regressor (Wooldridge 2010). We compared two different models, in this case, our original simple logistic regression model versus a random effect model, (i.e. a mixed effect logistic regression model), using the package plm in R (Croissant and Millo 2008). The result is significant, $x^2(34, N = 354{,}539) = 325.2$, $p < .001$.

However, the estimated coefficients of both models are similar or close for many variables, especially those with significant effects, and the earlier findings of the simple logistic regression still stand for the mixed effects model. In other words, while there is omitted variable bias, our results can be considered robust.

2.3.2.2.2. Interactions

We ran interactions between all the failure types and all relational factors (Table 5). Among these interactions, seven are significant. The two only positive interactions are those of call center and customer support and members ($\beta = .3$, $p < .001$) and shipped equipment issues and members ($\beta = .65$, $p < .05$). All the others are negative. They are the interactions of other experience issues and members ($\beta = -.32$, $p < .05$), service onboard and the top tier ($\beta = -.48$, $p < .05$), service onboard and the upper mid-tier ($\beta = -.45$, $p < .05$), call center and customer support and the upper mid-tier ($\beta = -.66$, $p < .01$), and finally, other bag issues and the lower mid-tier (β

= -.57, $p < .05$). The pseudo-R squared of this logistic regression with interactions is .039, log likelihood = -42315.19, BIC= 85678.23 and AIC = 84794.39.

Table 2-5: Simple logistic regression with interactions on vigilant complaints

Concepts	Constructs	β	z
Process-based failures	Service at airport	1.07	16.51***
	Service onboard	0.63	7.79***
	Delays	0.11	1.81 .
	Other experience issues	0.14	1.39
Result-based failures	Missing bags	-0.69	-7.34***
	Damaged bags	-1.0	-9.17***
	Shipped equipment issues	-0.71	-2.67**
	Other bag issues	-0.37	-4.47***
	Refund requests	-0.52	-7.03***
	Call center and customer support	-0.15	-1.84 .
Group actions	Number of people affected	0.06	4.53***
Moral-based failures	Disability-related	0.37	2.93**
Double deviation	Enumeration of complaints over the same flight	0.03	2.8**
	Count of complaints of the same flight from the same consumer	0.05	7.39***
	Duplicate complaints	0.3	6.45***
Belonging	Member (vs being non-member)	-0.06	-0.77
VIP status	Top tier	-0.11	-0.61
	Upper mid-tier	0.01	0.06
	Lower mid-tier	0.06	0.36
Control variables	Temperature of the destination location	0.003	4.09***
	Precipitation rate of the destination location	0.02	1.5
	Temperature of the departure location	0.003	3.61***
	Precipitation rate of the departure location	-0.02	-1.43
	Month of the complaint	-0.01	-3.59***
	Year of the complaint	-0.04	-5.28***
	Female ratio of the traveling group	-0.02	-0.83
	Ratio of unidentified gender of the traveling group	0.9	1.96 .
	Complainer being of unidentified gender (vs other genders)	-0.07	-0.75
	Daily journey duration in the last three years	0.0008	0.91

Concepts	Constructs	β	z
	Duration of the journey related to the complaint	0.000005	3.9***
	Minimum age among passengers traveling in groups	-0.002	-3.19***
	Ticket costs	-0.00002	-2.69***
	Number of tickets	-0.02	-4.72***
Interactions	Service at airport *member	-0.01	-0.16
	Service onboard*member	-0.09	-0.92
	Delays*member	-0.07	-0.95
	Other experience issues*member	-0.32	-2.41*
	Missing bags*member	-0.004	-0.03
	Damaged bags*member	-0.29	-1.96 .
	Shipped equipment issues*member	0.65	2.06*
	Other bag issues*member	-0.02	-0.16
	Refund requests*member	-0.006	-0.06
	Call center and customer support*member	0.30	3.01***
	Number of people affected*member	0.002	0.09
	Disability-related*member	0.04	0.23
	Service at airport *top tier	-0.26	-1.44
	Service onboard*top tier	-0.48	-2.38*
	Delays*top tier	-0.14	-0.85
	Other experience issues*top tier	0.14	0.64
	Missing bags*top tier	0.46	1.41
	Damaged bags*top tier	0.11	0.28
	Shipped equipment issues*top tier	-0.86	-0.84
	Other bag issues*top tier	-0.46	-1.66 .
	Refund requests*top tier	0.09	0.34
	Call center and customer support*top tier	0.03	0.11
	Number of people affected*top tier	0.06	1.03
	Disability-related*top tier	0.82	1.31
	Service at airport*upper mid-tier	-0.13	-0.83
	Service onboard*upper mid-tier	-0.45	-2.38*
	Delays*upper mid-tier	-0.03	-0.22
	Other experience issues*upper mid-tier	0.08	0.38
	Missing bags*upper mid-tier	0.14	0.5
	Damaged bags*upper mid-tier	-0.46	-1.18
	Shipped equipment issues*upper mid-tier	-0.57	-0.92
	Other bag issues*upper mid-tier	0.43	1.89 .
	Refund requests*upper mid-tier	-0.43	-1.87 .
	Call center and customer support*upper mid-tier	-0.66	-2.85**
	Number of people affected*upper mid-tier	-0.02	-0.29
	Disability-related*upper mid-tier	0.57	1.27
	Service at airport *lower mid-tier	-0.009	-0.06

Concepts	Constructs	β	z
	Service onboard*lower mid-tier	-0.24	-1.26
	Delays*lower mid-tier	0.05	0.31
	Other experience issues*lower mid-tier	0.16	0.67
	Missing bags*lower mid-tier	-0.02	-0.07
	Damaged bags*lower mid-tier	-0.41	-0.99
	Shipped equipment issues*lower mid-tier	0.80	1.71
	Other bag issues*lower mid-tier	-0.57	-2.52*
	Refund requests*lower mid-tier	-0.11	-0.54
	Call center and customer support*lower mid-tier	0.3	1.48
	Number of people affected*lower mid-tier	-0.007	-0.12
	Disability-related*lower mid-tier	-0.42	-0.79
Intercept	Intercept	75.7	5.04***

2.3.2.2.3. Another approach to the data

We also tested another approach to the data, with a dummy variable for all process-failure factors and another for all result-failure factors. Again, in this sample, as these two dummy variables are strongly corelated (-.84), we kept only the variable for process-related failure factors. In the logistic regression, process-related failures have a strong positive relationship with vigilant complaints ($\beta = .97$, $p < .001$). The previous findings for the relevant hypotheses still stand.

Among the interactions, however, when we use only process-related failures as the dummy variable and remove all other process-related and result-related factors, we do not have any significant interactions between the variable representing process-related failures and the relational factors.

2.3.2.3. Discussion of study 2

Confirmations of H2 – H8 underline the importance of both relational characteristics (i.e., membership vs non-memberships and membership rankings) and failure types to determine how complainers use their resources and make complaint decisions.

2.3.2.3.1. The role of unpredictability in drainage of resources

Among process-based failures, service at the airport had a stronger positive relationship with vigilant complaints than service on board does. Compared with services onboard, there are more combined factors at the airport for which consumers are not able to predict, e.g., check-in issues, availability of services and assistance at the airport, issues due to long transit time, etc. This unpredictability draws more resources of the customers, making them even more likely to give up investing in their relationship with the firm.

Among the results-related factors, damaged bags had a strongest negative relationship with vigilant complaints. This context has a clear result and the unpredictability level is in general lower than other result-based factors, leading customers more likely to invest in resources to achieve desirable solutions. Other factors, such as refund requests and missing bags might have more unknown issues (e.g. whether and when the bags would or did finally arrive, the appropriate level or eligibility of refunding), which draw more resources of the customers, making them less conciliatory.

2.3.2.3.2. Comparing results of study 1 and study 2

Some factors which have strong relationships with the probability of the complaints being vigilant in study 2 do not have significant relationship with change in purchase value (e.g, enumeration of complaints over the same flight and service at the airport). In some cases, this might indicate the possibility of vigilant complaints being the mediator (Li et al. 2007) in the relationship between these factors and change in purchase value and these are full mediations. Nevertheless, this speculation needs to be checked on the same sample (i.e., only among loyalty members for both studies) and with linear regression (See Appendix D) instead of mixed linear regression being used for study 1.

However, another possibility could also be different relationships of these variables with vigilant complaints and purchase value. For example, missing bags have negative relationship with the probability of vigilant complaints but also negative relationships with purchase value. This indicates that in these cases, even though consumers are more likely to adopt conciliatory tones to achieve solutions, they could still end up punishing the firms.

Some other variables show similar unintuitive effects. Among the control variables, the minimum age among passengers traveling in the group has a negative relationship with the probability of a complaint being vigilant. In other words, the less this minimum age is, the more likely the complaint is vigilant. Thus, customers traveling with very young children are also more likely to turn vigilant in case of service failures. However, as discussed in study 1, this variable also has a negative relationship with purchase value change, which means customers with very young children will increase purchase value in the future. In other words, it seems that even though these consumers become upset more easily, they also forgive more easily. Similarly, duration of the journey related to the complaint also has a positive relationship with vigilant complaints, indicating a drain of resources of the complainer of long trips. However, these complainers are more likely to stick with the firm in the future, with positive relationship with change in purchase value.

Some other variables, however, show consistent relationships at both stages. Ticket costs have a negative relationship with vigilant complaints, probably indicating higher levels of seats (e.g., business class or first class). This shows that the complainers probably have more resources available and are thus less likely to file vigilant complaints. And these complainers at the later stage also increase purchase value. Similarly, higher level members are less likely to be vigilant and later on, also increase purchase value.

2.3.2.3.3. Discussion of interactions

Among the seven significant interactions between failure types and relational factors, most of them are negative, indicating that members, especially higher ranked members, are often less likely to be vigilant in diverse contexts. However, in two exceptions to this rule, members are more likely to send vigilant complaints than non-members, when it comes to call center and customer support and shipped equipment issues. This might indicate that while members often invest more resources in their relationships with the firm, in a couple of contexts, they also expect more efforts from the firms. Specifically, in these cases, the entitlement (Boyd Iii and Helms 2005) is greater than the perceived benefits of the relationship, which ends up driving up the likelihood of vigilance. It should be noted that the interaction between call center and customer support and upper mid-tier is negative, and the main effects of both this context and shipped equipment issues before interaction analyses are both insignificant, indicating different trends among different groups of customers.

2.4. General discussion

2.4.1. Justice theory versus the theory of conservation of resources

Justice theory has long played a central role in service failure and recovery research, focusing on three types of justice (distributive, procedural and interactional justice) (e.g. Cai and Qu 2018). For example, depending on the types of failures (Grégoire et al 2010) and the relationships between the firm and the customers, customers could expect different types of fairness (Grégoire and Fisher 2008). However, in this research, we use the theory of conservation of resources, as it helps us to build our arguments in three aspects.

First, it allows us to explain the different mechanisms of using and protection of resources between vigilant and conciliatory complainers. By going beyond the justice perspective of the two schemas (Grégoire et al 2019), we prepare the groundwork for the links between each type of complaints with different stages of consumer experience. This promises to expand theory on consumers' experience in services through the resource usage, resource investment and earning resources. Challenges in research using longitudinal data is not only methodological, but also theoretical, as linking different stages of consumer behavior through theory requires smooth transition. As explained further below, the flexibility of resource theory makes this easier than using justice theory, when the dynamics become more complex.

Second, the resource theory allows us to connect the different amounts of resources with different levels in constructs, e.g., different levels of membership and with members and non-members. Among the variables, we have the number of people affected, which is also challenging to explain its relationship with the likelihood of a complaint being vigilant through the more abstract justice theory. But this relationship is quite straightforward when viewed in terms of resources.

Third, backed up with rich literature on stress and social adaptations (Hobfoll 1989, 2002), the lenses of resources also give us the ability to examine the wellbeing motivations of consumers. This understanding could encourage more research on service failures affecting vulnerable population. More vulnerable consumers have fewer resources and are thus more sensitive to any perceived loss than other consumer groups. Furthermore, by using the resource perspective in service failures and at different stages of the complaint process, we expand the possibility of helping firms support better customers' resources, and thus also their wellbeing (Doane et al. 2012). This also allows firms to develop long-term approaches with customers at different stages of the relationship.

2.4.2. Complex dynamics among constructs

Among the randomly collected public complaints by Meire et al (2024), vigilant complaints accounted for as much as 39% complaints, while in our data of private complaints, vigilant complaints account for only 2.7% of the complaints. Even with the relatively modest number of vigilant private complaint, their effect on purchase value is considerable. This also indicates possible differences in motivations between private and public conciliatory/vigilant complaints.

The rich dataset also allows us to detect complex dynamics between constructs. As indicated in discussion of study 1 and study 2, some variables (e.g. duration of the journey related to the complaint) are positively associated with the likelihood of vigilant complaints but at the later stages with an increase in purchase value. Or vice versa, some are negatively associated with the likelihood of vigilant complaints (e.g. missing bags) but at the later stage with a decrease in purchase value. This indicates the diverse routes of different variables at each stage of service failure and recovery journey.

In addition, many interactions between relational constructs and failure types have negative relationships with vigilant complaints, reflecting the buffering effect of membership and especially higher ranks on different failure factors. However, two interactions of membership have positive relationships with vigilant complaints. In these two cases, call center and customer support and shipped equipment issues, members' sense of entitlement (Fisk 2010) might reflect the competition for limited resources.

Furthermore, as discussed earlier, unpredictability at some stage of the service failure journey might also increase the likelihood of a complaint of becoming vigilant, due to the increased stress. Temporary unpredictability could lead to anxiety, as these consumers anticipate negative emotions and possible threats to their security (Lake and LaBar 2011).

2.4.3. Managerial implications

Identifying antecedents of vigilant complaints help firms to proactively work on these fronts to reduce the appearance of vigilant complaints in the future, for example encouraging more non-members to become members and lower-ranked members to become higher-ranked members. This could be done through promotions of the benefits or allowing customers to enjoy the benefits of members or higher-ranked members for a brief period.

Furthermore, different people have different amounts of resources, and they perceive the loss of resources also differently. Thus, attention should also be paid to vulnerable people who also get more affected in case of service failures, due to their possession of fewer resources. This explains why among factors associated with vigilance are complaints related to disability issues and to having young children in the traveling groups. Depending on their profiles, firms could offer social resources to these complainers. For example, if there are young children in the reservation, they could offer little gifts for these children along with the compensations.

In general, loyalty membership and higher ranks in membership help to reduce vigilance and do so in diverse failure contexts. Apart from investing in benefit programs and the promotions of these benefits for members and higher-ranked members, as they will encourage customers to stay conciliatory, firms also need to pay attention to the few exceptional contexts where membership could also backfire on them.

Identifying the reduction of purchase value in dollar terms means that firms could also measure the financial risks of vigilant complaints and are able to compare them against other types of risks. This tangibility also allows for a more extensive assessment in the future related to other people in these complainers' circles, for example those who accompanied them on the relevant flights or other flights and whether there were similar trends in their behavior.

2.4.3. Limitations and future research

For the study on the consequence of vigilant complaints, our sample is limited to loyalty members. Future research could compare behavior for non-members and members after sending vigilant complaints in the post-recovery stage. It is possible that vigilant non-members will be even more avoidant toward the firm than vigilant members.

In addition, our data do not include the text of responses to consumers. Future research could examine the effects of different responses to vigilant complaints. It is possible responses with more empathy in combination with certain types of compensations will help vigilant complainers recover more resources and thus react more positively to the solutions.

Furthermore, in this research, we only analyzed English complaints. While our sample is likely to include English speakers from different cultural backgrounds, it is possible that complaints in other languages have distinctive characteristics compared to complaints in English. For example, complainers using less popular languages might appreciate recovery efforts in their mother tongues, as like the effect of loyalty membership, they might feel belonging to the same community as the firm's representatives speaking their languages.

As we discussed earlier, vigilant complaints are more common on public platforms than in the private setting, as vigilant complainers in public settings seek to reach also other audience beyond the firms in indirect revenge. Future research could analyze the behavior of private complainers' behavior in public. In what circumstances are they going to send vigilant complaints both in public and private and in what circumstances will they only send vigilant complaints in one type of setting? Research on private versus public complaints and their convergent and divergent routes promises to bring important contributions to the field.

Finally, a variable that we do not have in our data is complainers' purpose for traveling. Depending on whether consumers had to pay out of their pockets for their trips (i.e., the trips having business or leisure purposes), they could have different reactions towards service failures. Consumers whose companies paid for their flights might be tied to that airline even when they were not happy with the quality of service. However, from the angle of resources, not having to pay for the tickets themselves can be considered a type of resource, which could make them less likely to send vigilant complaints.

This type of service captivity is different than the scenarios where consumers do not have the option to leave the provider due to their lack of resources. A typical scenario is in the health services where patients might have no choice but to stick with their current doctor/hospital, as they cannot afford to switch to a better one. Future research could examine how consumers react to service failures in different forms of service captivity.

References

Adams, J. S., & Freedman, S. (1976). Equity Theory Revisited: Comments and Annotated Bibliography. In *Advances in Experimental Social Psychology* (Vol. 9, pp. 43–90). Elsevier. https://doi.org/10.1016/S0065-2601(08)60058-1

Aggarwal, P. (2004). The Effects of Brand Relationship Norms on Consumer Attitudes and Behavior. *Journal of Consumer Research*, *31*(1), 87–101. https://doi.org/10.1086/383426

Antonides, G., & Ranyard, R. (2017). Mental Accounting and Economic Behaviour. In R. Ranyard (Ed.), *Economic Psychology* (1st ed., pp. 123–138). Wiley. https://doi.org/10.1002/9781118926352.ch8

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*(25), 3083–3107. https://doi.org/10.1002/sim.3697

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, *10*(2), 150–161. https://doi.org/10.1002/pst.433

Aziz, R., & Rahman, Z. (2022). Brand hate: a literature review and future research agenda. *European Journal of Marketing*, *56*(7), 2014–2051. https://doi.org/10.1108/EJM-03-2021-0189

Basso, K., & Pizzutti, C. (2016). Trust Recovery Following a Double Deviation. *Journal of Service Research*, *19*(2), 209–223. https://doi.org/10.1177/1094670515625455

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

Bechwati, N. N., & Morrin, M. (2003). Outraged Consumers: Getting Even at the Expense of Getting a Good Deal. *Journal of Consumer Psychology*, *13*(4), 440–453. https://doi.org/10.1207/S15327663JCP1304_11

Beverland, M. B., Kates, S. M., Lindgreen, A., & Chung, E. (2010). Exploring consumer conflict management in service encounters. *Journal of the Academy of Marketing Science*, *38*(5), 617–633. https://doi.org/10.1007/s11747-009-0162-0

Bies, R. J., Barclay, L. J., Tripp, T. M., & Aquino, K. (2016). A Systems Perspective on Forgiveness in Organizations. *Academy of Management Annals*, *10*(1), 245–318. https://doi.org/10.5465/19416520.2016.1120956

Boyd Iii, H. C., & Helms, J. E. (2005). Consumer entitlement theory and measurement. *Psychology and Marketing*, *22*(3), 271–286. https://doi.org/10.1002/mar.20058

Cai, R., & Qu, H. (2018). Customers’ perceived justice, emotions, direct and indirect reactions to service recovery: Moderating effects of recovery efforts. *Journal of Hospitality Marketing & Management*, *27*(3), 323–345. https://doi.org/10.1080/19368623.2018.1385434

Choi, S., & Mattila, A. S. (2008). Perceived controllability and service expectations: Influences on customer reactions following service failure. *Journal of Business Research*, *61*(1), 24–30. https://doi.org/10.1016/j.jbusres.2006.05.006

Croissant, Y., & Millo, G. (2008). Panel Data Econometrics in *R* : The **plm** Package. *Journal of Statistical Software*, *27*(2). https://doi.org/10.18637/jss.v027.i02

Cullinane, N., & Dundon, T. (2006). The psychological contract: A critical review. *International Journal of Management Reviews*, *8*(2), 113–129. https://doi.org/10.1111/j.1468-2370.2006.00123.x

Curci, A., Lanciano, T., Soleti, E., & Rimé, B. (2013). Negative emotional experiences arouse rumination and affect working memory capacity. *Emotion*, *13*(5), 867–880. https://doi.org/10.1037/a0032492

Danaher, P. J., Sajtos, L., & Danaher, T. S. (2020). Tactical use of rewards to enhance loyalty program effectiveness. *International Journal of Research in Marketing*, *37*(3), 505–520. https://doi.org/10.1016/j.ijresmar.2020.02.005

De Haan, E., Kannan, P. K., Verhoef, P. C., & Wiesel, T. (2018). Device Switching in Online Purchasing: Examining the Strategic Contingencies. *Journal of Marketing*, *82*(5), 1–19. https://doi.org/10.1509/jm.17.0113

Doane, L. S., Schumm, J. A., & Hobfoll, S. E. (2012). The Positive, Sustaining, and Protective Power of Resources: Insights from Conservation of Resources Theory. In K. Törnblom & A. Kazemi (Eds.), *Handbook of Social Resource Theory* (pp. 301–310). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-4175-5_19

Dorotic, M., Verhoef, P. C., Fok, D., & Bijmolt, T. H. A. (2014). Reward redemption effects in a loyalty program when customers choose how much and when to redeem. *International Journal of Research in Marketing*, *31*(4), 339–355. https://doi.org/10.1016/j.ijresmar.2014.06.001

Du, J., Fan, X., & Feng, T. (2014). Group Emotional Contagion and Complaint Intentions in Group Service Failure: The Role of Group Size and Group Familiarity. *Journal of Service Research*, *17*(3), 326–338. https://doi.org/10.1177/1094670513519290

Fetscherin, M. (2019). The five types of brand hate: How they affect consumer behavior. *Journal of Business Research, 101*, 116–127. https://doi.org/10.1016/j.jbusres.2019.04.017

Fisk, G. M. (2010). "I want it all and I want it now!" An examination of the etiology, expression, and escalation of excessive employee entitlement. *Human Resource Management Review*, *20*(2), 102–114. https://doi.org/10.1016/j.hrmr.2009.11.001

Glazer, A., Kanniainen, V., & Poutvaara, P. (2010). Firms' ethics, consumer boycotts, and signalling. *European Journal of Political Economy*, *26*(3), 340–350. https://doi.org/10.1016/j.ejpoleco.2009.12.002

Goldfarb, A., Tucker, C., & Wang, Y. (2022). Conducting Research in Marketing with Quasi-Experiments. *Journal of Marketing*, *86*(3), 1–20. https://doi.org/10.1177/00222429221082977

Grégoire, Y., & Fisher, R. J. (2008). Customer betrayal and retaliation: when your best customers become your worst enemies. *Journal of the Academy of Marketing Science*, *36*(2), 247–261. https://doi.org/10.1007/s11747-007-0054-0

Grégoire, Y., Ghadami, F., Laporte, S., Sénécal, S., & Larocque, D. (2018). How can firms stop customer revenge? The effects of direct and indirect revenge on post-complaint responses. *Journal of the Academy of Marketing Science*, *46*(6), 1052–1071. https://doi.org/10.1007/s11747-018-0597-2

Grégoire, Y., Laufer, D., & Tripp, T. M. (2010). A comprehensive model of customer direct and indirect revenge: understanding the effects of perceived greed and customer power. *Journal of the Academy of Marketing Science, 38*(6), 738–758. https://doi.org/10.1007/s11747-009-0186-5

Grégoire, Y., Legoux, R., Tripp, T. M., Radanielina-Hita, M.-L., Joireman, J., & Rotman, J. D. (2019a). What Do Online Complainers Want? An Examination of the Justice Motivations and the Moral Implications of Vigilante and Reparation Schemas. *Journal of Business Ethics*, *160*(1), 167–188. https://doi.org/10.1007/s10551-018-3850-1

Grégoire, Y., Legoux, R., Tripp, T. M., Radanielina-Hita, M.-L., Joireman, J., & Rotman, J. D. (2019b). What Do Online Complainers Want? An Examination of the Justice Motivations and the Moral Implications of Vigilante and Reparation Schemas. *Journal of Business Ethics*, *160*(1), 167–188. https://doi.org/10.1007/s10551-018-3850-1

Grégoire, Y., Tripp, T. M., & Legoux, R. (2009). When Customer Love Turns into Lasting Hate: The Effects of Relationship Strength and Time on Customer Revenge and Avoidance. *Journal of Marketing*, *73*(6), 18–32. https://doi.org/10.1509/jmkg.73.6.18

Gui, R., Meierer, M., Schilter, P., & Algesheimer, R. (2023). **REndo** : Internal Instrumental Variables to Address Endogeneity. *Journal of Statistical Software*, *107*(3). https://doi.org/10.18637/jss.v107.i03

Gunarathne, P., Rui, H., & Seidmann, A. (2018). When Social Media Delivers Customer Service: Differential Customer Treatment in the Airline Industry. *MIS Quarterly*, *42*(2), 489–520. https://doi.org/10.25300/MISQ/2018/14290

Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, Preventing, and Mitigating Online Firestorms in Brand Communities. *Journal of Marketing*, *83*(3), 1–21. https://doi.org/10.1177/0022242918822300

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513–524. https://doi.org/10.1037/0003-066X.44.3.513

Hobfoll, S. E. (2002). Social and Psychological Resources and Adaptation. *Review of General Psychology*, *6*(4), 307–324. https://doi.org/10.1037/1089-2680.6.4.307

Hobfoll, S. E., Freedy, J., Lane, C., & Geller, P. (1990). Conservation of Social Resources: Social Support Resource Theory. *Journal of Social and Personal Relationships*, *7*(4), 465–478. https://doi.org/10.1177/0265407590074004

Katja Wiemer-Hastings, K., & Xu, X. (2005). Content Differences for Abstract and Concrete Concepts. *Cognitive Science*, *29*(5), 719–736. https://doi.org/10.1207/s15516709cog0000_33

Khamitov, M., Grégoire, Y., & Suri, A. (2020). A systematic review of brand transgression, service failure recovery and product-harm crisis: integration and guiding insights. *Journal of the Academy of Marketing Science*, *48*(3), 519–542. https://doi.org/10.1007/s11747-019-00679-1

Kim, J.-S., & Frees, E. W. (2007). Multilevel Modeling with Correlated Effects. *Psychometrika*, *72*(4), 505–533. https://doi.org/10.1007/s11336-007-9008-1

Kim, P. H., Ferrin, D. L., Cooper, C. D., & Dirks, K. T. (2004). Removing the Shadow of Suspicion: The Effects of Apology Versus Denial for Repairing Competence- Versus Integrity-Based Trust Violations. *Journal of Applied Psychology*, *89*(1), 104–118. https://doi.org/10.1037/0021-9010.89.1.104

Lake, J. I., & LaBar, K. S. (2011). Unpredictability and Uncertainty in Anxiety: A New Direction for Emotional Timing Research. *Frontiers in Integrative Neuroscience*, *5*. https://doi.org/10.3389/fnint.2011.00055

Lambert, N. M., Stillman, T. F., Hicks, J. A., Kamble, S., Baumeister, R. F., & Fincham, F. D. (2013). To Belong Is to Matter: Sense of Belonging Enhances Meaning in Life. *Personality and Social Psychology Bulletin*, *39*(11), 1418–1427. https://doi.org/10.1177/0146167213499186

Leliveld, M. C., Van Dijk, E., & Van Beest, I. (2012). Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive injustice. *European Journal of Social Psychology*, *42*(2), 135–140. https://doi.org/10.1002/ejsp.872

Li, Y., Schneider, J. A., & Bennett, D. A. (2007). Estimation of the mediation effect with a binary mediator. *Statistics in Medicine*, *26*(18), 3398–3414. https://doi.org/10.1002/sim.2730

Meire, M., Grégoire, Y., Nguyen, C.M., Nepomuceno, M.V. (2024). Identifying and Recovering Online Complainers: A Combination of Machine Leaning and Experimental Evidence. Working paper

Nguyen, C. M., Nepomuceno, M. V., Grégoire, Y., & Legoux, R. (2023). Striking the Right Notes: Long- and Short-Term Financial Impacts of Musicians' Charity Advocacy Versus Other Signaling Types. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-023-05555-1

Obeidat, Z. M., Xiao, S. H., Qasem, Z. A., Dweeri, R. A., & Obeidat, A. (2018). Social media revenge: A typology of online consumer revenge. *Journal of Retailing and Consumer Services*, *45*, 239–255. https://doi.org/10.1016/j.jretconser.2018.09.010

Ringberg, T., Odekerken-Schröder, G., & Christensen, G. L. (2007). A Cultural Models Approach to Service Recovery. *Journal of Marketing*, *71*(3), 194–214. https://doi.org/10.1509/jmkg.71.3.194

Roskes, M., Elliot, A. J., Nijstad, B. A., & De Dreu, C. K. W. (2013). Avoidance Motivation and Conservation of Energy. *Emotion Review*, *5*(3), 264–268. https://doi.org/10.1177/1754073913477512

Simon, B. (2011). Not going to Starbucks: Boycotts and the out-scouring of politics in the branded world. *Journal of Consumer Culture*, *11*(2), 145–167. https://doi.org/10.1177/1469540511402448

Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. *Journal of Marketing Research*, *36*(3), 356. https://doi.org/10.2307/3152082

Stakhovych, S., & Tamaddoni, A. (2020). Mix&Match: A Resource-Based Complaint Recovery Framework for Tangible Compensation. *Journal of Service Research*, *23*(3), 337–352. https://doi.org/10.1177/1094670519898521

Suresh, P., & Chawla, V. (2022). The Burden of Double Deviation in Services: A Systematic Review and Research Agenda. *International Journal of Consumer Studies*, *46*(5), 1919–1941. https://doi.org/10.1111/ijcs.12836

Van Vaerenbergh, Y., & Orsingher, C. (2016). Service Recovery: An Integrative Framework and Research Agenda. *Academy of Management Perspectives*, *30*(3), 328–346. https://doi.org/10.5465/amp.2014.0143

Vermeer, S. A. M., Araujo, T., Bernritter, S. F., & Van Noort, G. (2019). Seeing the wood for the trees: How machine learning can help firms in identifying relevant electronic word-of-mouth in social media. *International Journal of Research in Marketing*, *36*(3), 492–508. https://doi.org/10.1016/j.ijresmar.2019.01.010

Ward, J. C., & Ostrom, A. L. (2006). Complaining to the Masses: The Role of Protest Framing in Customer-Created Complaint Web Sites. *Journal of Consumer Research*, *33*(2), 220–230. https://doi.org/10.1086/506303

Weiner, B., Amirkhan, J., Folkes, V. S., & Verette, J. A. (1987). An attributional analysis of excuse giving: Studies of a naive theory of emotion. *Journal of Personality and Social Psychology*, *52*(2), 316–324. https://doi.org/10.1037/0022-3514.52.2.316

Weitzl, W. J., & Hutzinger, C. (2019). Rise and fall of complainants' desires: The role of pre-failure brand commitment and online service recovery satisfaction. *Computers in Human Behavior*, *97*, 116–129. https://doi.org/10.1016/j.chb.2019.02.029

Wohl, M. J. A., & Branscombe, N. R. (2009). Group Threat, Collective Angst, and Ingroup Forgiveness for the War in Iraq. *Political Psychology*, *30*(2), 193–217. https://doi.org/10.1111/j.1467-9221.2008.00688.x

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, Mass: MIT Press.

Zhang, J., Zhang, Y., & Lu, L. (2023). The double-edged sword of employee forgiveness: How forgiveness motives steer forgiveness toward interpersonal citizenship behaviors and interpersonal deviance. *Business Ethics, the Environment & Responsibility*, *32*(4), 1247–1261. https://doi.org/10.1111/beer.12570

Zhang, Z., Kim, H. J., Lonjon, G., & Zhu, Y. (2019). Balance diagnostics after propensity score matching. *Annals of Translational Medicine*, *7*(1), 16–16. https://doi.org/10.21037/atm.2018.12.10

Zourrig, H., Chebat, J.-C., & Toffoli, R. (2015). "In-group love and out-group hate?" A cross cultural study on customers' revenge, avoidance and forgiveness behaviors. *Journal of Business Research*, *68*(3), 487–499. https://doi.org/10.1016/j.jbusres.2014.09.014

Web appendices

Appendix A

Coding book for complaint tweets

We focus on the context of luxury hotels, an area expecting high-quality services and quick resolutions to consumer complaints.

Vigilante tweet definition

By mentioning the firms in the tweets, the vigilantes use Twitter to put public pressure on them and punish them for the failures. There are four possible reasons for their decisions:

(1) Loyalty to the firms is answered with what they perceive as betrayal.
(2) They speak up on behalf of someone else who suffers from the perceived injustice from the firms or show support/build pressure by telling their own experience
(3) Multiple efforts to reach out to the firms do not lead to expected solutions
(4) They see the firms as being immoral and deliberate in the failures.

The language of vigilant tweets often demonstrate anger, often depicted by strong emotional or accusative words. They sometimes capitalize all letters in some words and use multiple exclamation marks (e.g. !!!).

Example 1:

@HiltonHotels should be ashamed of themselves charging $300 a night for this raggedy hotel. It's location doesn't even warrant this price tag EVEN IN downtown DC

Example 2:

@MarcEvanJackson @Marriott @FairfieldHotels I'm sorry this happened to you! Having your hotel room opened in the middle of the night is an absolute nightmare and 100% unacceptable!!

Vigilantes often use words showing anger, such as *appalled, should be ashamed, shameful, unacceptable, outrageous, garbage, pathetic, horrible, offensive, nightmare and horrified.* They can also sometimes use swear words.

To distinguish the occasional use of strong words in conciliatory messages and those in vigilant messages, the ones in vigilant messages tended to include more than one strong word in their complaints:

Example 3:

*@Hyatt @HyattConcierge @HyattPlaceDC_WH @HyattPlaceDC having **one of the worst hotel experiences** I have ever had. **Terrible place** to bring my guests from overseas*

Sometimes, this anger is apparent through irony. Below are three examples:

Example 4:

By the way if you still take the risk of staying with @HiltonHotels don't use the app. Please don't check in until you actually arrive and get a key card. Otherwise if you arrive late to a dirty room, you'll be blamed and called a liar and be charged and expected to stay in it.

Example 5:

@HiltonHotels it is truly remarkable your inability to put me in contact with someone from your fraud department to explain the inadequate training that your team members possess and the clear violation of PCI-DSS. I can only shake my head at this point.

Example 6:

@hotelshopuk @hotelassoc @HotelsUK @HiltonHotels @TUIGroup @allhoteldeals @bookingcom @TripAdvisor Why do hotels tuck the quilt in all around the mattress? You'd have to do a cartoon shuffle to slither into these 'envelopes'. Stop doing it, no one sleeps 'strapped in'.

Example 7:

@AirlineFlyer @Marriott @Hyatt @Verizon @TMobile Why do you think the Amtrak Guest Rewards program is so lame? No real partners just a hotel and car consolidator. It's basically an inward facing program closed off from the outside.

Another evidence of irony is the use of metaphors:

Example 8:

*Exampl@HyattAhmedabad AC doesn't work. Service is bad. So plan your stay accordingly. In last 18 hours, my food, sleep all messed up at this hotel with poor service **as icing on the cake** @HyattConcierge @Hyatte 8:*

Example 9:

*@HyattConcierge No hot water and water to shower at @ParkHyattDC - not nice 🚿. Had to do the **French shower way** 😰*

Example 10:

*@HyattConcierge @HyattConcierge, I have sent you a direct message, and provided you with the record of my conversation with the GM of your Mohican Sun property. How about you **get off your duff** and actually do something helpful, instead of trying to take this off-line to avoid looking bad?*

They also often accuse firms of greed and lack of morals, with such words as *rip-off, fraud, fraudulent, scam, corrupt, liar, crook, irresponsible, dupe and cheat.*

They can show disgust in their messages: *hate, loathe, disgusted, garbage, and repulsive.*

They can use vocabulary showing irony: *lame, cheeky and pathetic.*

And finally, they tend to advise the others to avoid using the firm's services: *please avoid, beware, not recommend, stay away, never staying again, #boycott.*

Conciliatory tweet definition

The conciliatory tweet users tend to use Twitter to resolve an issue with the company in a thoughtful manner. There are five possible reasons for their choice of using conciliatory tweets.

(1) They are currently in difficulty and need the firms' support to resolve it, for example unable to book or cancel a room
(2) They seek a solution for an issue that already took place in the past.
(3) They give feedback on an issue that took place in the past, to help the firm improve the services
(4) They join other Twitter users to seek the firm's solution to a similar issue.
(5) They show support to some other Twitter users who complain about the firms, but they do this in a factual and restrained manner.

Conciliatory tweets tend to describe disappointment but also demonstrate hope for the firm's support.

Example 1:

@HiltonHotels There are 2 unresolved issues: #90229115 Booked a handicap room weeks in advance. Got to the hotel to find it was just booked at 5pm that evening. I drop in from NJ with my elderly disabled parents. We had to leave and book a Marriott hotel.

Example 2:

@MarriottVacClub @MarriottBonvoy @Marriott On hold for over 1 hour and my 3rd/4th day trying to reach your office to cancel a reservation now only 24 hours away. Anything you can do over twitter DM?

Conciliatory tweets often use words indicating disappointment and/or sadness*: unfortunate, misfortune, discouraging, misleading, embarrassing, underwhelmed, dissatisfied, and frustrating.*

They also describe the unsatisfactory services, while often refraining from accusing the firms of being intentional in the failures. Conciliatory tweets also make efforts to stay factual, rather than using an aggressive or exaggeration style.

Thus, the choice of words is more careful and deliberate than that of the vigilantes, e.g., *unable to, unprofessional, abrasive, unreasonable, unfriendly, impolite, defective, inconsiderate, malfunctioning, inedible, dismissive, condescending, inconsistent, not helpful, less than accommodating, and indifference.*

Even when their experience was very unpleasant, a complaint could still be considered conciliatory. This often happens when they want the hotels to improve or get a solution, and their intention is not to damage the hotel brand. Thus, such words as *incredibly disappointed, a terrible experience*, *awful experience* can still be found in conciliatory complaints:

Example 3

Incredibly disappointed with @HiltonHotels @HiltonHHonors failure to follow their own stated price match policy and as a result, are expecting me to pay hundreds of dollars more than I should. What happened to customer service?

Example 4

Hey @HiltonHotels I stayed at Penn's Landing this weekend. Left my jacket in the closet. Called front desk got xfer to housekeeping and left a VM. No call back. How do I get my jacket back? Terrible customer service.

In the following example, the user uses quite strong words, but their intention is to help the hotel to improve "*DC and Hilton need to get this figured out*":

Example 5

This is disgusting every day. A liquid stream of waste flowing across the sidewalk smelling horrible. DC and Hilton need to get this figured out. @MayorBowser @HiltonHotels @CapitalHilton @DCDPW @DDOTDC @councilofdc https://t.co/JnJ3o1ORdz

Another strong expression can be found even in conciliatory complaints is *This is unacceptable,* if the user expresses their frustration to find a solution *We need to talk @HiltonHotels*

Example 6

We need to talk @HiltonHotels The front desk just gave someone a key to my room because he told them the wrong room number. No identification, no questions. They just walked into my room while I was in bed. This is unacceptable.

In addition, their messages might indicate efforts to communicate and reach solutions with firms, e.g., *please help, anything you can do, please DM, #fixit, get someone on the line, any responses, who can I talk to.*

Another feature of conciliatory complaints is that there might be some good/redeeming features about the brand despite the unpleasant experience:

Example 7

I’ve come to the conclusion that no matter how nice the hotel you stay at there’s always something broken @HiltonHotels https://t.co/hsAlNs4nsh

It should be noted also that as tweets are informal, users can sometimes use slangs in conciliatory words. For example, the word “shit” does not always indicate anger in informal English and can mean instead “stuff”. In the following conciliatory example, the user said *fix yo shit*, which means *fix your stuff.* Most of the time, though, “shit” is a vigilant word.

Example 8

@HyattConcierge Y'alls in room wifi isnt working (day 2) at hyatt Santa Clara.I would like to make you my home base but this isnt working.Yes I can connect, yes I get an IP, no I cant do an nslookup or get beyond the default GW for room wifi. I hate I have to do this. ***Fix U shit***

Appendix B: Descriptive statistics

2.B1. Descriptive statistics for the before-correction sample for regression

	Mean	Std	Min	Max
Purchase value [7]	160.86	868.83	-16741.8	29090.76
Change in redemption frequency	3.30	17.80	-343	596
Type	0.03	0.16	0	1
Mileage (loyalty points)	0.02	0.17	0	6
Promotion codes	0.34	0.51	0	9
Giftcard, voucher, ecoupon or credit	0.20	0.44	0	9
Cash or refund	0.37	0.68	0	16
Process-related failures	0.59	0.49	0	1
Result-related failures	0.41	0.49	0	1
Service at airport	0.08	0.27	0	1
Service onboard	0.11	0.32	0	1
Delays	0.32	0.47	0	1
Other experience issues	0.09	0.28	0	1
Missing bags	0.04	0.19	0	1
Damaged bags	0.08	0.27	0	1
Shipped equipment issues	0.01	0.09	0	1
Other bag issues	0.13	0.34	0	1
Refund requests	0.14	0.34	0	1
Call center and customer support	0.14	0.35	0	1
Number of people affected	1.22	0.67	1	27
Disability-related	0.01	0.07	0	1
Enumeration of complaints over the same flight	1.31	0.91	1	58
Count of complaints of the same flight from the same consumer	1.50	1.27	1	58
Duplicate complaints	0.05	0.22	0	1
Top tier	0.20	0.40	0	1
Upper mid-tier	0.23	0.42	0	1
Lower mid-tier	0.15	0.35	0	1
Temperature of the destination location	14.33	13.43	-46.44	49.73
Precipitation rate of the destination location	0.12	0.93	0	34.91

[7] As indicated in this research, the variable purchase value is created from the product of redemption frequency and the average value in dollars per redemption.

Temperature of the departure location	14.27	13.91	-46.25	50.34
Precipitation rate of the departure location	0.12	0.85	0	42.535
Month of the complaint	6.52	3.19	1	12
Year of the complaint	2020.82	1.47	2019	2022
Female ratio of the traveling group	0.41	0.41	0	1
Ratio of unidentified gender of the traveling group	0.00	0.02	0	1
Complainer being of unidentified gender	0.02	0.14	0	1
Daily journey duration in the last three years	18.74	26.97	-1.20548	593.063
Duration of the journey related to the complaint	1644.69	1994.15	-2006	46283
Minimum age among passengers traveling in the group	43.22	17.07	0	101
Ticket costs	1542.06	2233.35	0	107900
Number of tickets	2.82	3.03	1	75
Case duration	18.62	43.45	-1	1577
Number of days affected by COVID	191.15	100.72	0	394

2.B2. Descriptive statistics of the after-correction sample for regression

	Mean	**Std**	**Min**	**Max**
Purchase value	146.67	836.34	-10103.67	26845.50
Change in redemption frequency	3.00	17.13	-207.00	550.00
Type	0.03	0.18	0.00	1.00
Mileage (loyalty points)	0.02	0.17	0.00	6.00
Promotion codes	0.33	0.51	0.00	9.00
Giftcard, voucher, ecoupon or credit	0.23	0.46	0.00	9.00
Cash or refund	0.30	0.65	0.00	16.00
Process-related failures	0.71	0.45	0.00	1.00
Result-related failures	0.29	0.45	0.00	1.00
Service at airport	0.10	0.30	0.00	1.00
Service onboard	0.14	0.35	0.00	1.00
Delays	0.39	0.49	0.00	1.00
Other experience issues	0.09	0.28	0.00	1.00
Missing bags	0.02	0.15	0.00	1.00
Damaged bags	0.02	0.15	0.00	1.00
Shipped equipment issues	0.01	0.08	0.00	1.00
Other bag issues	0.11	0.31	0.00	1.00

Refund requests	0.09	0.29	0.00	1.00
Call center and customer support	0.14	0.34	0.00	1.00
Number of people affected	1.25	0.72	1.00	27.00
Disability-related	0.01	0.08	0.00	1.00
Enumeration of complaints over the same flight	1.33	0.89	1.00	30.00
Count of complaints of the same flight from the same consumer	1.52	1.27	1.00	58.00
Duplicate complaints	0.06	0.23	0.00	1.00
Top tier	0.19	0.39	0.00	1.00
Upper mid-tier	0.21	0.41	0.00	1.00
Lower mid-tier	0.15	0.36	0.00	1.00
Temperature of the destination location	14.57	13.31	-44.24	49.73
Precipitation rate of the destination location	0.13	0.98	0.00	34.91
Temperature of the departure location	14.55	13.79	-46.25	50.34
Precipitation rate of the departure location	0.12	0.79	0.00	41.49
Month of the complaint	6.53	3.17	1.00	12.00
Year of the complaint	2020.71	1.49	2019.00	2022.00
Female ratio of the traveling group	0.40	0.41	0.00	1.00
Ratio of unidentified gender of the traveling group	0.00	0.02	0.00	1.00
Complainer being of unidentified gender	0.02	0.14	0.00	1.00
Daily journey duration in the last three years	19.51	28.35	-1.21	593.06
Duration of the journey related to the complaint	1678.99	2051.43	-2006.00	46283.00
Minimum age among passengers traveling in the group	43.21	17.05	0.00	101.00
Ticket costs	1493.86	1973.24	0.00	83220.00
Number of tickets	2.79	2.85	1.00	55.00
Case duration	18.78	45.78	-1.00	1577.00
Number of days affected by COVID	187.25	100.18	0.00	394.00

2.B3. Descriptive statistics for the logistic regression

	Mean	Std	Min	Max
Vigilant complaints	0.03	0.16	0	1
Process-based failures	0.50	0.50	0	1
Result-based failures	0.50	0.50	0	1
Service at airport	0.06	0.24	0	1

Service onboard	0.07	0.26	0	1
Delays	0.32	0.46	0	1
Other experience issues	0.05	0.23	0	1
Missing bags	0.07	0.25	0	1
Damaged bags	0.09	0.28	0	1
Shipped equipment issues	0.01	0.11	0	1
Other bag issues	0.16	0.37	0	1
Refund requests	0.17	0.37	0	1
Call center and customer support	0.16	0.36	0	1
Number of people affected	1.28	0.84	1	82
Disability-related	0.01	0.09	0	1
Enumeration of complaints over the same flight	1.39	1.13	1	58
Count of complaints of the same flight from the same consumer	1.56	1.27	1	58
Duplicate complaints	0.06	0.23	0	1
Member (vs being non-member)	0.66	0.48	0	1
Top tier	0.05	0.22	0	1
Upper mid-tier	0.06	0.25	0	1
Lower mid-tier	0.05	0.22	0	1
Temperature of the destination location	14.33	13.82	-46.48	50.01
Precipitation rate of the destination location	0.13	0.88	0	42.124
Temperature of the departure location	14.44	14.32	-46.25	50.34
Precipitation rate of the departure location	0.13	0.85	0	42.535
Month of the complaint	6.84	3.42	1	12
Year of the complaint	2020.87	1.52	2019	2023
Female ratio of the traveling group	0.50	0.40	0	1
Ratio of unidentified gender of the traveling group	0.00	0.02	0	1
Complainer being of unidentified gender (vs other genders)	0.01	0.12	0	1
Daily journey duration in the last three years	8.24	15.48	-3.11872	593.063
Duration of the journey related to the complaint	1893.79	2108.91	-2851	57575
Minimum age among passengers traveling in groups	39.45	18.82	0	102
Ticket costs	1566.75	2528.87	0	193250.5
Number of tickets	3.21	3.82	1	156

Appendix C: 2C. Linear regression on purchase value

	Before correction		After correction	
Construct	**β**	**t value**	**β**	**t value**
Type	-40.06	-2.05*	-40.83	-2.17*
Mileage (loyalty points)	-39.29	-2.03*	-30.19	-1.49
Promotion codes	-3.49	-0.52	1.80	0.25
Giftcard, voucher, ecoupon or credit	-21.08	-2.66**	-14.54	-1.77 .
Cash or refund	-14.03	-2.67**	-15.36	-2.61*
Service at airport	-5.51	-0.37	-10.21	-0.7
Service onboard	-10.52	-0.7	-19.39	-1.28
Delays	-0.83	-0.06	-11.08	-0.83
Other experience issues	12.41	0.78	5.08	0.3
Missing bags	-66.26	-3.26**	-73.01	-2.91**
Damaged bags	-41.14	-2.41*	-48.64	-1.81 .
Shipped equipment issues	-46.11	-1.28	-55.81	-1.32
Other bag issues	-22.16	-1.45	-24.54	-1.43
Refund requests	-0.51	-0.04	-2.30	-0.14
Call center and customer support	-13.81	-0.92	-25.94	-1.65
Number of people affected	-16.06	-3.21**	-15.21	-2.99**
Disability-related	28.89	0.65	6.43	0.15
Enumeration of complaints over the same flight	4.66	1	1.35	0.26
Count of complaints of the same flight from the same consumer	6.49	1.95 .	7.26	2.03*
Duplicate complaints	-18.03	-1.06	-20.31	-1.14
Top tier	430.98	42.2***	406.10	36.51***
Upper mid-tier	178.68	20.97***	169.40	18.09***
Lower mid-tier	41.76	4.36***	36.56	3.57***
Temperature of the destination location	-0.73	-2.85**	-0.49	-1.77 .
Precipitation rate of the destination location	1.58	0.46	2.28	0.66
Temperature of the departure location	-0.99	-4.04***	-1.06	-3.98***
Precipitation rate of the departure location	8.09	2.19*	-0.68	-0.16
Month of the complaint	-20.03	-18.96***	-19.22	-16.7***
Year of the complaint	84.17	32.27***	82.76	29.66***
Female ratio of the traveling group	-27.48	-3.48**	-27.25	-3.19**
Ratio of unidentified gender of the traveling group	-67.33	-0.38	-31.08	-0.17
Complainer being of unidentified gender	4.28	0.18	4.84	0.19
Daily journey duration in the last three years	-4.15	-27.96***	-3.93	-25.76***
Duration of the journey related to the complaint	0.01	6.09***	0.01	4.68***

Minimum age among passengers traveling in the group	-1.46	-7.56***	-1.49	-7.11***
Ticket costs	0.01	5.98***	0.01	5.25***
Number of tickets	-7.94	-6.64***	-7.31	-5.29***
Case duration	0.02	0.2	-0.01	-0.07
Number of days affected by COVID	0.06	1.7 .	0.04	1.08
Intercept	-169700	-32.22***	-166900	-29.61***
Adjusted R squared		0.06781		0.06803

Appendix D: Correlations of variables

2.D1. Correlations of variables in logistic regression on vigilant complaints

a. Part 1

Variable	1	2	3	4	5	6	7	8
1. Type	1.00							
2. Service at airport	0.08	1.00						
3. Service onboard	0.03	-0.07	1.00					
4. Delays	0.02	-0.06	-0.15	1.00				
5. Other experience issues	0.00	-0.06	-0.07	-0.16	1.00			
6. Missing bags	-0.02	-0.07	-0.07	-0.18	-0.07	1.00		
7. Damaged bags	-0.04	-0.08	-0.09	-0.21	-0.07	-0.08	1.00	
8. Shipped equipment issues	-0.01	-0.03	-0.03	-0.07	0.04	-0.03	-0.03	1.00
9. Other bag issues	-0.02	-0.08	-0.12	-0.29	-0.10	-0.12	-0.13	0.18
10. Refund requests	-0.03	-0.09	-0.10	-0.30	-0.05	-0.12	-0.13	-0.05
11. Call center and customer support	-0.01	-0.11	-0.12	-0.21	-0.10	-0.12	-0.13	-0.05
12. Number of people affected	0.02	0.01	-0.01	0.12	0.01	-0.05	-0.08	-0.02
13. Disability-related	0.02	0.10	0.03	-0.01	0.002	-0.02	-0.02	0.03
14. Enumeration of complaints over the same flight	0.02	-0.02	-0.05	-0.01	-0.06	0.00	-0.03	-0.01
15. Count of complaints of the same flight from the same consumer	0.02	-0.02	-0.07	-0.05	-0.06	0.06	-0.03	-0.01
16. Duplicate complaints	0.02	-0.01	-0.04	-0.17	-0.06	0.09	-0.02	0.01
17. Member	-0.003	0.03	0.07	0.02	0.06	-0.09	-0.01	0.00
18. Lower mid-tier	0.001	0.01	0.03	-0.004	0.02	-0.01	-0.01	0.00
19. Upper mid-tier	-0.01	0.01	0.05	-0.004	0.04	-0.02	-0.01	-0.01
20. Top tier	0.00	0.02	0.07	0.004	0.06	-0.03	-0.02	-0.01
21. Temperature of the destination location	0.01	-0.02	-0.01	-0.001	-0.02	0.03	-0.05	-0.01
22. Precipitation rate of the destination location	0.002	0.001	-0.002	0.004	0.00	0.01	-0.001	0.001
23. Temperature of the departure location	0.01	-0.01	-0.004	0.01	-0.04	0.03	0.02	-0.01
24. Precipitation rate of the departure location	-0.003	-0.002	-0.003	0.003	0.003	-0.001	0.001	-0.002
25. Month of the complaint	-0.01	-0.03	-0.02	-0.04	-0.05	0.03	-0.02	-0.01
26. Year of the complaint	-0.02	-0.01	-0.09	-0.22	-0.10	0.19	0.10	0.03
27. Female ratio of the traveling group	-0.0005	0.00	-0.01	-0.02	-0.03	0.00	0.02	0.00
28. Ratio of unidentified gender of the traveling group	0.002	0.00	-0.002	-0.001	-0.002	0.000	0.002	0.000
29. Complainer being of unidentified gender	-0.002	0.00	-0.002	-0.003	-0.003	0.001	0.005	0.01
30. Daily journey duration in the last three years	0.004	0.02	0.07	0.05	0.08	-0.05	-0.03	-0.01
31. Duration of the journey related to the complaint	0.01	-0.02	-0.04	0.04	0.00	0.04	0.00	-0.02
32. Minimum age among passengers traveling in the group	0.003	0.05	0.05	0.02	0.06	-0.05	-0.04	-0.06
33. Ticket costs	-0.005	-0.02	0.02	-0.04	0.01	0.05	0.01	0.01
34. Number of tickets	-0.005	-0.02	-0.05	0.05	-0.02	0.02	-0.01	0.02

b. Part 2:

Variable	9	10	11	12	13	14	15	16
9. Other bag issues	1.00							
10. Refund requests	-0.05	1.00						
11. Call center and customer support	0.71	-0.17	1.00					
12. Number of people affected	-0.07	0.04	-0.08	1.00				
13. Disability-related	-0.01	-0.02	-0.02	0.01	1.00			
14. Enumeration of complaints over the same flight	0.04	0.00	0.07	0.09	0.00	1.00		
15. Count of complaints of the same flight from the same consumer	0.03	0.02	0.07	0.02	0.00	0.66	1.00	
16. Duplicate complaints	-0.11	-0.02	0.05	0.01	-0.01	0.22	0.26	1.00
17. Member	-0.05	-0.01	-0.05	-0.02	-0.02	-0.04	-0.02	-0.04
18. Lower mid-tier	-0.01	0.00	-0.01	0.00	-0.01	-0.01	-0.01	-0.01
19. Upper mid-tier	-0.01	-0.03	-0.01	-0.03	-0.01	-0.02	-0.02	-0.01
20. Top tier	-0.02	-0.06	-0.01	-0.03	-0.02	-0.03	-0.03	-0.02
21. Temperature of the destination location	0.05	-0.01	0.06	0.04	0.00	0.05	0.05	0.04
22. Precipitation rate of the destination location	0.002	-0.002	-0.002	0.003	0.001	0.000	-0.002	-0.002
23. Temperature of the departure location	0.02	-0.02	0.00	0.03	0.01	0.06	0.05	0.04
24. Precipitation rate of the departure location	0.0002	-0.002	0.0002	-0.003	-0.002	0.00001	0.002	0.001
25. Month of the complaint	0.004	0.01	-0.01	-0.01	0.00	0.03	0.02	0.01
26. Year of the complaint	0.14	0.08	0.11	0.00	-0.02	0.17	0.21	0.05
27. Female ratio of the traveling group	0.01	0.03	0.01	0.02	0.03	0.03	0.02	0.01
28. Ratio of unidentified gender of the traveling group	0.003	0.003	0.002	-0.004	-0.002	-0.001	-0.001	-0.001
29. Complainer being of unidentified gender	0.003	0.005	0.0002	0.001	-0.001	-0.002	0.000	-0.002
30. Daily journey duration in the last three years	-0.05	-0.08	-0.02	-0.04	-0.02	-0.05	-0.04	-0.03
31. Duration of the journey related to the complaint	0.01	-0.08	0.04	0.02	-0.02	0.05	0.06	0.03
32. Minimum age among passengers traveling in the group	-0.07	0.01	-0.06	-0.12	0.06	-0.08	-0.06	-0.03
33. Ticket costs	0.03	-0.05	0.03	0.15	-0.01	0.12	0.02	0.02
34. Number of tickets	0.00	-0.02	0.00	0.26	-0.01	0.20	0.08	0.02

c. Part 3

Variable	17	18	19	20	21	22	23	24	25
17. Member	1.00								
18. Lower mid-tier	0.17	1.00							
19. Upper mid-tier	0.19	-0.06	1.00						
20. Top tier	0.17	-0.05	-0.06	1.00					
21. Temperature of the destination location	-0.10	-0.02	-0.03	-0.04	1.00				
22. Precipitation rate of the destination location	0.002	-0.001	-0.004	0.001	-0.01	1.00			
23. Temperature of the departure location	-0.12	-0.03	-0.04	-0.04	0.35	-0.01	1.00		

24. Precipitation rate of the departure location	0.002	-0.004	-0.0005	-0.001	-0.01	0.004	-0.01	1.00	
25. Month of the complaint	-0.03	-0.01	0.00	-0.01	0.16	-0.01	0.14	-0.01	1.00
26. Year of the complaint	-0.04	0.01	-0.01	-0.02	0.02	0.00	0.03	0.00	-0.09
27. Female ratio of the traveling group	-0.05	-0.03	-0.08	-0.14	0.03	0.00	0.03	0.00	0.00
28. Ratio of unidentified gender of the traveling group	0.01	-0.002	-0.003	-0.002	-0.01	0.00	-0.01	0.01	0.01
29. Complainer being of unidentified gender	0.09	0.01	-0.01	-0.02	-0.01	0.00	-0.01	0.00	0.00
30. Daily journey duration in the last three years	0.21	0.04	0.18	0.48	-0.06	-0.01	-0.06	0.00	-0.01
31. Duration of the journey related to the complaint	-0.11	-0.03	-0.04	-0.02	0.11	-0.01	0.11	-0.01	-0.02
32. Minimum age among passengers traveling in the group	0.11	0.03	0.06	0.07	-0.02	-0.01	-0.03	0.00	0.03
33. Ticket costs	-0.04	0.00	0.02	0.05	0.08	-0.01	0.08	-0.01	0.01
34. Number of tickets	-0.06	-0.01	-0.03	-0.03	0.07	0.00	0.06	-0.01	-0.02

d. Part 4

Variable	26	27	28	29	30	31	32	33	34
26. Year of the complaint	1.00								
27. Female ratio of the traveling group	0.03	1.00							
28. Ratio of unidentified gender of the traveling group	0.01	-0.02	1.00						
29. Complainer being of unidentified gender	0.02	-0.01	0.07	1.00					
30. Daily journey duration in the last three years	-0.21	-0.15	0.00	-0.02	1.00				
31. Duration of the journey related to the complaint	-0.06	-0.01	-0.01	-0.02	0.17	1.00			
32. Minimum age among passengers traveling in the group	-0.10	-0.03	0.00	-0.02	0.12	-0.01	1.00		
33. Ticket costs	0.03	0.00	-0.01	-0.01	0.03	0.20	-0.08	1.00	
34. Number of tickets	0.06	0.03	-0.01	-0.01	-0.01	0.24	-0.20	0.44	1.00

2.D2. Correlations of variables in regressions on changes in purchase value

a. Part 1

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Change in purchase value	1.00														
Type	-0.01	1.00													
Mileage (loyalty points)	-0.003	0.01	1.00												
Promotion codes	-0.01	0.00	-0.08	1.00											
Giftcard, voucher, ecoupon or credit	0.01	0.03	-0.01	-0.28	1.00										
Cash or refund	-0.002	-0.03	-0.01	0.08	-0.08	1.00									
Service at airport	0.01	0.06	0.02	-0.02	0.01	-0.09	1.00								
Service onboard	0.004	0.002	0.04	0.10	0.00	-0.15	-0.13	1.00							
Delays	-0.03	-0.01	-0.01	-0.03	0.22	-0.05	-0.12	-0.29	1.00						
Other experience issues	0.01	-0.01	0.09	-0.09	-0.04	-0.06	-0.10	-0.13	-0.25	1.00					
Missing bags	0.01	0.00	-0.02	0.07	-0.07	-0.05	-0.05	-0.06	-0.12	-0.05	1.00				
Damaged bags	0.003	-0.01	-0.02	0.02	-0.07	0.10	-0.05	-0.06	-0.12	-0.04	-0.02	1.00			
Shipped equipment issues	-0.004	-0.01	-0.01	-0.02	-0.03	0.03	-0.03	-0.03	-0.06	0.05	-0.01	-0.01	1.00		
Other bag issues	0.01	-0.01	-0.05	-0.01	-0.09	0.26	-0.07	-0.14	-0.28	-0.11	-0.06	-0.05	0.13	1.00	
Refund requests	0.01	-0.02	-0.02	0.15	-0.06	0.25	-0.08	-0.11	-0.26	-0.06	-0.05	-0.05	-0.03	-0.02	1.00
Call center and customer support	0.01	0.00	-0.05	-0.07	-0.11	0.22	-0.13	-0.16	-0.21	-0.12	-0.06	-0.06	-0.03	0.68	-0.12
Number of people affected	-0.01	0.00	0.01	0.01	0.07	0.08	-0.01	-0.03	0.10	0.01	-0.04	-0.04	-0.02	-0.07	0.04
Disability-related	-0.01	0.02	0.00	-0.02	0.06	-0.01	0.07	0.00	-0.02	0.00	-0.01	-0.01	0.03	0.00	-0.01
Enumeration of complaints over the same flight	0.03	0.01	-0.03	-0.08	0.01	0.01	-0.04	-0.08	0.01	-0.06	0.00	0.00	-0.01	0.05	-0.01
Count of complaints of the same flight from the same consumer	0.03	0.01	-0.03	-0.07	0.00	-0.01	-0.04	-0.08	-0.01	-0.06	0.05	0.01	-0.01	0.04	0.01
Duplicate complaints	0.01	0.01	-0.03	-0.14	-0.11	-0.11	-0.03	-0.06	-0.19	-0.08	0.12	0.01	0.02	-0.09	0.00
Lower mid-tier	-0.02	0.01	-0.02	0.01	0.00	0.04	-0.02	-0.02	-0.01	-0.01	0.01	0.01	0.00	0.02	0.03
Upper mid-tier	0.04	0.00	0.00	0.01	0.00	-0.01	0.01	0.02	-0.01	0.00	0.02	-0.01	-0.01	0.00	-0.02
Top tier	0.10	0.00	0.10	-0.05	0.04	-0.05	0.03	0.09	-0.02	0.05	-0.01	-0.03	-0.01	-0.03	-0.08
Temperature of the destination location	0.00	0.00	-0.01	-0.04	0.06	0.05	-0.02	-0.02	-0.02	-0.04	0.06	-0.01	-0.01	0.08	0.01

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Precipitation rate of the destination location	0.004	-0.002	0.002	0.01	-0.001	-0.001	-0.001	-0.001	0.002	0.001	0.002	0.01	0.003	0.01	-0.004
Temperature of the departure location	0.00	0.00	-0.01	0.00	0.08	0.03	-0.01	0.00	0.00	-0.06	0.06	0.04	0.00	0.04	0.01
Precipitation rate of the departure location	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	-0.01	-0.01	0.00	-0.01	0.00
Month of the complaint	-0.07	-0.01	0.01	-0.05	0.04	0.00	-0.03	-0.01	-0.05	-0.06	0.02	0.00	-0.02	0.03	0.01
Year of the complaint	0.18	0.00	-0.03	-0.06	0.08	0.10	0.03	-0.06	-0.13	-0.07	0.12	0.07	0.00	0.16	0.09
Female ratio of the traveling group	-0.03	0.00	-0.04	0.02	0.00	0.04	0.01	-0.01	-0.02	-0.03	0.00	0.02	0.00	0.02	0.04
Ratio of unidentified gender of the traveling group	0.001	0.001	-0.003	0.01	-0.01	0.00	-0.01	-0.002	-0.002	-0.005	0.01	0.01	0.00	0.01	0.00
Complainer being of unidentified gender	0.003	-0.01	-0.01	0.002	0.002	0.02	-0.01	-0.02	0.00	-0.01	0.02	0.01	0.00	0.02	0.02
Daily journey duration in the last three years	-0.09	0.00	0.09	-0.04	0.02	-0.07	0.01	0.06	0.04	0.05	-0.04	-0.03	-0.01	-0.07	-0.10
Duration of the journey related to the complaint	-0.01	0.00	0.00	-0.04	0.02	0.00	-0.02	-0.03	0.03	0.02	0.02	0.01	-0.01	0.01	-0.05
Minimum age among passengers traveling in the group	-0.04	0.00	0.03	-0.03	-0.01	-0.07	0.03	0.05	0.00	0.05	-0.02	-0.04	-0.04	-0.07	-0.03
Ticket costs	0.03	0.00	0.03	-0.02	0.01	0.01	-0.01	0.06	-0.06	0.03	0.04	0.00	0.01	0.02	-0.05
Number of tickets	-0.01	0.01	-0.01	-0.02	0.05	0.07	-0.02	-0.07	0.06	-0.01	-0.01	-0.01	0.01	-0.01	0.00
Case duration	0.01	0.01	-0.02	0.04	0.13	0.05	-0.02	-0.07	0.07	-0.05	0.08	-0.02	-0.01	0.03	0.01
Number of days affected by COVID	0.05	-0.01	-0.02	-0.04	-0.01	-0.03	0.01	-0.04	-0.14	0.06	0.01	0.02	-0.01	0.04	0.07

b. Part 2

Variable	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Call center and customer support	1.00															
Number of people affected	-0.08	1.00														
Disability-related	-0.02	0.02	1.00													
Enumeration of complaints over the same flight	0.06	0.08	0.00	1.00												
Count of complaints of the same flight from the same consumer	0.06	0.03	0.03	0.64	1.00											
Duplicate complaints	0.06	0.00	0.00	0.25	0.26	1.00										
Lower mid-tier	0.01	0.05	0.00	0.01	0.01	0.01	1.00									
Upper mid-tier	0.01	-0.02	-0.02	0.00	0.00	0.01	-0.22	1.00								
Top tier	-0.02	-0.04	-0.02	-0.05	-0.05	-0.03	-0.21	-0.25	1.00							
Temperature of the destination location	0.08	0.03	-0.01	0.08	0.09	0.04	0.01	0.01	-0.03	1.00						
Precipitation rate of the destination location	-0.003	0.002	0.000	-0.01	-0.01	-0.01	0.00	-0.01	0.00	-0.01	1.00					
Temperature of the departure location	-0.01	0.03	0.004	0.09	0.08	0.03	0.02	0.01	-0.02	0.27	0.00	1.00				
Precipitation rate of the departure location	0.00	0.01	-0.005	-0.003	0.000	0.004	-0.01	0.00	0.00	0.01	0.00	-0.01	1.00			
Month of the complaint	-0.01	-0.03	0.00	0.07	0.04	0.02	0.00	0.02	0.01	0.23	-0.02	0.21	-0.01	1.00		
Year of the complaint	0.10	0.03	-0.02	0.18	0.19	0.07	0.05	0.05	-0.05	0.16	0.00	0.16	0.00	0.01	1.00	
Female ratio of the traveling group	0.01	0.07	0.03	0.04	0.03	0.02	0.02	-0.04	-0.19	0.04	0.00	0.03	0.00	-0.01	0.04	1.00

Variable	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Ratio of unidentified gender of the traveling group	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	-0.02
Complainer being of unidentified gender	0.01	0.01	0.00	0.01	0.01	0.01	0.01	-0.04	-0.05	0.02	0.01	0.02	-0.01	0.00	0.05	0.01
Daily journey duration in the last three years	-0.03	-0.07	-0.02	-0.07	-0.07	-0.04	-0.12	0.01	0.46	-0.08	-0.01	-0.08	0.00	-0.01	-0.34	-0.18
Duration of the journey related to the complaint	0.04	0.01	-0.02	0.04	0.05	0.00	-0.02	0.00	0.05	0.02	0.00	0.01	0.00	-0.03	-0.08	-0.01
Minimum age among passengers traveling in the group	-0.06	-0.15	0.05	-0.08	-0.07	-0.03	-0.03	0.05	0.08	-0.04	0.00	-0.05	0.00	0.03	-0.13	-0.04
Ticket costs	0.03	0.13	-0.01	0.06	0.02	0.01	-0.01	0.06	0.13	0.08	-0.01	0.07	-0.01	0.03	0.00	0.01
Number of tickets	0.00	0.30	0.00	0.16	0.11	0.02	0.03	-0.01	-0.03	0.05	0.00	0.03	0.00	-0.03	0.05	0.07
Case duration	0.02	0.04	0.00	0.04	0.05	-0.01	0.03	0.01	-0.09	0.07	0.00	0.07	0.00	0.05	0.14	0.03
Number of days affected by COVID	-0.01	-0.01	-0.01	0.00	0.02	0.00	0.02	0.01	-0.01	-0.03	0.02	-0.04	0.02	0.06	0.28	0.01

c. Part 3

Variable	32	33	34	35	36	37	38	39	40
Ratio of unidentified gender of the traveling group	1.00								
Complainer being of unidentified gender	0.09	1.00							
Daily journey duration in the last three years	-0.01	-0.06	1.00						
Duration of the journey related to the complaint	-0.01	-0.01	0.24	1.00					
Minimum age among passengers traveling in the group	-0.01	-0.07	0.12	0.05	1.00				
Ticket costs	-0.01	-0.02	0.08	0.28	-0.02	1.00			
Number of tickets	-0.01	-0.01	0.00	0.29	-0.18	0.29	1.00		
Case duration	0.01	0.02	-0.10	0.00	-0.05	0.00	0.04	1.00	
Number of days affected by COVID	0.00	0.01	-0.09	-0.02	-0.03	-0.04	0.01	-0.01	1.00

Appendix E: 2.E. Linear regression before and after propensity matched scoring on change in redemption frequency

Variables	Before correction β	Before correction t value	After correction β	After correction t value
Type	-0.88	-2.13*	-0.84	-2.17*
Mileage (loyalty points)	-0.91	-2.20*	-0.62	-1.49
Promotion codes	-0.05	-0.33	0.04	0.25
Giftcard, voucher, ecoupon or credit	-0.43	-2.57*	-0.30	-1.77 .
Cash or refund	-0.32	-2.90**	-0.31	-2.61*
Service at airport	-0.14	-0.45	-0.40	-1.28
Service onboard	-0.22	-0.70	-0.50	-1.43
Delays	-0.01	-0.02	-0.23	-0.83
Other experience issues	0.33	0.96	0.10	0.30
Missing bags	-1.40	-3.29**	-1.00	-1.81 .
Damaged bags	-0.83	-2.30*	-0.53	-1.65
Shipped equipment issues	-1.01	-1.35	-1.14	-1.32
Other bag issues	-0.45	-1.38	-1.50	-2.91*
Refund requests	-0.001	-0.002	-0.05	-0.14
Call center and customer support	-0.29	-0.91	-0.21	-0.70
Number of people affected	-0.33	-3.11**	-0.31	-2.99**
Disability-related	0.55	0.59	0.13	0.15
Enumeration of complaints over the same flight	0.11	1.09	0.03	0.26
Count of complaints of the same flight from the same consumer	0.13	1.87 .	0.15	2.03*
Duplicate complaints	-0.36	-1.01	-0.42	-1.14
Top tier	9.03	42.06***	8.32	36.51***
Upper mid-tier	3.73	20.87***	3.47	18.09***
Lower mid-tier	0.88	4.40***	0.75	3.57***
Temperature of the destination location	-0.01	-2.59**	-0.01	-1.77 .
Precipitation rate of the destination location	0.03	0.37	0.05	0.66
Temperature of the departure location	-0.02	-4.25***	-0.02	-3.98***
Precipitation rate of the departure location	0.17	2.14*	-0.01	-0.16
Month of the complaint	-0.41	-18.49***	-0.39	-16.70***
Year of the complaint	1.74	31.71***	1.70	29.66***
Female ratio of the traveling group	-0.54	-3.24***	-0.56	-3.19**
Ratio of unidentified gender of the traveling group	-1.33	-0.36	-0.64	-0.17
Complainer being of unidentified gender	0.14	0.28	0.10	0.19
Daily journey duration in the last three years	-0.09	-27.18***	-0.08	-25.76***
Duration of the journey related to the complaint	0.0002	6.09***	0.00	4.68***
Minimum age among passengers traveling in the group	-0.03	-7.47***	-0.03	-7.11***
Ticket costs	0.0002	6.00***	0.0002	5.25***
Number of tickets	-0.16	-6.51***	-0.15	-5.29***

Variables	Before correction		After correction	
	β	t value	β	t value
Case duration	0.0004	0.28	-0.0001	-0.07
Number of days affected by COVID	0.001	1.36	0.001	1.08
Intercept	-3516.35	-31.66***	-3419.09	-29.61***
Adjusted R squared	0.06799		0.068	

Appendix F: Standardized mean differences after propensity score matching

To measure the balance of the covariates between the vigilant and conciliatory groups before and after propensity score matching, we use the standardized mean difference (SMD), which is the mean difference between the treatment (i.e. the vigilant group) and the control (i.e. the conciliatory groups) divided by the pooled standard deviation Zhang et al., 2019), with the often used threshold of .1 to determine whether there is balance in a covariate. The table below shows the SMD of each covariate before and after correction. Among the covariates before correction, 6 out of 32 variables have SMD > .1, while only one covariate after correction (service at the airport) has SMD >.1. Furthermore, the SMD of service at the airport after correction is .27, lower than that before correction .34. In other words, propensity score matching helps the covariates to achieve decent balance.

2.F. Standardized mean differences of the covariates between vigilant and conciliatory groups before and after correction

Covariate	Standardized mean differerence before correction	Standardized mean difference after correction
Service at airport	**0.34**	**0.269**
Service onboard	**0.101**	0.01
Delays	0.077	0.064
Other experience issues	0.044	0.052
Missing bags	0.091	0.014
Damaged bags	**0.309**	0.049
Shipped equipment issues	0.072	0.056
Other bag issues	**0.138**	0.072
Refund requests	**0.224**	0.093
Call center and customer support	0.027	0.017
Number of people affected	0.063	0.021
Disability-related	0.091	0.077
Enumeration of complaints over the same flight	0.073	0.053

Count of complaints of the same flight from the same consumer	0.077	0.062
Duplicate complaints	0.073	0.046
Top tier	0.034	0.021
Upper mid-tier	0.04	0.007
Lower mid-tier	0.045	0.033
Temperature of the destination location	0.021	0.003
Precipitation rate of the destination location	0.003	0.012
Temperature of the departure location	0.035	0.014
Precipitation rate of the departure location	0.012	0.007
Month of the complaint	0.034	0.035
Year of the complaint	**0.101**	0.025
Female ratio of the traveling group	0.002	0.012
Ratio of unidentified gender of the traveling group	0.009	0.006
Complainer being of unidentified gender (vs other genders)	0.035	0.034
Daily journey duration in the last three years	0.034	0.006
Duration of the journey related to the complaint	0.034	0.016
Minimum age among passengers traveling in groups	0.015	0.015
Ticket costs	0.022	0.002
Number of tickets	0.022	0.033

Conclusion

The two research areas charity advocacy and service failures and recovery are both mature with rich literature behind them. With rapid changes in the marketplace and in society, they both need out-of-the-box thinkings to make further important contributions to the fields. This requires the use of more sophisticated methods, the consideration for complex relationship dynamics and a long-term vision. In both essays, textual analysis through deep learning plays a central role in the development of new constructs. Both pay attention to intricate relationships through time. And both are motivated by social well-being.

Deep learning in the identification of new constructs

While textual analysis is becoming more and more popular in marketing, deep learning for textual analysis is relatively new and mostly found in methodological papers. Meanwhile, mainstream approaches for text face limitations, most notably being unfeasible for big data (e.g. quantitative content analysis) or lower levels of accuracy (e.g. lexicon-based approaches like LIWC). By combining deep learning for textual analysis with advanced quantitative approaches, we uncover diverse aspects of field data and make novel theoretical contributions to two different areas in marketing, while also ensuring the robustness of our findings.

In essay one, we classified charity-related messages, along with mission-related messages and non-mission-related messages, making comparisons of their impacts possible. Charity messages have positive total effects on sales in both short and long terms. So do mission messages. However, non-mission messages have significant direct effects on sales in the short and long terms, though in opposite directions. That is, the short-term direct effect of non-mission signaling on sales is positive, but its long-term direct effect is negative. We explain this surprising result by arguing that the authenticity of non-mission related signals may decrease with repetition, which is the opposite for charity signaling.

Furthermore, by studying charity-advocating artists, we contribute to the understanding of the actors in the non-profit-related network by examining the role of an additional intermediary (i.e., supporting artists). The exploration of a finer-grained conceptualization of the different actors involved in charity donation is important as non-profit research keeps growing as a field.

In the second essay, textual analysis helped us to identify and thus, be able to contrast vigilant complaints with conciliatory complaints at different stages of the service failure and recovery journey. We examine the antecedents leading customers to become vigilantes or conciliators. This allows us to validate the two complaint types and prevent the occurrence of vigilante complaints in the future. Also, by identifying their consequence – purchase value, we also linked two complaint types to an objective measure with real financial impacts:

Different perspectives of time

Furthermore, working with big data allows us to study complex theoretical aspects. Specifically, we are interested in different aspects of time, i.e., the intermittent versus continuous nature of activities (essay 1) and the episodic fluctuations in consumer perceptions and experience (essay 2).

In essay one, we examined short and long-term mediation effects of charity-related messages and other types of messages on sales through social media engagement, using multilevel mediation. By analyzing simultaneously short- and long-term effects in essay one, we provide new insights into reconciling previous conflicting results. Indeed, we find that among musicians, charity signaling has a *positive indirect effect* on sales through social media engagement in the short and long terms. The indirect effect of charity signaling on sales (through engagement) is lower than the impacts of other types of signaling in the short term. Yet, in the long term, it is higher than other types of signaling, including mission signaling, despite the latter's focus on artists' core business.

In essay two, by studying the antecedents and consequences of vigilant and conciliatory complaints, we detect complex dynamics between constructs at different stages of the customer experience. Some variables (e.g. duration of the journey related to the complaint) are positively associated with the likelihood of vigilant complaints but at the later stages with an increase in purchase value. Or vice versa, some are negatively associated with the likelihood of vigilant complaints (e.g. missing bags) but at the later stage with a decrease in purchase value. This indicates the diverse routes of different variables at each stage of service failure and recovery journey.

Social well-being

Also, despite belonging to different contexts, both essays aim to contribute to different aspects of well-being. As an important part of sustainability, well-being issues are becoming more important than ever, in the aftermath of the COVID epidemic and with complex global environmental and social challenges around us. Our research helps brands and firms to achieve their financial goals while also making social good an important part of their strategies.

The first essay, with its focus on charity advocacy among human brands, supports efforts for community well-being. Previous literature examined the role of celebrities for advocacy mainly through their usefulness for the causes. However, to make this relationship long-lasting, it is important to consider their interests. As musicians might hesitate to support social causes, our research encourages them not only to do this openly but also do this on a regular basis. By linking charity advocacy with musicians' long-term financial performance, we propose a win-win situation for musicians and non-profits. In other words, we build a case for transformative prosocial actions to ensure sustainable social improvements.

The second essay, by using resource literature, instead, examines the role of consumer resources, including their mental health resources, in their decisions. Individuals, depending on failure types and existing relationships with the firm, perceive the loss of resources differently. These perceptions could also change at different stages of the customer journey. The essay also brings attention to vulnerable people, who get more affected in case of service failures, due to their possession of fewer resources. This perspective, which is distinct from the familiar justice perspective in service failure literature, allows firms to develop strategies with well-being at the center, and thereby indirectly contributing to a happy community.

Thus, by going beyond the disciplinary boundaries and relying on methodological advances and society-based motivations, both essays have theoretical and managerial contributions for their respective areas with new constructs, dynamics and perspectives.