Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

______________________________ _______________

Ning Zhong Date

The Dynamics of Textual Content on Social Media

By

Ning Zhong
Doctor of Philosophy

Business

David A. Schweidel
Advisor

T. Clifton Green
Committee Member

Jagdish Sheth
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

The Dynamics of Textual Content on Social Media

By

Ning Zhong
M.Sc., Georgia Institute of Technology, 2011

Advisor: David A. Schweidel, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Business
2019

Abstract

The Dynamics of Textual Content on Social Media
By Ning Zhong

While social media have emerged as open sources of insights for both marketing researchers and practitioners, much of the work on the dynamics in social media activity has focused on numeric metrics such as volume and valence. Exiting literature on the user-generated content (UGC) on social media has begun to explore its potential to yield marketing insights, but little has been done to consider how the textual content on social media may shift over time. The goal of this dissertation work is to find out how the text-based UGC on social media evolves over time by extending the topic modeling framework of latent Dirichlet allocation (LDA) in three empirical scenarios. In the first essay, a discrete-state dynamic topic model that incorporates multiple latent changepoints is developed to capture the underlying shifts of textual content that relates to a brand on social media around an event, such as a brand crisis, a new product release, or breaking news. This model may be used by marketers to actively monitor online conversations surrounding their own brands by detecting changes in the topics discussed on social media. In the second essay, a continuous-state dynamic topic model is proposed to examine the evolution of topic prevalence and evaluations in customer reviews on a multi-generational product. The findings show that the concerns of the review contributors at the early stage of product lifecycle are different from those of the review contributors at the later stage.

The Dynamics of Textual Content

By

Ning Zhong
M.Sc., Georgia Institute of Technology, 2011

Advisor: David A. Schweidel, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Business
2019

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my committee chair, Professor David A. Schweidel, who introduced me to the area of text mining and social media analytics. He continually and convincingly conveyed a spirit of adventure with regard to research and scholarship. Without his guidance, encouragement, and persistent help, this dissertation would not have been possible.

In addition, I would like to thank my committee members, Professor Jagdish Sheth and Professor T. Clifton Green, who advised from different angles that enriched my thinking. They provided insightful suggestions in writing the dissertation and finding my first academic job.

Moreover, I would like to thank to Goizueta Foundation and Laney Graduate School for their financial support granted through predoctoral fellowship. I would like to thank my fellow doctoral students for their feedback, cooperation and friendship. I would also like to express my gratitude to the staff that have taken great care of us for years, especially Ms. Allison Gilmore and Ms. Sonya Owens.

Last but not least, nobody has been more important to me in the pursuit of my PhD degree than my family members. Particularly, I would like to thank my parents, Mr. Gankun Zhong and Ms. Qiaoxia Shu, for supporting me spiritually throughout writing this dissertation and my life in general.

TABLE OF CONTENTS

INTRODUCTION 1
ESSAY 1
Capturing Changes in Social Media Content: A Multiple Latent Changepoint Topic Model 6
- Abstract 6
- Introduction 7
- Related Literature 9
- Model Development 13
- Empirical Applications 20
- Discussion 32
- References 36
- Appendix A.1. MCMC Algorithm for LDA-LC 55
- Appendix A.2. Simulation Study 59
- Appendix A.3. Most Relevant Words and Prevalence of Topics 62

ESSAY 2
The Evolution of Online Reviews: A Dynamic Topic Model for Multiple Text Streams 66
- Abstract 66
- Introduction 67
- Related Literature 70
- Data 72
- Model Development 75
- Results 81
- Conclusion 88
- Reference 92
- Appendix A.1. Collapsed Gibbs Sampler for SLDA 106
- Appendix A.2. MCMC Algorithm for SLDA-MS 108
- Appendix A.3. In-sample Model Fit 112
- Appendix A.4. Results and Discussions of the Remaining Categories 113

INTRODUCTION

Unlike traditional media that are characterized by unidirectional dissemination from publishers to receivers, social media allow multidirectional interactions among users that provoke community engagement, information exchange, and opinion evolution. The changing patterns present on social media disclose the degree of public attention, contain rich emotions of the users, and respond instantly to external shocks. The dynamic nature of social media activity itself provides marketers both opportunities and challenges. Specialized research on the dynamics of social media activity is important in marketing for at least the following reasons. From the aspect of customer behaviors, studying the social media activity reveals how the user-generated content (UGC) is built up and how the associated sentiment evolves over time (e.g. Godes and Silva 2012; Moe and Schweidel 2012). From the aspect of business operations, understanding the dynamics present in social media activity is not only crucial to manage brands either on a daily basis across platforms (Schweidel and Moe 2014) or in brand crises (Borah and Tellis 2016), but also essential to maintain customer relationship with interference (Ma et al. 2015).

While numerous studies on social media have concentrated on numeric metrics such as volume and valence, very few of them provide insights into what consumers literally express in online conversations with regard to brand perception and customer satisfaction. In particularly, user-generated content (UGC), e.g. customer reviews, blogs, discussion forums, tweets, is a valuable resource that can be further exploited to investigate how consumers evaluate brands over time. The explosive growth of both social media and social listening platforms has facilitated companies to listen from their consumers in unprecedented ways, among which simply monitoring volume and valence provides an intuitive grasp of how social media activity varies from one day to another. Notably, recent literature began to take advantage of mining tools for unstructured data to learn covered information from social media. Particularly, tracking specific phrases and co-mentions of brands help deepen the companies' understanding of customers' requirements (Lee and Bradlow 2011) and market structure (Netzer et al. 2012). To retrieve richer information about brands from social media, marketing researchers initiated the applications and extensions of topic models such as latent Dirichlet allocation, or LDA (Blei et al. 2003), to manifest latent topics from UGC (e.g. Tirunillai and Tellis 2014; Büschken and Allenby 2016), yet one critical shortcoming of traditional LDA is its assumption of static topic compositions. There is evidently a research gap between the demand for analyzing the dynamics of textual content in social media activity and the lack of appropriate text mining tools that accommodate such demand.

The goal of this dissertation is to shed some lights on discovering how and when the textual content of online opinions surrounding brands or products alters over time. In this dissertation, two generative statistical models are developed to showcase the dynamics of textual content present on social media by extending standard LDA framework (Blei et al. 2003) and its Gibbs sampler (Griffiths and Steyvers 2004). The first model detects when and by how much online conversations shift around an event; the second model examines how the textual content and associated ratings of the user-generated reviews for a multi-generational product evolve over time.

The first essay of this dissertation investigates the dynamics of topics present in social media posts before and after an event. We develop a discrete-state dynamic topic model that can be used to identify shifts in the content of social media posts. We extend standard LDA (Blei et al. 2003) to a modeling framework by incorporating multiple latent changepoints and allowing for the prevalence of topics to vary before and after each changepoint. In doing so, we leverage the information of when a social media post is made by assuming that the content of social media posts contributed at different times may focus on different topics. While there have been recent advances in the marketing literature in topic modeling (e.g., Tirunillai and Tellis 2014; Büschken and Allenby 2016), to the best of our knowledge we are among the first to incorporate the time at which social media content is contributed into the topic model.

We apply our modeling framework to the textual data from social media posts mentioning the Volkswagen brand around the emissions testing scandal in 2015, the Burger King brand around the introduction of the Angriest Whopper in 2016, and the Under Armour brand around the massive data breach in 2018. The proposed model identifies multiple changepoints in each of these three studies, including the date on which the focal event, i.e., the scandal, the new product release, and the data breach, occurs. We show that the model outperforms the dynamic topic model (Blei and Lafferty 2006) in both cross-validation and long-term forecasting. Moreover, given the unique identities of the posters, we are able to investigate which factor, the shift of discussions among the existing posters or the change of poster base, drives the shifts of content.

In addition, as we demonstrate in our empirical application, our modeling approach can also be applied by marketers to actively monitor social media conversations surrounding brands and identify when the content of these conversations shifts by re-estimating the proposed model using rolling time windows. Doing so could provide brand managers with an early indication of potential problems, such as consumer reactions to a new product or adverse reactions to new marketing campaigns. In addition to identifying conversational shifts as brands enter a crisis, it may also enable them to identify when they are emerging from a crisis based on changes in social media conversations.

The second essay of this dissertation examines the dynamics of topic proportions and their corresponding contributions to the overall ratings in user-generated product reviews. We establish a continuous-state dynamic topic model that describes the evolution of topic prevalence and associated evaluations in the online customer reviews over the lifecycle of a multi-generational product. We develop a Gibbs sampler for the supervised latent Dirichlet allocation model (Blei and McAuliffe 2007), or sLDA, that jointly models text and ratings of customer reviews. We then extend the modeling framework to a dynamic supervised LDA with multiple text streams, each of which origins from a generation of the product. The proposed model takes into account of the interdependence of topic prevalence and evaluations across multiple review streams and the serial dependence of topic prevalence and evaluations within each review stream over time. While some marketing literature has documented rating evolutions (e.g. Godes and Silver 2012; Moe and Schweidel 2012) and others incorporate the static effect of ratings to the review text (e.g. Büschken and Allenby 2016), to the best of our knowledge our study is among the first to jointly model the evolution of textual content and numeric ratings over product lifecycle.

We implement our model to a corpus of Amazon reviews on the three generations of the Nest learning thermostat ranging from February 2012 to February 2017. The proposed model assumes that: (1) these three generations of the product share the same pool of latent topics; (2) the prevalence and evaluations of topics for each generation of the product at each time point is dependent on the counterpart at the previous time point via random walks; and (3) the evolutions of each topic's prevalence and evaluation across all generations are correlated. We show that the proposed model has better predictive ability than the static model. And our results demonstrate that the popularity of the review topics for each generation of the product gradually shifts as time elapses. In general, the early reviews concentrated more on a smaller set of major topics, whereas the later reviews were more diversified in topic coverage. Moreover, the early review contributors commented more on the features of the product whereas the later review contributors were inclined to share their user experience. In addition, the evaluations of topics for each generation also exhibits temporal pattern over time. For instance, early review contributors appreciated more on the features and bill saving ability of the product, whereas later review contributors regarded highly of the thermostat's easiness of installation and daily user experience.

The findings of evolving topic prevalence and evaluations reveal the fact that the concerns of customers who made purchase at the early stage of product lifecycle are quite different from those who purchased at the later stage of product lifecycle. Such a phenomenon lasts both within each generation and over all generations. As an analysis tool, the proposed model facilitates managers not only to investigate how consumer preference shifts over time within each generation of product, but also to gauge the maturity of consumer population over generations.

ESSAY 1

Capturing Changes in Social Media Content: A Multiple Latent Changepoint Topic Model

Abstract

While social media has emerged as a popular source of insights for both researchers and practitioners, much of the work on the dynamics in social media has focused on common metrics such as volume and sentiment. In this research, we develop a changepoint model to capture the underlying shifts in social media content. We extend latent Dirichlet allocation (LDA), a topic modeling approach, by incorporating multiple latent changepoints through a Dirichlet process hidden Markov model that allows for the prevalence of topics to differ before and after each changepoint without requiring prior knowledge about the number of changepoints. We demonstrate our modeling framework using social media posts from brand crises (Volkswagen's 2015 emissions testing scandal and Under Armour's 2018 data breach) and a new product launch (Burger King's 2016 launch of the Angriest Whopper). We show that our model identifies shifts in the conversation surrounding each of these events and outperforms both static and other dynamic topic models. We demonstrate how the model may be used by marketers to actively monitor conversations around their brands, including distinguishing between changes in the conversation arising from a shift in the contributor base and underlying changes in the topics discussed by contributors.

Keywords: Social media, Changepoint models, Text analysis, Topic models

Introduction

Firms are increasingly interested in extracting marketing insights from social media data. The CMO Survey reported that more than 40% of surveyed companies are planning to invest in social listening and social media analytics.[1] According to Forrester Research's evaluation of leading social listening platforms, the top uses for social listening platforms included monitoring the brand and brand health, measuring campaign success and better understanding customers (Ngo and Pilecki 2016).

While much research on social media has focused on quantifiable measures (i.e., structured data) such as volume and sentiment, these metrics offer limited insights into how customers perceive the brand. For brands that monitor social media conversations in which the brand and its competitors are mentioned, volume and sentiment provide a convenient summary of how social media activity varies from one day to the next. Such measures have been the focus of research investigating the dynamics in online conversations (e.g., Godes and Silva 2012; Moe and Schweidel 2012; Schweidel and Moe 2014; Xiong and Bharadwaj 2014; Ma et al. 2015; Borah and Tellis 2016; Fossen and Schweidel 2017), with limited work considering dynamics in the textual content of social media posts.

In this research, we contribute to the growing research on social media dynamics by developing a modeling framework to examine how the content of social media contributions may shift over time. As the foundation of our model, we employ latent Dirichlet allocation (LDA; Blei et al. 2003), a popular topic modeling framework that assumes documents are comprised of latent topics and that words are associated with different topics (e.g., Tirunillai and Tellis 2014; Büschken and Allenby 2016; Puranam et al. 2017). We combine the topic modeling framework with a hidden Markov model (HMM; e.g., Chib 1998; Netzer et al. 2008; Schweidel et al. 2012)

[1] http://www.slideshare.net/christinemoorman/the-cmo-survey-highlights-and-insights-feb-2016

that assumes the content of posts occurring before and after each changepoint may differ and does not require prior knowledge about the number of changepoints.

To the best of our knowledge, this research is the first to identify the presence and timing of latent changepoints in topics discussed in social media messages. In doing so, our research complements prior work that examines differences in online conversations before and after specific known events such as the launch of a new public policy (e.g., Puranam et al. 2017), and reveals how conversations evolve afterward. We demonstrate how our modeling approach can provide managers with an indication of when underlying shifts in social media conversations have occurred and investigate whether they stem from the participation of new contributors or changes in the topics discussed by previous contributors. Based on the reactions of previous and new contributors, marketers can then determine the appropriate response.

We demonstrate our modeling framework by applying it to three empirical contexts. We first investigate social media conversations mentioning Volkswagen around the time of its emissions testing scandal (Woodyard 2015). While the timing of the event is known *after the fact*, our model is able to identify the changes in the conversation as they occur. Our analysis not only reveals that topics related to this brand crisis became more prevalent in social media posts mentioning the brand, but also sheds light on those topics that were less frequently mentioned following the changepoint. Having demonstrated the ability of the model to detect changepoints in the conversation, we then apply our model to two events that did not garner the same degree of media attention -- Under Armour's 2018 data breach and Burger King's launch of the Angriest Whopper in 2016. Using Twitter data for these two contexts, we demonstrate how brands can employ our framework to assess if the observed conversation dynamics are due to shifts in the prevalence of topics among prior contributors or changes to the contributor base.

The contribution of this research lies in extending the use of topic models in marketing to capture changes in conversations that may occur over time. We develop a topic model with multiple latent changepoints to identify when and to what extent the shifts in the underlying topics mentioned occur. In doing so, we contribute to the literature that has investigated social media dynamics but primarily focused on metrics such as volume and sentiment (e.g., Godes and Silva 2012; Moe and Schweidel 2012; Borah and Tellis 2016). Our research also contributes to the growing body of research in marketing that seeks to leverage unstructured textual data for the purposes of marketing insights (e.g., Lee and Bradlow 2011; Netzer et al. 2012; Tirunillai and Tellis 2014) by incorporating latent regimes into the data generating process. We illustrate the managerial relevance of the modeling framework by using it to detect conversational shifts and distinguish between two potential causes for these shifts, which may affect how marketers choose to react.

The remainder of this manuscript proceeds as follows. We next review the related literature. We then describe our modeling framework and the metrics of interest to managers. We proceed to discuss the datasets employed in our empirical analysis. We then present our empirical findings and conclude with a discussion of potential areas for future research.

Related Literature

We draw on multiple streams of work that have largely evolved independently of each other. We first discuss related work into the dynamics of consumers' social media activity. We then discuss research in marketing that has delved into user-generated content with a focus on text analytics methods. We then discuss the limited research that has examined the textual dynamics and the marketing literature on which we draw to develop a discrete-state model of social media content.

Social Media Dynamics

While prior research has examined temporal patterns in social media activity, much of this stream has focused on metrics such as volume and sentiment. For example, Godes and Silva (2012) use product-level data to investigate the temporal and sequential evolution of online product ratings. Using individual-level data on online product reviews, Moe and Schweidel (2012) model a user's decision of whether or not to contribute a review, as well as the sentiment of the review. The authors demonstrate dynamics in users' incidence and evaluation decisions arising from heterogeneity across users. Schweidel and Moe (2014) also document the presence dynamics in the sentiment expressed and the venue to which social media posts are contributed.

Research has also viewed product reviews as means by which early purchasers may provide potential buyers with more information than was initially available. Kuksov and Xie (2010) examine the impact that product reviews may have on the firm's pricing decisions. Sun (2012) looks at the impact of a high variance in previously contributed reviews as providing information to consumers when the average ratings is low, as this may indicate that the product appeals to some customers but not all. Moe and Trusov (2011) also examine how previously contributed reviews affect sales. In doing so, they decompose the effects of previous reviews into a direct effect on sales and an indirect effect through their impact on subsequent reviews.

Understanding the dynamics present in social media activity is essential to maintaining the brand, sensing market, and managing customer relationships. Schweidel and Moe (2014) demonstrate that the analysis of social media data can yield a measure of brand health that is a leading indicator of survey-based metrics. Looking at how brand perceptions within an entire industry may shift, Borah and Tellis (2016) investigate the dynamics surrounding social media conversations following product recalls and find evidence of negative spillover effects. To spot market trends, Du and Kamakura (2012) propose a dynamic factor-analytic model that teases out the cross-brands common trend underlying Google trends data. In terms of managing relationships, Ma et al. (2015) make use of an HMM to characterize customers' relationship with the firm and their social media activity that mentions the firm. The authors show how their model can inform the firm's decision of for which customers it will intervene to improve the relationship.

As our discussion of the extant research on social media dynamics reveals, there is considerable interest in identifying social media dynamics so that brands may take appropriate actions. Despite some exceptions (e.g. Liu et al. 2016; Liu et al. 2017), much of the extant research has focused on metrics such as volume, sentiment and variation, which has not sought to examine the dynamics in the content of social media posts.

Text Analysis of User-Generated Content

Text analysis has become more popular in the marketing literature in recent years. Researchers have applied text analytic methods to data arising from social tags (Nam and Kannon 2014; Nam et al. 2017), microblogs (Culotta and Cutler 2016), forums (Netzer et al. 2012), online reviews (Tirunillai and Tellis 2014; Büschken and Allenby 2016), and search data (Ringel and Skiera 2016; Liu and Toubia 2018). To extract meaningful content from text, some work takes advantage of word-level or phrase-level text to cluster and summarize topics by similarity or co-occurrence. For example, Archak et al. (2011) incorporate the text of product reviews into a consumer choice model. Lee and Bradlow (2011) develop an automated approach to use the text of product reviews to conduct marketing research. Their approach begins by identifying specific phrases that are present in an online review and then grouping together keywords that are similar to each other. In a similar fashion, Netzer et al. (2012) look at the frequency with which brands are co-mentioned, a high tendency for brands to be co-mentioned as indicative of the brands being competitive in the minds of consumers.

More recently, marketing researchers have begun to use and extend LDA (Blei et al. 2003), a topic modeling framework. LDA assumes that for each token position within a document, a latent topic is drawn. Conditional on the topic drawn, a word is then drawn from a vocabulary. Tirunillai and Tellis (2014) build upon the LDA framework by simultaneously identifying the dimensions of brands that are discussed in user-generated product reviews and the sentiment associated with the dimension. Büschken and Allenby (2016) extend LDA by developing a sentence-constrained LDA that they show provides a superior model fit compared to the standard LDA model in prediction. To investigate the consumers' content preference, Liu and Toubia (2018) develop a hierarchical dual LDA that relates the topics in search queries to the topics in webpages of top search results. Other researchers have used LDA to investigate how the content of social media messages has changed over time, specifically in regard to known events. For example, Puranam et al. (2017) use LDA to investigate the effect of new regulations on the content of online reviews.

Though temporal changes have been examined using LDA-based models, such approaches are limited in their ability to detect emerging shifts in the content of social media messages and alert managers. While Tirunillai and Tellis (2014) examine the frequency with which brands are associated with specific dimensions and how this frequency shifts over time, they assume a time-invariant process governing the content of reviews. That is, they assume that the parameters governing the prevalence of topics do not change over time. Similarly, Puranam et al. (2017) assume a time-invariant data generating process and conduct a difference-in-difference analysis to estimate the impact of the regulations. Though the authors examine how topic prevalence differs before and after an event, such an approach can only be applied after the fact when the event is known. This limits the ability of managers to react to shifts in the topics being mentioned.

To the best of our knowledge, there has been limited research that incorporates dynamics into the underlying process that governs textual content. Blei and Lafferty (2006) propose a dynamic topic model in which the hyperpriors of the LDA model follow a random walk. Puranam et al. (2017) test this model and find little evidence of topic evolution in their context. In the context of customer relationship management, researchers have discussed the flexibility afforded by capturing state dependence through discrete states rather than through a continuous specification (e.g., Netzer et al. 2008). While allowing for the continuous evolution of the hyperpriors (e.g., Blei and Lafferty 2006) and the use of latent changepoints both allow one to incorporate dynamics into an LDA-based model, we favor the use of latent changepoints for the ease with which it can be implemented using Gibbs sampling (e.g., Griffiths and Steyvers 2004) and the interpretability that it affords. In particular, our approach facilitates comparing the prevalence of topics across all the regimes. As we demonstrate, this provides managers with a convenient means of assessing when and to what extent the conversation has shifted.

Model Development

In this section, we describe the LDA-based modeling framework that we develop to detect shifts in the content of social media comments. We begin by briefly introducing the standard LDA model (e.g., Blei et al. 2003; Griffiths and Steyver 2004) that is at the core of our model. Building upon this foundation, we then embed a multiple latent changepoint model to accommodate the temporal nature by which social media posts are made. Our resulting LDA with Multiple Latent Changepoints (LDA-MLC) model allows for the content of social media posts to manifest from multiple underlying regimes.

Latent Dirichlet Allocation (LDA)

LDA is a generative statistical model in the field of topic modeling that portrays collections of discrete data, such as a text corpus, by a finite mixture of unobserved clusters, such as underlying topics (Blei et al., 2003). It allows each document in a text collection to be described by a mixture of topics, with each word being attributed to a topic with different weight. We refer to literature on natural language processing and define the following terms: a *word* is a set of English characters in a vocabulary generated from a document collection. A *token* is a chopped instance of a sequence of characters in documents, which in our case can be a word or an acronym. A *document*, or social media post in our empirical analysis, is a sequence of tokens which may consist of repeated words. To provide an illustration, consider the following post about the Volkswagen emissions cheating scandal:

> "Volkswagen has said that in some cases, the cars can be fixed by reprogramming the software. But in other cases, Volkswagen may need to install new hardware. Matthias Müller, the Volkswagen chief executive, has not ruled out giving some

customers new vehicles if repairs are not possible. Hang on to your tdi's...this could be the biggest rfd deal ever!!!"

The document is comprised of four sentences. The word "Volkswagen" appears three times in the document: the first token of the first sentence, the fifth token of the second sentence and the fourth token of the third sentence.

For each document *d* in the document collection, LDA assumes the following generative process such that:

1. Choose $\theta^{(d)} \sim \text{Dirichlet}(\boldsymbol{\alpha})$
2. For each token *i* in document *d*

(a) Choose a topic $z_i \sim \text{Multinomial}\left(\theta^{(d)}\right)$

(b) Choose a word $w_i \mid z_i \sim \text{Multinomial}\left(\phi^{(z_i)}\right)$

where $\theta^{(d)}$ is a document-specific probability vector over topics that denotes topic proportions in document *d*. The topic-specific probability vector $\phi^{(z)}$ for topic *z* consists of the probabilities with which words are drawn for a token associated with topic *z*. Griffiths and Steyvers (2004) propose the use of Gibbs sampling for LDA by incorporating a Dirichlet prior on $\phi^{(j)}$ such that $\phi^{(j)} \sim \text{Dirichlet}(\boldsymbol{\beta})$. Rather than explicitly estimating θ and ϕ, they evaluate the posterior distribution over the assignment of words w to topics z. As detailed by Griffiths and Steyvers (2004), since the hyperpriors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are conjugate to the multinomial distribution and are predetermined, the joint distribution $P(\boldsymbol{w}, \boldsymbol{z}) = P(\boldsymbol{w}|\boldsymbol{z})P(\boldsymbol{z})$ can be computed by integrating out θ and ϕ, where

$$P(\boldsymbol{w}|\boldsymbol{z}) = \left(\frac{\Gamma(W\beta)}{\Gamma(\beta)^W}\right)^J \prod_{j=1}^{J} \frac{\prod_w \Gamma\left(n_j^{(w)} + \beta\right)}{\Gamma\left(n_j^{(\cdot)} + W\beta\right)} \tag{1}$$

$$P(\boldsymbol{z}) = \left(\frac{\Gamma(T\alpha)}{\Gamma(\alpha)^J}\right)^D \prod_{d=1}^{D} \frac{\prod_j \Gamma\left(n_j^{(d)} + \alpha\right)}{\Gamma\left(n_\cdot^{(d)} + J\alpha\right)} \tag{2}$$

in which $n_j^{(w)}$ counts the number of times word *w* is assigned to topic *j* in the whole collection, $n_j^{(\cdot)}$ counts the number of tokens assigned to topic *j* in the whole collection, *W* is the size of vocabulary, $n_j^{(d)}$ counts the number of tokens assigned to topic *j* in document *d*, $n_\cdot^{(d)}$ counts the number of all tokens in document *d*, *J* is the number of topics, *D* is the number of documents, and $\Gamma(\cdot)$ is a gamma function. The conditional distribution of token *i* being assigned to topic *j* can be represented as (Griffiths and Steyvers 2004):

$$P(z_i = j | \boldsymbol{z}_{-i}, \boldsymbol{w}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \propto \frac{n_{-i,j}^{(w_i)} + \beta}{n_{-i,j}^{(\cdot)} + W\beta} \left(n_{-i,j}^{(d_i)} + \alpha\right) \tag{3}$$

where d_i and w_i denotes the document and word index of token *i* respectively, $n_{-i,j}^{(w_i)}$ counts the number of times word *w* is assigned to topic *j*, excluding the current token *i*, while $n_{-i,j}^{(\cdot)}$ counts the number of times any token is assigned to topic *j*, excluding the current token *i*; $n_{-i,j}^{(d_i)}$ counts the number of times topic *j* is assigned to all tokens in document *d*, excluding the current token *i*, while $n_{-i,\cdot}^{(d_i)}$ counts the number of all tokens in document *d*, excluding the current token *i*. The estimates of word distribution by topic and topic distribution by document can be obtained by

$$\hat{\phi}_j^{(w)} = \frac{n_j^{(w)} + \beta}{n_j^{(\cdot)} + W\beta} \tag{4}$$

$$\hat{\theta}_j^{(d)} = \frac{n_j^{(d)} + \alpha}{n_\cdot^{(d)} + J\alpha} \tag{5}$$

where $n_j^{(w)}$ counts the number of times word w is assigned to topic j, while $n_j^{(\cdot)}$ counts the number of times any token is assigned to topic j; $n_j^{(d)}$ counts the number of times topic j is assigned to any tokens in document d, while $n_{\cdot}^{(d)}$ counts the number of all tokens in document d.

Latent Dirichlet Allocation with Multiple Latent Changepoints (LDA-MLC)

LDA assumes that the topic proportions of all documents in a text collection are drawn from a single underlying Dirichlet distribution, thereby ignoring the temporal nature of social media activity. Either exogenous shocks, such as brand crises, new product introductions and public policy changes, or endogenous evolution, such as user-generated conversations, may change the frequency with which topics are discussed in social media comments. If such events are known a priori, we may simply divide the documents into two sets: those documents that were contributed before the event and those documents that were contributed afterward (e.g., Puranam et al. 2017). If the event is unknown to us a priori, we cannot divide the data into two regimes. Moreover, the topic proportions may continue to change in the aftermath of such an event, requiring more than "before" and "after" regimes to capture temporal variation in conversations.

To overcome this limitation and detect changepoints as they emerge, we assume that the frequency with which topics are discussed may change over time. We allow for one or more underlying regimes from which the prevalence with which topics are discussed in a document, indicated by the vector of Dirichlet parameter $\boldsymbol{\alpha}$, is drawn, while fixing the word distribution of each topic within the observation time window. We assume that the prevalence with which topics are discussed between each latent changepoint may differ. In our LDA-MLC model, we assume that there are multiple regimes that govern the content of social media posts, and that the regime at time t (denoted s_t) evolves according to an HMM (e.g. Chib 1998; Netzer et al. 2008). Specifically, we assume the following generative process:

1. Choose a regime $k \sim P(s_t|s_{t-1})$, where s_t is the discrete state at time t, t=1,2,…,T. Documents contributed at time t belong to regime k
2. For each document d in regime k, the collection of which is denoted $\mathcal{C}_k$ with $k \in \mathbb{Z}^+, k \leq T$, choose $\theta^{(d)} \sim \text{Dirichlet}\left(\alpha_0 \boldsymbol{m}^{(k)}\right)$
3. For each token i in document d

(a) Choose a topic $z_i \sim \text{Multinomial}\left(\theta^{(d)}\right)$

(b) Choose a word $w_i \mid z_i \sim \text{Multinomial}\left(\phi^{(z_i)}\right)$, where $\phi^{(z)} \sim \text{Dirichlet}(\boldsymbol{\beta})$

To allow for any possible number of changepoints ranging from 0 to $T-1$, we use Dirichlet process hidden Markov model (DP-HMM) that allows for one-step forward transitions through the discrete state space. Specifically, the transition matrix can be represented by

$$P = \begin{bmatrix} p_{11} & p_{12} & 0 & \cdots \\ 0 & p_{22} & p_{23} & \cdots \\ \vdots & \vdots & \ddots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{1}$$

From state k, the only possible paths are to remain in state k or transition to the next state k+1 (i.e. $p_{kk} + p_{k,k+1} = 1$). We assume a transition matrix that only permits forward transitions to characterize the evolution of brand-related social media posts, as has been used to understand the evolution of household life cycles (e.g., Du and Kamakura 2006). The transition probability is given by:

$$p(s_t = k|s_{t-1} = j) = \begin{cases} \dfrac{n_{jj} + \nu}{n_{jj} + \nu + \lambda} & k = j \\ \dfrac{\lambda}{n_{jj} + \nu + \lambda} & s_t \text{ takesa new state} \end{cases} \tag{7}$$

where n_{jj} denotes the counts of self-transitions that have occurred so far in the current iteration for state k, λ is the hyperprior that controls the tendency to remain in the current state, and ν is the hyperprior that controls the tendency to transition to the next state (See Appendix A.1.). Importantly, the DP-HMM does not require an a priori specification of the number of changepoints (Ko et al. 2015). This makes it particularly well-suited for our examination of the topic evolution in social media conversations, as we may infer the number of conversational regimes from the data.

In addition to Dirichlet symmetric hyperprior $\boldsymbol{\beta}$ for the word distribution, we take the Dirichlet concentration parameter α_0 for topic sparsity as predetermined (Griffiths and Steyvers, 2004) but estimate the vector of probability measure $\boldsymbol{m}^{(k)}$ (i.e., topic prevalence) for each regime *k*. Thus, the joint distribution of word $\boldsymbol{w}$ and topic assignment $\boldsymbol{z}$ can be represented by:

$$P(\boldsymbol{w}, \boldsymbol{z}|\alpha_0 \boldsymbol{m}, \boldsymbol{\beta}, \boldsymbol{s}) = P(\boldsymbol{w}|\boldsymbol{z}, \boldsymbol{\beta}) \prod_{t=1}^{T} \prod_{d \in \mathcal{C}_t} P\left(\boldsymbol{z}^{(d)}|\alpha_0 \boldsymbol{m}^{(s_t)}\right) \tag{2}$$

where $\boldsymbol{s}$ denotes the sequence of state s_t at each time *t*, $\boldsymbol{z}^{(d)}$ denotes the topic assignments of tokens in document *d*, $\mathcal{C}_t$ denotes the document collection at time *t*. We apply the following hybrid algorithm to draw topic assignments $\boldsymbol{z}$, probability measure $\boldsymbol{m}$, and state $\boldsymbol{s}$ in turn (See Appendix A.1. for details):

1. $\boldsymbol{z} \mid \boldsymbol{w}, \boldsymbol{m}, \boldsymbol{s}$
2. $\boldsymbol{m} \mid \boldsymbol{z}, \boldsymbol{s}$
3. $\boldsymbol{s} \mid \boldsymbol{m}, \boldsymbol{z}$

The proposed LDA-MLC model nests a number of cases that are worth noting. Our model nests the scenario in which there is a single regime (i.e., $s_1 = s_2 = \cdots = s_T$). The resulting time-invariant model, consistent with that used by Puranam et al. (2017), differs from the standard Gibbs sampling procedure for LDA model (Griffiths and Steyvers 2004) in that we estimate the Dirichlet parameter governing the distribution of topics, $\alpha_0 \boldsymbol{m}$, from the data rather than assuming a diffuse prior. We refer to this as a "modified LDA" model.[2] In addition to nesting a static LDA model, our approach allows topics to flexibly evolve. If new topics were to emerge late in an observation period, this would manifest as a topic having a high prevalence only in the later regimes and a low (or 0) prevalence in earlier regimes. The LDA-MLC model also allows for topics to emerge and fade, with the conversation returning to its previous content. This would manifest as a regime shift (e.g., from regime 1 to regime 2) when the new topic emerges, followed by another regime shift (e.g., from regime 2 to regime 3) when the topic becomes less popular. In regime 3, we would observe topic proportions similar to that of regime 1, allowing us to conclude that the topic proportions have returned to their prior levels. We next describe the social media data we employ in our empirical analysis to demonstrate the predictive ability of the LDA-MLC model and how it may be used to detect conversational changes.

Empirical Applications

Data

To illustrate our LDA-MLC model, we first apply it to social media data pertaining to the Volkswagen automotive brand around the time of the emissions testing scandal, which arose in September 2015. We use Crimson Hexagon, a popular social listening platform, to download the

[2] To ensure that the LDA-MLC model and the modified LDA model are identified, we conducted simulations using both of these models. The results of the simulation study are presented in Appendix A.2. The results of our simulation show that the LDA-MLC model is able to detect abrupt changes in in topic prevalence, but that it is not suited for detecting more subtle shifts in topic prevalence.

text of social media comments on blogs and discussion forums.[3] We construct a query to identify social media posts on blogs and discussion forums that contain the phrases "Volkswagen" or "VW" (case insensitive). Our goal is to demonstrate the LDA-MLC model's ability to identify the shifts in the content of social media messages. We therefore pull comments before and after the news of the emissions testing scandal broke on September 18, 2015, spanning a period from September 4, 2015 to October 1, 2015. We anticipate identifying a changepoint on September 18, 2015. As we will show, other changepoints are detected both before and after this particular date, demonstrating the model's ability to capture both minor and major changepoints.

We prepare the data for analysis by preprocessing the textual data by lowercasing and tokenizing characters and then removing punctuations (e.g. "/", "@", "|", etc.), numbers, stop words (e.g. "the", "a", "and", etc.), and URL links. In addition, we use the WordNet Lemmatizer (Miller 1995) to minimize the redundancy of words by taking their root form only (e.g. "cars" stems from "car", and thus all occasions of "cars" and "car" are treated as the same). As a common practice in natural language processing, we use sparsity=0.99 to cut off infrequently used words in the documents and provide the final set of words that will be analyzed. In doing so, we focus on the words that appear in at least 1% of the documents, consisting of a vocabulary with 1,081 words. This results in a total of 132,165 documents, from which we extract a 12.5% random sample.

Examining the volume of comments prior to Volkswagen's scandal on September 18, 2015, we see that the daily average during our observation period (2,086 posts/day) is considerably lower than that after news of the scandal broke (8,538 posts/day). Interestingly, the

[3] Due to the limits on the number of posts for which the full text can be exported from Twitter via Crimson Hexagon, we use data from blogs and discussion forums for which downloads are limited to 10,000 posts per day. When this limit is reached, which occurs for Volkswagen for the time period of September 21-24, a random sample of 10,000 posts are collected and downloaded.

volume of comments reaches its peak not on the day of the initial news, but a few days later between September 21 (11,335 posts) and September 24 (12,834 posts). Beyond fluctuations in volume, one would expect that the content of social media messages would differ before and after news of Volkswagen's emission test cheating was released. Dividing our observation period into the days before and after the event, we present word clouds for the two periods in Figure 1. Figure 1 (a) is generated from a random sample of 1,000 posts between September 11, 2015, to September 17, 2015, and Figure 1 (b) is generated from a random sample of the same size between September 18, 2015, and September 24, 2015.

[Insert Figure 1 Here]

We see that posts before the scandal focused on Volkswagen models (e.g. "Jetta", "Golf", etc.), car brands (e.g. "Audi", "Ford", etc.), and product features (e.g. "engine", "parts", etc.). In sharp contrast, the posts after the scandal are dominated by the scandal-related words (e.g. "emissions", "diesel", "cheating", etc.). While Figure 1 provides model-free evidence for the changes in the content of social media posts, we can only produce these results knowing on what date to divide the data. We next present the results of the LDA-MLC model that can alert managers as to when such shifts occur.

Model Comparison

We estimate the model on the textual data of mixed social media posts (i.e. forum, blogs, reviews) that mention Volkswagen brand ranging from September 4, 2015 to October 1, 2016, during which the emissions testing scandal was reported. In contrast to studies that have used academic articles (Blei et al. 2003; Griffiths and Steyvers 2004) and product reviews (Tirunillai and Tellis 2014; Büschken and Allenby 2016) as the document corpus, social media conversations are characterized by highly diversified content surrounding brands (e.g., Schweidel and Moe 2014). We optimize the symmetric Dirichlet hyperprior for word distributions, $\boldsymbol{\beta}$, with the data, as suggested by Wallach et al. (2009). The Dirichlet hyperprior $\alpha_0\boldsymbol{m}$ works to smooth the count of tokens and a large value of α_0 would smooth out the topic proportions of short documents (Wallach et al. 2009). Given the short length of social media posts, we set the concentration parameter of Dirichlet hyperprior for topic distributions, $\alpha_0 = 5$. We increase the number of topics until the change of perplexity for the cross-validation sample flattens (Zhao et al. 2015), resulting in our use of 50 topics.[4] Perplexity is a widely used predictive metric in machine learning based on marginal likelihood (Blei 2003; Grün and Hornik 2011).

To assess the performance of the LDA-MLC model, we use 80% of the social media posts of our sample within the one-month time windows which ranges from September 4, 2015 to October 1, 2015 for calibration, the rest 20% of our sample within the same time window for cross-validation, and the posts in the following 5 days from October 2, 2015 to October 6, 2015 for forecasting. We use the estimates of the modified LDA as the guidance of dispersed starting points to run multiple chains and record the estimate that yields the highest perplexity (Blei et al. 2003). We compare our model with (1) Blei and Lafferty's (2006) dynamic topic model (DTM) that assumes the hyperpriors of topic proportions and word distribution follow multivariate normal random walks, and (2) the modified LDA that assumes the hyperpriors of topic proportions are time-invariant (e.g., Puranam et al. 2017).

In Table 1, we summarize the model performance of the LDA-MLC model, the DTM, and the modified LDA model by reporting the log marginal likelihood of the calibration samples

[4] Although we adopt unsupervised learning on Volkswagen-related text, one can always implement a semi-supervised model, such as by seeding specific words into pre-labeled topics if they know which topics are of interest a priori (e.g., Tirunillai and Tellis 2014; Puranam et al. 2017).

and the holdout samples. In conducting our holdout analysis, we focus on the model results obtained using 50 topics[5] and its forecasting for the social media posts from the following day (October 2, 2015), the following 2 days (from October 2, 2015 to October 3, 2015), through the following 5 days (from October 2, 2015 to October 6, 2015).

[Insert Table 1 Here]

Although the DTM outperforms the LDA-MLC model and the modified LDA model with regard to the in-sample fit, the LDA-MLC model dominates both the DTM and the modified LDA in the cross-validation, suggesting that the DTM overfits the calibration data and performs poorer in predicting unseen documents. In addition, while the DTM outperforms the LDA-MLC model in forecasting when considering social media posts from just the following day, which confirms Blei and Lafferty's (2006) model comparison results, both the LDA-MLC model and the modified LDA model outperform DTM in any longer date ranges. We next present the detailed results of the LDA-MLC model applied to the Volkswagen social media data.

LDA-MLC Results

The LDA-MLC model detects eight changepoints, or nine regimes, in the content of social media posts mentioning the Volkswagen brand during the one-month time window with a posterior mean probability in excess of 99.99%. We apply the relevance metric that balances the word distribution of topics with the word frequency in the text collection using weight λ = 0.6 (Sievert and Shirley 2014). To illustrate the manner in which the topics shift, in Table 2 we present the five most prevalent topics in each regime (e.g., the topics corresponding to the five largest elements of $\boldsymbol{m}^{(k)}$ in regime k), along with a selection of the most relevant words associated with

[5] Holdout performance was assessed varying the number of topics from 10 to 100 in increments of 10 topics. For all the number of topics considered, the relative performance of the three models is consistent, with the LDA-MLC model outperforming DTM, which in turn outperforms the modified LDA model.

the topics.[6]

[Insert Table 2 Here]

As shown in Table 2, the two most prevalent topics are "Daily Activities", which contains frequent words (e.g. "good," "great," and "love") in daily conversations throughout the calibration time window. Prior empirical research applying the standard LDA model to social media posts has also identified topics that are comprised of words focused on daily activity and personal musings (e.g., Zhang et al. 2016). This topic accounts for approximately 19% of the content of the social media posts before September 18, 2015 in our data. In addition, the most prevalent topics before September 18, 2015 include vehicle usage ("User Experience," which is marked by words such as "drive," "buy," and "own") and maintenance ("Maintenance," which is marked by words such as "oil," "engine," and "filter").

The prevalence of "Daily Activities" drops to approximately 13% after the fifth changepoint that occurs at the beginning of September 18, 2015 on which the Environmental Protection Agency (EPA) accuses Volkswagen of deliberately deploying "defeat devices" on 482,000 cars sold in the United State to mislead emission tests that would show the vehicles to violate the Clean Air Act standards (Woodyard 2015). This event is captured by topic "EPA Accusation" (marked by words such as "EPA," "air," and "defeat") with a prevalence of 8.4%, and topic "Fine" (characterized by words such as "fine," "recall," "issue," and "pay") with the prevalence of 7.0%. Thereafter, as shown in Table 2, the emissions scandal-related topics become more popular.

To characterize the nature of these changes in social media content, we identify those topics that experience the largest changes in prevalence (in magnitude) after each identified

[6] See Appendix A.4. for the prevalence of all the 50 topics over regimes and the top 20 most prevalent words associated with each topic.

changepoint compared to its prevalence prior to the changepoint. We present the three topics that experience the largest increases and decreases, respectively, in prevalence and a sample of most relevant words of these topics in Table 3.

[Insert Table 3 Here]

Before September 18, 2015, we identify three changepoints in Volkswagen-related social media comments. For instance, the prevalence of the topic “Concept Cars” (characterized by words including “Frankfurt,” “concept,” and “design”) increases by 1.8% after September 14, 2015, while the topic “Classics” (represented by words such as “bug,” “bus,” and “beetle”) decreases by 1.5%. Such changes in prevalence enable our model to identify the dates on which the content of social media conversations shift.

The fourth changepoint identified as September 18, 2015 is characterized by an increase in scandal-related topics “EPA Accusation” (+8.2%), “Fine” (+6.8%), and “EPA Accusation Details” (+4.1%). The content of the conversation changes a few days later after the fifth changepoint on September 21, 2015, with the “Scandal” (+4.7%), “Fuel” (+1.7%), and “Question” (+1.6%) increasing while the prevalence of “EPA Accusation” (-4.5%) subsides. This follows Volkswagen first admitting to cheating and formally apologizing (e.g., Woodyard 2015). The focus on the “Leadership” (+2.3%) topic begins to increase following the sixth changepoint on September 23, 2015 changepoint, when a change of Volkswagen’s management occurred (Woodyard 2015). There is also increased discussion of the “German Economy” topic (+0.9%), consistent with the negative spillover across brands identified by Borah and Tellis (2016). With the seventh identified changepoint on September 27, 2015, “Warnings” (+1.7%) from Bosch about the abuse of their emissions testing software becomes popular while the prevalence of other scandal-related topics like “Leadership” (-2.0%) and “Fine” (-1.9%) begin to diminish. On

September 30, 2015 where the last changepoint is identified, “Daily Activities” (+3.3%) begins to increase in prevalence while the scandal-related topics such as “Warnings” (-1.5%) and “Scandal” (-2.0%) continue to decline.

Calculating Eqn. (10) in Appendix A.1. for each document that emerges on each day, we illustrate how the average proportions of topics related to the emissions testing scandal shift over time in Figure 2 (a) in contrast to the volume and sentiment proportions of comments provided by Crimson Hexagon in Figure 2 (b).

[Insert Figure 2 Here]

The topics related to the EPA’s accusation of Volkswagen’s cheating emerge first following the fourth changepoint on September 18, 2015. After Volkswagen formally admitted its wrongdoing on September 20, 2015, social media posts shift their focus from the EPA’s announcement to Volkswagen’s behavior and begin to call it a “Scandal”. In this context, we see that the breaking news (i.e., “EPA Accusation” and “Leadership”) emerges and fades relatively quickly, while the details and consequences surrounding the brand crisis (i.e., “EPA Accusation Details”, “Fuel”, and “Fine”) grow slowly and linger longer. The discussions of “Emissions” and “Scandal” emerge at the onset of the crisis and continue to exist in prevalence while other scandal-related topics begin to decline. This may suggest a longer-term impact on brand associations in the wake of the brand crisis. In contrast, the volume of comments does not shift drastically until September 21 when the content has already changed. Moreover, while major changes can be clearly detected by the LDA-MLC model, there is no discernible trend in sentiment after the initial report on September 18. Obviously, the LDA-MLC model provides more accurate and profound details than simply monitoring volume and sentiment.

Changes in Topic Prevalence vs. the Contributor Base: Burger King and Under Armour

The Volkswagen empirical context demonstrates our model's ability to identify changes in the topics discussed on social media. The dynamics observed in topic prevalence at the aggregate level may arise from two distinct sources. First, it may be that the topics mentioned by the same contributors shift over time. Second, the changes in topic prevalence observed in aggregate may arise not from changes in contributors' topic prevalences but from changes in the contributor base. That is, it is possible that a regime shift we detect at the aggregate level arises from new contributors who mention different topics than the contributors in the previous regime.

To investigate these alternative explanations for changes in topic prevalence, as well as demonstrate the proposed model's performance to brands that were not attracting the same level of attention as Volkswagen, we collected two datasets from Twitter: (1) 75,778 posts about Burger King from March 15, 2016 to April 11, 2016, during which time a new product (The Angriest Whopper) was released, and (2) 54,045 posts about Under Armour from March 15, 2018 to April 11, 2018, during which time a data breach was disclosed. We use Crimson Hexagon to pull the text of twitter posts that contains case-insensitive key words "Burger King" or "BurgerKing" for the former study and "Under Armour", "Under Armor", "UnderArmour", or "UnderArmor" for the latter study. Using a sparsity of 0.99975 yields a vocabulary of 1,899 unique words for the Burger King study and 2,471 unique words for the Under Armour study. The usernames associated with Twitter posts enable us to identify the content contributed by each user over time, thus allowing us to distinguish between the two sources of observed dynamics. The Burger King investigation applies our modeling framework to a scenario in which the focal event (the release of a new product) is considerably less newsworthy than Volkswagen's emissions testing scandal and Under Armour's data breach.

The optimal numbers of topics for both studies are 60 with $\alpha_0 = 1$ and optimized symmetric $\boldsymbol{\beta}$. In both studies, we use 80% of all data for calibration and the rest for validation. As was the case in the Volkswagen analysis, the LDA-MLC model outperforms both the modified LDA and DTM in cross-validation in these two empirical studies, which we report in Table 4.

[Insert Table 4 Here]

We present the most shifted topics of these two studies at each changepoint in Table 5 and Table 6.

[Insert Table 5 Here]

[Insert Table 6 Here]

Despite the smaller change in volume compared to the Volkswagen analysis, the LDA-MLC model accurately detects the focal events of Burger King's new product launch on March 29, 2016 and Under Armour's data breach on March 29, 2018. It also detects other shifts such as the prank call that makes Burger King's workers to break the store windows on April 9, 2016 and the Under Armour All-America Camp's Orlando stop on March 25, 2018.

To examine the sources of the observed dynamics, we divide contributors into two groups by their posting behavior. One group consists of the contributors who post both before and after the focal event (which we term "enduring contributors"), while the other group consists of contributors who post either before or after the focal event (which we term "transient contributors"). Table 7 provides the distribution of contributors over the two groups.

[Insert Table 7 Here]

The higher ratio of posts-to-authors from those contributors who post both before and after the focal event (3.67 posts/author for Burger King and 12.49 for Under Armour) compared to the ratio among those who only post before or after the focal event (1.14 posts/author for Burger King and 1.22 for Under Armour) suggests that the enduring contributors are more engaged with the brand compared to the transient contributors.

We plot the average topic proportions given by Eqn. (10) in Appendix A.1. for the topics relevant to the focal events for each group of contributors and for all contributors in Figure 3 (a) and (b), respectively, in contrast to the volume and sentiment proportions of comments.

[Insert Figure 3 Here]

Comparing the content of posts from enduring contributors to the transient contributors who post after the launch of The Angriest Whopper, we see that they exhibit similar topic proportions over time, except for "Red Buns". This suggests that both changes in topic prevalence and changes in the contributor base contribute to the observed dynamics. We also note that, in contrast to the Volkswagen analysis, we detect shifts in the content despite volume and sentiment remaining relatively stable throughout the observation period. In the case of Under Armour, however, we find that the enduring contributors are much less likely to discuss the data breach than the transient contributors who only post after the data breach. This suggests that changes in the contributor base are driving the dynamics observed to a large extent.

Detecting Conversational Changepoints

To illustrate the ability of the LDA-MLC model to detect conversational shifts in social media conversations surrounding a brand, we estimate the model using a rolling window of one week throughout the observation period. We run the model with (1) no changepoint by setting the initial discrete state to one regime and (2) at most one changepoint by setting the initial state to two regimes, enabling us to ascertain if the addition of a changepoint is warranted by calculating the Bayes factor. In studying Volkswagen, for example, we estimate the LDA-MLC model on one week of data, shift the window one day forward and re-estimate the model, repeating this process until the end of the rolling window occurs on October 7, 2015. We begin our analysis on August 22, 2015, yielding an additional week of data that we use to determine a reporting threshold that we apply to our observation period from September 4, 2015 to October 7, 2015. Based on our examination of the perplexity change for one week of data from the beginning of our observation period, the optimal number of topics is 20 topics with $\alpha_0 = 5$ in the study of Volkswagen and 40 topics with $\alpha_0 = 1$ in the studies of Burger King and Under Armour, all using optimized symmetric $\boldsymbol{\beta}$.

Each day of our observation horizon appears in seven rolling windows. We report in Figure 4 the number of rolling windows in which the start of a given date has the highest probability within the rolling window of being a changepoint. A changepoint is identified when a certain difference of the holdout sample LMD between one changepoint model and no changepoint model is achieved. Referring to their first weeks' results, we set the threshold of difference to 10 for Volkswagen, 50 for Burger King, and 50 for Under Armour, respectively. This is akin to setting the control limit in CuSum control charts (e.g., Granjon 2014). As the changepoint is identified between days (e.g., between day 1 and 2, between day 2 and 3, etc.), this ranges from 0 to 6 counts for a given date. We also identify the dates that are recognized as the first post-changepoint day the first time they enter the rolling window. In the example of Volkswagen, for the week spanning September 12 – September 18, 2015, if the changepoint is identified as occurring at the end of September 17, then September 18 is the first post-changepoint day and is identified as the first time that it is included in a rolling window.

[Insert Figure 4 Here]

The solid bars indicate that the date is identified as the first post-changepoint day the first time it enters the rolling window. For such detection to occur, there must be a significant shift in the content of the social media posts on this day compared to the previous 6 days of the rolling window. For example, as shown in Figure 4, both the changepoint on September 18, 2015 in the Volkswagen study and the changepoint on March 29, 2018 in the Under Armour study are detected on the first day that it is included in a rolling window. The release date of the Angriest Whopper in Burger King study is March 29, 2016. For the rolling window ending on March 29, 2016, the changepoint detected occurs on March 27, 2016, which is the first time that we observe an increase in average proportion of a topic related to the Angriest Whopper as shown in Figure 3 (a) due to pre-release trials. While the LDA-MLC model is able to detect the changes in the social media conversations related to Volkswagen and Under Armour, which exhibit a spike in volume corresponding to the focal events, the analysis of Burger King suggests that a sufficient volume of conversation is necessary to detect the conversational shifts in a timely fashion.

Discussion

In this research, we develop a topic model that discretely partitions social media posts based on the posting time to identify the shifts in the content of social media posts. We build upon the popular LDA modeling framework by incorporating a Dirichlet process hidden Markov model to allow for multiple latent changepoints, with topic prevalence varying before and after these changepoints. Our topic model with multiple latent changepoints allows us to capture a number of temporal patterns in conversation topics, including topic proportions remaining constant over time, temporary changes in topic prevalence that return to their prior levels, and topic proportions that continue to shift over time. While there have been recent advances in the marketing literature in topic modeling (e.g., Tirunillai and Tellis 2014; Büschken and Allenby 2016; Liu and Toubia 2018), to the best of our knowledge we are among the first to incorporate the time at which social media content is contributed into a topic modeling framework.

As we illustrate, our modeling approach can be used by marketers to actively monitor social media conversations surrounding brands and identify when the content of these conversations shifts. By considering the Bayes factor to compare the analysis with no changepoint to that with one changepoint during a given timeframe, managers can distinguish minor and major shifts in social media topics. This can provide them with an early indication of potential problems, such as consumer reactions to a new product or adverse reactions to new marketing campaigns. In addition to identifying conversational shifts as brands enter a crisis, it may also enable them to identify when they are emerging from a crisis based on changes in social media conversations. In doing so, marketers can make use of social media data to assess how perceptions of their brands are changing over time.

Our research also demonstrates the importance of looking beyond common metrics in monitoring social media conversations. While the volume and sentiment of comments is easily tracked and reported in social media listening tools, we show in our analysis of Burger King that changes in volume and sentiment need not occur for changes in the conversations to occur. Moreover, the shift in topics relating to Volkswagen's emission testing scandal occurs several days before the volume of social media posts peak. By jointly monitoring shifts in structured data such as volume, sentiment, and the unstructured content of social media posts, marketers can gain insight into what is driving a sudden influx of brand-related social media activity, whether it is a product launch (e.g., Dellarocas et al. 2007), advertising campaign (e.g., Fossen and Schweidel 2019) or customer service failure (e.g., Ma et al. 2015), which can inform how they choose to respond.

Related to this, our research also informs managers as to which consumers are contributing posts that focus on different topics. We show that changes to the contributor base as well as shifts in the topics discussed by earlier contributors both contribute to the observed shifts in social media conversations. This decomposition of the aggregate comments into contributor groups can be crucial for marketers to understand. Contributions of brand-related social media posts may serve as a proxy for consumers' engagement with the brand (e.g., Malthouse et al. 2013). Differentiating between comments from consumers who may already be loyal to the brand and those who have only recently joined the conversation and may be less loyal to the brand will enable brands to assess how perceptions may differ between these groups and take appropriate actions. As an illustration, while Nike's advertising featuring Colin Kaepernick drew calls for a boycott, online sales and the company's stock price ultimately surged.[7]

Our analysis highlights the importance of identifying not only how is the content of the conversation changing but also what is driving it, as this will inform marketers' next steps. There are a number of ways in which the LDA-MLC modeling framework could be extended. From a methodological perspective, conventional LDA-based models ignore the ordering of words and semantic relations under the assumption of "bag-of-words". One way to relax such an assumption is to extend our model to an n-gram one. An alternative is to develop a model in which the words are presented in vectors (e.g. Mikolov et al. 2013). It is also worthwhile to explore approaches that may reduce the computational burden. While we rely on MCMC to estimate the model, researchers using variational inference may be able to extend our work and apply it to larger datasets without the need for sampling (Blei et al. 2017). From a substantive perspective, if data were available for the same user across multiple social media platforms, our

[7] https://www.cnbc.com/2018/09/14/nikes-kaepernick-ad-should-fuel-sales-as-retailer-knows-its-consumer.html

modeling approach could be extended to examine how users may manage their online reputation based on the content contributed to particular platforms and the audience they intend to reach. While we incorporate dynamics in conversational content into the LDA-MLC model via latent changepoints, other factors may also contribute to conversation dynamics. For example, marketing efforts undertaken by a firm or its competitors may change the content that is discussed. Taking account of the interactions between marketing mix and online conversations would further tease apart the sources of the conversation dynamics. Given the interest in understanding the dynamics associated with structured user-generated content such as volume and sentiment, we hope that our work encourages future efforts to understand the dynamics of content and the potential applications of such methods for marketing insights.

References

Archak, Nikolay, Anindya Ghose and Panagiotis G. Ipeirotis (2011), "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science*, 57(8), 1485-1509.

Blei, David M., Andrew Y. Ng and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, Vol. 3, 993-1022.

Blei, David M. and John D. Lafferty (2006), "Dynamic Topic Models," *Proceedings of the 23rd International Conference on Machine Learning*, Pittsburgh, PA.

Blei, David M., Alp Kucukelbir and Jon D. McAuliffe (2017), "Variational Inference: A Review for Statisticians," *Journal of the American Statistical Association*, 518 (112), 859-877.

Borah, Abhishek and Gerard J. Tellis (2016), "Halo (Spillover) Effects in Social Media: Do Product Recalls of One Brand Hurt or Help Rival Brands?" *Journal of Marketing Research*, 53 (2), 143-160.

Büschken, Joachim and Greg M. Allenby (2016), "Sentence-Based Text Analysis for Customer Reviews," *Marketing Science*, forthcoming.

Chib, Siddhartha (1998), "Estimation and Comparison of Multiple Change-point Models," *Journal of econometrics*, 86 (2), 221-241.

Connor, Robert J. and James E. Mosimann (1969), "Concepts of Independence for Proportions with a Generalization of the Dirichlet Distribution," *Journal of the American Statistical Association*, 64(325), 194-206.

Culotta, Aron and Jennifer Cutler (2016), "Mining Brand Perceptions from Twitter Social Networks," *Marketing science*, 35 (3), 343-362.

Dellarocas, Chrysanthos, Zhang, Xiaoquan.Michael and Neveen F. Awad (2007), "Exploring the value of online product reviews in forecasting sales: The case of motion pictures," *Journal of Interactive Marketing*, 21(4), 23-45.

Du, Rex Y. and Wagner A. Kamakura (2006), "Household Life Cycles and Lifestyles in the United States," *Journal of Marketing Research*, 43 (1), 121-132.

Du, Rex Y. and Wagner A. Kamakura (2012), "Quantitative Trendspotting," *Journal of Marketing Research*, 49 (4), 514-536.

Fossen, Beth L. and David A. Schweidel (2017), "Television Advertising and Online Word-of-Mouth: An Empirical Investigation of Social TV Activity," *Marketing Science*, 36 (1), 105-123.

Fossen, Beth L. and David A. Schweidel (2019), "Social TV, Advertising and Sales: Are Social Shows Good for Advertisers?" *Marketing Science*, forthcoming.

Godes, David and Jose C. Silva (2012), "Sequential and Temporal Dynamics of Online Opinion," *Marketing Science*, 31 (3), 448-473.

Granjon, Pierre (2013), "The CuSum Algorithm - A Small Review," *working paper*.

Griffiths, Thomas L. and Mark Steyvers (2004), "Finding Scientific Topics," *Proceedings of the National Academy of Sciences*, Vol. 101, Suppl 1, 5228-5235.

Grün Bettina and Kurt Hornik (2011), "Topicmodels: An R Package for Fitting Topic Models," *Journal of Statistical Software*, 40 (13), 1-30.

Heinrich, Gregor (2009), "Parameter estimation for text analysis," Technical Report, Fraunhofer IGD.

Liu, Jia and Olivier Toubia (2018), "A Semantic Approach for Estimating Consumer Content

Preferences from Online Search Queries," *Marketing Science*, 37 (6), 885-1052.
Ko, Stanley I. M., Terence T. L. Chong, and Pulak Ghosh (2015), "Dirichlet process hidden Markov multiple change-point model," *Bayesian Analysis*, 10 (2), 275-296.
Kuksov, Dmitri and Ying Xie (2010), "Pricing, Frills, and Customer Ratings," *Marketing Science*, 29 (5), 925-943.
Lee, Thomas Y. and Eric T. Bradlow (2011), "Automated Marketing Research Using Online Customer Reviews," *Journal of Marketing Research*, 48 (October), 881-894.
Liu, Liu, Daria Dzyabura and Natalie Mizik (2017), "Visual Listening In: Extracting Brand Image Portrayed on Social Media," *Working Paper*.
Liu, Xiao, Param V. Singh and Kannan Srinivasan (2016), "A Structured Analysis of Unstructured Big Data Leveraging Cloud Computing," *Marketing Science*, 35 (3), 363-388.
Ma, Liye, Baohong Sun and Sunder Kekre (2015), "The Squeaky Wheel Gets the Grease – An Empirical Analysis of Customer Voice and Firm Intervention on Twitter," *Marketing Science*, 34 (5), 627-645.
Malthouse, Edward C., Michael Haenlein, Bernd Skiera, Egbert Wege and Michael Zhang (2013), "Managing Customer Relationships in the Social Media Era: Introducing the Social CRM House," *Journal of Interactive Marketing*, 27 (4), 270-280.
Mikolov, Tomas, Kai Chen, Greg Corrado and Jeffrey Dean (2013), "Efficient Estimation of Word Representations in Vector Space," arXiv preprint arXiv:1301.3781.
Miller, George A. (1995), "WordNet: A Lexical Database for English," *Communications of the ACM*, 38 (11), 39-41.
Minka, Thomas P. (2000), "Estimating a Dirichlet distribution," Technical report, MIT.
Moe, Wendy W. and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*, 48 (3), 444-456.
Moe, Wendy W. and David A. Schweidel (2012), "Online Product Opinions: Incidence, Evaluation, and Evolution," *Marketing Science*, 31(3), 372-386.
Nam, Hyoryung, and P.K. Kannan (2014), "Informational Value of Social Tagging Networks," *Journal of Marketing*, 78 (4), 21–40.
Nam, Hyoryung, Yogesh V. Joshi and P. K. Kannan (2017), "Harvesting Brand Information from Social Tags," *Journal of Marketing*, 81 (4), 88-108.
Netzer, Oded, James M. Lattin and V. Srinivasan (2008), "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science*, 27 (2), 185-204.
Netzer, Oded, Ronen Feldman, Jacob Goldenberg and Moshe Fresko (2012), "Mine Your Own Business: Market-Structure Surveillance through Text Mining," *Marketing Science*, 31(3), 521-543.
Ngo, Samantha and Mary Pilecki (2016), "The Forrester Wave: Enterprise Social Listening Platforms, Q1 2016," *Forrester*, March 2, accessed at https://www.forrester.com/report/The+Forrester+Wave+Enterprise+Social+Listening+Platforms+Q1+2016/-/E-RES122523.
Puranam, Dinesh, Vishal Narayan, and Vrinda Kadiyali (2017), "The Effect of Calorie Posting Regulation on Consumer Opinions: A Flexible Latent Dirichlet Allocation Model with Informative Priors," *Marketing Science*, 36 (5), 726-746.
Ringel, Daniel M., and Bernd Skiera (2016), "Visualizing Asymmetric Competition among More than 1,000 Products Using Big Search Data," *Marketing Science*, 35 (3), 511–34.
Schweidel, David A., Eric T. Bradlow and Peter S. Fader (2011), "Portfolio Dynamics for Customers of a Multiservice Provider," *Management Science*, 57 (3), 471-486.

Schweidel, David A. and Wendy W. Moe (2014), “Listening in on Social Media: A Joint Model of Sentiment and Venue Format Choice,” *Journal of Marketing Research*, 51 (August), 387-402.

Sievert, Carson and Kenneth E. Shirley (2014), “LDAvis: A Method for Visualizing and Interpreting Topics,” *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, June, 63-70.

Sun, Monic (2012), “How Does the Variance of Product Ratings Matter?” *Management Science*, 58 (4), 696-707.

Tirunillai, Seshadri and Gerard J. Tellis (2014), “Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation,” *Journal of Marketing Research*, 51 (4), 463-479.

Wallach, Hanna M., David M. Mimno and Andrew McCallum (2009), "Rethinking LDA: Why Priors Matter," *Advances in Neural Information Processing Systems*, 1973-1981.

Wong, Tzu-Tsung (1998), “Generalized Dirichlet Distribution in Bayesian Analysis,” *Applied Mathematics and Computation*, 97(2-3), 165–181.

Woodyard, Chris (2015), “Chronology: How VW’s Emissions Scandal Has Mushroomed,” *USA Today*, November 20, access at http://www.usatoday.com/story/money/cars/2015/11/20/vw-volkswagen-chronology-emissions/76122812/.

Xiong, Guiyang and Sundar Bharadwaj (2014), “Prerelease Buzz Evolution Patterns and New Product Performance,” *Marketing Science*, 33(3), 401-421.

Zhang, Yuchi, Wendy W. Moe and David A. Schweidel (2017), “Modeling the Role of Message Content and Influencers in Social Media Rebroadcasting,” *International Journal of Research in Marketing*, 34 (1), 100-119.

Zhao, Weizhong, James J. Chen, Roger Perkins, Zhichao Liu, Weigong Ge, Yijun Ding and Wen Zou (2015), "A Heuristic Approach to Determine an Appropriate Number of Topics in Topic Modeling," *BMC Bioinformatics*, 16(13), S8.

Table 1. Model Comparison by LMD for Volkswagen Study

Study	Type	Date Range	Modified LDA	LDA-MLC	DTM
Volkswagen	In-sample	9/4/2015 – 10/1/2015	-2977159.87	-2978832.72	-2914559.00
	Cross-validation	9/4/2015 – 10/1/2015	-330833.25	-330418.26	-343997.00
	Forecasting	10/2/2015	-201053.38	-201068.93	-200155.61
		10/2/2015 – 10/3/2015	-355720.34	-355706.26	-387920.17
		10/2/2015 – 10/4/2015	-481066.02	-481128.18	-538126.62
		10/2/2015 – 10/5/2015	-687807.52	-687944.56	-784961.83
		10/2/2015 – 10/6/2015	-908100.33	-908307.77	-1047813.21

Table 2. Most Relevant Topics in Each Regime of Volkswagen Study

Regime	Date Range	Most Prevalent Topics	Topic Label	Sample of Most Relevant Words	Prevalence
1	9/4/2015 - 9/6/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	18.8%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	6.8%
		50	"Maintenance"	oil, engine, pump, filter, belt, timing, water, valve, fuel, pressure	5.7%
		27	"Brakes & Tires"	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt	5.5%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	5.5%
2	9/7/2015 - 9/8/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	20.1%
		27	"Brakes & Tires"	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt	6.8%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	6.7%
		20	“Classics"	bug, bus, van, beetle, original, paint, color, type, love, black	5.1%
		33	"Family"	city, family, trip, kid, town, summer, day, road, park, travel	4.8%
3	9/9/2015 - 9/13/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	18.7%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	7.2%
		27	"Brakes & Tires"	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt	5.8%
		50	"Maintenance"	oil, engine, pump, filter, belt, timing, water, valve, fuel, pressure	5.4%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	4.9%
4	9/14/2015 - 9/17/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	17.9%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	5.6%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	5.4%
		27	"Brakes & Tires"	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt	5.3%
		47	"Package"	rear, seat, wheel, steering, sport, speed, trim, design, interior, suspension	4.2%
5	9/18/2015 - 9/20/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	15.3%
		38	"EPA Accusation"	epa, air, clean, defeat, agency, software, device, recall, protection, california	8.4%
		42	"Fine"	fine, recall, epa, issue, wonder, pay, government, huge, owner, penalty	7.0%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	5.1%
		22	"Question"	people, write, wrong, know, public, point, truth, happen, bad, trust	4.8%

Table 2. Most Relevant Topics in Each Regime of Volkswagen Study (Cont.)

Regime	Date Range	Most Prevalent Topics	Topic Label	Sample of Most Relevant Words	Prevalence
6	9/21/2015 - 9/22/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	12.8%
		42	"Fine"	fine, recall, epa, issue, wonder, pay, government, huge, owner, penalty	7.7%
		22	"Question"	people, write, wrong, know, public, point, truth, happen, bad, trust	6.5%
		46	"Scandal"	scandal, company, billion, german, carmaker, winterkorn, diesel, worldwide, software, euro	5.9%
		44	"EPA Accusation Details"	test, testing, emission, software, pass, cheat, mode, epa, real, road	5.0%
7	9/23/2015 - 9/26/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	12.5%
		22	"Question"	people, write, wrong, know, public, point, truth, happen, bad, trust	7.0%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	5.4%
		44	"EPA Accusation Details"	test, testing, emission, software, pass, cheat, mode, epa, real, device	4.8%
		42	"Fine"	fine, recall, epa, issue, wonder, pay, government, huge, owner, penalty	4.7%
8	9/27/2015 - 9/29/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	12.7%
		22	"Question"	people, write, wrong, know, public, point, truth, happen, bad, trust	6.7%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	5.8%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	4.7%
		46	"Scandal"	scandal, company, billion, german, carmaker, winterkorn, diesel, worldwide, software, euro	4.1%
9	9/30/2015 - 10/1/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	16.0%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	6.4%
		22	"Question"	people, write, wrong, know, public, point, truth, happen, bad, trust	5.6%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	5.0%
		44	"EPA Accusation Details"	test, testing, emission, software, pass, cheat, mode, epa, real, device	3.3%

Table 3. Topics with Largest Increases and Decreases in Prevalence of Volkswagen Study

Changepoint	**Date**	**Most Shifted Topics**	**Topic Label**	**Sample of Most Relevant Words**	**Prevalence Change**
1	9/7/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	1.3%
		27	"Brake & Tires"	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt	1.3%
		45	"Incidents"	man, police, tell, white, left, road, day, wife, black, morning	1.1%
		31	"Brands"	audi, porsche, skoda, seat, group, luxury, bmw, brand, vag, bentley	-1.0%
		19	"Durability"	long, term, level, number, time, current, fact, high, potential, risk	-1.2%
		50	"Maintenance"	oil, engine, pump, filter, belt, timing, water, valve, fuel, pressure	-1.4%
2	9/9/2015	50	"Maintenance"	oil, engine, pump, filter, belt, timing, water, valve, fuel, pressure	1.1%
		25	"Model"	tdi, jetta, passat, golf, wagon, beetle, trade, sell, dealer, wife	0.8%
		31	"Brands"	audi, porsche, skoda, seat, group, luxury, bmw, brand, vag, bentley	0.7%
		27	"Brakes & Tires"	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt	-1.0%
		32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	-1.4%
		33	"Family"	city, family, trip, kid, town, summer, day, road, park, travel	-1.4%
3	9/14/2015	43	"Concept Cars"	frankfurt, concept, tiguan, suv, motor, bentley, design, version, renault, generation	1.8%
		48	"Competition"	sales, india, market, fiat, plant, automotive, brand, auto, year, chrysler	1.0%
		3	"Electric Cars"	electric, tesla, model, apple, car, battery, range, technology, project, charge	1.0%
		50	"Maintenance"	oil, engine, pump, filter, belt, timing, water, valve, fuel, pressure	-1.4%
		20	"Classics"	bug, bus, van, beetle, original, paint, color, type, love, black	-1.5%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	-1.8%
4	9/18/2015	38	"EPA Accusation"	epa, air, clean, defeat, agency, software, device, recall, protection, california	8.2%
		42	"Fine"	fine, recall, epa, issue, wonder, pay, government, huge, owner, penalty	6.8%
		44	"EPA Accusation Details"	test, testing, emission, software, pass, cheat, mode, epa, real, device	4.1%
		47	"Package"	rear, seat, wheel, steering, sport, speed, trim, design, interior, suspension	-2.7%
		43	"Concept Cars"	frankfurt, concept, tiguan, suv, motor, bentley, design, version, renault, generation	-3.0%
		27	"Brakes & Tires"	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt	-3.2%

Table 3. Topics with Largest Increases and Decreases in Prevalence of Volkswagen Study (Cont.)

Changepoint	Date	Most Shifted Topics	Topic Label	Sample of Most Relevant Words	Prevalence Change
5	9/21/2015	46	"Scandal"	scandal, company, billion, german, carmaker, winterkorn, diesel, worldwide, software, euro	4.7%
		15	"Fuel"	diesel, gasoline, clean, petrol, engine, fuel, powered, technology, dirty, emission	1.7%
		22	"Question"	people, write, wrong, know, public, point, truth, happen, bad, trust	1.6%
		16	"Forums"	post, https, forum, thread, info, click, view, site, watch, link	-1.2%
		32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	-2.5%
		38	"EPA Accusation"	epa, air, clean, defeat, agency, software, device, recall, protection, california	-4.5%
6	9/23/2015	4	"Leadership"	ceo, board, winterkorn, executive, matthias, company, chief, chairman, porsche, resign	2.3%
		17	"German Economy"	german, germany, europe, industry, country, european, crisis, auto, euro, scandal	0.9%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	0.8%
		38	"EPA Accusation"	epa, air, clean, defeat, agency, software, device, recall, protection, california	-1.2%
		46	"Scandal"	scandal, company, billion, german, carmaker, winterkorn, diesel, worldwide, software, euro	-1.5%
		42	"Fine"	fine, recall, epa, issue, wonder, pay, government, huge, owner, penalty	-2.9%
7	9/27/2015	30	"Warnings"	bosch, software, illegal, vag, group, production, engineer, components, arm, reported	1.7%
		26	"User Experience"	car, drive, buy, own, reliable, year, owner, months, good, price	0.8%
		33	"Family"	city, family, trip, kid, town, summer, day, road, park, travel	0.7%
		38	"EPA Accusation"	epa, air, clean, defeat, agency, software, device, recall, protection, california	-1.2%
		42	"Fine"	fine, recall, epa, issue, wonder, pay, government, huge, owner, penalty	-1.9%
		4	"Leadership"	ceo, board, winterkorn, executive, matthias, company, chief, chairman, porsche, resign	-2.0%
8	9/30/2015	32	"Daily Activities"	good, lot, time, bit, pretty, better, work, great, love, stuff	3.3%
		16	"Forums"	post, https, forum, thread, info, click, view, site, watch, link	1.0%
		18	"Service"	service, dealer, warranty, dealership, customer, car, experience, excellent, extended, manager	0.9%
		22	"Question"	people, write, wrong, know, public, point, truth, happen, bad, trust	-1.1%
		30	"Warnings"	bosch, software, illegal, vag, group, production, engineer, components, arm, reported	-1.5%
		46	"Scandal"	scandal, company, billion, german, carmaker, winterkorn, diesel, worldwide, software, euro	-2.0%

Table 4 Model Comparison by LMD for Burger King and Under Armour Studies

Study	Type	Date Range	Modified LDA	LDA-MLC	DTM
Burger King	In-sample	3/15/2016 – 4/11/2016	-1383374.54	-1395962.95	-1218600.46
	Cross-validation	3/15/2016 – 4/11/2016	-169547.47	-167733.95	-193556.94
	Forecasting	4/12/2016	-53354.20	-53523.30	-49037.72
		4/12/2016 – 4/13/2016	-130591.39	-132285.54	-161933.53
		4/12/2016 – 4/14/2016	-190524.48	-193229.29	-248243.47
		4/12/2016 – 4/15/2016	-242864.67	-246796.74	-324928.10
		4/12/2016 – 4/16/2016	-275617.18	-280375.63	-375376.14
Under Armour	In-sample	3/15/2018 – 4/11/2018	-1587204.26	-1582654.20	-1503786.58
	Cross-validation	3/15/2018 – 4/11/2018	-190057.98	-188115.45	-209508.81
	Forecasting	4/12/2018	-77303.23	-76784.72	-75493.82
		4/12/2018 – 4/13/2018	-133305.33	-132411.31	-164648.93
		4/12/2018 – 4/14/2018	-188782.11	-187638.26	-252769.92
		4/12/2018 – 4/15/2018	-260952.30	-260064.67	-348013.12
		4/12/2018 – 4/16/2018	-315654.70	-315229.05	-416875.46

Table 5. Topics with Largest Increases and Decreases in Prevalence of Burger King Study

Changepoint	Date	Most Shifted Topics	Topic Label	Sample of Most Relevant Words	Prevalence Change
1	3/16/2016	6	"Family Tie"	jr, somebody, yell, grandkids, ai, order, papaw, pawpaw, whopper, king	7.5%
		28	"American Values"	wendys, finna, checker, game, semester, deal, mcdonalds, racism, american, en	3.8%
		17	"Justin Bieber"	like, video, act, shit, try, marry, justin, date, look, girl	3.7%
		11	"Daily"	go, work, time, like, drive, way, lunch, say, last, make	-3.0%
		56	"Joke"	burbers, subject, woud, king, cashier, trust, optin, message, best, burger	-7.4%
		37	"Fondness"	mcm, bite, ask, eat, dog, hot, peep, day, inner, savage	-11.4%
2	3/20/2016	52	"Hotdog Compliment"	pretty, good, tell, hotdog, try, deal, tale, guy, live, secret	10.4%
		34	"Beer"	beer, select, location, move, rap, official, app, science, seungkwan, lover	2.6%
		47	"Sarcasm"	unfollow, right, order, dog, hot, weber, mike, delete, walking, ahead	1.4%
		17	"Justin Bieber"	like, video, act, shit, try, marry, justin, date, look, girl	-3.5%
		37	"Fondness"	mcm, bite, ask, eat, dog, hot, peep, day, inner, savage	-4.2%
		6	"Family Tie"	jr, somebody, yell, grandkids, ai, order, papaw, pawpaw, whopper, king	-5.9%
3	3/22/2016	47	"Sarcasm"	unfollow, right, order, dog, hot, weber, mike, delete, walking, ahead	8.8%
		3	"Ad Resurface"	brussels, ad, attack, rally, response, symbol, defiance, terrorist, unlikely, advert	6.6%
		48	"Robery"	comment, manager, public, officer, boycottburgerking, stand, fall, let, mock, cop	4.0%
		60	"Eating"	eat, dog, block, hot, hotdog, trust, grill, mark, people, buy	-3.4%
		32	"Gilled Hotdog"	dog, hot, commercial, sell, grilled, grill, hotdog, chili, try, disgust	-4.6%
		52	"Hotdog Compliment"	pretty, good, tell, hotdog, try, deal, tale, guy, live, secret	-9.7%
4	3/24/2016	42	"Rolling Blunt"	roll, blunt, stream, shitwire, arrogance, treatise, longform, platform, content, pro	16.4%
		60	"Eating"	eat, dog, block, hot, hotdog, trust, grill, mark, people, buy	2.9%
		32	"Gilled Hotdog"	dog, hot, commercial, sell, grilled, grill, hotdog, chili, try, disgust	2.8%
		43	"Buyout"	burgerkingmx, technology, acqui-hired, brewster, fullcontact, steal, owner, rbi, team, read	-5.1%
		3	"Ad Resurface"	brussels, ad, attack, rally, response, symbol, defiance, terrorist, unlikely, advert	-6.2%
		47	"Sarcasm"	unfollow, right, order, dog, hot, weber, mike, delete, walking, ahead	-9.5%
5	3/26/2016	11	"Daily"	go, work, time, like, drive, way, lunch, say, last, make	6.8%
		5	"Crown"	crown, work, cow, teacher, shoutout, prom, bih, end, pin, school	2.7%
		8	"Competitors"	food, fast, mcdonalds, wendys, tacobell, kfc, popeyes, arbys, subway, box	1.9%
		26	"Card Trading"	trading, card, bid, pokemon, gold-plated, follow-up, whoppe, charizard, market, halloween-themed	-1.0%
		48	"Robbery"	comment, manager, public, officer, boycottburgerking, stand, fall, let, mock, cop	-3.7%
		42	"Rolling Blunt"	roll, blunt, stream, shitwire, arrogance, treatise, longform, platform, content, pro	-13.9%

Table 5. Topics with Largest Increases and Decreases in Prevalence of Burger King Study (Cont.)

Changepoint	Date	Most Shifted Topics	Topic Label	Sample of Most Relevant Words	Prevalence Change
6	3/29/2016	50	"Angriest Whopper Release"	whopper, angriest, angry, red, new, spicy, review, event, try, blockbuster	13.0%
		18	"Angriest Whopper Descriptions"	sauce, bake, bun, angriest, whopper, hot, red, new, so-called, fiery	9.7%
		13	"Red Buns"	sizzle, face, red, bright, bun, worried, customer, release, certain, heat	5.8%
		6	"Family Tie"	jr, somebody, yell, grandkids, ai, order, papaw, pawpaw, whopper, king	-2.4%
		32	"Gilled Hotdog"	dog, hot, commercial, sell, grilled, grill, hotdog, chili, try, disgust	-4.2%
		11	"Daily"	go, work, time, like, drive, way, lunch, say, last, make	-9.3%
7	3/31/2016	11	"Daily"	go, work, time, like, drive, way, lunch, say, last, make	4.1%
		29	"Fried Chicken"	fry, chicken, smell, good, introduce, new, true, sandwich, shake, believe	3.5%
		39	"Colorful Buns"	color, bun, black, red, poop, green, different, colored, chill, turn	3.4%
		50	"Angriest Whopper Release"	whopper, angriest, angry, red, new, spicy, review, event, try, blockbuster	-5.1%
		18	"Angriest Whopper Descriptions"	sauce, bake, bun, angriest, whopper, hot, red, new, so-called, fiery	-5.7%
		13	"Red Buns"	sizzle, face, red, bright, bun, worried, customer, release, certain, heat	-5.8%
8	4/3/2016	53	"Food Safety"	antibiotic, raise, petition, meat, sign, stop, use, wendys, tell, mcdonalds	4.9%
		11	"Daily"	go, work, time, like, drive, way, lunch, say, last, make	3.1%
		44	"Chicken Nuggets"	nugget, chicken, friend, fake, separate, occasion, fry, casually, cameraman, cool	1.5%
		29	"Fried Chicken"	fry, chicken, smell, good, introduce, new, true, sandwich, shake, believe	-2.7%
		50	"Angriest Whopper Release"	whopper, angriest, angry, red, new, spicy, review, event, try, blockbuster	-2.7%
		60	"Eating"	eat, dog, block, hot, hotdog, trust, grill, mark, people, buy	-5.0%
9	4/8/2016	38	"Prank Call"	prank, worker, window, smash, caller, restaurant, trick, minnesota, bust, convince	19.0%
		27	"Promotion"	admit, tried, promotion, amaze, ashamed, pay, love, dog, hot, belly	14.3%
		20	"Prank Consequence"	break, store, employee, trick, window, vandalism, instruct, explosive, far, shawnee	7.0%
		53	"Food Safety"	antibiotic, raise, petition, meat, sign, stop, use, wendys, tell, mcdonalds	-4.8%
		33	"Comments"	burger, king, go, veggie, trash, nasty, taco, breakfast, eat, mcdonalds	-5.8%
		11	"Daily"	go, work, time, like, drive, way, lunch, say, last, make	-11.2%

Table 6. Topics with Largest Increases and Decreases in Prevalence of Under Armour Study

Changepoint	Date	Most Shifted Topics	Topic Label	Sample of Most Relevant Words	Prevalence Change
1	3/16/2018	57	"Unleash Chaos T-shirt"	unleashchaos, wewill, victory, charlotte, historic, ready, national, championship, irish, round	17.2%
		2	"Retro"	retro, teen, wolf, kanye, bash, yeezy, know, look, go, end	10.7%
		56	"All-American Camp Invitation"	invite, bless, iwill, game, all-american, camp, play, receive, honor, american	3.4%
		10	"Relax"	waitstaff, treatment, alternate, random, contender, billowy, tyrannical, bsn, cash, fat	-3.4%
		26	"Positivity"	like, wear, love, nike, make, good, look, think, go, need	-8.2%
		25	"Stephen Curry"	curry, pi, steph, day, stephen, shoe, restock, all-white, strongest, surprisingly	-19.1%
2	3/18/2018	31	"Unreleased"	unreleased, swag, retrievernation, hook, game, tonight, genius, thank, shoe, curry	9.1%
		22	"Build The Belief"	build, belief, strong, enjoy, stay, chasegreatnesscollection, projectrock, gear, available, new	5.0%
		3	"Campaign"	projectrock, representative, movement, disruptive, drop, authentic, build, chasegreatnesscollection, inspiration, chasegreatness	3.5%
		56	"All-American Camp Invitation"	invite, bless, iwill, game, all-american, camp, play, receive, honor, american	-4.8%
		2	"Retro"	retro, teen, wolf, kanye, bash, yeezy, know, look, go, end	-10.6%
		57	"Unleash Chaos T-shirt"	unleashchaos, wewill, victory, charlotte, historic, ready, national, championship, irish, round	-17.6%
3	3/23/2018	31	"Unreleased"	unreleased, swag, retrievernation, hook, game, tonight, genius, thank, shoe, curry	4.3%
		19	"Misty Copeland"	create, prop, groundbreaking, campaign, papa, principal, dancer, barrier, dedication, ballet	3.9%
		27	"Fleece Hoodie"	icon, hoodie, caliber, camo, fleece, fade, zip, sale, pullover, blocked	2.0%
		1	"All-American Camp"	camp, uaallamerica, today, great, quarterback, pierre, orlando, talented, all-america, compete	-2.8%
		56	"All-American Camp Invitation"	invite, bless, iwill, game, all-american, camp, play, receive, honor, american	-3.4%
		47	"All-American Camp Miami"	miami, camp, wr, recruit, db, espnjr, visit, all-american, dl, talent	-3.5%
4	3/25/2018	1	"All-American Camp"	camp, uaallamerica, today, great, quarterback, pierre, orlando, talented, all-america, compete	7.1%
		51	"Gears"	champ, umbc, gear, rate, discount, uniform, money, thank, apparel, receive	4.6%
		21	"Hoodie"	hoodie, fleece, armour, stack, boy, shop, rival, available, camo, full-zip	4.1%
		19	"Misty Copeland"	create, prop, groundbreaking, campaign, papa, principal, dancer, barrier, dedication, ballet	-4.0%
		26	"Positivity"	like, wear, love, nike, make, good, look, think, go, need	-5.4%
		31	"Unreleased"	unreleased, swag, retrievernation, hook, game, tonight, genius, thank, shoe, curry	-13.8%
5	3/27/2018	8	"Gift Card"	enter, hour, follow, madness, win, retweet, chance, card, gift, giveaway	10.8%
		32	"Aaron Judge"	judge, aaron, bat, wristband, glove, sign, cleat, nms, adidas, boycottmd	5.0%
		26	"Positivity"	like, wear, love, nike, make, good, look, think, go, need	3.1%
		44	"Willie Taggart"	select, camp, taggart, willie, football, junior, lacrosse, wave, boston, diamond	-3.0%
		51	"Gears"	champ, umbc, gear, rate, discount, uniform, money, thank, apparel, receive	-3.2%
		1	"All-American Camp"	camp, uaallamerica, today, great, quarterback, pierre, orlando, talented, all-america, compete	-6.8%

Table 6. Topics with Largest Increases and Decreases in Prevalence of Under Armour Study (Cont.)

Changepoint	Date	Most Shifted Topics	Topic Label	Sample of Most Relevant Words	Prevalence Change
6	3/29/2018	18	"Data Breach Announcement"	breach, data, affect, million, user, myfitnesspal, account, say, break, security	43.3%
		37	"Cybersecurity	account, hack, say, myfitnesspal, million, breach, compromise, reuters, massive, data	14.6%
		20	"Data Breach Details"	app, steal, hacker, fitness, health, popular, hack, diet, myfitnesspal, data	5.1%
		35	"Fit"	amazon, sleeve, short, xl, check, size, nwt, loose, large, shirt	-4.9%
		26	"Positivity"	like, wear, love, nike, make, good, look, think, go, need	-10.3%
		8	"Gift Card"	enter, hour, follow, madness, win, retweet, chance, card, gift, giveaway	-12.3%
7	3/31/2018	26	"Positivity"	like, wear, love, nike, make, good, look, think, go, need	6.8%
		2	"Retro"	retro, teen, wolf, kanye, bash, yeezy, know, look, go, end	4.6%
		35	"Fit"	amazon, sleeve, short, xl, check, size, nwt, loose, large, shirt	4.0%
		6	"Compromised Data"	email, address, username, nutrition, hash, password, food, compromise, data, info	-2.3%
		37	"Cybersecurity"	account, hack, say, myfitnesspal, million, breach, compromise, reuters, massive, data	-8.0%
		18	"Data Breach Announcement"	breach, data, affect, million, user, myfitnesspal, account, say, break, security	-35.4%
8	4/3/2018	56	"All-American Camp Invitation"	invite, bless, iwill, game, all-american, camp, play, receive, honor, american	4.2%
		55	"Sales"	walmart, threshold, crony, kushner, depot, benefiting, misdemeanor, furthermore, loss, crime	3.4%
		54	"Mark Richt"	aquinas, richt, theu, in-depth, solomon, anthony, thomas, mark, decision, lb	3.1%
		2	"Retro"	retro, teen, wolf, kanye, bash, yeezy, know, look, go, end	-4.6%
		37	"Cybersecurity"	account, hack, say, myfitnesspal, million, breach, compromise, reuters, massive, data	-4.9%
		18	"Data Breach Announcement"	breach, data, affect, million, user, myfitnesspal, account, say, break, security	-6.4%
9	4/5/2018	38	"Order"	shipping, order, free, available, men, shop, new, tank, coolswitch, thermocline	7.5%
		42	"Jordan Spieth"	master, tour, cap, jordan, spieth, green, sign, giveaway, animal, slipper	4.2%
		26	"Positivity"	like, wear, love, nike, make, good, look, think, go, need	1.8%
		54	"Mark Richt"	aquinas, richt, theu, in-depth, solomon, anthony, thomas, mark, decision, lb	-3.1%
		55	"Sales"	walmart, threshold, crony, kushner, depot, benefiting, misdemeanor, furthermore, loss, crime	-3.6%
		56	"All-American Camp Invitation"	invite, bless, iwill, game, all-american, camp, play, receive, honor, american	-3.6%
10	4/8/2018	38	"Order"	shipping, order, free, available, men, shop, new, tank, coolswitch, thermocline	23.0%
		1	"All-American Camp"	camp, uaallamerica, today, great, quarterback, pierre, orlando, talented, all-america, compete	5.0%
		49	"Rory McIlroy"	prime, rory, january, mcilroy, announcement, spieth, cover, happen, barely, sign	2.2%
		35	"Fit"	amazon, sleeve, short, xl, check, size, nwt, loose, large, shirt	-2.6%
		40	"Style"	logo, short-sleeve, tee, freedom, shirt, flag, men, shop, boxed, tonal	-4.0%
		26	"Positivity"	like, wear, love, nike, make, good, look, think, go, need	-4.9%

Table 6. Topics with Largest Increases and Decreases in Prevalence of Under Armour Study (Cont.)

Changepoint	Date	Most Shifted Topics	Topic Label	Sample of Most Relevant Words	Prevalence Change
11	4/10/2018	52	"College Sports"	year, nike, adidas, michigan, kansas, ucla, college, ohio, louisville, texas	6.1%
		53	"Tuition"	tuition, vacation, bookstore, card, beach, gift, amazing, extra, prize, pm	4.3%
		48	"Andre Ingram"	ingram, andre, year-old, nba, debut, league, curry, make, uahovr, year	4.1%
		49	"Rory McIlroy"	prime, rory, january, mcilroy, announcement, spieth, cover, happen, barely, sign	-2.2%
		38	"Order"	shipping, order, free, available, men, shop, new, tank, coolswitch, thermocline	-3.9%
		1	"All-American Camp"	camp, uaallamerica, today, great, quarterback, pierre, orlando, talented, all-america, compete	-4.6%

Table 7. User Distribution by Posting Behavior

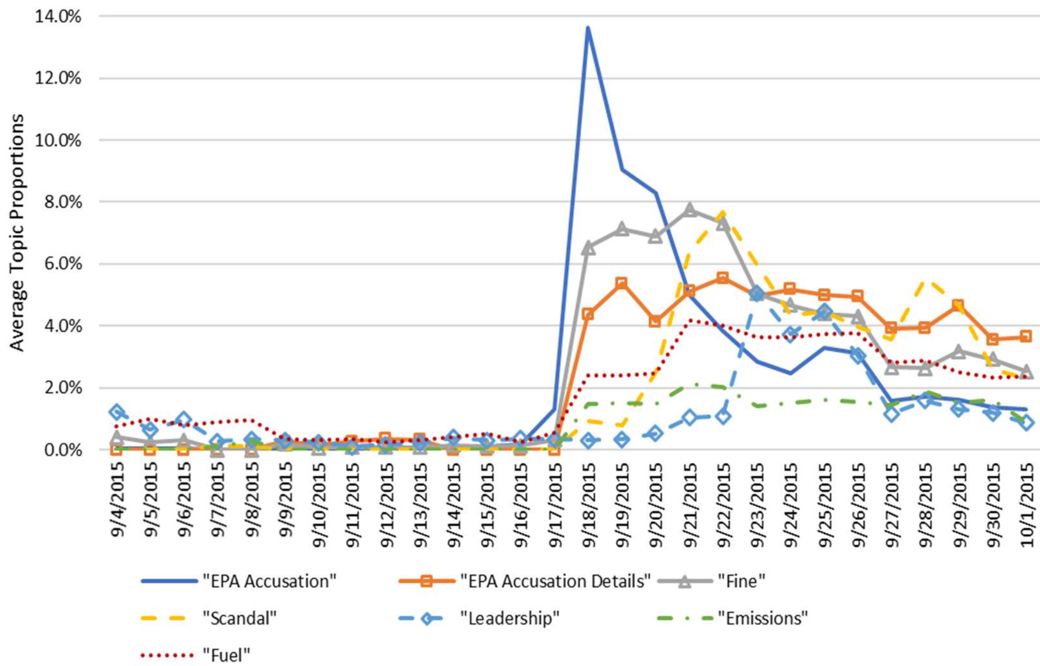
Study	Posting Behavior	Number of Authors	Number of Posts
(a) Burger King	Post Before Only	25,131	28,370
	Post After Only	31,345	36,190
	Post Both Before and After	3,059	11,218
(b) Under Armour	Post Before Only	9,781	11,845
	Post After Only	16,799	20,638
	Post Both Before and After	1,726	21,562

Figure 1. Comparison of Word Clouds for Volkswagen's Emissions Testing Scandal

(a) Word Cloud before Scandal

(b) Word Cloud after Scandal

Figure 2. Changing Pattern of Scandal-Related Topics vs. Volume and Sentiment



No: The number of posts that mention Volkswagen brand between Day 18 and Day 21 are censored due to the restrictions of Crimson Hexagon.

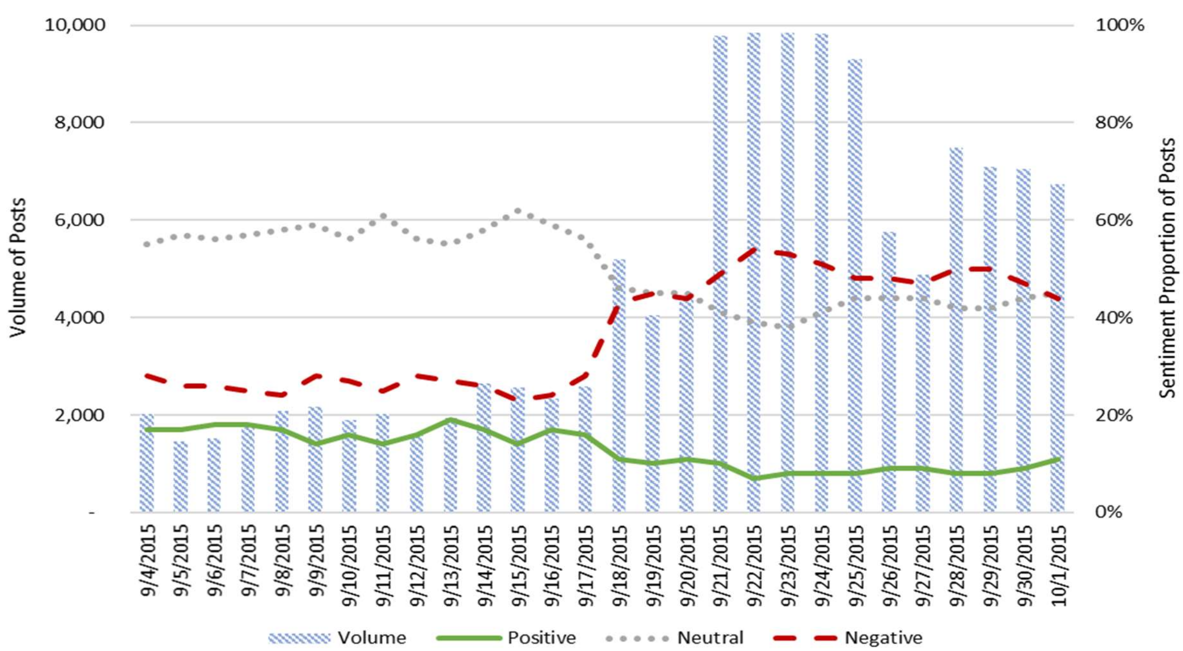
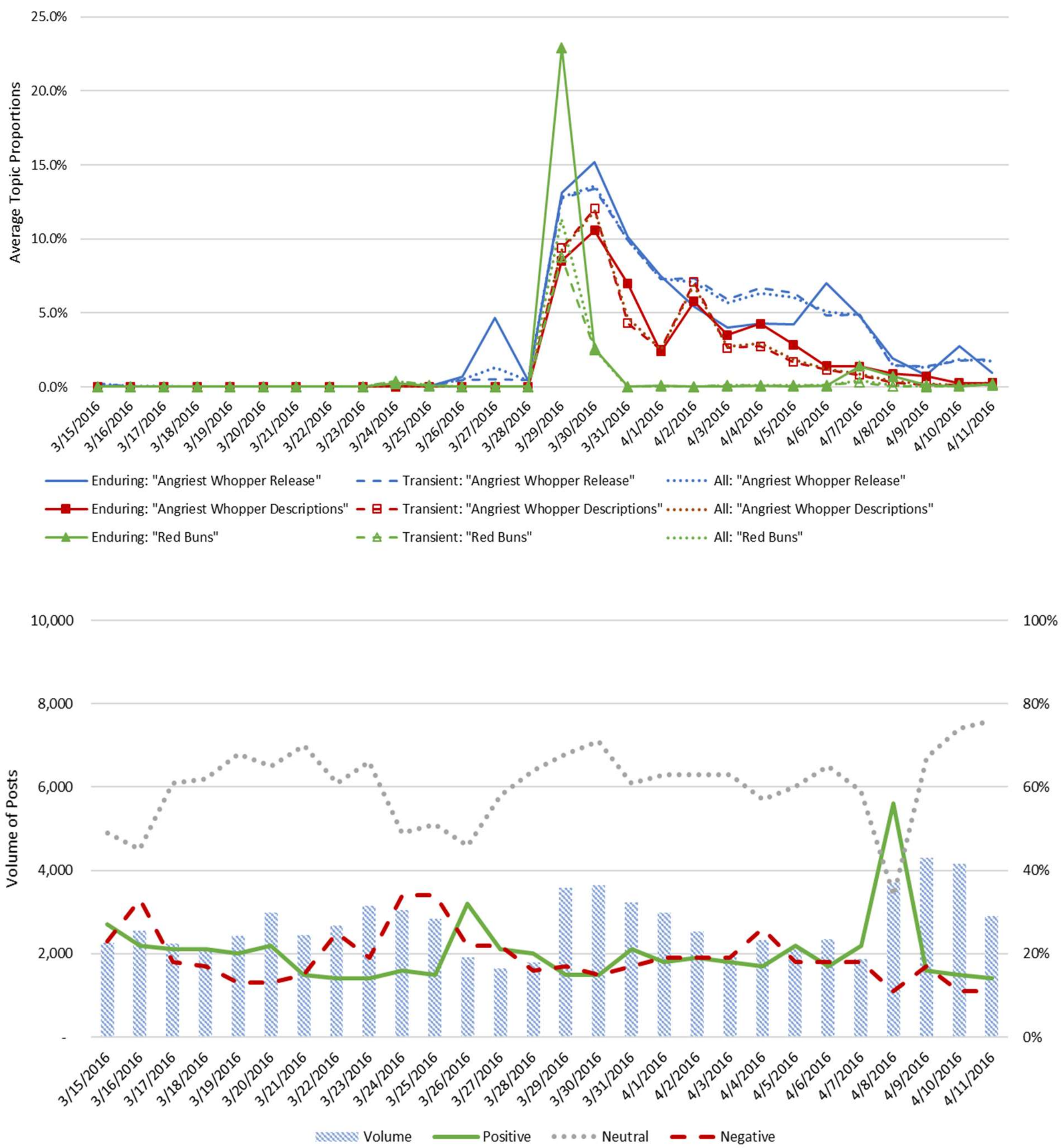
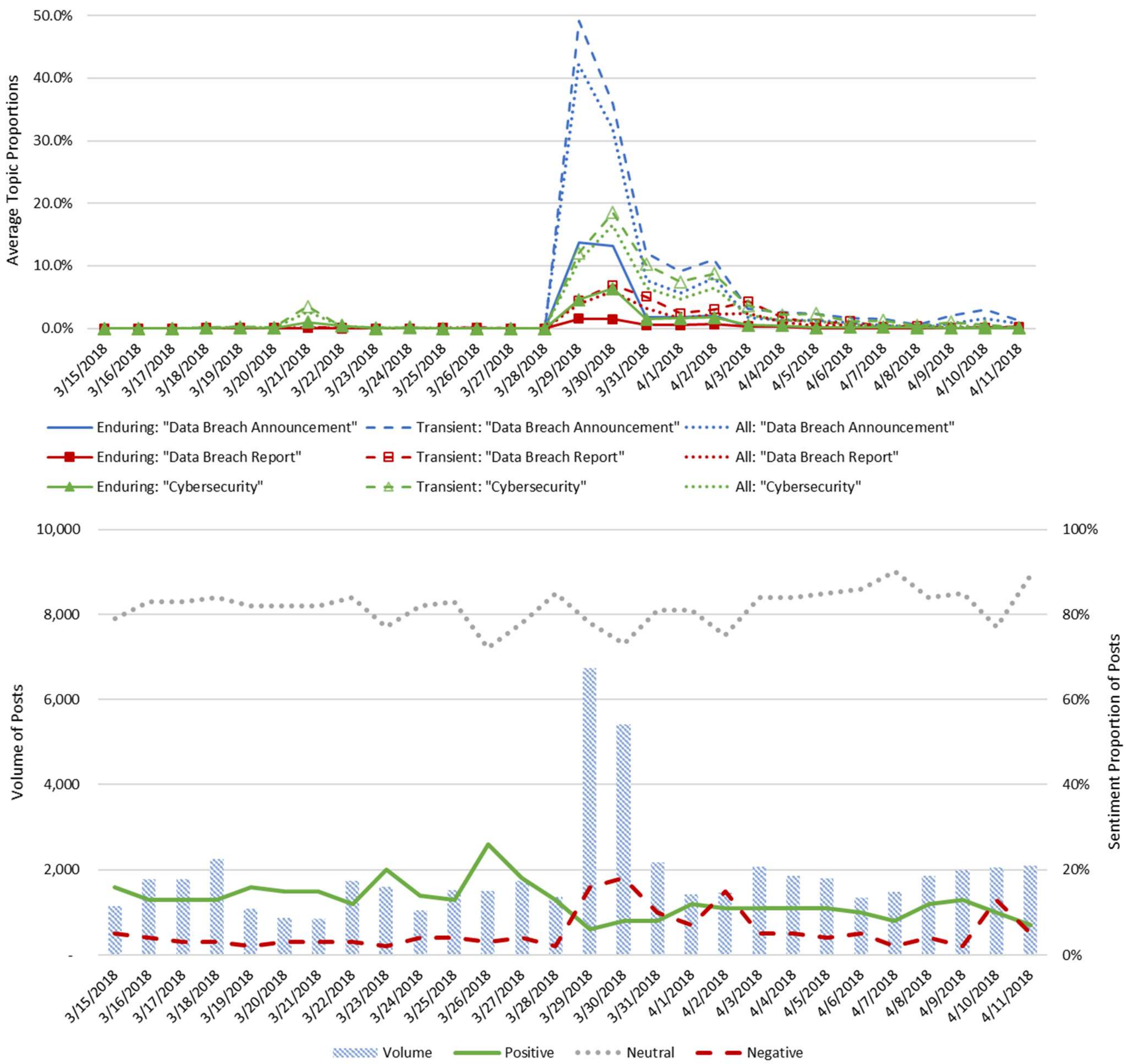


Figure 3. Average Topic Proportions of Event-Related Topics

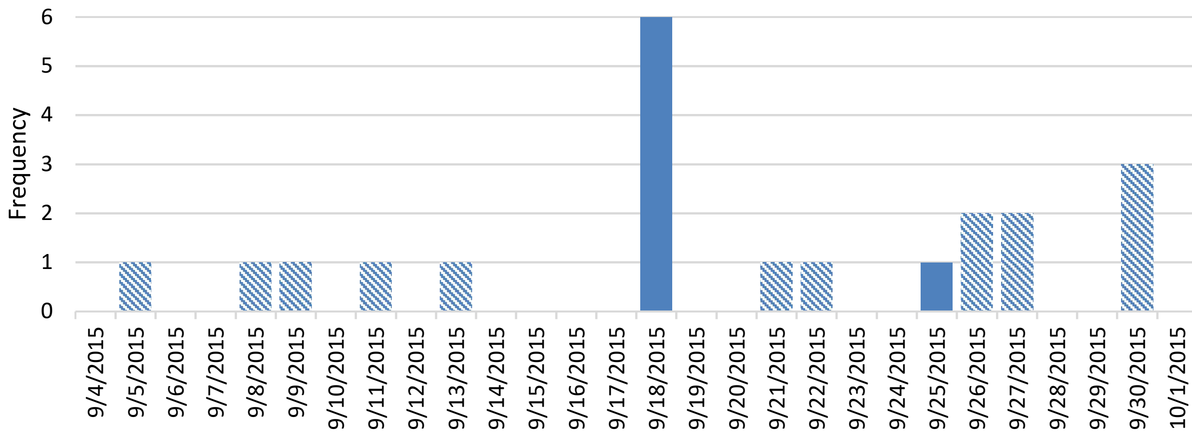


(a) Burger King

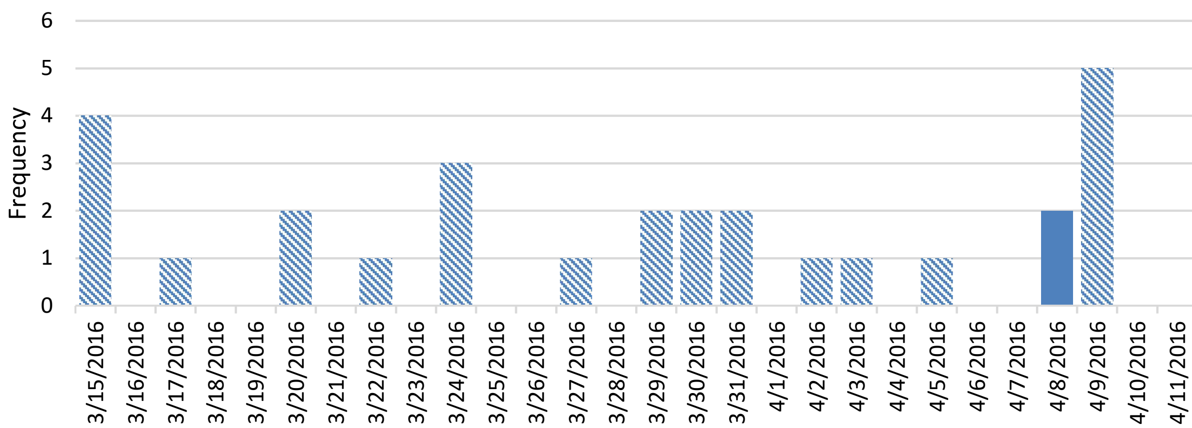


(b) Under Armour

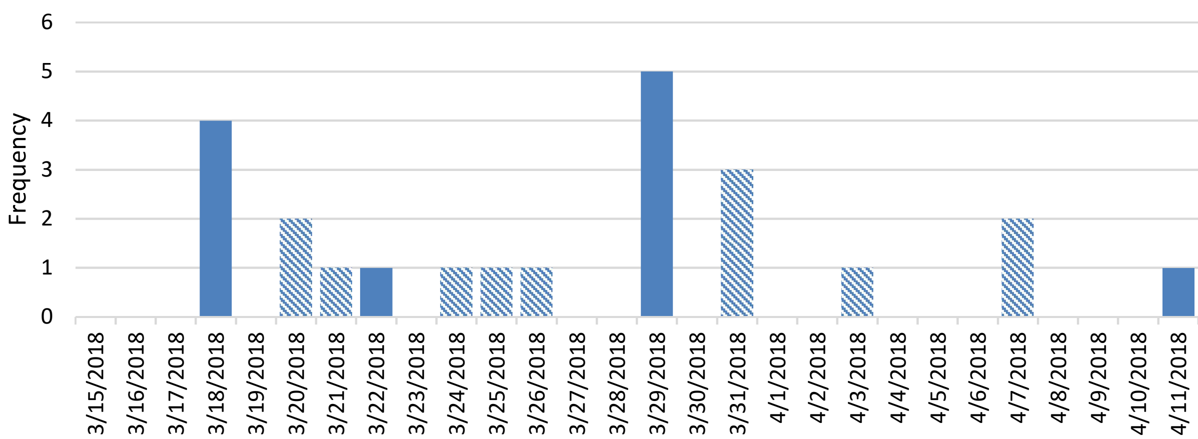
Figure 4. Identified Changepoints in Rolling Window Analysis



(a) Volkswagen



(b) Burger King



(c) Under Armour

Appendix A.1. MCMC Algorithm for LDA-LC

The Hybrid algorithm for LDA-MLC integrates Bayesian changepoint detection with Gibbs sampling for LDA via the following three steps:

1. $\boldsymbol{z} \mid \boldsymbol{w}, \boldsymbol{m}, \boldsymbol{s}$

For any document $d \in C_t$ where state $s_t = k$, the conditional distribution of token *i* being assigned to topic *j* can be represented as

$$P(z_i = j|\boldsymbol{z}_{-i}, \boldsymbol{w}) \propto \frac{P\left(z_i = j, \boldsymbol{z}_{-i}, \boldsymbol{w}|\alpha_0 \boldsymbol{m}^{(k)}, \boldsymbol{\beta}\right)}{P(\boldsymbol{z}_{-i}, \boldsymbol{w}_{-i}|\alpha_0 \boldsymbol{m}^{(k)}, \boldsymbol{\beta})} = \frac{n_{-i,j}^{(w_i)} + \beta}{n_{-i,j}^{(\cdot)} + W\beta}\left(n_{-i,j}^{(d_i)} + \alpha_0 m_j^{(k)}\right) \tag{3}$$

Similar to Gibbs sampling procedure for LDA proposed by Griffiths and Steyvers (2004), the estimates of word distribution of topics ϕ can be estimates by Eqn. (4) and the estimates of topic distribution of documents θ can be obtained by

$$\hat{\theta}_j^{(d)} = \frac{n_j^{(d)} + \alpha_0 m_j^{(k)}}{n_\cdot^{(d)} + \alpha_0} \qquad \text{if } d \in C_k \tag{4}$$

2. $\boldsymbol{m} \mid \boldsymbol{z}, \boldsymbol{s}$

Given the fact that $\alpha_0 \boldsymbol{m}^{(k)}$ for any regime *k* is involved in $p(\boldsymbol{w}|\boldsymbol{z}, \boldsymbol{\beta})$ and $p(\boldsymbol{\beta})$ in Eqn. (8), probability measures of the hyperpriors, $\boldsymbol{m}$, can be sampled separately given the other parameters. According to the generative process of LDA, for each document *d* topic proportion θ is drawn from Dirichlet distribution with parameter $\boldsymbol{\alpha}$. This compound distribution is a Dirichlet-multinomial distribution, or Pólya distribution (Minka 2000). Therefore, the likelihood of $\alpha_0 \boldsymbol{m}$ conditional on topic assignments $\boldsymbol{z}$ and state $\boldsymbol{s}$ is given by

$$L(\alpha_0 \boldsymbol{m}|\boldsymbol{z},\boldsymbol{s}) = P(\boldsymbol{z}|\alpha_0\boldsymbol{m},\boldsymbol{s}) = \prod_{t=1}^{T}\prod_{d\in\mathcal{C}_t}\left(\frac{\Gamma(\alpha_0)}{\Gamma\left(\mathrm{n}_{\cdot}^{(d)}+\alpha_0\right)}\prod_{j=1}^{J}\frac{\Gamma\left(n_j^{(d)}+\alpha_0 m_j^{(s_t)}\right)}{\Gamma\left(\alpha_0 m_j^{(s_t)}\right)}\right) \quad (5)$$

where $\mathrm{n}_j^{(d)}$ counts the number of tokens assigned to topic *j* in document *d*, and $\mathrm{n}_{\cdot}^{(d)}$ counts the total number of all tokens in document *d*. Given the changing number of regimes *k* over iterations and conditional independence of $\boldsymbol{m}^{(k)}$ for each regime *k*, we use MAP to estimate probability measure $\boldsymbol{m}$ on Eqn. (11) for every 20 iterations (Minka 2000; Heinrich 2009).

3. $\boldsymbol{s} \mid \boldsymbol{m}, \boldsymbol{z}$

As we assume that the shifts of regimes follow a hidden Markov chain in which the number of states is endogenously determined, the transition matrix with infinite dimension is represented by

$$P = \begin{bmatrix} p_{11} & p_{12} & 0 & \cdots \\ 0 & p_{22} & p_{23} & \cdots \\ \vdots & \vdots & \ddots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (6)$$

At a certain state *k*, the only possible paths are self-transition and transition to the next state *k*+1 (i.e. $p_{kk} + p_{k,k+1} = 1$). The likelihood of $\boldsymbol{s}$ conditional on topic assignment $\boldsymbol{z}$ and probability measure $\boldsymbol{m}$ can be written by

$$\begin{aligned} L(\boldsymbol{s}|\alpha_0\boldsymbol{m},\boldsymbol{z}) = p(\boldsymbol{z}|\boldsymbol{s},\alpha_0\boldsymbol{m}) &= \prod_{t=1}^{T}\prod_{d\in\mathcal{C}_t} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_t)}\right) \\ &= \prod_{t=1}^{T}\prod_{d\in\mathcal{C}_t}\left(\frac{\Gamma(\alpha_0)}{\Gamma\left(\mathrm{n}_{\cdot}^{(d)}+\alpha_0\right)}\prod_{j=1}^{J}\frac{\Gamma\left(n_j^{(d)}+\alpha_0 m_j^{(s_t)}\right)}{\Gamma\left(\alpha_0 m_j^{(s_t)}\right)}\right) \end{aligned} \quad (7)$$

Ko et al. (2015) use Dirichlet process prior for transition probabilities and show that s_t can be drawn via Gibbs Sampling in turn for $t = 1,2,\dots,T$ by

$$p(s_t|S_{t-1},S^{t+1},\alpha_0\boldsymbol{m},\boldsymbol{z}) \propto p(s_t|s_{t-1},S_{t-2})p(s_{t+1}|s_t,S^{t+2})\prod_{d\in\mathcal{C}_t} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_t)}\right) \quad (14)$$

where $S_{t-2} = (s_1, s_2, \dots, s_{t-2})$ and $S^{t+2} = (s_{t+2}, s_{t+2}, \dots, s_T)$. This method requires the state s_t be sampled if and only if $s_{t-1} \neq s_{t+1}$. Therefore, the general form of conditional posterior of state s_t can be represented by

$$p(s_t|S_{t-1}, S^{t+1}, \alpha_0\boldsymbol{m}, \boldsymbol{z}) \\ \propto \begin{cases} \dfrac{n_{kk}+\lambda}{n_{kk}+\nu+\lambda} \cdot \dfrac{\nu}{n_{kk}+1+\nu+\lambda} \cdot \prod\limits_{d\in C_1} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_t)}\right) & \text{if } s_t \text{ not changed} \\ \dfrac{\nu}{n_{kk}+\nu+\lambda} \cdot \dfrac{n_{k+1,k+1}+\lambda}{n_{k+1,k+1}+\nu+\lambda} \cdot \prod\limits_{d\in C_1} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_{t+1})}\right) & \text{if } s_t = s_{t+1} \end{cases} \tag{8}$$

where n_{kk} denotes the counts of self-transitions that have occurred so far in the current iteration for state k, λ is the hyperprior that controls the tendency to linger in a state, and ν is the hyperprior that controls the tendency to join the next state. Noticeably, for the state of the starting point, s_1, the posterior is

$$p(s_1|S^2, \alpha_0\boldsymbol{m}, \boldsymbol{z}) \\ \propto \begin{cases} \dfrac{\lambda}{\nu+\lambda} \cdot \dfrac{\nu}{\nu+\lambda} \cdot \prod\limits_{d\in C_1} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_1)}\right) & \text{if } s_1 \text{ not changed} \\ \dfrac{\nu}{\nu+\lambda} \cdot \dfrac{n_{s_2,s_2}+\lambda}{n_{s_2,s_2}+\nu+\lambda} \cdot \prod\limits_{d\in C_1} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_2)}\right) & \text{if } s_1 = s_2 \end{cases} \tag{9}$$

And for the state of ending point, s_T, the posterior is

$$p(s_T|S_{T-1}, \alpha_0\boldsymbol{m}, \boldsymbol{z}) \\ \propto \begin{cases} \dfrac{n_{s_{T-1},s_{T-1}}+\lambda}{n_{s_{T-1},s_{T-1}}+\nu+\lambda} \cdot 1 \cdot \prod\limits_{d\in C_1} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_{T-1})}\right) & \text{if } s_T = s_{T-1} \\ \dfrac{\nu}{n_{s_{T-1},s_{T-1}}+\nu+\lambda} \cdot 1 \cdot \prod\limits_{d\in C_1} p\left(\boldsymbol{z}^{(d)}|\alpha_0\boldsymbol{m}^{(s_T)}\right) & \text{if } s_T \text{ not changed} \end{cases} \tag{10}$$

In our research, Dirichlet process hyperpriors λ and ν are estimated via Metropolis-Hasting (MH). Specifically, with the number of states specified in each iteration, the hidden Markov model reduces to the generalized Dirichlet distribution (Connor and Mosimann 1969; Wong 1998). The posterior of λ and ν is

$$p(\lambda, \nu|\boldsymbol{s}) \propto \mathrm{Gamma}(a_\lambda, b_\lambda)\mathrm{Gamma}(a_\nu, b_\nu) \prod_k \frac{\nu\Gamma(\lambda+\nu)}{\Gamma(\lambda)} \frac{\Gamma(n_{kk}+\lambda)}{\Gamma(n_{kk}+1+\lambda+\nu)} \tag{11}$$

where $\mathrm{Gamma}(a_\lambda, b_\lambda)$ is the prior for λ and $\mathrm{Gamma}(a_\nu, b_\nu)$ is the prior for ν. We set $a_\lambda = b_\lambda = a_\nu = b_\nu = 1$.

Appendix A.2. Simulation Study

To examine the validity of our hybrid algorithm, we create a dataset according to the generative process of LDA-MLC with the number of documents, $D_t = 300$, $t = 1,2,\ldots,12$, the size of vocabulary, $W = 100$, the number of tokens in each document d, $n^{(d)}$ ~ Unif $(10,90)$, the number of topics, $J = 5$, the concentration parameter for topic distributions of documents, $\alpha_0 = 5$, and the symmetric Dirichlet parameter for word distributions of topics, $\beta = 0.5$. Two changepoints are assumed to occur at the beginning of $t = 5$ and $t = 9$. The true values and the parameter estimates under each model specification are shown in Table A1. The hybrid algorithm is conducted for 5,000 iterations after 5,000 burn-in periods. Since we estimate $\boldsymbol{m}$ by ML, report the sample standard errors to show its stability after the burn-in period (Ko et al. 2015).

[Insert Table A1 Here]

LDA-MLC correctly identifies the two changepoints that we preset. As shown in Table A1, nearly all the true values are contained within the 95% confidence interval for the LDA-MLC, demonstrating our ability to recover the true parameter values.

In addition to estimating the LDA-MLC model, we also estimate the modified LDA model. As the data were generated from a three-regime model, we would expect the topic proportions inferred under a single regime model to fall between the true topic proportions from the two regimes. This is exactly what is observed in Table A1. For example, the Topic 5 has a proportion of 10% in regime 1, 20% in regime 2, and 30% in regime 3. Whereas the LDA-MLC model captures this shift, the modified LDA model estimates a topic proportion of 21% all over the time.

The LDA-MLC model is capable of detecting a changepoint in which topic proportions vary by up to 10% from one regime to another, as depicted in the scenario presented in Table A1.

To assess the ability of the LDA-MLC model to detect smaller shifts in topic prevalence from one regime to the next, we conduct two additional simulations in which we vary the maximum range of topic prevalence $\boldsymbol{m}$. These results are reported in Table A2.

[Insert Table A2 Here]

As shown in Table A2, when the maximum range of $\boldsymbol{m}$ is 5%, (i.e. Topic 1 shifts from 0.25 to 0.20 at the first changepoint and shifts from 0.20 to 0.15 at the second changepoint), both changepoints ($t = 5$ and $t = 9$) are correctly identified by LDA-MLC. To contrast, when the maximum range of $\boldsymbol{m}$ is as small as 1%, (i.e. Topic 1 changes from 0.21 to 0.20 at the first changepoint and from 0.20 to 0.19 at the second changepoint), there is no changepoint detected by LDA-MLC. Thus, while the LDA-MLC model is able to detect abrupt changes in topic prevalence (e.g., changes of 5% and 10%), this simulation study suggests that it is not well-suited to capture gradual changes.

Table A1: Parameter Recovery of Simulation Analysis

Regime	Parameter	True Value	*Modified LDA* Estimate (S.E.)	*LDA-MLC* Estimate (S.E.)
1	$m_1^{(1)}$	0.3000	0.2012 (0.0041)	0.3126 (0.0047)
	$m_2^{(1)}$	0.2500	0.1986 (0.0056)	0.2524 (0.0063)
	$m_3^{(1)}$	0.2000	0.1932 (0.0046)	0.2045 (0.0062)
	$m_4^{(1)}$	0.1500	0.2085 (0.0039)	0.1236 (0.0042)
	$m_5^{(1)}$	0.1000	0.1985 (0.0049)	0.1068 (0.0043)
2	$m_1^{(2)}$	0.2000	-	0.1899 (0.0037)
	$m_2^{(2)}$	0.2000	-	0.1956 (0.0056)
	$m_3^{(2)}$	0.2000	-	0.2069 (0.0058)
	$m_4^{(2)}$	0.2000	-	0.2053 (0.0046)
	$m_5^{(2)}$	0.2000	-	0.2024 (0.0046)
3	$m_1^{(3)}$	0.1000	-	0.1055 (0.0039)
	$m_2^{(3)}$	0.1500	-	0.1487 (0.0053)
	$m_3^{(3)}$	0.2000	-	0.2002 (0.0061)
	$m_4^{(3)}$	0.2500	-	0.2519 (0.0053)
	$m_5^{(3)}$	0.3000	-	0.2937 (0.0055)

Table A2: Sensitivity Analysis of LDA-MLC

		5% Change		*1% Change*	
Regime	Parameter	True Value	Estimate (S.E.)	True Value	Estimate (S.E.)
1	$m_1^{(1)}$	0.2500	0.2550 (0.0062)	0.2100	0.1909 (0.0048)
	$m_2^{(1)}$	0.2000	0.1929 (0.0049)	0.2000	0.1999 (0.0043)
	$m_3^{(1)}$	0.2000	0.2066 (0.0056)	0.2000	0.2020 (0.0043)
	$m_4^{(1)}$	0.2000	0.1996 (0.0057)	0.2000	0.1947 (0.0047)
	$m_5^{(1)}$	0.1500	0.1458 (0.0058)	0.1900	0.2126 (0.0061)
2	$m_1^{(2)}$	0.2000	0.1957 (0.0059)	0.2000	-
	$m_2^{(2)}$	0.2000	0.2022 (0.0050)	0.2000	-
	$m_3^{(2)}$	0.2000	0.2002 (0.0053)	0.2000	-
	$m_4^{(2)}$	0.2000	0.2045 (0.0054)	0.2000	-
	$m_5^{(2)}$	0.2000	0.1974 (0.0061)	0.2000	-
3	$m_1^{(3)}$	0.1500	0.1611 (0.0056)	0.1900	-
	$m_2^{(3)}$	0.2000	0.2120 (0.0051)	0.2000	-
	$m_3^{(3)}$	0.2000	0.2004 (0.0058)	0.2000	-
	$m_4^{(3)}$	0.2000	0.1956 (0.0058)	0.2000	-
	$m_5^{(3)}$	0.2500	0.2309 (0.0067)	0.2100	-

Appendix A.3. Most Relevant Words and Prevalence of Topics

Volkswagen		Topic Prevalence $m_j^{(k)}$ for topic j in regime k								
Topic	**Most Relevant Words**	**Regime 1**	**Regime 2**	**Regime 3**	**Regime 4**	**Regime 5**	**Regime 6**	**Regime 7**	**Regime 8**	**Regime 9**
1	gear, manual, polo, clutch, dsg, tsi, speed, torque, shift, automatic, engine, stick, drive, second, tire, traffic, power, own, tdi, dual	2.1%	1.8%	2.4%	1.8%	0.8%	0.6%	0.5%	0.6%	0.9%
2	blog, group, scandal, technical, commercial, brand, passenger, confirm, worldwide, working, solution, reveal, generation, plan, boss, detail, announc, statement, latest, fully	0.0%	0.0%	0.2%	0.0%	0.0%	0.4%	0.8%	1.5%	1.4%
3	electric, tesla, model, apple, car, battery, range, technology, project, charge, plug, hybrid, vehicle, launch, auto, future, google, company, suv, build	0.3%	0.2%	0.4%	1.4%	0.7%	0.5%	0.6%	0.5%	0.6%
4	ceo, board, winterkorn, executive, matthias, company, group, chief, chairman, porsche, resigned, management, meeting, fresh, scandal, boss, member, wednesday, head, named	0.6%	0.3%	0.2%	0.4%	0.4%	1.0%	3.4%	1.3%	0.9%
5	pollution, climate, carbon, air, health, co2, death, nox, change, nitrogen, energy, pollutant, smog, emit, global, reduce, real, human, tax, clean	0.0%	0.0%	0.1%	0.2%	0.5%	1.2%	1.8%	1.5%	1.6%
6	passed, turbo, care, health, condition, join, provide, subaru, arrive, hopefully, annual, travel, sell, eye, senior, canada, driven, excellent, community, standard	0.0%	0.3%	0.7%	0.2%	0.2%	0.2%	0.1%	0.1%	0.3%
7	president, obama, news, state, york, government, united, war, house, washington, reuters, party, visit, thursday, political, times, minister, country, police, china	0.7%	0.7%	0.4%	0.7%	1.1%	1.3%	1.6%	1.4%	1.1%
8	nox, urea, injection, exhaust, co2, euro, meet, system, diesel, fuel, tank, cycle, reduce, tdi, limit, mile, engine, output, solution, nitrogen	0.0%	0.1%	0.0%	0.0%	1.4%	1.8%	1.4%	1.5%	1.1%
9	australia, gold, london, stage, job, week, silver, round, night, water, generation, live, tuesday, match, central, answer, day, final, daily, stand	0.4%	0.6%	0.8%	1.1%	0.4%	0.2%	0.5%	0.7%	0.7%
10	business, study, social, university, facebook, age, school, technology, media, book, marketing, life, help, young, series, content, word, history, online, american	1.8%	1.9%	1.3%	1.6%	1.2%	1.4%	2.1%	2.1%	1.9%
11	sensor, ecu, bank, cylinder, factory, control, low, exhaust, high, power, torque, boost, position, stock, engine, fuel,	1.3%	0.3%	0.4%	0.6%	0.6%	0.4%	0.3%	0.4%	0.3%

	ignition, upgrade, intake, response									
12	battery, wire, box, mode, mini, max, power, output, ground, control, tank, heat, range, charge, hot, built, low, black, kit, auto	1.6%	1.9%	1.3%	1.4%	1.0%	0.6%	0.4%	0.7%	0.8%
13	money, cash, stock, tax, price, buy, pay, market, sell, free, paid, credit, fund, finance, buying, selling, trading, cost, trade, purchase	1.5%	1.2%	1.5%	1.8%	1.8%	3.3%	3.2%	3.0%	2.8%
14	action, class, fraud, lawsuit, criminal, law, suit, filed, legal, court, firm, employee, justice, case, corporate, company, department, deception, case, federal	0.3%	0.2%	0.2%	0.2%	1.6%	1.9%	2.1%	1.7%	1.6%
15	diesel, gasoline, clean, petrol, engine, fuel, powered, technology, dirty, cleaner, truck, europe, emission, gas, efficiency, efficient, america, modern, market, pollution	0.8%	0.8%	0.4%	0.4%	2.4%	4.1%	3.7%	2.8%	2.4%
16	post, https, forum, thread, info, click, view, site, watch, link, subject, list, search, contact, video, website, edit, facebook, content, google	3.7%	4.0%	3.9%	3.5%	2.6%	1.4%	1.8%	1.9%	2.9%
17	german, germany, europe, industry, country, european, crisis, auto, euro, scandal, economy, largest, berlin, biggest, government, political, main, scale, market, minister	0.9%	0.7%	0.8%	1.2%	1.0%	1.7%	2.6%	2.5%	1.8%
18	service, dealer, warranty, dealership, customer, car, experience, excellent, extended, manager, happy, repair, local, deal, sales, product, problem, tell, street, fixed	2.4%	3.5%	3.1%	2.9%	1.9%	1.7%	1.5%	1.9%	2.8%
19	long, term, level, number, time, current, fact, high, potential, based, risk, case, short, position, example, note, process, simply, specific, large	5.5%	4.3%	4.9%	5.6%	5.1%	4.9%	5.4%	5.8%	6.4%
20	bug, bus, van, beetle, original, paint, color, type, love, black, part, project, body, window, blue, panel, super, white, late, red	4.5%	5.1%	4.7%	3.2%	1.7%	1.1%	0.9%	1.3%	2.0%
21	apple, source, phone, computer, tool, install, software, open, release, user, store, update, access, version, support, file, hardware, tech, upgrade, touch	1.2%	1.0%	1.3%	1.7%	0.9%	0.9%	0.8%	1.1%	1.4%
22	people, write, wrong, know, public, point, truth, happen, bad, trust, worse, question, fact, course, person, case, care, law, big, blame	2.1%	2.2%	1.9%	2.0%	4.8%	6.5%	7.0%	6.7%	5.6%
23	fuel, economy, gas, mpg, mileage, power, efficiency, efficient, highway, lower, engine, cost, torque, better, higher, average, low, price, diesel, increase	1.3%	1.8%	1.1%	1.6%	2.6%	3.0%	2.4%	2.5%	1.6%
24	online, order, buy, cheap, weight, canada, free, loss, australia, website, sale, purchase, usa, delivery, best, anti,	1.0%	0.8%	1.0%	0.6%	0.8%	0.1%	0.4%	0.6%	0.7%

	store, body, internet, product									
25	tdi, jetta, passat, golf, wagon, beetle, trade, sell, dealer, wife, buy, sedan, tsi, buy, mileage, car, sold, model, tiguan, sale	2.7%	1.8%	2.7%	2.9%	3.0%	2.8%	2.1%	2.4%	2.4%
26	car, drive, buy, owned, reliable, year, owner, months, good, price, insurance, fun, wife, garage, feel, test, buying, vehicle, time, purchase	6.8%	6.7%	7.2%	5.4%	4.2%	3.1%	3.9%	4.7%	5.0%
27	part, kit, oem, brake, rear, installed, interior, arm, fit, bolt, tire, plate, set, stock, spring, condition, black, clutch, original, wheel	5.5%	6.8%	5.8%	5.3%	2.0%	1.1%	1.2%	1.5%	2.1%
28	photo, image, sept, file, credit, wolfsburg, town, sunday, france, city, plant, left, factory, germany, week, day, people, war, built, west	0.4%	0.5%	0.2%	0.4%	0.4%	0.2%	0.3%	0.5%	0.4%
29	ford, honda, toyota, nissan, mercedes, mpg, mazda, bmw, benz, hyundai, volvo, sedan, subaru, hybrid, renault, premium, class, series, passenger, chrysler	2.5%	2.0%	2.0%	2.7%	1.1%	1.1%	0.9%	1.2%	1.1%
30	bosch, software, illegal, vag, group, production, engineer, component, arm, report, internal, letter, warning, know, fiat, media, testing, german, computer, paper	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%	0.2%	2.0%	0.4%
31	audi, porsche, skoda, seat, group, luxury, bmw, brand, vag, bentley, mercedes, owned, car, engine, quality, white, suspension, head, affect, automotive	2.4%	1.4%	2.1%	2.5%	1.7%	1.6%	1.3%	1.9%	1.3%
32	good, lot, time, bit, pretty, better, work, great, love, stuff, bad, big, thought, hard, nice, couple, guess, people, long, best	18.8%	20.1%	18.7%	17.9%	15.3%	12.8%	12.5%	12.7%	16.0%
33	city, family, trip, kid, town, summer, day, road, park, travel, parking, van, ride, room, event, live, bus, food, night, south	4.4%	4.8%	3.3%	2.3%	1.6%	1.3%	1.3%	2.0%	1.7%
34	safety, vehicle, system, recall, motor, report, include, auto, technology, traffic, highway, driver, automatic, benz, analysis, national, repair, security, toyota, industry	0.5%	0.5%	0.8%	0.8%	0.4%	0.1%	0.6%	0.8%	0.6%
35	team, racing, race, red, season, win, game, second, weekend, title, lead, final, round, renault, driver, sport, year, track, record, field	1.4%	1.5%	1.5%	1.5%	1.6%	0.6%	0.8%	0.9%	0.7%
36	posted, video, location, google, email, media, file, reuters, cut, photo, apple, delivery, united, latest, details, powered, record, view, beat, california	0.5%	0.0%	0.2%	0.2%	0.2%	0.2%	0.3%	0.2%	0.1%
37	china, percent, chinese, growth, global, data, market, rate, economic, demand, investors, year, bank, dollar, financial, economy, gold, week, fell, investment	0.4%	0.5%	0.4%	0.8%	0.7%	1.8%	2.0%	1.7%	1.6%
38	epa, air, clean, defeat, agency, software, device, recall,	0.1%	0.0%	0.0%	0.2%	8.4%	3.9%	2.7%	1.5%	1.2%

	protection, california, testing, diesel, violation, resource, carb, pollution, audi, pollutant, nitrogen, oxide									
39	code, fault, light, key, switch, ignition, address, shop, unit, number, sensor, turn, replaced, radio, steering, work, check, start, door, driver	1.6%	1.5%	2.1%	1.7%	0.8%	0.7%	0.7%	1.0%	1.3%
40	gti, golf, focus, subaru, mini, turbo, car, door, package, automatic, faster, manual, ford, consider, current, dsg, special, smaller, track, fun	2.3%	2.5%	2.8%	3.2%	1.3%	0.9%	0.6%	0.9%	0.9%
41	bmw, european, transport, road, test, report, limit, lab, diesel, europe, times, real, testing, council, auto, legal, mercedes, oxide, higher, standard	0.3%	0.9%	0.3%	0.6%	0.5%	1.3%	1.8%	0.8%	0.7%
42	fine, recall, epa, issue, wonder, pay, government, huge, owner, penalty, forced, billion, sold, doubt, happen, bring, big, vehicle, damage, usa	0.3%	0.0%	0.1%	0.1%	7.0%	7.7%	4.7%	2.9%	2.7%
43	frankfurt, concept, tiguan, suv, motor, bentley, design, version, renault, generation, production, hybrid, friendly, model, range, reveal, based, petrol, launch, interior	1.2%	1.4%	1.9%	3.7%	0.6%	0.3%	0.2%	0.2%	0.3%
44	test, testing, emission, software, pass, cheat, mode, epa, real, road, device, code, car, defeat, vehicle, equipment, normal, engine, control, computer	0.0%	0.0%	0.3%	0.0%	4.1%	5.0%	4.8%	3.9%	3.3%
45	man, police, tell, white, left, road, day, wife, black, morning, friend, turn, ask, called, feel, start, sign, head, house, blue	2.6%	3.7%	3.1%	2.3%	1.4%	1.6%	1.4%	1.6%	2.3%
46	scandal, company, billion, german, carmaker, winterkorn, diesel, worldwide, software, euro, ceo, automaker, monday, admitted, trust, tuesday, regulator, rigging, probe, germany	0.0%	0.1%	0.0%	0.0%	1.1%	5.9%	4.4%	4.1%	2.2%
47	rear, seat, wheel, steering, sport, speed, trim, design, interior, suspension, driver, package, cross, leather, standard, control, drive, roof, automatic, version	3.8%	3.8%	4.1%	4.2%	1.4%	0.9%	0.9%	1.0%	1.2%
48	sales, india, market, fiat, plant, automotive, brand, auto, year, chrysler, company, percent, unit, sell, america, north, production, month, american, worker	1.7%	1.1%	1.5%	2.5%	1.9%	1.7%	1.8%	1.3%	1.9%
49	news, reading, read, green, continue, originally, diesel, article, filed, scandal, comment, email, press, source, campaign, published, website, minute, marketing, sale	0.2%	0.3%	0.6%	0.6%	1.3%	1.7%	2.0%	2.0%	1.9%
50	oil, engine, pump, filter, belt, timing, water, valve, fuel, pressure, tank, intake, change, running, air, replace, check, block, plug, cold	5.7%	4.3%	5.4%	4.1%	2.3%	1.4%	1.4%	1.9%	2.7%

ESSAY 2

The Evolution of Online Reviews: A Dynamic Topic Model for Multiple Text Streams

Abstract

Although user-generated content (UGC) has become a valuable source of data for marketers to understand brand perceptions, customer satisfaction, and market structure, research on the dynamics of UGC has predominantly focused on structured data such as the volume and valence of posts. To examine the evolution of text-based UGC over time, we analyze a corpus of 15,727 time-stamped online reviews of a multi-generational high-tech product posted between 2012 and 2017. The corpus consists of multiple text streams, each containing the sequence of reviews for one generation of the product. We develop a continuous-state dynamic topic model that jointly models texts and ratings to accommodate both the serial dependence of topic prevalence and its relationship to ratings within each text stream. The model also accommodates the correlation between topic prevalence and its relationship to ratings across text streams. Our results show that earlier reviews are more homogeneous than later reviews with respect to the topics discussed. Moreover, earlier reviews are more clinical and feature-oriented, whereas later reviews are more emotional and experience-related. Our modeling approach enables marketers to identify which topics are top of mind for consumers and how topics are evaluated at different stages of the product lifecycle.

Keywords: User-generated content, Text analysis, Topic models, Product lifecycle

Introduction

Consumers have access to an enormous amount of user-generated content (UGC) on e-commerce websites and user forums that has substantial impact on their perceptions and behaviors. In a recent survey conducted by Nielsen, nearly two thirds of the respondents reported trusting consumer opinions posted online[8]. Online reviews affect consumers' decision-making by assisting them in identifying the products that match their personal needs and requirements (Chen and Xie 2008). The evolution of online reviews not only reflects shifting consumer preferences (Decker and Trusov 2010) but may also affect potential future customers. Much research has investigated the presence of dynamics in UGC for individual products, which may arise due to a number of factors such as the impact of previous ratings (e.g., Moe and Trusov 2011; Godes and Silva 2012) or changing contributor base (e.g., Li and Hitt 2008; Moe and Schweidel 2012). In contrast, research has not explored the dynamics of UGC in the context of multi-generational products. Examining such dynamics would enable managers understand the idiosyncrasies and commonalities of consumer perceptions of each product generation, facilitating decisions related to the marketing activities and the design of new products.

In addition to its focus on individual products, the literature on the dynamics of UGC predominantly investigates structured data such as the volume and valence of online opinions (e.g., Godes and Silva 2012; Moe and Schweidel 2012; Schweidel and Moe 2014; Xiong and Bharadwaj 2014; Ma et al. 2015; Borah and Tellis 2016; Fossen and Schweidel 2016). Research has shown that UGC contains information that is not captured by these metrics (e.g., Liu et al. 2016). While research has suggested that the posted ratings in reviews are expected to change over time, the shifts in metrics such as volume, valence and their variance offer little insight into

---

[8] *https://www.nielsen.com/content/dam/nielsenglobal/apac/docs/reports/2015/nielsen-global-trust-in-advertising-report-september-2015.pdf.*

how consumers' perceptions of the product are evolving. In the case of high-tech products, the textual content of earlier reviews is likely to differ from that of later reviews, as early adopters are opinion leaders in the diffusion process whose views may impact later adopters (e.g., Iyengar et al. 2011; Goldenberg et al. 2006; Rogers 2003). The failure to consider the dynamics in the content of online customer reviews can lead to dissatisfied consumers due to a mismatch between attribute information and customer preference (Chen and Xie 2008).

In this research, we contribute to the growing literature on social media dynamics by developing a modeling framework to investigate how the content of multiple text streams evolves over time. We demonstrate the proposed methodology using a dataset of 15,727 online reviews from Amazon.com of a multi-generational high-tech product from February 2012 to February 2017. This research investigates the dynamics in content both within and across product generations and tackles the following research questions: (1) How does the mixture of topics mentioned by consumers change over time and across generations? (2) How does the composition of content (which we refer to as "topic prevalence") evolve over time and across generations? (3) How does the relationship between content and ratings (which we refer to as "topic evaluation") shift over time and across generations?

We develop a modeling framework that builds upon supervised latent Dirichlet allocation (SLDA; Blei and McAuliffe 2007), a topic modeling framework that assumes documents are comprised of latent topics and words are associated with different topics, and that ratings are associated with these topics. In contrast to topic models that have appeared in the marketing literature and have been applied to a single corpus (e.g., Tirunillai and Tellis 2014; Büschken and Allenby 2016), our modeling framework accounts for each review being associated with a particular generation of the product and posted at a particular time. Thus, our corpus consists of multiple text streams that evolve in parallel and may be related to each other. Given the structure of the data, our modeling framework not only considers the serial dependence of topic prevalence and topic evaluation within each text stream, but also allows for the correlation across text streams. We find that the content of earlier reviews systematically differs from that of later reviews. Specifically, both within and across generations, earlier reviews are more similar with regards to the mixture of topics, whereas later reviews are more diverse. Moreover, earlier reviews are more clinical and feature-oriented, whereas later reviews are more emotional and experience-related. We also find that the evaluation of feature-related topics decreases more over time than the evaluation of experience-related topics.

The contribution of this research is twofold. First, it adds to the marketing literature that leverages unstructured data to understand customer preferences (e.g., Lee and Bradlow 2011; Netzer et al. 2012) by developing a continuous-state dynamic topic model that captures the evolution of textual content across multiple text streams. Beyond the application to customer reviews of multi-generational products, the model can also be used to capture the shifts in textual content surrounding a brand across multiple products or social media platforms (e.g., Schweidel and Moe 2014). Second, we contribute to the literature on social media dynamics that has primarily focused on metrics such as volume and valence (e.g., Godes and Silva 2012; Moe and Schweidel 2012; Borah and Tellis 2016). By extending this literature to examine textual dynamics, we gain insights into how the prevalence and the evaluation of each topic in online reviews change over time. In combining the text of reviews with ordinal ratings (e.g., Büschken and Allenby 2016), we provide marketing managers with insight into the underlying drivers of shifts in ratings and how they may change over time.

The remainder of this manuscript proceeds as follows. We review the literature on customer review dynamics and text analysis in the next section. We then introduce the dataset used in our empirical analysis and develop our modeling framework. Lastly, we present our empirical findings and conclude with a discussion on limitations and future research.

Related Literature

Dynamics in Online Reviews

Research has recognized the impact of online reviews on product sales (Chevalier and Mayzlin 2006; Li and Hitt 2008; Duan et al. 2009; Moe and Trusov 2011), box office revenue (Liu 2006; Chintagunta et al. 2010), a firm's pricing decision (Kuksov and Xie 2010), quality perception (Sun 2012), and customer purchase behavior (Zhao et al. 2013).

As customer reviews are contributed over time, their content or valence indicated by ratings may demonstrate temporal patterns. Although online reviews record customers' personal experiences with the product, prior research has shown that the evolution of review ratings is also subject to the impact of earlier ratings owing to a bandwagon effect or differentiation behavior (e.g., Schlosser 2005; Moe and Trusov 2011; Lee et al. 2015). Moreover, research has found that the volume of online word-of-mouth or previous reviews is a driver of the dynamics in review ratings (e.g., Godes and Mayzlin 2004; Godes and Silva 2012). In addition to the previous reviews that have been contributed, researchers have identified shifts in the posting population as another explanation for the temporal variation in product reviews (e.g., Li and Hitt 2008; Moe and Schweidel 2012). While there continues to be interest in understanding the drivers of dynamics in online reviews, extant literature focuses primarily on metrics such as the volume and ratings, consequently ignoring the textual content of what is discussed and of concern to review contributors over time. In this research, we seek to examine the dynamics present in UGC by means of time-stamped text, which provides deeper insights for both researchers and practitioners.

Text Analysis of User-generated Content

Text analysis has emerged as an important tool in the marketing literature in recent years. Some work converts unstructured data to numeric measures by decomposing text into individual words or phrases and then clustering them by their similarity or concurrence (e.g., Archak et al. 2011; Lee and Bradlow 2011; Netzer et al. 2012). More recently, marketing researchers have begun to apply and extend latent Dirichlet allocation (LDA; Blei et al. 2003), a hierarchical topic modeling framework, to user-generated textual content. Tirunillai and Tellis (2014) build upon the LDA framework by simultaneously identifying the unobserved dimensions of brands that are discussed in product reviews and the valence associated with each dimension. While they examine how the frequency with which brands are associated with specific dimensions shifts over time in a post hoc analysis, the authors assume that the parameters governing the prevalence of topics do not change over time. Büschken and Allenby (2016) develop a sentence-constrained LDA assuming that all words within each sentence consist of just one topic. Applying their model to a corpus of user-generated reviews of hotels, they show that their sentence-constrained LDA provides a superior model fit in prediction over the standard LDA model. In addition, they associate review ratings with text via the proposal distribution of each review's topic proportions. Puranam et al. (2017) use a posterior analysis of LDA to investigate how the textual content of online customer reviews of restaurants changes before and after a known policy change mandating all restaurants to post calorie information on their menus.

Although marketing researchers have examined temporal changes of topic proportions by means of LDA-based models, such approaches are limited in their ability to model the inherent dynamics underlying the text of UGC. Specifically, nearly all marketing research related to LDA assumes that the parameters governing the data-generating process are time-invariant (Tirunillai and Tellis 2014; Büschken and Allenby 2016) and do not consider how those parameters may shift over time. To the best of our knowledge, our research is among the first in marketing to incorporate dynamics into the underlying process that governs the textual content of documents and explore how this may impact product ratings. While prior research on LDA has studied the context of multiple text streams, research has either assumed the underlying parameters that govern the topic prevalence to be time-invariant (e.g., Paul and Girju 2009) or omits key ways in which the topics across different text streams may be related to each other (e.g., Hong et al. 2011). We build upon the spirit of Blei and Lafferty's (2006, 2007) topic models by allowing topics to be correlated across streams and for the serial dependence of topics within each stream.

Data

To examine how the composition and evaluation of content in customer reviews shift over time, we study the publicly available online reviews for the 1st, 2nd, and 3rd generations of the Nest learning thermostat. We collect the raw text of customer reviews from Amazon.com ranging from February 2012 to February 2017.[9]

Referring to the natural language processing (NLP) literature, we define the following commonly-used terms before introducing the steps of text preprocessing. A *word* is a distinct set of English characters in the vocabulary generated from a collection of documents. A *token* is a

[9] The 1st generation of the Nest learning thermostat was released on October 25, 2011, but the earliest available customer reviews on Amazon is on February 22, 2012. Despite this, since the 2nd generation and 3rd generations were released on October 2, 2012 and September 1, 2015, respectively, our dataset covers the duration since their introductions.

chopped instance of a character sequence in documents, which can be a word or acronym. A *document*, or a customer review in our case, is a meaningful sequence of tokens that may be composed of repeated words. A *corpus*, or a text collection, is the set of all documents under analysis. As an illustration, consider the following customer review on the Nest thermostat:

> "This thermostat is beautiful on my wall, easy to use and saves energy on my heating bill and I can't wait to see the savings on my cooling bill. I can control it from anywhere, via my iPhone or browser. It is the best programmable thermostat out there. I would definitely buy it again for my next home."

The above document consists of four sentences. The word "thermostat" appears twice in the document: the second token in the first sentence and the sixth token in the third sentence.

We preprocess all text through several steps. First, strings of dollar sign (e.g., "$", "$$$", etc.) are replaced by abbreviation "USD" because review contributors often use the dollar sign symbol to refer to price, cost, or money. Second, other punctuations (e.g., "/", "@", "|", etc.), numbers, stop words (e.g., "the", "a", "and", etc.), and URL links are removed from the text. Third, to minimize redundancy, we follow the Porter Stemmer method (Porter 2006) and only take the root form of each word (e.g., "thermostats" stems from "thermostat", and thus all occasions of "thermostats" and "thermostat" are treated as the same). Moreover, we replace informal abbreviations with their formal forms (e.g., "gen" is an informal abbreviation of "generation", so we replace "gen" with "generation"). Last, we use sparsity ≥ 0.995 to cut off infrequent words in the corpus, which is a widespread practice in NLP. By doing so, we focus on the unique words that appear in at least 0.5% documents of the corpus. This results in 15,727 reviews and 792 unique words in the final corpus. Our corpus is comprised of three text streams, each of which holds the sequence of time-stamped reviews for one generation of the Nest thermostat.

Figure 1 depicts the trend of review volumes by months after preprocessing.

[Insert Figure 1 Here]

The volume of customer reviews of the later generations of the product is considerably higher than the volume of reviews of the earlier generations, which is consistent with the growth of the smart thermostat market.[10] Moreover, we observe increases in review volume around the holiday season (i.e., November, December, and January) for each generation. Figure 2 shows the average number of tokens per review by month after preprocessing.

[Insert Figure 2 Here]

According to Figure 2, the length of customer reviews, measured by the number of tokens, decreases both within a generation and across generations. This may suggest a potential shrinkage of contributed content per review over time. On average, the first months after the release of both the 2nd and 3rd generations have the longest reviews within their respective generations, indicating that earlier buyers tend to comment more than customers who make later purchases. This is consistent with diffusion of innovation theory which suggests that early adopters are more willing to embrace new technologies than late adopters and act as opinion leaders (Rogers 2003). Figure 3 shows the average rating of the reviews by month.

[Insert Figure 3 Here]

The average overall rating of each generation gradually decreases over time, which is consistent with extant literature (Li and Hitt 2008; Godes and Silva 2009), especially after the release of the next generation. While Figures 2 and 3 provide model-free evidence for potential changes in the content and evaluation of customer reviews over the product lifecycle, it does not shed light on how the content of reviews may be changing over time. We therefore develop a continuous-state

---

[10] See *http://www.grandviewresearch.com/industry-analysis/smart-thermostat-market*.

dynamic topic model that captures the evolution of content and evaluation.

Model Development

Gibbs Sampler for Supervised LDA (SLDA)

LDA is a generative statistical model in natural language processing (NLP) that explains a corpus by a set of unobserved topics (Blei et al. 2003). It allows each document in the corpus to be described by a finite mixture of these topics and each word in the corpus to be associated with each topic with different frequencies. In the case of online reviews, however, authors post both a textual review and an ordinal rating. Blei and McAuliffe (2008) propose a process that jointly generates unstructured text and the structured response variable; namely, supervised latent Dirichlet allocation (or SLDA), which directly exploits the empirical topic proportions calculated by topic assignments of all tokens in a document as the variables to predict the document-level response variable. The generative process of SLDA can be described as:

1. For topic *j*, draw word distribution $\phi^j \sim Dir(\beta)$
2. For document *d*, draw topic proportion $\theta^d \sim Dir(\alpha)$
3. For each token *i* in document *d*
   (a) Draw topic assignment $z_{di} \sim Mult(\theta^d)$
   (b) Draw word $w_{di}|z_{di} \sim Mult(\phi^{z_i})$
4. Draw document *d*'s response variable $y_d|\boldsymbol{z}_{1:N^d}, \boldsymbol{\gamma}, \sigma^2 \sim N(\overline{\boldsymbol{z}}_d\boldsymbol{\gamma}, \sigma^2)$

where N^d is the number of tokens in document *d*, $\overline{\boldsymbol{z}}_d = \frac{1}{N^d}\sum_{i=1}^{N^d} z_{di}$ is the vector of topic proportions computed by averaging topic assignment of each token in document *d*, $\boldsymbol{\gamma}$ is the vector of parameters corresponding to each topic's proportion, and σ^2 is the variance parameter of response variable y_d. Since the predictors in the response model are empirical topic proportions computed by the fragment of tokens assigned to each topic, there are components of $\bar{\boldsymbol{z}}$ always sum to one. Therefore, the original intercept term of the response model is dropped for identification. The only difference between LDA and SLDA is Step 4 of the generative process (Blei et al. 2003; Griffiths and Steyvers 2004). This framework is particularly flexible in marketing research because the response variable can be continuous (e.g., durations), count (e.g., sales, website visits), categorical (e.g., brand choices), or ordinal data (e.g., ratings, sales ranks).

Conditional on the draws of topic assignment $\boldsymbol{z}$, the observation of word $\boldsymbol{w}$ and response $\boldsymbol{y}$ are independent with each other. Given topic assignment $\boldsymbol{z}$, Dirichlet parameter $\boldsymbol{\alpha}$ can be drawn from a Pólya distribution (Minka 2000), while topic proportion coefficient $\boldsymbol{\gamma}$ can be drawn from an ordinal probit model in the case of ordinal response $\boldsymbol{y}$ (Greene 2012). The joint distribution of words and topics, $P(\boldsymbol{w}, \boldsymbol{z}|\boldsymbol{\alpha}, \boldsymbol{\beta}) = P(\boldsymbol{w}|\boldsymbol{z}, \boldsymbol{\beta})P(\boldsymbol{z}|\boldsymbol{\alpha})$, can be computed by integrating out θ and ϕ such that

$$P(\boldsymbol{w}|\boldsymbol{\beta}, \boldsymbol{z}) = \left(\frac{\Gamma(W\beta)}{\Gamma(\beta)^W}\right)^J \prod_{j=1}^{J} \frac{\prod_w \Gamma\left(N_j^w + \beta\right)}{\Gamma\left(N_j + W\beta\right)} \tag{12}$$

$$P(\boldsymbol{z}|\boldsymbol{\alpha}) = \left(\frac{\Gamma(J\alpha)}{\Gamma(\alpha)^J}\right)^D \prod_{d=1}^{D} \frac{\prod_j \Gamma\left(N_j^d + \alpha\right)}{\Gamma(N^d + J\alpha)} \tag{13}$$

where N_j^w records the number of times word w is assigned to topic j in the whole corpus, N_j records the number of tokens assigned to topic j in the whole corpus, W is the size of vocabulary generated from the whole corpus, N_j^d records the number of tokens assigned to topic j in document d, N^d records the total number of tokens in document d, J is the number of predetermined topics, D is the number of documents in the whole corpus, and $\Gamma(\cdot)$ is a gamma function. The ordinal probit can be expressed as

$$v_d = \bar{\boldsymbol{z}}_d \boldsymbol{\gamma} + \zeta_d \tag{14}$$

$$y_d = \begin{cases} 1 & \text{if } v_d \leq \tau_1 \\ 2 & \text{if } \tau_1 < v_d \leq \tau_2 \\ \vdots & \\ R & \text{if } \tau_{R-1} < v_d \end{cases} \tag{15}$$

where v_d is the unobserved continuous variable underlying the observed ordinal response y_d for document *d*, $\zeta_d \sim N(0, \sigma^2)$, 1 is the lowest rating, and *R* is the highest rating. Note that for identification, we constrain first threshold $\tau_1 = 0$ and the variance $\sigma^2 = 1$ for the unobserved variable $\boldsymbol{v}$ (e.g., Greene 2012). Thus, the joint distribution of word $\boldsymbol{w}$, topic assignment $\boldsymbol{z}$, ordinal response $\boldsymbol{y}$, and unobserved continuous rating $\boldsymbol{v}$ can be represented by

$$P(\boldsymbol{w}, \boldsymbol{z}, \boldsymbol{y}, \boldsymbol{v} | \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\tau}) = P(\boldsymbol{w} | \boldsymbol{z}, \boldsymbol{\beta}) P(\boldsymbol{y} | \boldsymbol{v}, \boldsymbol{\tau}) P(\boldsymbol{v} | \boldsymbol{\gamma}, \boldsymbol{z}) P(\boldsymbol{z} | \boldsymbol{\alpha}) \tag{16}$$

The conditional distribution of token *i* in document *d* being assigned to topic *j* is represented by

$$p(z_{di} = j | \boldsymbol{w}, \boldsymbol{z}_{-di}, \boldsymbol{y}, \boldsymbol{v}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \propto \frac{N_{-di,j}^{w} + \beta}{N_{-di,j} + W\beta} \left(N_{-d\ ,j}^{d} + \alpha\right) \exp\left(-\frac{\left(v_d - \bar{\boldsymbol{z}}_{d, z_{di}=j} \boldsymbol{\gamma}\right)^2}{2}\right) \tag{17}$$

where $N_{-di,j}^{w}$ records the number of times word *w* is assigned to topic *j* excluding token *i* in document *d*, $N_{-di,j}$ records the number of times any token is assigned to topic *j* excluding token *i* in document *d*, $N_{-di,j}^{d}$ records the number of times topic *j* is assigned to all tokens in document *d* excluding token *i* in document *d*. Moreover, $\bar{\boldsymbol{z}}_{d, z_{di}=j} = \frac{1}{N^d} \left(\sum_{i'=1, i' \neq i}^{N^d} z_{di'} + z_{di}^{j} \right)$, in which z_{di} , $i' \neq i$, is the topic assignment of any token in document *d* excluding token *i*, and z_{di}^{j} is the tentative topic assignment of token *i* in document *d* to topic *j* (see Appendix A.1). Particularly, if response variable $\boldsymbol{y}$ were continuous, v_d in Eqn. (6) would be replaced by y_d. Estimates of word distribution by topic, $\hat{\phi}_j^{w}$, and topic distribution by document, $\hat{\theta}_j^{d}$, can be calculated as

$$\hat{\phi}_j^w = \frac{N_j^w + \beta}{N_j + W\beta} \tag{18}$$

$$\hat{\theta}_j^d = \frac{N_j^d + \alpha}{N^d + J\alpha} \tag{19}$$

During sampling, $\boldsymbol{v}, \boldsymbol{\gamma}$, and $\boldsymbol{\tau}$ are updated after each iteration given current topic assignments $\boldsymbol{z}$ and the observed ordinal response $\boldsymbol{y}$ via Gibbs sampler with data augmentation (Albert and Chib 1993). Without the rejection step, this pure Gibbs sampler for SLDA is potentially faster to converge than Metropolis-Hastings algorithm for similar models (Büschken and Allenby 2016). The collapsed Gibbs sampler in Eqn. (6) is part of the Markov Chain Monte Carlo (MCMC) algorithm for dynamic SLDA with multiple text streams, which we develop in the next section.

MCMC for Dynamic Supervised LDA with Multiple Text Streams

We assume that the prevalence of topics in UGC such as online reviews exhibit temporal variation. For multi-generational products, the prevalence of topics on which reviewers comment both within and across generations may shift over time. Therefore, we regard all the reviews of each generation as a sequence of a text stream so that the topic prevalence of each stream is governed by its own parameters, but the set of topics is common across all streams. In doing so, we extend SLDA to accommodate topic dynamics across multiple text streams. Specifically, we allow for the prevalence of each topic to evolve within a given stream and for there to be a correlation in the prevalence of topics across streams. For parsimony, we parameterize the Dirichlet hyperprior for topic *j* at time *t* across streams by a multivariate random walk such that

$$\boldsymbol{x}_{j,t} = \boldsymbol{x}_{j,t-1} + \boldsymbol{\xi}_{j,t} \tag{20}$$

$$\boldsymbol{\alpha}_{j,t} = \exp(\boldsymbol{x}_{j,t}) \tag{21}$$

where $\boldsymbol{x}_{j,t} = \left(x_{j,t}^1, \dots, x_{j,t}^K\right)'$ is a vector of parameters for text stream $k = 1, 2, \dots, K$ and

$\boldsymbol{\xi}_{j,t} \sim MVN(\mathbf{0}_{K\times 1}, \boldsymbol{\Omega}_j)$ is a topic-specific random disturbance term, $\boldsymbol{\alpha}_{j,t} = (\alpha_{j,t}^1, \dots, \alpha_{j,t}^K)'$ is the vector of Dirichlet hyperprior for topic proportions of documents in text stream 1, 2, …, *K*, and the exponential transformation is used to guarantee the positivity of Dirichlet hyperprior. As the first moment of Dirichlet distribution, the ratio $\alpha_{j,t}^k / \sum_{j=1}^{J} \alpha_{j,t}^k$ is defined as the prevalence of topic *j* in stream *k* at time *t*. Given the similarity of mathematical forms between topic prevalence and market share models (Cooper et al. 1988, pp. 26-27), we borrow the concept of brand attraction to define the value of $\alpha_{j,t}^k$ as the attraction of topic *j* in stream *k* at time *t*.

Similar to the Dirichlet hyperprior for topic proportions, the coefficient for any topic *j*'s proportion at time *t* in the response model is assumed to follow a multivariate random walk process such that

$$\boldsymbol{\gamma}_{j,t} = \boldsymbol{\gamma}_{j,t-1} + \boldsymbol{\varepsilon}_{j,t} \tag{22}$$

where $\boldsymbol{\gamma}_{j,t} = (\gamma_{j,t}^1, \dots, \gamma_{j,t}^K)'$ is a vector of parameters for text stream $k = 1, 2, \dots, K$ and $\boldsymbol{\varepsilon}_{j,t} \sim MVN(\mathbf{0}_{K\times 1}, \boldsymbol{\Sigma}_j)$ is a topic-specific random disturbance term. In this setting, the evaluation of any topic *j* is serially dependent within the same text stream, and the evaluation of the same topic are correlated across all text streams. The generative process of the model is described as

1. For topic *j*, draw a word distribution $\phi^j \sim Dir(\beta)$;
2. For document *d* in text stream *k* at time *t*, draw a vector of topic proportions $\theta^{kd} \sim Dir(\boldsymbol{\alpha}_t^k)$, where $\boldsymbol{\alpha}_t^k = (\alpha_{1,t}^k, \dots, \alpha_{J,t}^k)$ is the vector of Dirichlet hyperprior for topic proportions of documents in text stream *k* at time *t*. The *j*-th dimensions of the Dirichlet parameters $\boldsymbol{\alpha}_t^k$ across all text streams *k* = 1, 2, …, *K* follow Eqn. (9) and Eqn. (10);
3. For each token *i* in document *d* of text stream *k* at time *t*:

   (a) Draw a topic assignment $z_{di} \sim Mult(\theta^{kd})$;

(b) Draw word $w_{di}|z_{di} \sim Mult(\phi^{z_{di}})$;

4. Draw document *d*'s response variable $y_d|z_{d,1:N^d}, \boldsymbol{\gamma}_t^k, \sigma^2 \sim N(\bar{\boldsymbol{z}}_d \boldsymbol{\gamma}_t^k, \sigma^2)$, where $\boldsymbol{\gamma}_t^k = (\gamma_{1,t}^k, \dots, \gamma_{J,t}^k)'$ is the vector of topic evaluations for documents in text stream *k* at time *t*. The *j*-th dimensions of $\boldsymbol{\gamma}_t^k$ across all text streams *k* = 1, 2, …, *K* follow Eqn. (11).

We take the Dirichlet symmetric hyperprior $\boldsymbol{\beta}$ for the word distribution as predetermined (Griffiths and Stevyers 2004), but we estimate the Dirichlet asymmetric hyperprior $\boldsymbol{\alpha}_t = (\alpha_{1,t}, \dots, \alpha_{J,t})$ for topic proportions of documents present at time *t*. The joint distribution of word $\boldsymbol{w}$, topic assignment $\boldsymbol{z}$, ordinal response $\boldsymbol{y}$, and unobserved variable $\boldsymbol{v}$ can be represented by:

$$P(\boldsymbol{w}, \boldsymbol{z}, \boldsymbol{y}, \boldsymbol{v} | \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \\ = P(\boldsymbol{w}|\boldsymbol{z}, \boldsymbol{\beta}) \prod_{t=1}^{T} \prod_{k=1}^{K} \prod_{d \in C_t^k} P(y_d | v_d, \boldsymbol{\tau}) P(v_d | \boldsymbol{\gamma}_t^k, \boldsymbol{z}_d) P(\boldsymbol{z}_d | \boldsymbol{\alpha}_t^k) \tag{23}$$

where C_t^k denotes the corpus of text stream *k* at time *t* and $\boldsymbol{z}_d$ denotes the topic assignments of tokens in document *d*. We apply the following Metropolis-Hasting (MH) sampling scheme to draw topic assignments $\mathbf{z}$, the Dirichlet hyperprior for topic proportions $\boldsymbol{\alpha}$, the covariance matrix for topic proportions $\boldsymbol{\Omega}$, rating thresholds $\boldsymbol{\tau}$, unobserved variables $\boldsymbol{v}$, topic evaluations $\boldsymbol{\gamma}$, and the covariance matrix for topic evaluations $\boldsymbol{\Sigma}$ in turn (see Appendix A.2 for details):

1. $\boldsymbol{z} \mid \boldsymbol{w}, \boldsymbol{\alpha}, \boldsymbol{v}, \boldsymbol{\gamma}$;
2. $\boldsymbol{\alpha} \mid \boldsymbol{z}, \boldsymbol{\Omega}$;
3. $\boldsymbol{\Omega} \mid \boldsymbol{\alpha}$;
4. $\boldsymbol{\tau} \mid \boldsymbol{y}, \boldsymbol{v}$;
5. $\boldsymbol{v} \mid \boldsymbol{y}, \boldsymbol{\gamma}, \boldsymbol{\tau}$;
6. $\boldsymbol{\gamma} \mid \boldsymbol{z}, \boldsymbol{v}, \boldsymbol{\Sigma}$;

7. $\boldsymbol{\Sigma} \mid \boldsymbol{\gamma}$.

In our empirical study, the three text streams begin at different times because the product generations were released sequentially. Rather than assuming independent random walks for each generation, we simplify the estimation procedure by assuming the existence of reviews at each time period in the text streams for the 2nd and the 3rd generations of product before they were actually released (See Figure 4). The parameters of these reviews (i.e., $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$) evolve following the same process as actual reviews, but do not enter the likelihood of the model.

Results

Model Fit

We set the Dirichlet hyperprior for word distributions, $\beta = 0.5$, to control the granularity and the number of the topics learned by LDA-based models (Griffiths and Steyvers 2004). In the following analysis, we set the number of topics $J = 20$, which explains the majority of variance in the rating model by taking the elbow in the curve of log marginal density (LMD).[11] We take dispersed starting points to run multiple chains and record the estimates that yield the highest marginal likelihood (e.g., Blei et al. 2010).

To validate the proposed model, we consider the task of predicting the text of reviews in the next month. We compare the predictive ability of the proposed model, i.e. the dynamic SLDA model with correlated multiple text streams, with (1) the static SLDA model with single text stream, (2) the dynamic SLDA model with single text stream, and (3) the static SLDA model with multiple text streams with 20 topics. For the single text stream models, we treat the online reviews of all the three generations as one text stream. We calibrate the model using data from

[11] See Appendix A.3 for the in-sample LMD, which contains a varying number of topics from 5 to 40 in increments of 5.

February 2012 through April 2016 to predict the text of the reviews contributed in May 2016. We then use the reviews contributed through May 2016 to predict the review text in June 2016. We continue this process through February 2017. For the dynamic models, the estimates of parameters in the last month of a calibration period are used to predict the text in the next month. For the static models, the estimates of parameters obtained from all previous months are used to predict the text in the next month. All the models are estimated with the same granularity, $\beta = 0.5$. We assess the fit of the holdout sample based on the LMD, which takes into account both text and ratings, and perplexity, which is a commonly-used measure of performance for LDA. Perplexity is equivalent to the inverse of the geometric mean of per-word likelihood; thus a lower perplexity score indicates a better generalization performance (Blei et al. 2003). Table 1 shows both the LMD and perplexity scores of the four models.

[Insert Table 1 Here]

The dynamic supervised LDA with multiple text streams outperforms the rest three models in predicting the next month's review text. This comparison result suggests that the prevalence of topics exhibits temporal patterns, therefore the results discussed hereafter are obtained from the proposed dynamic model based on the data ranging from February 2012 to February 2017.

The Mixture of Topics in Reviews

The concentration parameter of a *j*-dimensional Dirichlet prior at time *t* is defined as the summation of topic attractions such that $\sum_{j=1}^{J} \alpha_{j,t}$ (Wallach et al. 2009). A Dirichlet prior with a larger value for the concentration parameter generates more similar multinomial distributions as vectors of topic proportions of each document. By contrast, a smaller value of concentration parameter generates more diverse multinomial distributions as vectors of topic proportions of each document. Intuitively, the higher the value of concentration parameter, the more similar the reviews are to each other with respect to the mixture of topics. In contrast, the lower the value of concentration parameter, the more diverse the reviews are from each other with regards to the mixture of topics. Figure 5 depicts the concentration parameter for each stream over time.

[Insert Figure 5]

Based on the decreasing trend of concentration parameters over generations, the earlier review contributors hold more homogeneous content, whereas later review contributors hold more heterogeneous content. Together with the model-free evidence of a decreasing trend of average review length over generations, as shown in Figure 2, this finding is consistent with existing literature on the diffusion of innovations suggesting that early adopters come from a smaller group of people who are usually the opinion leaders that bridge the gap between abstract product descriptions and practical use, whereas later adopters come from a larger group of people who are mostly opinion followers in a social network (Rogers 2003; Iyengar et al. 2011).

Review Content by Topics

Table 2 presents all 20 topics estimated by the dynamic model. To better describe the content of topics, we apply the relevance metric that balances the word distribution of topics with the word frequency in the text collection using weight $\lambda = 0.6$ (Sievert and Shirley 2014) and present the 30 most relevant words for each topic.

[Insert Table 2 Here]

To facilitate the discussion of our results, we re-index all the topics in Table 2 and classify them into seven categories:

- “Features” (i.e., Topic 1 “Display”, Topic 2 “Eco”, Topic 3 “Interface”, Topic 4 “Learning”, Topic 5 “Schedule”, and Topic 6 “System”) that describe the functional attributes of the Nest thermostat;

- "Setup" (i.e., Topic 7 "Installation", Topic 8 "Instruction", and Topic 9 "Power") that focuses on installation and getting started with the product;
- "Praise" (i.e., Topic 10 "Compliment") that compliments the product;
- "Experience" (i.e., Topic 11 "Ease of Use" and Topic 12 "User Experience") that talks about the experiential attributes of the product;
- "Criticism" (i.e., Topic 13 "Compatibility", Topic 14 "Connection", Topic 15 "Temperature Feels", and Topic 16 "Unit Failure") that focuses on complaints and issues when using the product;
- "Financial" (i.e., Topic 17 "Price" and Topic 18 "Saving") that discusses cost-effectiveness of the product; and
- "Product Line" (i.e., Topic 19 "Integration" and Topic 20 "Repeat Purchase") that talks about product's integration with other products and upgrading to a newer generation.

The last eight rows of Table 2 show the average of topic prevalence and topic evaluation of each generation over time. These two metrics demonstrate the evolution of content over generations.

The category "Features" includes Topic 1 "Display" which comments on the display system of the Nest thermostat, Topic 2 "Eco" which presents the energy saving function, Topic 3 "Interface" which comments on the design of the user interface used to control the product, Topic 4 "Learning" which describes how the product builds a schedule to adjust the temperature by learning users' behavior[12], Topic 5 "Schedule" which mainly describes the auto-schedule[13] and auto-away[14] functions, and Topic 6 "System" which discusses the HVAC system that works with the thermostat. According to the average prevalence of topics in the "Features" category

[12] *https://nest.com/support/article/How-to-maximize-energy-savings-with-your-Nest-Learning-Thermostat*.
[13] *https://nest.com/support/article/How-does-Auto-Schedule-learn*.
[14] *https://nest.com/blog/2011/11/18/what-is-auto-away/*.

shown in Table 2, the review contributors of the earlier generations of the product are more likely to comment on the functional attributes than those of the later generations. This finding is consistent with the characteristics of early adopters in the diffusion process of innovations who are more willing to take the risk of experimenting with new technologies than later adopters (Rogers 2003). One exception to this pattern is Topic 1 "Display" whose prevalence for the 3rd generation (2.6%) is higher than previous generations (1.4% and 1.6%). This finding corresponds with the improvements made to the display for the 3rd generation, which includes enlarging the screen size by 40% and increasing the resolution to make the display more readable from a distance.[15] The evaluation of Topic 1 "Display" also reveals improved satisfaction with the display design in the 3rd generation (0.87) compared to the 1st (0.50) and the 2nd generations (0.25). The second exception is Topic 5 "System", caused when Nest increased the number of signal connectors on its thermostat in the 2nd generation, enabling the thermostat to work with additional types of HVAC systems, which results in more feedback on this feature in the 2nd generation reviews.[16] To elaborate on how the topic prevalence evolves over time, we plot the temporal trend of the prevalence of topics in the "Features" category in Figure 6.

[Insert Figure 6 Here]

Figure 6a reveals that the improved design of the display for the 3rd generation of the product stimulates a high prevalence of Topic 1 "Display" at the beginning of the 3rd generation's release. The effect lasted about 6-12 months and even brought up the prevalence of the same topic in reviews of the earlier two generations. The prevalence of Topic 2 "Eco" (Figure 6b), Topic 3 "Interface" (Figure 6c), and Topic 4 "Learning" (Figure 6d) decreases within each generation. Comparatively, the discussion about Topic 4 "Schedule" (Figure 6e) does not change much over

[15] *https://nest.com/uk/support/article/What-s-new-in-the-3rd-generation-Nest-Learning-Thermostat*.
[16] *https://nest.com/support/pro/article/Help-with-installation-and-set-up*.

time, suggesting that once novel features such as auto-schedule and auto-away may now be seen as core functions by consumers. Figure 6f shows that the prevalence of Topic 5 “System” on the 1st generation stays relatively stable, while the prevalence of Topic 5 on the 2nd and 3rd generations declines. This is because the 2nd and the 3rd generations are compatible with 95% heating and cooling systems in use as compared to the 1st generation’s compatibility of 75% (Hildenbrand 2017). Reviews of the 1st generation continue to talk about this topic owing to its poor compatibility. The high prevalence of this topic at the beginning of the 2nd generation also confirms this improvement to some extent.

Figure 7 exhibits the trend of the evaluation of topics in the “Features” category.

[Insert Figure 7 Here]

In general, the evaluation of most feature-related topics declines over time within each generation of product (Figure 7b). A striking exception is Topic 1 “Display” (Figure 7a) whose evaluation on the 1st and 2nd generations remains relatively stable, while the evaluation on the 3rd generation increases, coinciding with the inclusion of a larger screen with higher resolution.

The prevalence of topics in the “Experience” category displays an opposite trend to the “Features” category. Specifically, the prevalence of both Topic 11 “Ease of Use” and Topic 12 “User Experience” increases from one generation to the next. As shown in Table 2, the experience-related topics in the 1st generation account for only 9.8% of review content (4.6% for Topic 11 and 5.2% for Topic 12) but then account for up to 33.8% of review content in the 3rd generation of the product (19.7% for Topic 11 and 14.1% for Topic 12). Compared with earlier generations’ review contributors who are excited about the product features and enthusiastic about explaining the installation of the product in detail, the review contributors of later generations are inclined to share their personal stories describing their use of the product and highlight the ease of using the product. Figure 8 shows the temporal pattern of topic prevalence in the "Experience" category, while Figure 9 shows the trend of the topic evaluation in the "Experience" category over time.

[Insert Figure 8 Here]

[Insert Figure 9 Here]

Figure 8 shows that the prevalence of both Topic 11 "Ease of Use" and Topic 12 "User Experience" gradually increases within each generation, except for Topic 12 for the 1st generation. The contrast between the "Feature" category and the "Experience" category complies with the previous literature, which claims that early and late adopters are from different customer bases with dissimilar preferences and tastes (Li and Hitt 2008).

The "Product Line" category contains Topic 19 "Integration", which talks about the integration of the Nest thermostat with Nest's other products such as Nest Protect and Nest Cam, and Topic 20 "Repeat Purchase", which talks about repeat purchases of either current or newer generation of Nest thermostat. Although the prevalence of Topic 20 "Repeat Purchase" remains relatively constant over generations, the prevalence of Topic 19 "Integration" is higher in later generations. Nest announced the 1st generation of their second product, the Nest Protect smoke and carbon monoxide detector, in October 2013 and then released a second generation in June 2015. Later, Nest released an indoor security camera, the Nest Cam, in June 2015, and an outdoor security camera, the Nest Cam Outdoor, in July 2016.[17] These products work together by communicating through Wi-Fi to prevent danger from spreading.[18] To investigate the effects of the introduction of new products on Nest thermostat, we plot the temporal trend of prevalence of Topic 19 "Integration" and Topic 20 "Repeat Purchase" in Figure 10.

[17] *https://nest.com/blog/*.
[18] *https://nest.com/smoke-co-alarm/meet-nest-protect/*.

[Insert Figure 10 Here]

The prevalence of Topic 19 "Integration" (Figure 10a) on the 2^{nd} generation of the Nest thermostat first began to rise in October 2013, when the first generation of Nest Protect was released, and continued to increase over time as additional products were introduced on the market. On the other hand, the prevalence of Topic 20 "Repeat Purchase" (Figure 10b) is at its highest at the beginning of the release of the 2^{nd} and the 3^{rd} generations of the Nest thermostat, indicating that satisfied customers of previous generations may purchase an upgrade when newer versions of the product become available. The findings in the "Product Line" category show that the content of customer reviews shifts in response to the introductions of new products, although Figure 11 reveals that the evaluations of these two topics hardly change over time.

[Insert Figure 11 Here]

The results pertaining to the remaining categories in Table 2 are discussed in Appendix A.4. Looking at the most relevant words associated with each topic, we find that the review contributors use more non-emotional nouns and verbs for topics that prevail early, such as the topics in the "Features" and "Setup" categories. In contrast, the review contributors use more emotional adjectives and adverbs for topics that prevail late, such as the topics in the "Experience" and "Praise" categories. This is consistent with the findings of prior literature that illustrates how later ratings tend to be more extreme than earlier ratings under certain circumstances (e.g., Moe and Trusov 2011; Moe and Schweidel 2012).

Conclusion

In this research, we investigate the dynamics present in the textual content of user-generated product reviews of a high-tech multi-generational product by developing a continuous-state dynamic topic model that considers the evolution of unobserved topics in multiple review streams. We build upon the SLDA modeling framework by allowing the prevalence and the evaluations of topics to change both within and across generations of a product in parallel over time. Unlike the previous literature that assumes the underlying governing process of the composition of topics in customer reviews to be static (e.g., Tirunillai and Tellis 2014; Büschken and Allenby 2016), our research takes advantage of the timestamp associated with each review to construct a multi-stream corpus that facilitates the discovery of temporal changes in topic prevalence as well as topic evaluations.

To the best of our knowledge, we are among the first in the marketing literature to examine the textual dynamics of multiple text streams. Since our modeling framework allows for both the commonalities and idiosyncrasies of topics across multiple text streams, our approach can be applied not only to the review streams of multi-generational products, but also to multiple review streams across a product line for a given brand or to reviews for different brands within the same category. Furthermore, it can be used to model the review and post streams of the same brand from multiple social media venues, as prior literature has found a substantial difference in expressed sentiment across venues (Schweidel and Moe 2014).

Our findings reveal that topic prevalence exhibits significant temporal variation both within a given generation and across generations. In our empirical context, we find that earlier reviews are more homogeneous to each other than later reviews with respect to the mixture of topics. We also find that earlier reviews are more objective and feature-oriented, whereas later reviews are more subjective and experience-related. Over time, the evaluation of feature-related topics declines more than the evaluation of experience-related topics. The deviations from the typical trends coincide with major product changes (e.g., Topic 1 “Display”), new product introduction (e.g., Topic 19 "Integration"), and temporary service outages (e.g., Topic 15 "Temperature Feels"), thus suggesting that online reviews provide a means of getting at the "voice of consumers" to assess their response to such events.

The findings of this research provide insights for marketers into what aspects related to a product concern customers at different stages of product lifecycle. By understanding such differences, marketers can enhance customers' evaluations of products by sending marketing signals that matches their preferences (Chen and Xie 2008), such as by varying advertising messages or the focus of marketing activities over the course of the product lifecycle. For instance, early advertisements in the product lifecycle may focus on product features while later advertisements may emphasize user experience. Last but not least, the deterioration of consumers' excitement with product features expressed in earlier reviews could also give marketers insights into the timing of their new product launches.

Our research has implications for consumers who rely on online reviews to inform their purchase decisions. For consumers seeking to acquire information about what to expect from the major functions of a product and how to get started with using it, they will find more relevant content in earlier reviews that emphasize the details of product features and setup instructions. In contrast, for consumers interested in assessing product quality and service quality, such content is more likely to be found in more recent reviews. In this case, consumers should be aware that later reviews tend to be more subjective and emotional, unlike earlier reviews that are more objective and clinical. In addition, given that the similarity of topics covered in reviews declines in later reviews, consumers who look at later reviews may need to read more reviews than those who look just at the earlier reviews in order to retrieve more comprehensive information.

Our research is also of relevance to e-commerce platforms on how the reviews are presented to consumers. While consumers have access to prior reviews, they must sort through the comments to extract information that is most relevant to them. The volume of information with which consumers are presented can increase their search costs and reduce their purchase intentions (e.g., Spenner and Freeman 2012; Branco et al. 2016). E-commerce platforms could aid consumers in acquiring relevant information from product reviews, such as by identifying those topics with higher prevalence or evaluation compared to the trajectory that they were on, and making reviews focusing on such topics more prominently.

One limitation of this research lies in the heterogeneity of review contributors. While we have identified the presence of dynamics in the textual content and evaluations across multiple streams, we are limited in our ability to speak to the underlying cause of the dynamics. One explanation is that the temporal pattern of topic prevalence and topic evaluation is due to contributors of later reviews choosing to differentiate themselves by discussing different topics compared to earlier reviewers. Alternatively, later reviewers may systematically differ from early reviewers in their preferences. To disentangle such explanations, multiple reviews contributed by the same reviewer are needed (e.g., Moe and Schweidel 2012). Extending our modeling framework to accommodate author-level effects, pending the availability of such data, is an area for future research. Our model assumes the existence of homogeneous thresholds and evaluation parameters at the same time for the ordinal probit model. To relax such an assumption, repeated observations of reviews contributed by the same reviewer at the same time are necessary. As we make use of data from a single product, research would also be warranted on a broad range of product categories as this would allow us to understand the extent to which the textual dynamics of consumer reviews are common or different across categories.

Reference

Albert, James H. and Siddhartha Chib. (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American statistical Association*, 88(422), 669-679.

Archak, Nikolay, Anindya Ghose and Panagiotis G. Ipeirotis (2011), "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science*, 57(8), 1485-1509.

Blei, David M., Andrew Y. Ng and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, Vol. 3, 993-1022.

Blei, David M. and John D. Lafferty (2006), "Dynamic Topic Models," *Proceedings of the 23rd International Conference on Machine Learning*, Pittsburgh, PA.

Blei, David M. and John D. Lafferty (2007) "A Correlated Topic Model of Science." The Annals of Applied Statistics, 1(1), 17-35.

Blei, David M. and Jon D. McAuliffe (2007), "Supervised Topic Models," *Neural Information Processing System*, 20, 121-128, MIT Press, Cambridge, MA.

Blei, David M., Thomas L. Griffiths and Michael I. Jordan (2010), "The Nested Chinese Restaurant Process and Bayesian Nonparametric Inference of Topic Hierarchies," *Journal of the ACM (JACM)*, 57 (2), 1-30.

Borah, Abhishek and Gerard J. Tellis (2016), "Halo (Spillover) Effects in Social Media: Do Product Recalls of One Brand Hurt or Help Rival Brands?" *Journal of Marketing Research*, 53(2), 143-160.

Branco, Fernando, Monic Sun and J. Miguel Villas-Boas (2016), "Too Much Information? Information Provision and Search Costs." *Marketing Science*, 35(4), 605-618.

Büschken, Joachim and Greg M. Allenby (2016), "Sentence-Based Text Analysis for Customer Reviews," *Marketing Science*, 35(6), 953-975.

Chen, Yubo and Jinhong Xie (2008), "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science*, 54(3), 477–491.

Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43(3), 345–354.

Chintagunta Pradeep K., Shyam Gopinath and Sriram Venkataraman (2010), "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation across Local Markets," *Marketing Science*, 29(5), 944–959.

Constine, Josh (2014), "Update: Nest Says Shut-off Heat Is Sometimes Its Fault, Also Pushes Thermostat 4.0.1 Firmware To Fix 4.0 Problems," *Tech Crunch*, Jan 6, accessed at: *https://techcrunch.com/2014/01/06/nest-4-0-firmware-battery-problems/.*

Cooper, Lee G., Masao Nakanishi and Jehoshua Eliashberg (1988), "Market-share Analysis: Evaluating Competitive Marketing Effectiveness," Boston: Kluwer Academic Publishers.

Decker, Reinhold and Michael Trusov (2010), "Estimating Aggregate Consumer Preferences from Online Product Reviews," *International Journal of Research in Marketing*, 27(4), 293-307.

Duan, Wenjing, Bin Gu and Andrew B. Whinston (2009), "Informational Cascades and Software Adoption on the Internet: An Empirical Investigation," *MIS Quarterly*, 33(1), 23–48.

Fossen, Beth L. and David A. Schweidel (2016), "Television Advertising and Online Word-of-Mouth: An Empirical Investigation of Social TV Activity," *Marketing Science*, forthcoming.

Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545–560.

Godes, David and Jose C. Silva (2012), "Sequential and Temporal Dynamics of Online Opinion," *Marketing Science*, 31(3), 448-473.

Goldenberg, Jacob, Donald R. Lehmann, Daniella Shidlovski, and Michal M. Barak (2006), "The Role of Expert versus Social Opinion Leaders in New Product Adoption," *Marketing Science Institute Report*, 06–124, Cambridge, MA.

Griffiths, Thomas L. and Mark Steyvers (2004), "Finding Scientific Topics," *Proceedings of the National Academy of Sciences*, Vol. 101, Suppl 1, 5228-5235.

Hildenbrand, Jerry (2017), "Nest Thermostat Gen 1 vs Gen 2 vs Gen 3: Comparison and Main Differences," *Android Central*, March 2, accessed at http://www.androidcentral.com/nest-gen-1-vs-gen-2-vs-gen-3-comparison-and-main-differences.

Hong, Liangjie, Byron Dom, Siva Gurumurthy, and Kostas Tsioutsiouliklis (2011), "A Time-dependent Topic Model for Multiple Text Streams," *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 832-840.

Iyengar, Raghuram, Christophe Van den Bulte, and Thomas W. Valente (2011), "Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science*, 30(2), 195-212.

Kuksov, Dmitri and Ying Xie (2010), "Pricing, Frills, and Customer Ratings," *Marketing Science*, 29(5), 925-943.

Lee, Thomas Y. and Eric T. Bradlow (2011), "Automated Marketing Research Using Online Customer Reviews," *Journal of Marketing Research*, 48(October), 881-894.

Lee, Young-Jin, Kartik Hosanagar and Yong Tan (2015), "Do I Follow My Friends or the Crowd? Information Cascades in Online Movie Ratings," *Management Science*, 61(9), 2241-2258.

Li, Xinxin and Lorin M. Hitt (2008), "Self-selection and Information Role of Online Product Reviews," *Information Systems Research*, 19(4), 456–474.

Liu, Xiao, Param Vir Singh and Kannan Srinivasan (2016), "A Structured Analysis of Unstructured Big Data by Leveraging Cloud Computing," *Marketing Science*, 35(3), 363-388.

Liu, Yong (2006), "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 70(3), 74–89.

Ma, Liye, Baohong Sun and Sunder Kekre (2015), "The Squeaky Wheel Gets the Grease – An Empirical Analysis of Customer Voice and Firm Intervention on Twitter," *Marketing Science*, 34(5), 627-645.

Minka, Thomas P. (2000), "Estimating a Dirichlet Distribution," Technical report, MIT.

Moe, Wendy W. and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*, 48(3), 444-456.

Moe, Wendy W. and David A. Schweidel (2012), "Online Product Opinions: Incidence, Evaluation, and Evolution," *Marketing Science*, 31(3), 372-386.

Netzer, Oded, Ronen Feldman, Jacob Goldenberg and Moshe Fresko (2012), "Mine Your Own Business: Market-structure Surveillance through Text Mining," *Marketing Science*, 31(3), 521-543.

Paul, Michael and Roxana Girju (2009), "Cross-cultural Analysis of Blogs and Forums with Mixed-collection Topic Models," *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, Vol. 3, 1408-1417.

Porter, Martin (2006), "The English (Porter2) Stemming Algorithm," January 10, accessed at: http://snowball.tartarus.org/algorithms/english/stemmer.html.

Puranam, Dinesh, Vishal Narayan, and Vrinda Kadiyali (2017), "The Effect of Calorie Posting Regulation on Consumer Opinions: A Flexible Latent Dirichlet Allocation Model with Informative Priors," *Marketing Science*, 36 (5), 726-746.

Rogers, Everett M. (2003) *The Diffusion of Innovations*, Fifth Edition, The Free Press, NewYork.

Schlosser, Ann (2005), "Posting versus Lurking: Communicating in a Multiple Audience Context," *Journal of Consumer Research*, 32(2), 260-265.

Schweidel, David A. and Wendy W. Moe (2014), "Listening in on Social Media: A Joint Model of Sentiment and Venue Format Choice," *Journal of Marketing Research*, 51(August), 387-402.

Sievert, Carson and Kenneth E. Shirley (2014), "LDAvis: A Method for Visualizing and Interpreting Topics," *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, June, 63-70.

Spenner, Patrick and Karen Freeman (2012), "To Keep Your Customers, Keep It Simple," *Harvard Business Review*, 90, 108–114.

Sun, Monic (2012), "How Does the Variance of Product Ratings Matter?" *Management Science*, 58(4), 696-707.

Tirunillai, Seshadri and Gerard J. Tellis (2014), "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, 51(4), 463-479.

Wallach, Hanna M., David Mimno, and Andrew McCallum (2009), "Rethinking LDA: Why Priors Matter," *Advances in Neural Information Processing Systems*, 22, 1973-1981, MIT Press, Cambridge, MA.

Xiong, Guiyang and Sundar Bharadwaj (2014), "Prerelease Buzz Evolution Patterns and New Product Performance," *Marketing Science*, 33(3), 401-421.

Zhao, Yi, Sha Yang, Vishal Narayan and Ying Zhao (2013), "Modeling Consumer Learning from Online Product Reviews," *Marketing Science*, 32(1),153-169

Table 1. LMD and the Perplexity of Next Month's Hold-out Sample for the Static and Dynamic Models

Month	Single Text Stream				Multiple Text Streams			
	LMD		Perplexity		LMD		Perplexity	
	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic
16-May	-50230	-50124	328.0	324.2	-50086	-50035	322.6	320.5
16-Jun	-53748	-53593	343.9	337.6	-53629	-53477	339.1	333.7
16-Jul	-86911	-86840	322.1	320.8	-86878	-86782	321.3	319.4
16-Aug	-83095	-82771	325.0	317.6	-83038	-82741	323.7	316.2
16-Sep	-58943	-58808	303.2	298.3	-58734	-58632	297.2	293.5
16-Oct	-51182	-50836	315.8	303.4	-50662	-50572	297.5	294.5
16-Nov	-71843	-71755	323.1	319.8	-71698	-71599	319.0	316.4
16-Dec	-124886	-124507	331.7	324.6	-124700	-124318	328.6	321.2
17-Jan	-129809	-129374	303.1	297.0	-129387	-128950	301.6	295.5
17-Feb	-94083	-93982	289.4	287.4	-93988	-93869	287.3	285.2

Table 2. 25 Most Relevant Words of Topics ($J = 20$)

Category	Features						Setup			Praise
Label	"Display"	"Eco"	"Interface"	"Learning"	"Schedule"	"System"	"Installation"	"Instructions"	"Power"	"Compliment"
Rank	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
1	display	energy	device	house	schedule	heat	plate	instruction	wire	product
2	temperature	usage	interface	turn	feature	system	wall	installation	common	great
3	current	saving	web	home	mode	pump	old	install	power	star
4	room	report	design	temperature	set	air	screw	video	charge	excellent
5	show	save	simple	heat	auto	fan	hole	nest	furnace	recommend
6	sensor	monthly	app	leave	thermostat	stage	wire	easy	battery	good
7	set	money	access	warm	learn	nest	cover	wiring	Nest	work
8	large	help	apps	bed	temperature	cool	paint	minute	voltage	awesome
9	number	home	wi-fi	adjust	time	HVAC	thermostat	thermostat	system	worth
10	humidity	data	password	able	home	run	include	follow	issue	quality
11	screen	history	user	cool	auto-away	humidifier	mount	setup	run	highly
12	actual	track	account	phone	manually	speed	installation	watch	connect	love
13	walk	consumption	apple	forget	day	furnace	label	step	HVAC	money
14	motion	cool	iPhone	love	setting	gas	box	diagram	red	amazing
15	option	leaf	setup	down	program	AC	Nest	old	problem	well
16	thermostat	cost	network	night	eco	zone	screwdriver	less	blue	best
17	read	efficient	IOS	know	work	forced	remove	website	hook	fantastic
18	big	reduce	iPad	cold	learning	cycle	connect	online	require	design
19	small	thermostat	mobile	day	want	compressor	plastic	breeze	work	expect
20	nest	heat	intuitive	thing	automatically	work	base	simple	cable	buy
21	outside	bill	android	arrive	adjust	condition	unit	clear	thermostat	definitely
22	light	nest	connect	app	week	thermostat	white	picture	wiring	ease
23	setting	long	control	set	function	control	look	YouTube	connection	thermostat
24	Ecobee	green	beautiful	AC	pretty	central	minute	box	unit	amaze
25	dial	control	easy	hot	turn	conditioner	small	direction	low	price
Prevalence										
Gen 1	1.4%	6.4%	6.9%	10.3%	6.5%	3.5%	3.5%	4.2%	2.9%	5.1%
Gen 2	1.6%	6.3%	4.8%	9.2%	5.2%	3.8%	2.6%	3.8%	3.3%	8.2%
Gen 3	2.6%	5.2%	2.1%	7.0%	4.2%	1.9%	2.2%	5.5%	1.5%	13.0%
Evaluation										

Gen 1	0.50	2.83	2.20	3.86	1.26	0.55	3.36	9.83	1.00	6.81
Gen 2	0.25	3.46	2.82	3.68	0.88	0.51	3.00	4.61	1.54	3.62
Gen 3	0.87	3.27	1.69	3.37	1.09	1.63	3.01	3.82	1.31	3.30

Category	Experience		Criticism				Financial		Product Line	
Label	"Ease of Use"	"User Experience"	"Compatibility"	"Connection"	"Temperature Feels"	"Unit Failure"	"Price"	"Savings"	"Integration"	"Repeat Purchase"
Rank	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19	Topic 20
1	easy	love	support	router	degree	unit	buy	bill	smoke	generation
2	install	thermostat	customer	update	temperature	Nest	thermostat	USD	detector	Nest
3	great	control	Nest	wi-fi	Nest	base	good	save	protect	upstairs
4	star	phone	service	software	set	fail	thing	month	Alexa	second
5	set up	Nest	tech	issue	house	problem	review	rebate	echo	unit
6	look	home	work	problem	cold	replacement	people	electric	Nest	downstairs
7	work	adjust	hour	Nest	heat	AC	Honeywell	money	home	first
8	super	temperature	return	battery	thermostat	replace	price	pay	Google	upgrade
9	product	easy	compatible	offline	turn	compressor	look	saving	alarm	floor
10	setup	able	unit	firmware	back	cause	want	year	automation	zone
11	love	smart	system	network	time	repair	USD	utility	integrate	third
12	app	install	person	wireless	hot	thermostat	nest	energy	camera	version
13	simple	app	email	connect	night	work	smart	company	integration	thermostat
14	far	recommend	issue	support	want	failure	Google	gas	Amazon	buy
15	operate	ability	receive	fix	day	damage	expensive	cost	smart	home
16	extremely	learn	compatibility	connection	hour	issue	cheap	electricity	device	purchase
17	nice	remotely	wait	connectivity	freeze	stop	year	worth	thermostat	happy
18	awesome	highly	wiring	internet	high	HVAC	well	last	add	model
19	installation	great	problem	disconnect	warm	back	piece	first	security	replace
20	program	absolutely	Amazon	work	reach	furnace	little	power	work	large
21	well	good	back	lose	keep	USD	a lot	summer	pair	display
22	minute	monitor	contact	device	cool	return	technology	price	fire	screen
23	intuitive	awesome	time	list	wake	start	know	buy	buy	order
24	advertise	best	answer	thermostat	winter	year	money	purchase	system	house
25	beautiful	ease	on hold	version	thing	heat	pretty	Nest	light	year
Prevalence										
Gen 1	4.6%	5.2%	7.5%	5.2%	5.5%	5.8%	6.5%	6.0%	0.6%	2.5%
Gen 2	7.4%	6.9%	5.8%	5.1%	5.0%	5.1%	6.7%	5.1%	1.3%	2.5%
Gen 3	19.7%	14.1%	2.7%	0.8%	1.6%	1.1%	2.6%	6.0%	3.5%	2.6%
Evaluation										

Gen 1	5.06	4.82	-1.42	-0.23	-1.55	-2.98	-1.69	4.08	6.67	4.73
Gen 2	4.17	4.50	-1.22	-0.72	-0.99	-2.31	0.07	3.32	1.56	3.28
Gen 3	3.84	4.70	-0.43	-0.49	-1.03	-1.59	1.10	3.60	2.90	2.84

Figure 1. The Volume of Reviews by Month

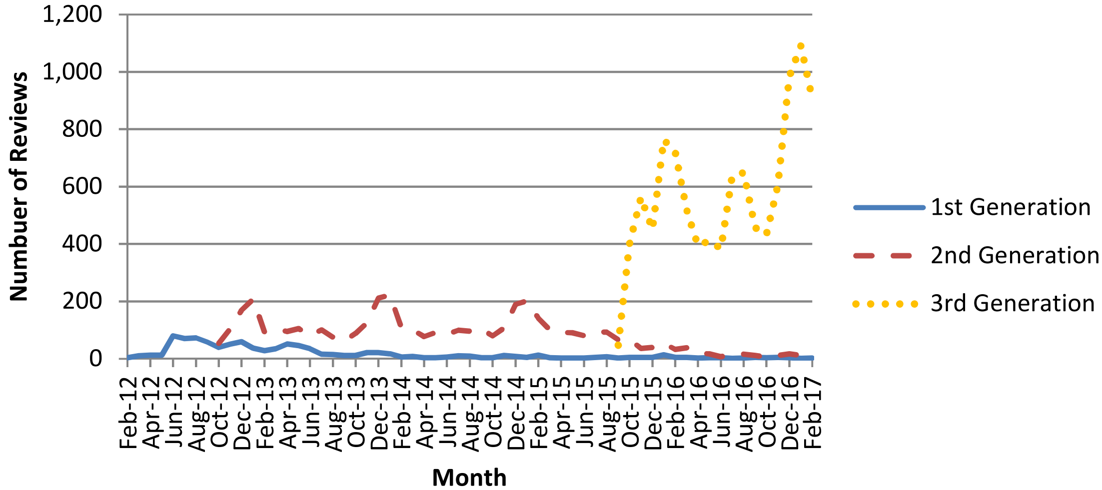


Figure 2. Average Number of Tokens by Month

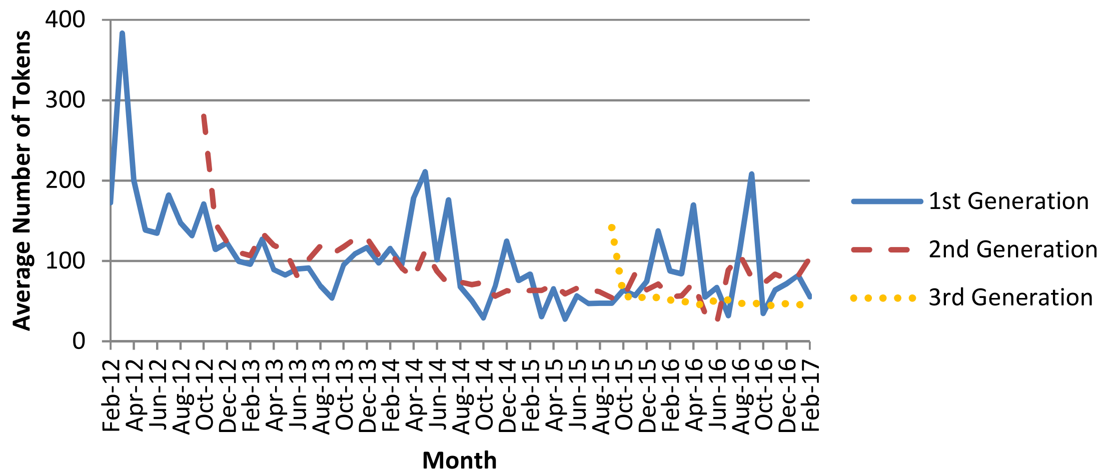


Figure 3. Average Rating by Month

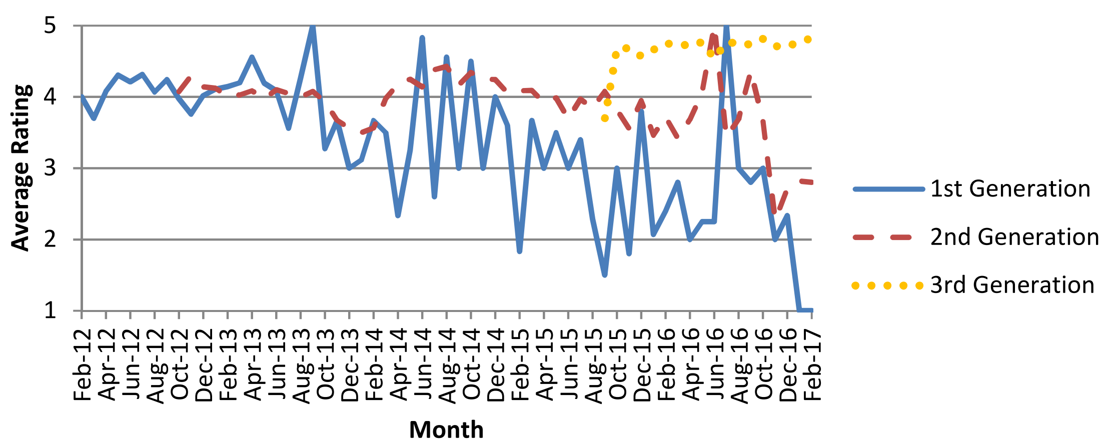


Figure 4. Multiple Review Streams with Different Start Times

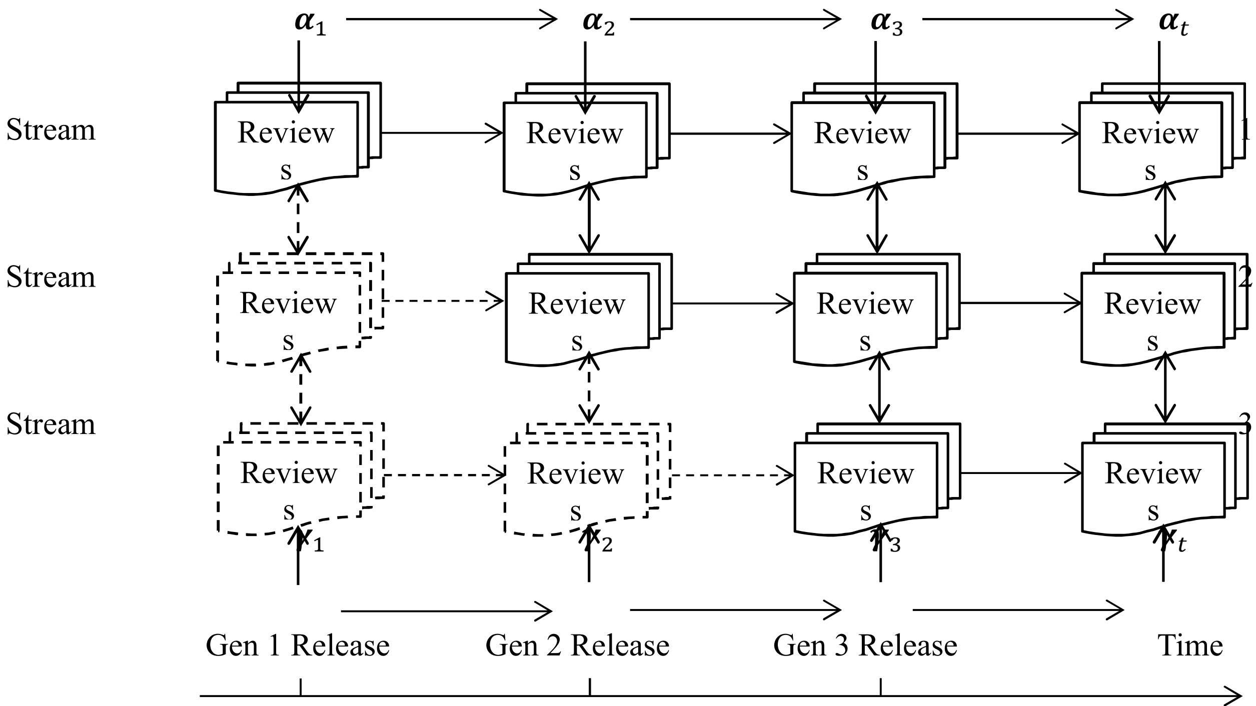


Figure 5. Trend of Concentration

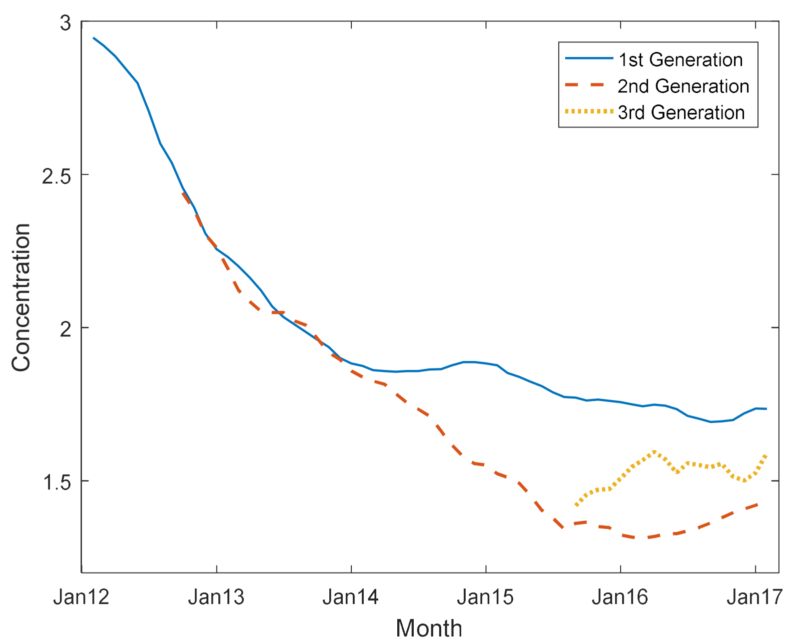
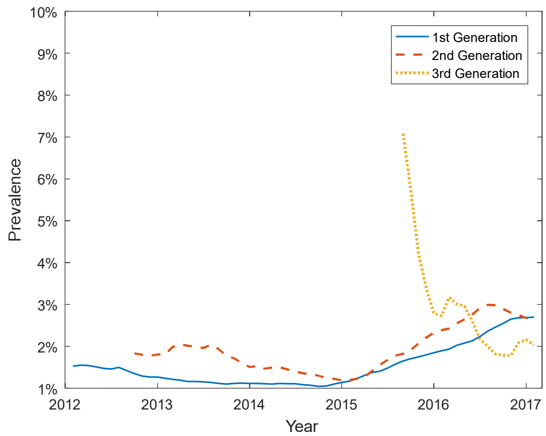
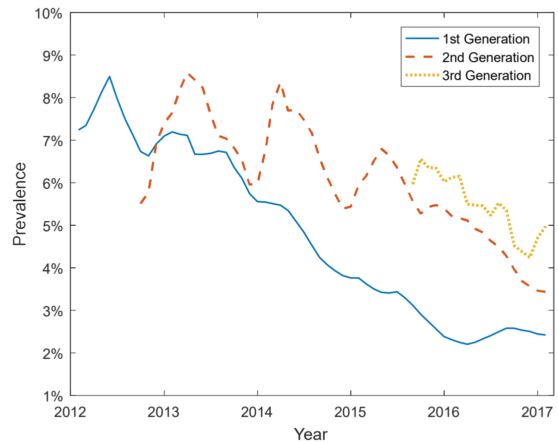


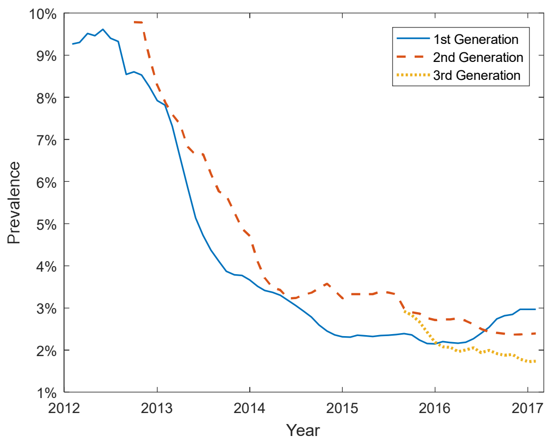
Figure 6. Prevalence of Topics in the “Features” Category



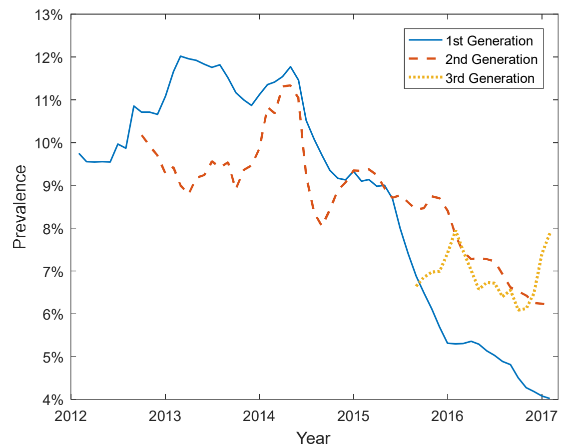
(a) Topic 1 “Display”



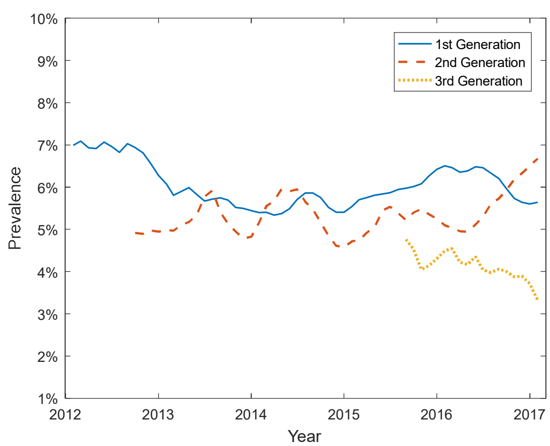
(b) Topic 2 “Eco”



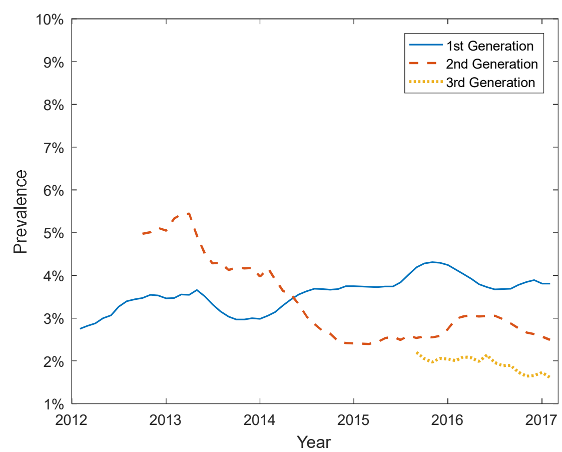
(c) Topic 3 “Interface”



(d) Topic 4 “Learning”

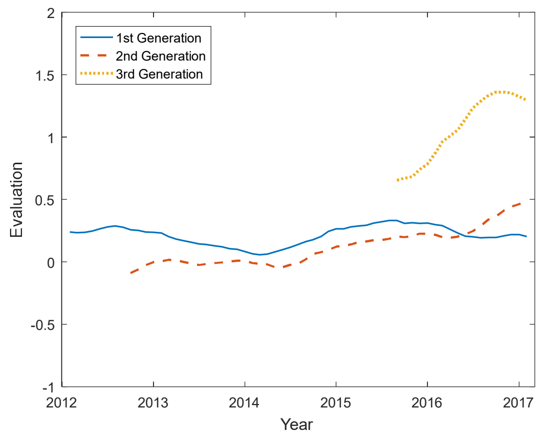


(e) Topic 5 “Schedule”

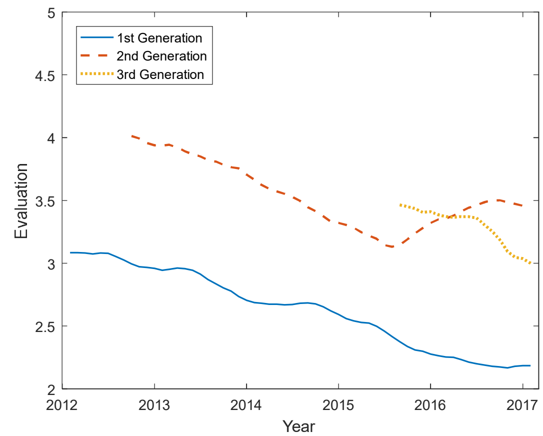


(f) Topic 6 “System”

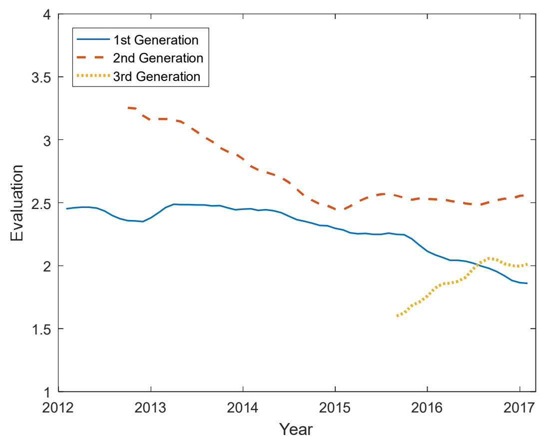
Figure 7. Evaluation of Topics in the “Features” Category



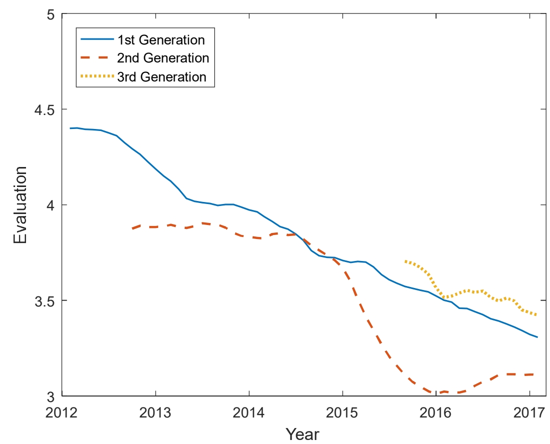
(a) Topic 1 “Display”



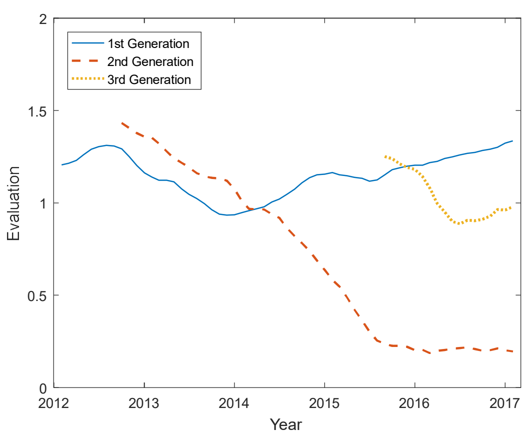
(b) Topic 2 “Eco”



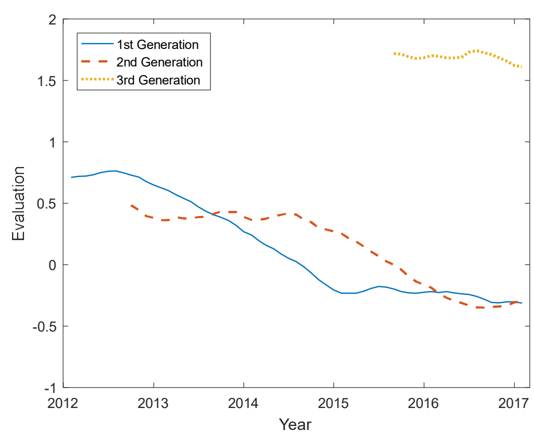
(c) Topic 3 “Interface”



(d) Topic 4 “Learning”

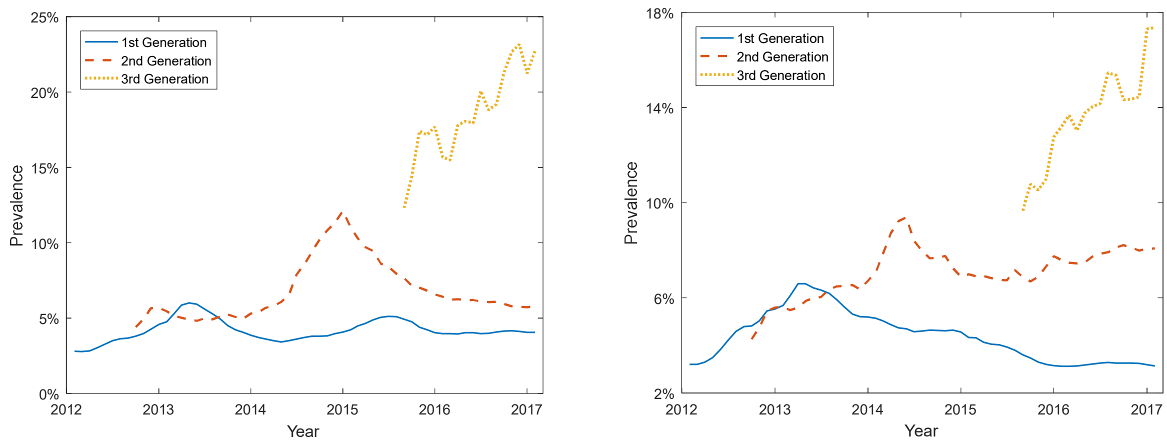


(e) Topic 5 “Schedule”



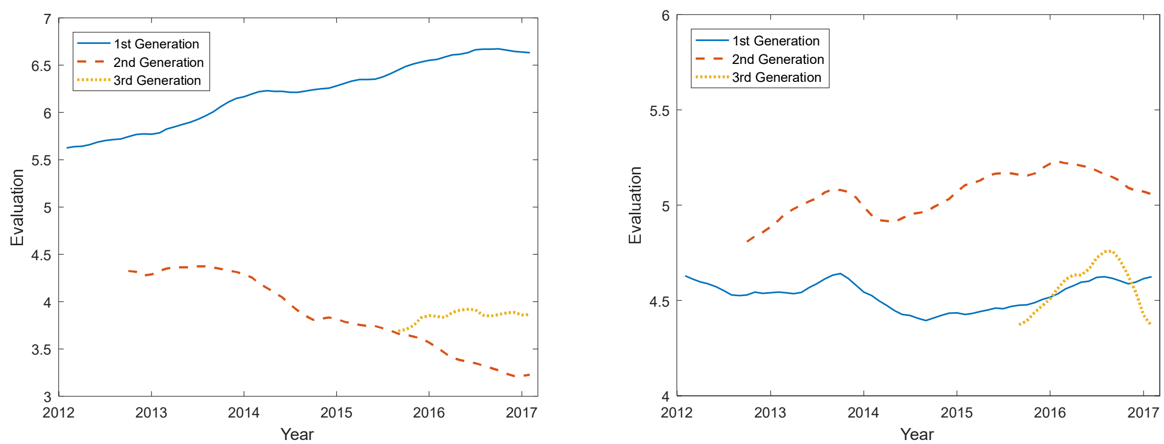
(f) Topic 6 “System”

Figure 8. Prevalence of Topics in the "Experience" Category



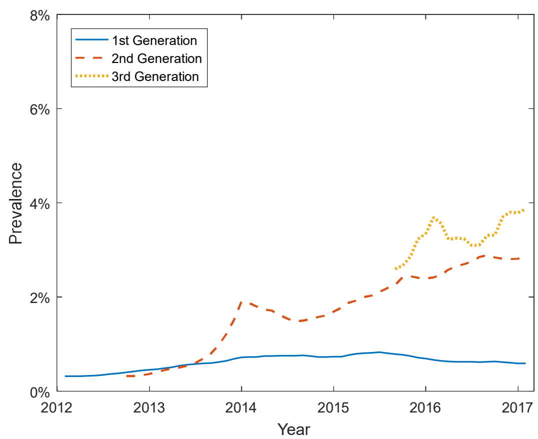
(a) Topic 11 "Ease of Use" (b) Topic 12 "User Experience"

Figure 9. Evaluation of Topics in the "Experience" Category

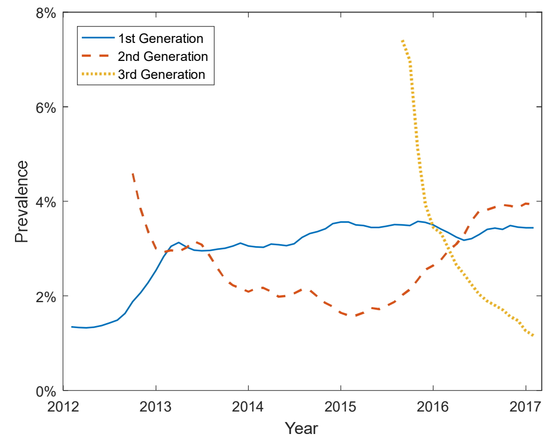


(a) Topic 11 "Ease of Use" (b) Topic 12 "User Experience"

Figure 10. Prevalence of Topics in the “Product Line” Category

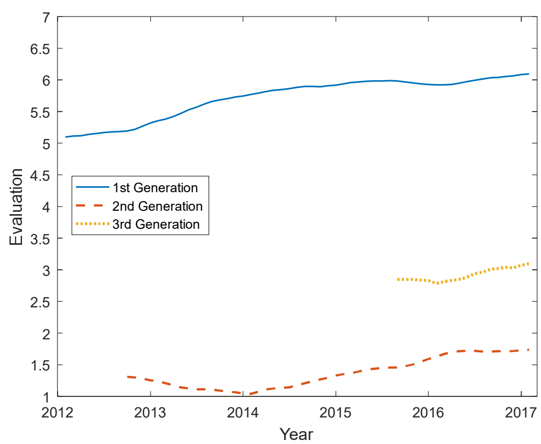


(a) Topic 19 “Integration”

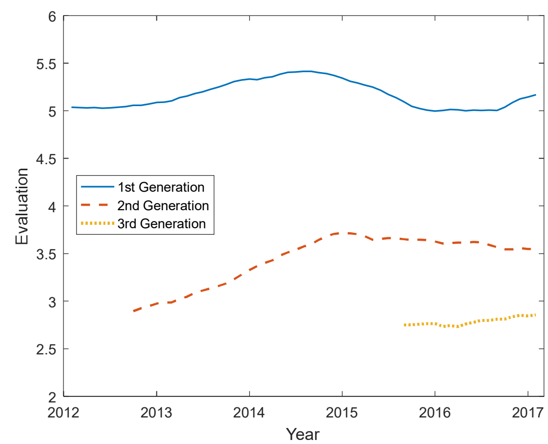


(b) Topic 20 “Repeat Purchase”

Figure 11. Evaluation of Topics in the “Product Line” Category



(a) Topic 19 “Integration”



(b) Topic 20 “Repeat Purchase”

Appendix A.1. Collapsed Gibbs Sampler for SLDA

The joint likelihood of observed word $\boldsymbol{w}$, unobserved topic $\boldsymbol{z}$, and continuous rating $\boldsymbol{v}$ can be expressed by

$$p(\boldsymbol{w},\boldsymbol{z},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) = p(\boldsymbol{w}|\boldsymbol{\beta},\boldsymbol{z})p(\boldsymbol{z}|\boldsymbol{\alpha})p(\boldsymbol{v}|\boldsymbol{\gamma},\boldsymbol{z})$$

$$\propto \prod_{j=1}^{J} \frac{\prod_w \Gamma\left(N_j^w+\beta\right)}{\Gamma\left(N_j+W\beta\right)} \prod_{d=1}^{D} \frac{\prod_j \Gamma\left(N_j^d+\alpha\right)}{\Gamma(N^d+J\alpha)} \prod_{d=1}^{D} \exp\left(-\frac{(v_d-\bar{\boldsymbol{z}}_d\boldsymbol{\gamma})^2}{2\sigma^2}\right)$$

To derive the expression to sample topic assignments of SLDA, we constrain the variance to 1 (Greene 2012) and assume no intercept (Blei and McAuliffe 2007) in ordinal probit for identification. By Bayes' theorem, the full conditional posterior of $\boldsymbol{z}_{di}$, the topic assignment of token i in document d, is given by

$$\begin{aligned}
p(\boldsymbol{z}_{di}|\boldsymbol{w},\boldsymbol{z}_{-di},\boldsymbol{y},\boldsymbol{v},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\tau}) &= \frac{p(\boldsymbol{w},\boldsymbol{z},\boldsymbol{y},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\tau})}{p(\boldsymbol{w},\boldsymbol{z}_{-di},\boldsymbol{y},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\tau})} \\
&= \frac{p(\boldsymbol{w},\boldsymbol{z},\boldsymbol{y},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\tau})}{p(\{\boldsymbol{w}_{di},\boldsymbol{w}_{-di}\},\boldsymbol{z}_{-di},\boldsymbol{y},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\tau})} \\
&= \frac{p(\boldsymbol{w},\boldsymbol{z},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})p(\boldsymbol{y}|\boldsymbol{v},\boldsymbol{\tau})}{p(\{\boldsymbol{w}_{di},\boldsymbol{w}_{-di}\},\boldsymbol{z}_{-di},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})p(\boldsymbol{y}|\boldsymbol{v},\boldsymbol{\tau})} \\
&\propto \frac{p(\boldsymbol{w},\boldsymbol{z},\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})}{p(\boldsymbol{w}_{-di},\boldsymbol{z}_{-d}\ ,\boldsymbol{v}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})} \\
&= \frac{p(\boldsymbol{w}|\boldsymbol{\beta},\boldsymbol{z})p(\boldsymbol{z}|\boldsymbol{\alpha})}{p(\boldsymbol{w}_{-d}|\boldsymbol{\beta},\boldsymbol{z}_{-di})p(\boldsymbol{z}_{-di}|\boldsymbol{\alpha})} \frac{p(\boldsymbol{v}|\boldsymbol{\gamma},\boldsymbol{z})}{p(\boldsymbol{v}|\boldsymbol{\gamma},\boldsymbol{z}_{-di})} \qquad (24)
\end{aligned}$$

where $\boldsymbol{z}_{di}$ denotes the topic assignments of token i in document d, $\boldsymbol{z}_{-di}$ denotes the topic assignments of all tokens except for token i in document d, $\boldsymbol{w}$ denotes the observed words, $\boldsymbol{v}$ denotes the continuous rating that can be augmented in the case of the ordinal probit with observed ordinal rating $\boldsymbol{y}$ and threshold $\boldsymbol{\tau}$, $\boldsymbol{\alpha}$ denotes the Dirichlet hyperprior for topic proportions of each document, $\boldsymbol{\beta}$ denotes the Dirichlet hyperprior for word distributions of each topic, and $\boldsymbol{\gamma}$ denotes the coefficients of the topic proportions in the ordinal probit. The first factor of Eqn. (13) has the same form of conditional posterior as standard LDA (Griffiths and Steyvers 2004). The denominator of the second factor is the same for all possible topic assignments of token *i* in document *d* and can thus be treated as a constant. As a result, the conditional distribution of token *i* in document *d* being assigned to topic *j* can be expressed by

$$p(z_{di} = j|\boldsymbol{w}, \boldsymbol{z}_{-di}, \boldsymbol{v}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \propto \frac{N^{w}_{-di,j} + \beta}{N_{-di,j} + W\beta}\left(N^{d}_{-di,j} + \alpha\right) \exp\left(-\frac{\left(v_d - \bar{\boldsymbol{z}}_{d,z_{di}=j}\boldsymbol{\gamma}\right)^2}{2}\right)$$

where $\bar{\boldsymbol{z}}_{d,z_{di}=j} = \frac{1}{N^d}\left(\sum_{i'=1,i'\neq i}^{N^d} z_{di} \ + z_{di}\right)$, in which $z_{di'}$, $i' \neq i$, is the topic assignment of any token in document *d* excluding token *i*; z_{di} is the tentative topic assignment of token *i* in document *d* to topic *j*.

To show how the model works in the context of multiple text streams, we demonstrate a simulation study with the following settings: Dirichlet hyperprior for word distribution β=0.3, the number of topics $J = 5$, the number of text streams $k = 3$, the number of time periods $T = 60$, the number of documents from each stream at each time period is 300, and the number of tokens within each document is 50. Noticeably, although Stream 1 starts from the beginning, Stream 2 starts at $t_2 = 21$ and Stream 3 starts at $t_3 = 41$. To simplify the demonstration, we only show the results of the static model. The true values, estimates, and 95% intervals of topic prevalence and topic evaluations parameters are shown in Table A.1.

[Insert Table A.1 Here]

As is shown by Table A.1, the proposed model is not only able to capture stream-specific topics (e.g., Topic 5 for Stream 3), but also able to capture disappearing (e.g., Topic 1) and emerging (e.g., Topic 5) topics in later-starting text stream.

Appendix A.2. MCMC Algorithm for SLDA-MS

The MCMC algorithm for SLDA-MS integrates temporal priors for topic proportions and topic evaluations across text streams with a Gibbs sampler for the SLDA via the following steps:

1. $\boldsymbol{z} \mid \boldsymbol{w}, \boldsymbol{\alpha}, \boldsymbol{v}, \boldsymbol{\gamma}$

For document *d* in text stream C^k at time *t*, namely $d \in C_t^k$, the conditional distribution of token *i* being assigned to topic *j* can be represented as

$$p(z_{di} = j | \boldsymbol{w}, \boldsymbol{z}_{-d}, \boldsymbol{v}, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \propto \frac{N_{-di,j}^{w} + \beta}{N_{-di,j} + W\beta} \left(N_{-di,j}^{d} + \alpha_{j,t}^{k}\right) \exp\left(-\frac{\left(v_d - \bar{\boldsymbol{z}}_{d,z_{di}=j}\boldsymbol{\gamma}\right)^2}{2}\right)$$

See Appendix A.1. for the details of the derivation. Similar to a Gibbs sampler procedure for LDA proposed by Griffiths and Steyvers (2004), the estimates of the word distribution of topics ϕ can be obtained by Eqn. (6) and the estimates of the topic distribution of documents θ for document $d \in C_t^k$ can be obtained by

$$\hat{\theta}_j^d = \frac{N_j^d + \alpha_{j,t}^k}{N^d + \sum_{j'=1}^{J} \alpha_{j',t}^k} \tag{25}$$

2. $\boldsymbol{\alpha} \mid \boldsymbol{z}, \boldsymbol{\Omega}$

Given the fact that $\boldsymbol{\alpha}$ are not involved in $P(\boldsymbol{w}|\boldsymbol{z}, \boldsymbol{\beta})$ and $P(\boldsymbol{y}|\boldsymbol{v}, \boldsymbol{\tau})P(\boldsymbol{v}|\boldsymbol{\gamma}, \boldsymbol{z})$ in Eqn. (12), the Dirichlet hyperpriors for topic proportions, $\boldsymbol{\alpha}$, can be sampled separately given $\boldsymbol{z}$. According to the generative process of LDA, for any document *d* in text stream *k* at time *t*, topic proportion θ^{kd} is drawn from a Dirichlet distribution with parameter $\boldsymbol{\alpha}_t^k$. This compound distribution is a Dirichlet-multinomial distribution, or Pólya distribution (Minka 2000). Therefore, the likelihood of $\boldsymbol{\alpha}$ conditional on topic assignments $\boldsymbol{z}$ is given by

$$L(\boldsymbol{\alpha}|\boldsymbol{z}) = P(\boldsymbol{z}|\boldsymbol{\alpha}) = \prod_{t=1}^{T} P(\boldsymbol{z}_t|\boldsymbol{\alpha}_t)$$

$$= \prod_{t=1}^{T}\prod_{k=1}^{K}\prod_{d\in C_t^k}\left(\frac{\Gamma\left(\sum_{j'=1}^{J}\alpha_{j',t}^{k}\right)}{\Gamma\left(N^d+\sum_{j'=1}^{J}\alpha_{j',t}^{k}\right)}\prod_{j=1}^{J}\frac{\Gamma\left(N_j^d+\alpha_{j,t}^{k}\right)}{\Gamma\left(\alpha_{j,t}^{k}\right)}\right) \tag{26}$$

where $\boldsymbol{z}_t$ denotes the topic assignments of tokens in documents generated at time *t*, $\boldsymbol{\alpha}_t$ denotes Dirichlet hyperpriors for topic proportions of all text streams at time *t*, C_t^k denotes the collection of documents in text stream *k* at time *t*, $\alpha_{j,t}^k$ denotes the scalar of the Dirichlet hyperprior for topic *j*'s proportion of documents in text stream *k* at time *t*, N_j^d records the number of tokens assigned to topic *j* in document *d*, and N^d records the total number of all tokens in document *d*. To sample $\boldsymbol{\alpha}_t$, we parameterize it as follows:

$$\pi\left(\boldsymbol{x}_{j,t}|\boldsymbol{x}_{j,-t},\boldsymbol{\Omega}\right) \sim \begin{cases} MVN\left(\boldsymbol{x}_{j,t+1},\boldsymbol{\Omega}\right) & \text{for } t=1 \\ MVN\left(\left(\boldsymbol{x}_{j,t-1}+\boldsymbol{x}_{j,t+1}\right)/2\,,\boldsymbol{\Omega}/2\right) & \text{for } t=2,\dots,T-1 \\ MVN\left(\boldsymbol{x}_{j,t-1},\boldsymbol{\Omega}\right) & \text{for } t=T \end{cases} \tag{27}$$

$$\boldsymbol{\alpha}_{j,t} = \exp\left(\boldsymbol{x}_{j,t}\right) \tag{28}$$

where $\boldsymbol{x}_{j,t} = \left(x_{j,t}^1,\dots,x_{j,t}^K\right)'$ denotes the vector of the log transformation of the Dirichlet hyperprior $\boldsymbol{\alpha}_{j,t} = \left(\alpha_{j,t}^1,\dots,\alpha_{j,t}^K\right)'$ for topic *j* across all text streams $k = 1,2,..,K$. Let $\boldsymbol{x}_t = \left(\boldsymbol{x}_{1,t},\dots,\boldsymbol{x}_{J,t}\right)$. For each iteration, sample $\boldsymbol{x}_t$ from a proposal density, $\text{vec}\left(\boldsymbol{x}_t^{\text{candidate}}\right) \sim MVN(\text{vec}(\boldsymbol{x}_t^{\text{current}}),\Delta_x\sigma_x^2 I)$. We then sample $\boldsymbol{x}_t^{\text{candidate}}$ by Eqn. (15) with an acceptance ratio given by

$$r\left(\boldsymbol{x}_t^{\text{candidate}},\boldsymbol{x}_t^{\text{current}}\right) = \frac{P\left(\boldsymbol{z}_t|\exp\left(\boldsymbol{x}_t^{\text{candidate}}\right)\right)\prod_{j=1}^{J}\pi\left(\boldsymbol{x}_{j,t}^{\text{candidate}}|\boldsymbol{x}_{j,-t},\boldsymbol{\Omega}\right)}{P(\boldsymbol{z}_t|\exp(\boldsymbol{x}_t^{\text{current}}))\prod_{j=1}^{J}\pi\left(\boldsymbol{x}_{j,t}^{\text{current}}|\boldsymbol{x}_{j,-t},\boldsymbol{\Omega}\right)}$$

3. $\boldsymbol{\Omega} \mid \boldsymbol{\alpha}$

Given the prior of the covariance matrix $\mathbf{\Omega}$, $\pi(\mathbf{\Omega}) \sim IW(\nu_{\Omega}, S_{\Omega})$, its posterior can be expressed by

$$P(\mathbf{\Omega}|\boldsymbol{\alpha}) \sim IW\left(\nu_{\Omega} + T, S_{\Omega} + \sum_{t=1}^{T}(\boldsymbol{x}_t - \boldsymbol{x}_{t-1})(\boldsymbol{x}_t - \boldsymbol{x}_{t-1})'\right) \tag{29}$$

4. $\boldsymbol{\tau} \mid \boldsymbol{y}, \boldsymbol{v}$

Let τ_r be the threshold of response $\boldsymbol{y}$ between category r and $r+1$, $r = 2, \dots, R-1$. Sample τ_r from a uniform distribution bounded between the maximum v_d in category r and the minimum v_d in category $r+1$ (Albert and Chib 1993) such that

$$P(\tau_r|\boldsymbol{y}, \boldsymbol{v}) \sim U(\max\{v_d : y_d = r\}, \min\{v_d : y_d = r+1\}) \tag{30}$$

5. $\boldsymbol{v} \mid \boldsymbol{y}, \boldsymbol{\gamma}, \boldsymbol{\tau}$

For any document $d \in C_t^k$, sample the unobserved variable v_d by a truncated normal distribution that is determined by the threshold τ_{r-1} and τ_r associated with $y_d = r$ (Albert and Chib 1993). Specifically, draw $u \sim U(0,1)$, then $v_d \sim TN_{(\tau_{r-1},\tau_r)}(\bar{\boldsymbol{z}}_d \boldsymbol{\gamma}_t^k, 1)$ such that

$$\begin{aligned} v_d = \bar{\boldsymbol{z}}_d \boldsymbol{\gamma}_t^k + \Phi^{-1}\Big(u\left(\Phi\left(\tau_r - \bar{\boldsymbol{z}}_d \boldsymbol{\gamma}_t^k\right) - \Phi\left(\tau_{r-1} - \bar{\boldsymbol{z}}_d \boldsymbol{\gamma}_t^k\right)\right) \\ + \Phi\left(\tau_{r-1} - \bar{\boldsymbol{z}}_d \boldsymbol{\gamma}_t^k\right)\Big) \end{aligned} \tag{31}$$

where $\Phi(\cdot)$ is the CDF of a standard normal distribution.

6. $\boldsymbol{\gamma} \mid \boldsymbol{z}, \boldsymbol{v}, \mathbf{\Sigma}$

Taking advantage of the augmented data $\boldsymbol{v}$, we can write the likelihood of $\boldsymbol{\gamma}$ conditioned on topic assignments $\boldsymbol{z}$ and unobserved variable $\boldsymbol{v}$ as

$$L(\boldsymbol{\gamma}|\boldsymbol{z},\boldsymbol{v}) = P(\boldsymbol{z},\boldsymbol{v}|\boldsymbol{\gamma}) = \prod_{t=1}^{T} P(\boldsymbol{z}_t,\boldsymbol{v}_t|\boldsymbol{\gamma}_t)$$
$$= \prod_{t=1}^{T}\prod_{k=1}^{K}\prod_{d\in C_{k,t}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(v_d - \bar{\boldsymbol{z}}_d\boldsymbol{\gamma}_t^k\right)^2\right) \tag{32}$$

where $\boldsymbol{v}_t$ denotes the set of all v_d in which $d \in C_t^{1:K}$. Sample $\boldsymbol{\gamma}_t$ as follows:

$$\pi(\boldsymbol{\gamma}_{j,t}|\boldsymbol{\gamma}_{j,-t},\boldsymbol{\Sigma}) \sim \begin{cases} MVN(\boldsymbol{\gamma}_{j,t+1},\boldsymbol{\Sigma}) & \text{for } t=1 \\ MVN((\boldsymbol{\gamma}_{j,t-1}+\boldsymbol{\gamma}_{j,t+1})/2,\boldsymbol{\Sigma}/2) & \text{for } t=2,\dots,T-1 \\ MVN(\boldsymbol{\gamma}_{j,t-1},\boldsymbol{\Sigma}) & \text{for } t=T \end{cases} \tag{33}$$

where $\boldsymbol{\gamma}_{j,t} = (\gamma_{j,t}^1,\dots,\gamma_{j,t}^K)'$ denotes the vector of topic evaluations. Let $\boldsymbol{\gamma}_t = (\boldsymbol{\gamma}_{1,t},\dots,\boldsymbol{\gamma}_{J,t})$. For each iteration, sample $\boldsymbol{\gamma}_t$ from a proposal density, $\text{vec}(\boldsymbol{\gamma}_t^{\text{candidate}}) \sim MVN(\text{vec}(\boldsymbol{\gamma}_t^{\text{current}}),\Delta_\gamma\sigma_\gamma^2 I)$. We then sample $\boldsymbol{\gamma}_t^{\text{candidate}}$ by Eqn. (21) with an acceptance ratio given by

$$r(\boldsymbol{\gamma}_t^{\text{candidate}},\boldsymbol{\gamma}_t^{\text{current}}) = \frac{P(\boldsymbol{z}_t,\boldsymbol{v}_t \mid \boldsymbol{\gamma}_t^{\text{candidate}})\prod_{j=1}^{J}\pi(\boldsymbol{\gamma}_{j,t}^{\text{candidate}}|\boldsymbol{\gamma}_{j,-t},\boldsymbol{\Sigma})}{P(\boldsymbol{z}_t,\boldsymbol{v}_t \mid \boldsymbol{\gamma}_t^{\text{current}})\prod_{j=1}^{J}\pi(\boldsymbol{\gamma}_{j,t}^{\text{current}}|\boldsymbol{\gamma}_{j,-t},\boldsymbol{\Sigma})}$$

7. $\boldsymbol{\Sigma} \mid \boldsymbol{\gamma}$

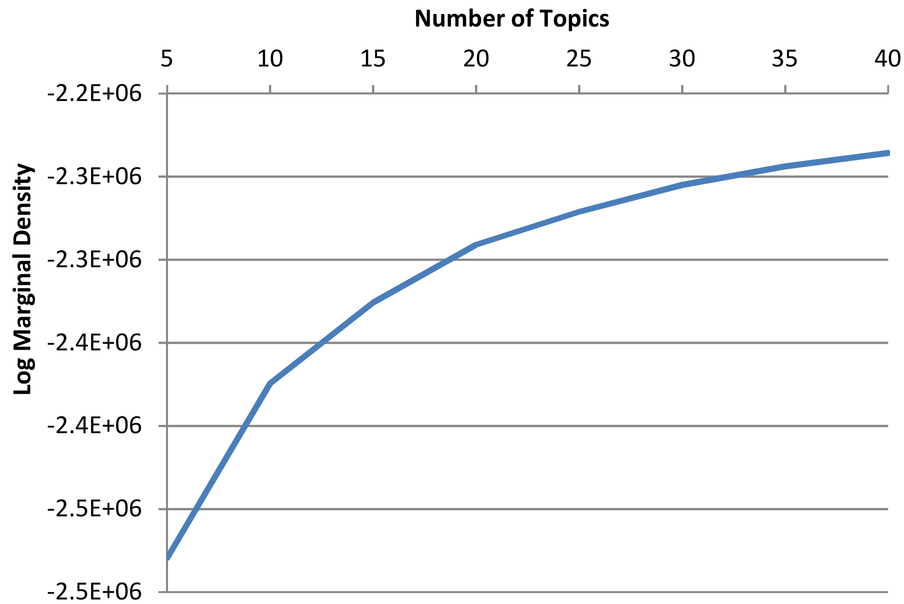
Given the prior of covariance matrix $\boldsymbol{\Sigma}$, $\pi(\boldsymbol{\Sigma}) \sim IW(\nu_{\boldsymbol{\Sigma}}, S_{\boldsymbol{\Sigma}})$, the posterior can be sampled by

$$P(\boldsymbol{\Sigma}|\boldsymbol{\gamma}) \sim IW\left(\nu_{\boldsymbol{\Sigma}} + T, S_{\boldsymbol{\Sigma}} + \sum_{t=1}^{T}(\boldsymbol{\gamma}_t - \boldsymbol{\gamma}_{t-1})(\boldsymbol{\gamma}_t - \boldsymbol{\gamma}_{t-1})'\right) \tag{34}$$

To reduce the autocorrelation, the estimates from Step 1 to Step 7 are sampled every 20 iterations.

Appendix A.3. In-sample Model Fit

Figure A.1. Log Marginal Density (LMD) by Number of Topics



Appendix A.4. Results and Discussions of the Remaining Categories

In the “Setup” category, as shown in Table 2, the average prevalence of Topic 7 “Installation” which introduces how to replace the old thermostat and install the Nest thermostat, and Topic 9 “Power” which describes how to the connect wires to Nest thermostat, is lower for the 3rd generation (2.2% and 1.5%) than in the previous two generations. By contrast, the average prevalence of Topic 8 “Instruction”, which addresses how to install the product by self-teaching, peaks at the 3rd generation (5.5%), indicating that the reviewer contributors of the earlier two generations mean to share more details about how to get started with the newly released product, whereas their counterparts of the 3rd generations refer later customers to the available video instructions for installation. Figure A.2 shows the temporal pattern of the prevalence of topics in the “Setup” category.

[Insert Figure A.2 Here]

Figure A.2 shows that all topics in the “Setup” Category became less prevalent within each generation over time, suggesting that earlier review contributors within each generation may be more enthusiastic about how they started using the product than later contributors. This trend may also reflect an increasing familiarity with the technology. The trends of the evaluations of topics in the “Setup” category are shown in Figure A.3.

[Insert Figure A.3 Here]

Unlike in the “Features” category, the evaluation of topics in the “Setup” category changes less and does not show a consistent trend of evolution.

The “Praise” category only contains Topic 10 “Compliment”, which commends the Nest thermostat in a holistic manner. The increasingly average topic prevalence over generations shown in Table 2 indicates that review contributors become more and more satisfied with the product in general. The trends of the topic prevalence and topic evaluation within each generation are shown in Figure A.4 and Figure A.5, respectively.

[Insert Figure A.4 Here]

[Insert Figure A.5 Here]

As shown in Figure A.4, the prevalence of Topic 10 “Compliment” on the 2nd and 3rd generations started at a low value for a couple of months before rising, suggesting that earlier review contributors of each generation were more cautious about praising the product than were later review contributors. Also, the release of the succeeding generation extinguished compliments on the preceding generation (Figure A.4). Given the stable topic evaluation (Figure A.5), we find that the release of a new generation product harms the ratings of the old generation product.

The “Criticism” category includes Topic 13 “Compatibility” which reports the issue of the Nest thermostat’s incompatibility with the old thermostat system, Topic 14 “Connection” which complains about the issue of Wi-Fi connection to the Nest thermostat, Topic 15 “Temperature Feels” which criticizes the disparity between the set temperature and “feels like” temperature, and Topic 16 “Unit Failure” which reports the failure of the product and its return or replacement of the product. In Table 2, we see declining trends of average topic prevalence in this category over generations, implying that these issues are attenuated with the release of newer generations. When we take a further look at the temporal pattern of prevalence shown in Figure A.6; however, we gain more insights into how the prevalence of these issues changes within each generation of product.

[Insert Figure A.6 Here]

The prevalence of Topic 13 “Compatibility” (Figure A.6a) on the 1st generation of product gradually increases over time, but stays low in the 2nd and the 3rd generations. This is due to the fact that Nest redesigned the 2nd generation thermostat to be compatible with 95% heating and cooling systems in use, more so than the 1st generation's compatibility of 75% (Hildenbrand 2017). With the release of the 2nd and 3rd generations, customers begin to have high expectations that their HVAC system would be compatible with the Nest thermostat, which leads to more complaints from the buyers of the 1st generation. In addition, both the 1st and the 2nd generations experience a bump of the prevalence of Topic 14 "Connection" (Figure A.6b) in the winter months of 2013 – 2014 because the release of the 4.0 firmware resulted in the loss of Wi-Fi connectivity (Constine 2014). Lastly, the prevalence of both Topic 15 "Temperature Feels" (Figure A.6c) and Topic 16 "Unit Failure" (Figure A.6d) on the 1st and the 2nd generation thermostats consistently increases but stays constant on the 3rd generations. This is either because later review contributors care more about product quality or else the product quality gets worsens with market expansion. Figure A.7 shows the temporal pattern of the evaluation of each topic in the "Criticism" category.

[Insert Figure A.7 Here]

In general, the curves of the topic evaluation in the "Criticism" category either decrease or flatten, indicating that review contributors have not shown tolerance to user issues over time. Given the increasing prevalence and non-increasing evaluation, the findings in the "Criticism" category warn the company that both product quality and service quality are critical to maintaining good review ratings in the long-term.

The "Financial" category encompasses Topic 17 "Price" in which review contributors opine on the fairness of the expense of acquiring a Nest thermostat, and Topic 18 "Savings" which reports the cost saved by rebate programs, gas usage, or electricity usage. As shown in Table 2, compared with the 1st and the 2nd generations of product, the 3rd generation possesses a lower average prevalence of Topic 17 “Price” most likely because the review contributors of the newest generation perceive higher value of the product than their counterparts of earlier generations. The increasing average evaluation of this topic over generations provides supporting evidence to the improvement of perceived value. By contrast, the evaluation of Topic 17 “Price” on the 1st generation is negative, mainly because review contributors of this generation believe the thermostat’s price is high and thus doubt the cost effectiveness of buying the product. Figure A.8 and Figure A.9 depicts the temporal trends of topic prevalence and topic evaluation in the “Financial” category, respectively.

[Insert Figure A.8 Here]

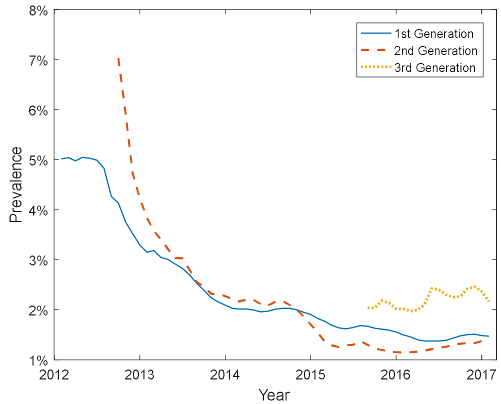
[Insert Figure A.9 Here]

Notably, the prevalence of Topic 17 “Price” (Figure A.8a) on the 1st generation rises after the 2nd generation is released. On the other hand, Figure A.9a shows that the evaluation of Topic 13 “Price” on the 1st generation started from a negative value and continues to decline over time. A possible explanation is that the earlier review contributors of the 1st generation doubted the value of the product at its original price. As the later generations are released on the market, the perceived value of the 1st generation causes more criticism and the trend declines even though the price is reduced. The declining evaluation of Topic 13 “Price” happens in the 2nd generation as well. As the difference of price between the two successive generations of the Nest thermostat is about $20, Nest may want to reconsider their pricing strategy across generations to improve their customers’ perceived value as well as review ratings.

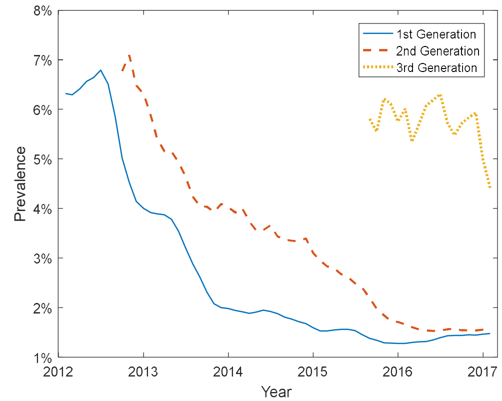
Table A.1 Simulation of Static SLDA with Multiple Text Streams

Stream	Topic	Parameters	True Value	Est	95% Interval	Parameter	True Value	Est	95% Interval
	1	α_1^1	1.65	1.64	[1.60, 1.68]	γ_1^1	1.50	1.44	[1.40, 1.48]
	2	α_2^1	1.00	1.00	[0.97, 1.03]	γ_2^1	0.50	0.47	[0.42, 0.52]
1	3	α_3^1	0.61	0.61	[0.59, 0.62]	γ_3^1	0.50	0.51	[0.45, 0.58]
	4	α_4^1	0.00	0.00	[0.00, 0.00]	γ_4^1	-1.00	-1.82	[-2.48, -1.02]
	5	α_5^1	0.00	0.00	[0.00, 0.00]	γ_5^1	0.50	0.82	[0.12, 1.48]
	1	α_1^2	1.00	1.01	[0.98, 1.04]	γ_1^2	1.00	0.93	[0.82, 1.04]
	2	α_2^2	1.65	1.67	[1.63, 1.72]	γ_2^2	1.00	0.95	[0.86, 1.03]
2	3	α_3^2	1.65	1.65	[1.60, 1.70]	γ_3^2	1.00	0.90	[0.81, 0.98]
	4	α_4^2	1.65	1.69	[1.65, 1.73]	γ_4^2	-0.50	-0.44	[-0.52, -0.35]
	5	α_5^2	0.00	0.00	[0.00, 0.00]	γ_5^2	1.00	1.71	[1.02, 2.40]
	1	α_1^3	0.00	0.00	[0.00, 0.00]	γ_1^3	0.50	0.88	[0.28, 1.47]
	2	α_2^3	0.61	0.63	[0.58, 0.66]	γ_2^3	0.00	-0.02	[-0.17, 0.15]
3	3	α_3^3	1.00	0.99	[0.93, 1.03]	γ_3^3	1.50	1.36	[1.24, 1.48]
	4	α_4^3	1.00	1.02	[0.96, 1.07]	γ_4^3	0.00	-0.07	[-0.18, 0.06]
	5	α_5^3	1.65	1.66	[1.58, 1.72]	γ_5^3	1.50	1.47	[1.28, 1.56]

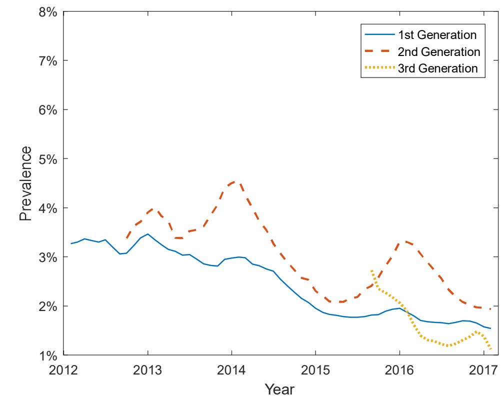
Figure A.2. Prevalence of Topics in the "Setup" Category



(a) Topic 7 "Installation"

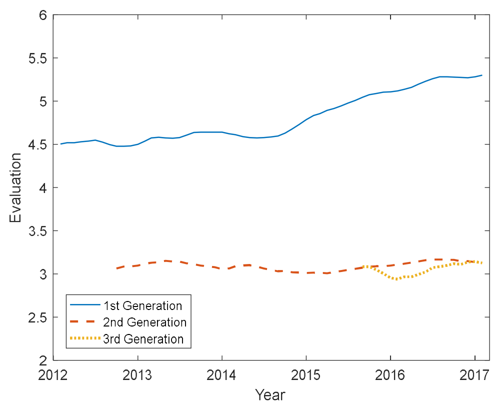


(b) Topic 8 "Instruction"

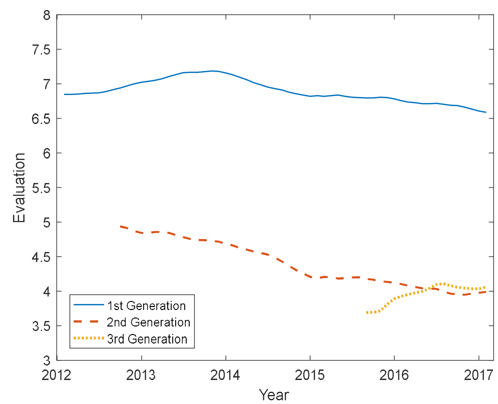


(c) Topic 9 "Power"

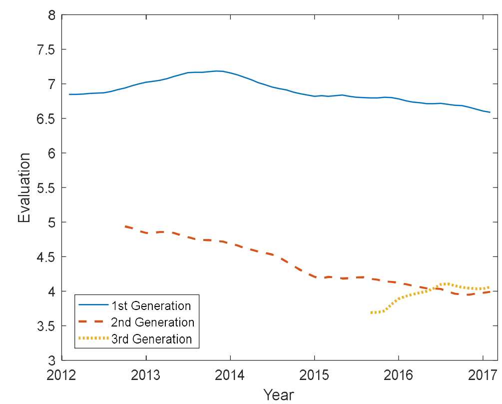
Figure A.3. Evaluation of Topics in the "Setup" Category



(a) Topic 7 "Installation"

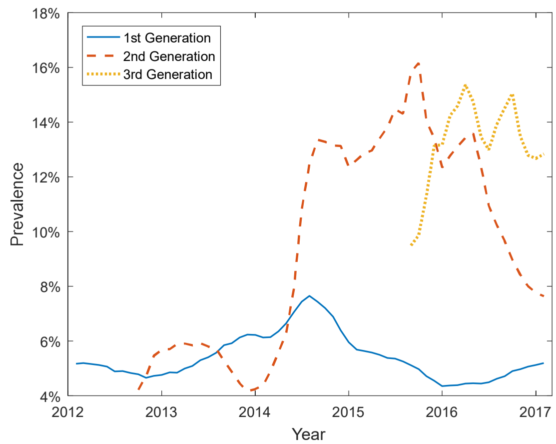


(b) Topic 8 "Instruction"



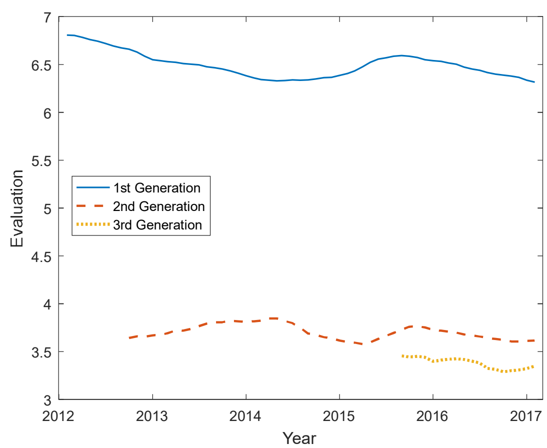
(c) Topic 9 "Power"

Figure A.4. Prevalence of Topics in the “Praise” Category



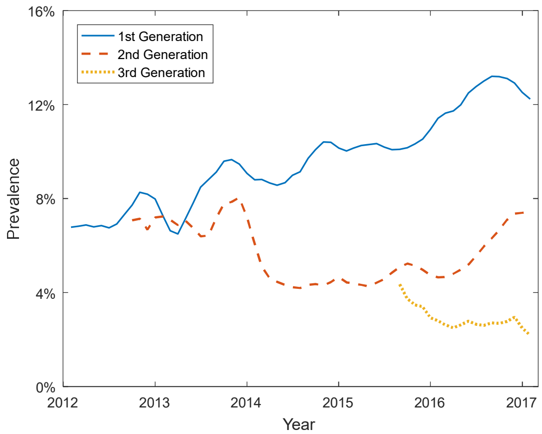
(a) Topic 10 “Compliment”

Figure A.5. Evaluation of Topics in the “Praise” Category

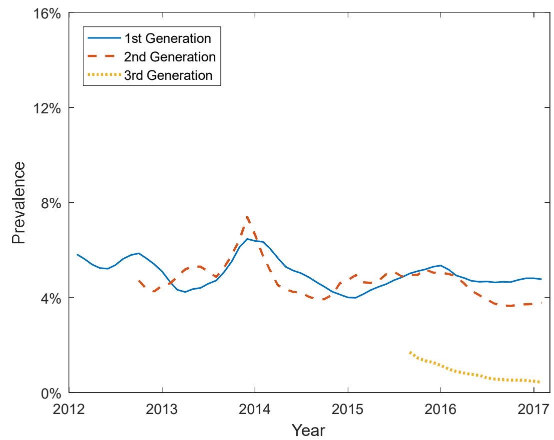


(a) Topic 10 “Compliment”

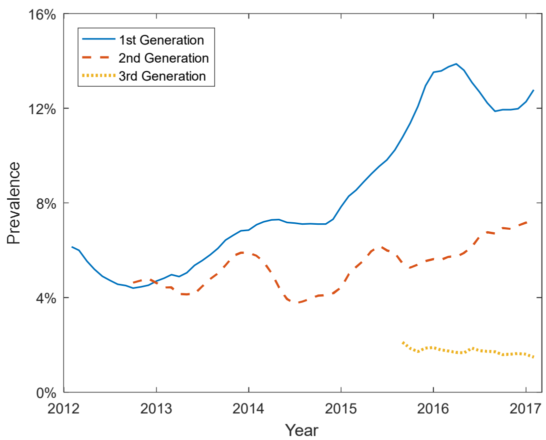
Figure A.6. Prevalence of Topics in the "Criticism" Category



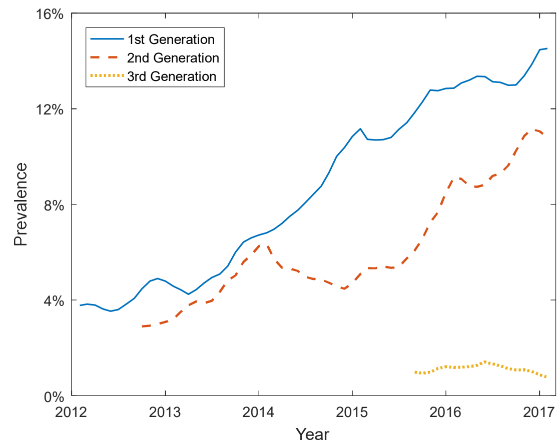
(a) Topic 13 "Compatibility"



(b) Topic 14 "Connection"

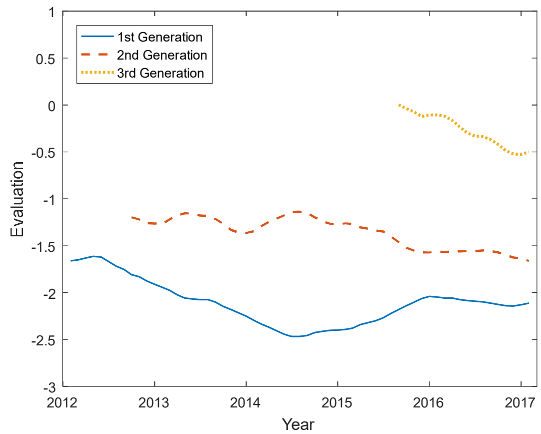


(c) Topic 15 "Temperature Feels"

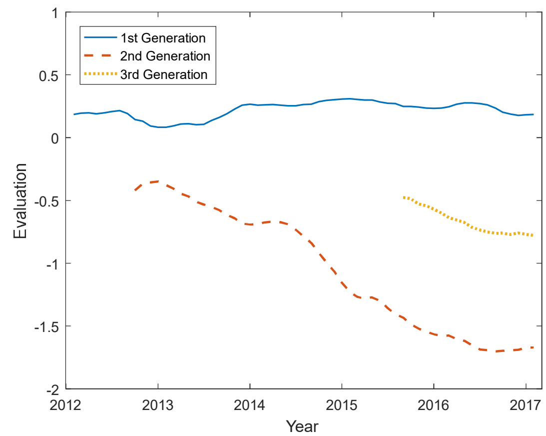


(d) Topic 16 "Unit Failure"

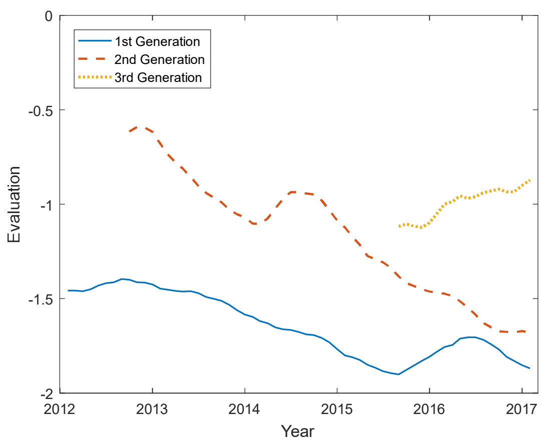
Figure A.7. Evaluation of Topics in the “Criticism” Category



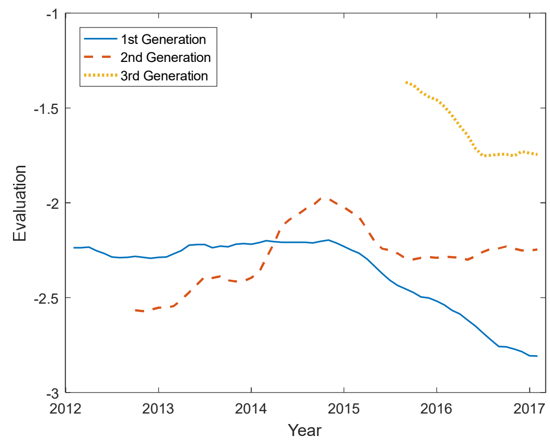
(a) Topic 13 “Compatibility”



(b) Topic 14 “Connection”

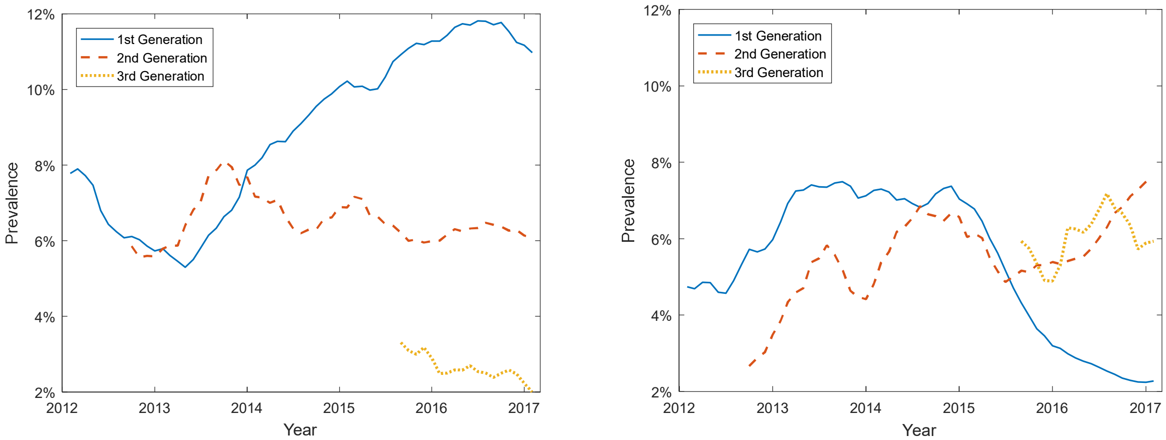


(c) Topic 15 “Temperature Feels”



(d) Topic 16 “Unit Failure”

Figure A.8. Prevalence of Topics in the “Financial” Category



(a) Topic 17 “Price” (b) Topic 18 “Savings”

Figure A.9. Evaluation of Topics in the “Financial” Category