POMS PRODUCTION AND OPERATIONS MANAGEMENT

Expertise and Identity Cues: The Dynamic Impact of Different Online WOM Sender Types on Product Demand

Journal:	*Production and Operations Management*
Manuscript ID	POM-Mar-24-OA-0368
Wiley - Manuscript type:	Original Article
Keywords:	Online Word of Mouth, Reviews, Critics, Message Sender, Demographics, Dynamic Linear Model

Abstract:	This study employs a Multivariate Dynamic Linear Modeling (DLM) approach to dissect how online reviews and critics distinguished by demographic factors like gender and ethnicity influence product demand dynamics over time. Analyzing data from two distinct categories—movies (hedonic) and cameras (utilitarian)—this research unveils the intricate roles of message source expertise and demographics in shaping consumer demand. There are several key findings. First, we find that both review and critic valences significantly boost demand, albeit with nuanced, time-sensitive effects that vary by product type. Critic valence exhibits enduring positive influence, contrasting with the transient impact of review valence in the movie sector. Second, our findings challenge conventional gender stereotypes; showing that the percentage of women reviewers and women critics have a positive impact on the product demand and the effect gets stronger over time for reviews written by women. However, we identify a concerning trend: reviews from minority groups depress sales figures, a bias somewhat alleviated when minority voices belong to expert critics. This points to the mitigating power of perceived expertise against racial biases since it provides a clue of competence. In addition, the interaction effect between reviewer demographics and review valence shows that the percentage of posts by women and minority group would strengthen the impact of valence on product sales except for the nonwhite reviewers. The findings related to the demographics stay consistent for both movie and camera category. Finally, our research provides important managerial and practical implications when it comes to recognizing the most influential strata of the reviewers or critics.

Expertise and Identity Cues: The Dynamic Impact of Different Online WOM Sender Types on Product Demand

ABSTRACT

This study employs a Multivariate Dynamic Linear Modeling (DLM) approach to dissect how online reviews and critics distinguished by demographic factors like gender and ethnicity influence product demand dynamics over time. Analyzing data from two distinct categories—movies (hedonic) and cameras (utilitarian)—this research unveils the intricate roles of message source expertise and demographics in shaping consumer demand. There are several key findings. First, we find that both review and critic valences significantly boost demand, albeit with nuanced, time-sensitive effects that vary by product type. Critic valence exhibits enduring positive influence, contrasting with the transient impact of review valence in the movie sector. Second, our findings challenge conventional gender stereotypes; showing that the percentage of women reviewers and women critics have a positive impact on the product demand and the effect gets stronger over time for reviews written by women. However, we identify a concerning trend: reviews from minority groups depress sales figures, a bias somewhat alleviated when minority voices belong to expert critics. This points to the mitigating power of perceived expertise against racial biases since it provides a clue of competence. In addition, the interaction effect between reviewer demographics and review valence shows that the percentage of posts by women and minority group would strengthen the impact of valence on product sales except for the nonwhite reviewers. The findings related to the demographics stay consistent for both movie and camera category. Finally, our research provides important managerial and practical implications when it comes to recognizing the most influential strata of the reviewers or critics.

Keywords: Online Word of Mouth; Reviews; Critics; Message Sender Expertise; Message Sender Demographics; Multivariate Dynamic Linear Model; Bayesian

1. INTRODUCTION

In any communication, there are three major components of message source or the sender of the message, the message receiver, and the message itself (Katz and Lazarsfeld 1955; Money, Gilly, and Graham 1998; Brown and Reingen 1987). We mainly focus on message senders in this research, and more specifically, investigate the message source from the two perspectives of expertise and demographics. Regarding expertise, the extant literature in the online word of mouth (OWOM) domain suggests two main message sources: customer reviews and professional critics. Reviewers are mostly individuals who share their product experience with other customers, however critics are independent publishers, with more expertise in evaluating the product performance. Reviewers are more likely to be targeted by brands and online platforms to act as community influencers, which reflects the highly connected nature of online consumers, however critics are professional entities with more expertise than individual reviewers and thus, less likely to be targeted by brands or platforms. Another important factor of any message sender is their demographic characteristics, for example their gender or ethnicity. Previous research has shown that revealing identity-related information improves the credibility of the online post[1], but does it hold even when the reviewer or critics belongs to a minority group (e.g., female; non-white)? And finally, understanding the impact of OWOM over time is vital for both researchers as well as managers, as information diffusion by critics and reviewers varies in the short term and long term and can influence consumers in different ways.

Online reviews, including those from critics and consumers, play a vital role in consumer decision-making. However, the two groups are different in their level of expertise that eventually leads to user reviews being spontaneously generated and critical reviews being more in-depth analysis about the quality of a product (Chiu et al. 2022). In addition to the higher level of experience, critics are mostly independent publishing agencies and less likely than individual reviewers to be targeted by brands to act as influencers. Also, critics typically are not incentivized to generate more posts and their goal is to produce a quality post that signals their expertise in the specific product domain or category. Hence, critics focus on improving their category experience and deepening their category knowledge instead of a general knowledge across different product categories.

In addition to the category expertise, message senders differ on identity-related dimensions including demographic characteristics resulting in heterogeneity among the group. Individuals span various ages, income levels, gender and racial groups, thus creating a diverse set of senders with different messages

[1] https://websitebuilder.org/blog/online-review-statistics/

to convey. Not all the reviewers are willing to reveal their identity due to privacy concerns, though previous research discusses scenarios where identity cues can affect consumer's decision purchase (Li and Liang 2022). In this research, we explore the role of gender and ethnicity as contributing factors to heterogeneity of message senders' influence on product sales. More specifically, we are interested in the interplay between domain expertise (critics vs reviewers) and belonging to a minority group (e.g., female; non-white). In the literature of Diversity, Equality and Inclusion (DEI), there is exploration of how the audience diversity can moderates the impact of social media content (Barasch and Berger 2014; Adamopoulos, Ghose, and Todri 2018) and documentation of the information bias against minority groups (Inci et al. 2017; Bhattacharya et al. 2020), but the heterogeneity of OWOM producers and the potential bias against content generated by minority groups have not been studied in current literature. Therefore, in this study, we take one step further to explore the interplay between expertise and the identity of minority groups.

For this area of research, it is vital to consider dynamic effects of OWOM as impact varies in both the short-term and long-term regarding product demand. Understanding how this phenomenon occurs over time offers insight into how different senders (reviewers and critics) may exert power over consumers and build lasting attitude shifts based on the information they share. We use a dynamic model to investigate how the effects related to message source expertise and its gender or ethnicity change over time. Specifically, we use a multivariate dynamic linear model (DLM) within a Bayesian estimation framework to understand the heterogeneous effects of message senders (reviewers and critics) on product sales across different minority strata based on gender and ethnicity (e.g., female; non-white). This is the first research to the best of our knowledge that addresses both aspects of expertise and identity cues of a message source in their influence on product demand. To obtain additional insights, our DLM framework allows us to explore the evolution of such effects over time while considering different heterogeneities among message senders.

To explore these research questions, we collected data for both hedonic and utilitarian products (movies and cameras). The movie dataset consists of the reviews, critics and sales performance for 93 movies that were screened in theaters in 2022. We collect their box office data from the-numbers.com, and their reviews and critics from rottentomatoes.com posted within 30 days of the release of the movie. The camera dataset is a panel data for 64 new cameras released in 2020, 2021 and 2022. We collected google trend search data as a proxy for camera sales, the consumers' review data from bhphotovideo.com and the critics data from articles under google review with a time range up to 6 months after introduction of the camera accounting for the longer product lifecycle compared to movies.

Our study makes two important contributions to the literature on OWOM. First, we show that both reviews and critics' opinions matter for the box office sales. Also, looking at the impacts of reviewers and critics over time we find that there is heterogeneity across different product categories: for movies, the effect of audience valence dies out after some time, but the impact of critics' valence stays more stable over time. For cameras, valence of reviews matters more compared with valence of critic and that the impacts of these two types of OWOM stays relatively stable over time. As presented in our contribution table 1, there are several studies about the impact of message source expertise (e.g., Hoskins et al. 2021; Chakravarty et al. 2010) however, none of them investigate the time-varying effects of reviewers and critics on sales over time. The two studies (Gelper et al. 2018; Gopinath et al. 2014) that use dynamic modeling consider only online forum posts and not critics.

Our work also contributes to the literature of Diversity, Equality, and Inclusion (DEI) by investigating the interplaying role of message source demographics (e.g., gender and ethnicity). When it comes to reviewer's identity, most of the past research explores if (or not) a reviewer reveals their real name, and how it affects content generation behavior, etc. (e.g., Pu et al. 2020; Forman et al. 2008). There are few articles looking at other aspects of reviewer's demographics. For example, Dellarocas et al. 2007 capture the entropy of gender distribution of reviewers as a control variable in their model. When it comes to ethnicity, there are only two experimental studies looking at the role of reviewer's race in review credibility or purchase intention (Azer et al. 2023; Lin and Xu 2017), but they only consider reviewers and not critics.

We are mainly interested to explore the interplay between being part of a minority group (e.g., women; non-white) and message source expertise. Considering the prevalent gender and racial stereotypes, one might expect message source expertise to act as a counter-stereotype clue and mitigate such biases. Interestingly, we didn't find a stereotype towards reviews or critics written by women in either movie or camera category, but we see that non-white critics have a much more positive influence on product sales than non-white reviewers, supporting the mitigating role of expertise in racial stereotype. Finally, we look at their effects on demand over time to track all the changes in both short-term and long-term. With respect to the demographic characteristics, the positive impact of women reviewers on product sales gets stronger, while women critics and critics from non-white experts stay stable over time. These findings provide helpful managerial insights when it comes to recognize the most influential groups of reviewers and critics at different stages after product introduction.

The rest of the paper is organized as follows. First, we provide a review of related previous literature, and then describe the data and the details of our empirical model and our estimation results. We then conclude and discuss managerial implications followed by future research paths in the final section.

2. Literature Review

2.1 Expertise of Message Source

Previous literature has mainly looked at two types of message source from the angle of expertise, customer reviews and professional critics. Most of the studies looking at the impact of customer reviews found a positive (negative) impact of valence (standard deviation) of prior community reviews (Godes & Mayzlin, 2009; Gopinath et al., 2014; Luo 2009; Moon et al. 2010) on product sales. In addition, previous research has documented a positive influence of customer reviews' volume (Dhar and Chang 2009; Duan et al. 2008; Godes and Mayzlin 2004; Liu 2006; Luo 2009; Gopinath et al. 2014) on a consumer's brand evaluations, trial, and purchase decisions.

Studies exploring the role of critics' valence have shown a positive impact on consumer attitudes towards the product (Chakravarty et al. 2010; Chen et al. 2012; Cox and Kaimann 2015). Although there is extensive research about the importance of online reviews on customer actions towards brand/product, there is only limited research about how differently audience and critics groups write reviews for products and the differential impact of their posts on product sales. For example, Deng 2020 and Kim et al. 2023 show that both types of reviews play distinct roles in influencing movie box office sales, with user reviews guiding perceptions of quality and critic reviews providing depth and context. They suggest that moviegoers place considerable trust in the opinions and analyses provided by critics, possibly perceiving them as less susceptible to bias compared to user reviews. This can be also due to how differently audience and critics groups write reviews for movies. Critics and consumers use different languages and techniques in their reviews. Critics provide detailed, in-depth analysis, while consumers use fewer words and simpler language[2].

Such distinction highlights the different roles these reviews play in influencing audience decisions, with critic reviews seen as more reliable and insightful than consumer reviews. In essence, these findings suggest that while audience reviews are crucial for gauging immediate reaction and emotional connection with audiences, critic reviews hold a significant place in shaping a movie's reputation and perceived quality, particularly among audiences seeking more nuanced insights. The Elaboration Likelihood Model (Petty & Cacioppo, 1986) proposes that the likelihood of an individual be impacted by external information sources

[2] https://relativeinsight.com/using-language-analysis-to-compare-critic-and-consumer-movie-reviews/#:~:text=Consumers%20seek%20excitement%20and%20suspense,the%20eyes%20of%20the%20viewer

depends on the degree to which the information is considered relevant. This suggests that consumers of online reviews overall are likely to go through the reviews and pay attention to the information shared with them. Therefore, the message source expertise and credibility is likely to play an important role, which is reflected in the consensus among most previous researchers that critics have a bigger impact than consumer reviews on product sales (e.g. Chakravarty et al. 2010; Cox and Kaimann 2015).

Table 1 shows how our research contributes to the extent literature in this domain. Although most of the articles use a static approach there are few articles exploring the time-varying effects of online reviews on product sales (e.g. Gelper et al. 2018; Gopinath et al. 2014). However, they do not investigate the differential dynamic effects of reviewers and critics.

Table 1: Contribution Table

	Experience of Message Source - Critics and User Reviews		Demographics of Message Source				Methodology - Empirical/ Experiment	Product Category		DV
			Gender		Ethnicity					
	Overall Impact	Time Varying Impact	Overall Impact	Time Varying Impact	Overall Impact	Time Varying Impact		Utilitarian	Hedonic	
Our Study	Yes	Yes	Yes	Yes	Yes	Yes	Empirical	Yes (Cameras)	Yes (Movies)	Sales, Google Search
Azer et al. 2023	User Reviews Only	No	No	No	Yes	No	Experiment	No	Yes	Review Credibility
Chakravarty et al. 2010	Yes	No	No	No	No	No	Both	No	Yes	Sales
Chen et al. 2012	Critics Only	No	No	No	No	No	Empirical	No	Yes	Firm Value
Chiu et al. 2022	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Cox and Kaimann 2015	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Dellarocas et al. 2007	Yes	No	Entropy of gender distr. Of reviewers	No	No	No	Empirical	No	Yes	Sales
Deng 2020	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Gelper et al. 2018	User Forum Posts Only	Yes	No	No	No	No	Empirical	No	Yes	Sales
Gopinath et al. 2014	User Forum Posts Only	Yes	No	No	No	No	Empirical	Yes	No	Sales
Hoskins et al. 2021	Yes	No	No	No	No	No	Empirical	No	Yes	Reviewer Rating
Kim et al. 2023	Yes	No	No	No	No	No	Both	No	Yes	Sales
Lin and Xu 2017	User Reviews Only	No	No	No	Yes	No	Experiment	No	Yes	Purchase Intention
Niraj et al. 2015	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Wang et al. 2015	Yes	No	No	No	No	No	Empirical	Yes	Yes	Sales

2.2 Identity of Message Source

There are several studies in the domain of social networks and OWOM looking at the impact of real name disclosure on user's content generation behavior in terms of posting volume or linguistic style of the posts (Pu et al. 2020; Kinler and Hoadley 2005; Leshed 2009; Cho et al. 2012; Paskuda and Lewkowicz 2017; Huang et al. 2017); finding that when the identities of online users are revealed, they tend to share information that is more socially acceptable, produce fewer critical remarks, and less often employ emotional or unfiltered language. Some other studies find a positive effect for such identity reveals on important outcomes e.g., customer satisfaction or product sales (Chung et al. 2023; Forman et al. 2008). Forman et al. 2008 find that reviewers who reveal their real names or their location have a stronger effect on the review helpfulness votes and subsequently product sales.

Reviewer's name is the most common identity cue that has been studied in previous articles. Among demographic characteristics, although gender and ethnicity are used frequently to develop homophily measures in the online community (e.g. Zhang and Hanks 2018; Kakar et al. 2018), there are few studies exploring their effects on product sales. For example, there is only one study controlling for the gender of the reviewer, measuring it as the entropy of gender distribution among the reviewers (Dellarocas et al. 2007). Another important demographic characteristic is the ethnicity of reviewers, which has received very limited attention in this domain. As presented in Table 1, there are only two experimental studies (Lin and Xu 2017; Azer et al. 2023) exploring the effect of reviewer's ethnicity on purchase intention and reviewer's credibility finding that white reviewers are perceived to be more credible than black or Asian ones.

Considering the prevalent gender and racial stereotypes, one of our main focuses in this research is the interplaying role of expertise and demographics of the message source on product demand. There are established literature about gender stereotypes showing that women compared to men get less credibility in the workplace (e.g., Eagly and Carli 2007), or in scientific fields (e.g., Knobloch-Westerwick et al. 2013). Others show evidence that women's reports of pain are more likely to be dismissed (e.g., Hoffmann and Tarzian 2001), or women's testimonies can be subject to greater scrutiny and skepticism (e.g. Starr 2012). Inci et al. (2017) demonstrates that, although having identical formal status, female executives are at a disadvantage to men in terms of access to inside knowledge. Bhattacharya et al. (2020) show women are at disadvantage when purchasing financial products: financial planners are significantly more likely to give undiversified financial advice to female than to male auditors. Similarly, there is evidence for racial stereotypes in different contexts. For example, black compared to white individuals are less likely to receive appropriate pain management (e.g. Hoffman et al. 2016), or to be treated fairly in justice system or educational settings (e.g. Eberhardt et al. 2004; Okonofua and Eberhardt 2015). Bhutta et al. (2022) find that minority mortgage applicants are less likely than white applicants to receive algorithmic approval from race-blind government automated underwriting systems (AUS).

To the best of our knowledge there is no empirical study looking at the impact of gender and ethnicity of reviewers on product sales. Moreover, no study has explored such stereotypes among reviewers and critics and how they vary over time. Although revealing some identity related information such as gender or ethnicity (in case reviewer belongs to a minority group such as women or non-white) can lead to stereotypes and perhaps decrease the credibility of the review consequently, it is not clear if such effect is the same for critics reviewers. In other words, it is possible that message source expertise has a mitigating role for the gender or racial stereotypes. There are relevant research showing that competence and expertise can be a counter-stereotypic clue, which can mitigate gender or racial stereotypes. For example, Heilman and Okimoto (2007) suggest that when women demonstrate clear competence in traditionally male domains, it can mitigate negative stereotypes and alter perceptions about women's abilities. Similarly, Owens and Lynch (2012) highlighted that black professionals who demonstrate high competence in their fields can alter perceptions held by their white colleagues, mitigating racial biases and stereotypes.

Thus, considering that message source expertise can be a potential mitigation for gender and racial stereotypes, we explore the differential effects of women critics vs women reviewers and non-white critics vs non-white reviewers on product demand in both hedonic (movies) and utilitarian (cameras) context. In addition, we study the time-varying effects of such interplay to see if the stereotypes or the mitigating role of expertise has a lasting effect on demand, or if it doesn't follow a stable pattern over time since more than one factor is involved (message source expertise and prevalent stereotypes).

3. Data and Measurement

The first dataset for this study consists of 93 movies that were in theaters in 2022. We select the movie category since they are a type of hedonic good for which both online reviews and critics should be highly relevant, and their sales are known to peak shortly after release (Gopinath et al. 2013). Hence, we collect daily data for a one-month period after the movie release date.

We collected movie box office data from the-numbers.com, which is a website that tracks box office revenue. In addition to the movie box office data, we collect the audience review and critics data from rottentomatoes.com, which is a website that aggregates movie reviews. For audience review, it provides a rating from 1-5. For critics, it has two categories "fresh" and "rotten", where "fresh" means a positive sentiment towards the movie and "rotten" means negative attitude towards the movie.

Similar to prior research, we define the valence of audience reviews ($VALENCE\_REV$) as the cumulative average rating of posted reviews and volume of audience reviews ($VOLUME\_REV$) as the total number of posted reviews. Similarly, for valence of critics, we define ($VALENCE\_CRIT$) as the percentage of "fresh" critics. We further break down the audience reviews and expert critics by their gender and racial groups: ($PERCENTAGE\_REV\_WOMEN$) and ($PERCENTAGE\_CRIT\_WOMEN$) are the percentage of reviews by women audience and percentage of women critics, respectively. Similarly, ($PERCENTAGE\_REV\_NONWHITE$) is the percentage of reviews by nonwhite audience and ($PERCENTAGE\_CRIT\_NONWHITE$) is the percentage of nonwhite critics.

In addition, we control for other factors that would affect the box office sales of the movies: AGE is the number of days since the release of the movie; number of theaters movies were screened $THEATER$; number of competing movies of the same genre $COMPETITION$ as well as weekend dummy $WEEKEND$. Table 2 shows the descriptive statistics for the different variables for the movie dataset.

Table 2: Descriptive Statistics (Movie Dataset)

Variables	Definition	Mean	Std. Dev.	Min	Max
BOX_OFFICE_{it}	Box office sales for movie i on day t	1800000	4770000	302	73941280
$VALENCE\_REV_{it}$	Average rating of reviews for movie i until day t	3.957	.628	2.167	5
$VOLUME\_REV_{it}$	Total number of audience reviews for movie i until day t	1814.182	3560.852	0	28486
$VALENCE\_CRIT_{it}$	Percentage of "fresh" critics ratings for movie i until day t	72.105	23.525	3.226	100
$VOLUME\_CRIT_{it}$	Total number of critics ratings for movie i until day t	167.007	97.888	1	424
$PERCENTAGE\_REV\_WOMEN_{it}$	Percentage of reviews by women for movie i until day t	10.863	5.828	0	35
$PERCENTAGE\_REV\_NONWHITE_{it}$	Percentage of reviews by nonwhite audience for movie i until day t	2.63	1.786	0	20
$PERCENTAGE\_CRIT\_WOMEN_{it}$	Percentage of women critics for movie i until day t	22.962	6.68	0	50
$PERCENTAGE\_CRIT\_NONWHITE_{it}$	Percentage of nonwhite critics for movie i until day t	5.357	2.346	0	12.5
AGE_{it}	Number of days since the release of movie i until day t	16	8.368	2	30
$THEATER_{it}$	Number of theaters movie i is aired on day t	2099.895	1541.849	2	4751
$COMPETITION_{it}$	Number of competing movies of the same genre as movie i screened on day t	3.777	2.781	0	13
$WEEKEND_{t}$	Whether day t is Friday, Saturday, or Sunday	.445	.497	0	1

The second data set is from a utilitarian product, cameras, to investigate any heterogeneity across product categories. The camera dataset for this study comprises of 64 new cameras introduced to the market in 2020, 2021, and 2022. Unlike movies, the sales of cameras typically take longer to peak after release. Therefore, we collected weekly data for a six-month period following the introduction date.

We collected Google Trends search data as a proxy for camera sales. Additionally, we gathered consumer reviews from bhphotovideo.com, a website that sells digital products similar to Amazon. Furthermore, we collected expert critics' reviews from Google Reviews. For both consumer reviews and critics ratings, valences are measured by ratings ranging from 1 to 5. We collected reviews and critics posted until six months after the camera was introduced.

We use the Google search of camera i at week t ($SEARCH$) as a proxy of camera sales. The other main variables are defined in a similar fashion as those in the movie category except for $VALENCE\_CRIT$, which is now measured in the 1-5 scale rather than the "fresh" and "rotten" in movie category. Similarly, we control for other factors that would affect the Google searches of the cameras: AGE is the number of days since the introduction of the product; number of competing cameras $COMPETITION$ as well as month dummy $MONTH$. Table 3 shows the descriptive statistics for the different variables for the camera dataset.

Table 3: Descriptive Statistics (Camera Dataset)

Variables	Definition	Mean	Std. Dev.	Min	Max
$SEARCH_{it}$	Average Google trend search for camera i in week t	10.547	8.381	0	82
$VALENCE\_REV_{it}$	Average rating of reviews for camera i until week t	4.641	.409	2	5
$VOLUME\_REV_{it}$	Total number of posted reviews for camera i until week t	33.745	68.693	1	407
$VALENCE\_CRIT_{it}$	Average rating of critics for camera i until week t	4.32	.341	3.188	5
$VOLUME\_CRIT_{it}$	Total number of critics posts for camera i until week t	20.828	12.681	2	84
$PERCENTAGE\_REV\_WOMEN_{it}$	Percentage of reviews by women for camera i until week t	26.126	22.164	0	93.75
$PERCENTAGE\_CRIT\_WOMEN_{it}$	Percentage of posts by women critics for camera i until week t	15.022	13.902	0	75
$PERCENTAGE\_REV\_NONWHITE_{it}$	Percentage of reviews by nonwhite reviewers for camera i until week t	29.187	24.052	0	100
$PERCENTAGE\_CRIT\_NONWHITE_{it}$	Percentage of posts by nonwhite critics for camera i until week t	14.771	13.032	0	75
AGE_{it}	Number of days since the introduction of camera i until week t	94.441	52.496	4	182
$COMPITITION_{it}$	Number of competing cameras for camera i in week t	11.678	4.434	0	20

4. Model Specification

There are different approaches to model social contagion process. The Bass diffusion model (Bass 1969) is a common social contagion model that considers several dynamic processes in the online communities' growth. Bass diffusion model has been used in marketing research for diffusion of innovation (Mahajan et al. 1990). We don't use the generalized Bass model for our research, since one of our main objectives is to estimate the time-varying effects of our key OWOM measures. Additionally, to obtain robust parameter estimation for the Bass model, it's required that the data includes the peak for the noncumulative adoption curve, which is not easy to assess a priori (Srinivasan and Mason 1986; Heeler and Hustad 1980). Prior researchers have also used vector autoregressive (VAR) approaches to model social contagion. VAR model assumes that endogenous variables can be explained by their own lagged variables and other endogenous lagged variables (Dekimpe and Hanssens 1999). Interpretation of VAR model variables can't be done by its own; we need to use impulse response functions (IRF) or conduct an elasticity analysis. Although using VAR model can be helpful in estimating short and long- term effects of variables, it ignores the dynamic parameter effects. In contrast, DLM accounts for evolution/nonstationary in the data, without data transformations, such as differencing. Van Herdee et al. (2004) discuss the benefits of DLM in comparison to other time-series models (e.g., VAR). First, the DLM parameters can be nonstationary. DLM allows for both a random walk with trend in the y_t and parameters through the observation and state equations. Second, DLM allows identifying the dynamic path of parameters over time as a function of their lags. Ataman et al. (2010) use DLM model to distinguish the short-term and long-term impacts of marketing strategies on sales.

Our econometric model is based on a multivariate DLM framework since it offers a flexible means to assess how the impacts of different OWOM measures evolve over time. Moreover, our model specification allows us to capture both cross-sectional heterogeneity (e.g., differences in parameters across demographics) and the longitudinal heterogeneity (e.g., changing parameters across time). Similar to prior researchers, first we model product demand as a function of explanatory variables in the observation equation. Second, we allow all the key parameters in the sales model to change over time in the state equation.

Observation Equations

For each movie i (i=1,…,N) and time t (t=1,…T), we have the observation equation system as below:

$$BOX\_OFFICE_{it} = \alpha 1_i + F11_t'\beta 11_t + F12_t'\beta 12 + v1_{it} \tag{1}$$

$$\begin{aligned} where\ F11 = [&VALENCE\_REV_{it-1}, VOLUME\_REV_{it-1},\ VALENCE\_CRIT_{it-1}, \\ &VOLUME\_CRIT_{it-1}, PERCENTAGE\_REV\_WOMEN_{it-1}, \\ &PERCENTAGE\_REV\_NONWHITE_{it-1}, PERCENTAGE\_CRIT\_WOMEN_{it-1}, \\ &PERCENTAGE\_CRIT\_NONWHITE_{it-1}] \\ F12 = [&AGE_{it}, COMPETITION_{it},\ THEATER_{it}, WEEKEND_{it}] \end{aligned}$$

$$VALENCE\_REV_{it} = \alpha 2_i + F21_t'\beta 21_t + F22_t'\beta 22 + v2_{it} \tag{2}$$

$$where\ F21 = [BOX\_OFFICE_{it-1}, VOLUME\_REV_{it-1}, VALENCE\_CRIT_{it-1}, VOLUME\_CRIT_{it-1}]$$
$$F22 = [AGE_{it}, COMPETITION_{it}, WEEKEND_{it}]$$

$$VOLUME\_REV_{it} = \alpha 3_i + F31_t'\beta 31_t + F32_t'\beta 32 + v3_{it} \tag{3}$$

$$where\ F31 = [BOX\_OFFICE_{it-1}, VALENCE\_REV_{it-1}, VALENCE\_CRIT_{it-1}, VOLUME\_CRIT_{it-1}]$$
$$F32 = [AGE_{it}, COMPETITION_{it}, THEATER_{it}, WEEKEND_{it}]$$

$$[v1_{it}, v2_{it}, v3_{it}] \sim MVT_v(0, V)$$

State Equations

$$\begin{aligned} \beta 11_t &= \beta 11_{t-1} + w_t^1, \ w_t^1 \sim MVN(0, W^1) \\ \beta 21_t &= \beta 21_{t-1} + w_t^2, \quad w_t^2 \sim MVN(0, W^2) \\ \beta 31_t &= \beta 31_{t-1} + w_t^3, \ w_t^3 \sim MVN(0, W^3) \end{aligned} \tag{4}$$

For the observation equations, $\alpha 1$- $\alpha 3$ are movie specific intercepts. These control for time-invariant unobserved product specific characteristics. F11, F21, F31 are matrices with independent variables with time-varying parameters $\beta 11_t$,$\beta 21_t$,$\beta 31_t$ respectively. F12, F22, F33 are matrices with independent variables with time-invariant parameters $\beta 12$,$\beta 22$,$\beta 32$ respectively. For camera analysis, we use a similar multivariate dynamic linear model structure with BOX_OFFICE_{it} replaced with Google search for movie i at time t ($SEARCH_{it}$). We no longer have $THEATER_{it}$ and $WEEKEND_{it}$. Instead we include $MONTH_t$ which controls for seasonality. The other variables are similar to the movie dataset. In the state equations, we allow the time-varying parameters to evolve using random walk specification. We estimate both the observation and the state equations jointly, assuming that the error terms in observation equations follow multivariate T distribution and that the error terms in state equations follow multivariate normal distribution. We place normal priors on all parameters of the observation equation. The state equation error covariance matrix is assumed to be diagonal, and we place an inverse Gamma prior on its diagonal elements.

5. Results

Our key research question is to investigate whether the source type of message sender has an impact on movie box office sales and the Google search of cameras. We look at both the overall impact of reviews and critics as well as the heterogenous impact of reviews and critics from different gender and racial groups.

The DLM model enables us to further explore how the impact of these message senders would change dynamically.

5.1 Impact of Review Valence and Reviewer Demographics

Overall Impact

Table 4A shows the overall impact. For time varying parameters these are based on estimates for all time periods. Column (1) shows the overall impact of reviews and critics on movie box office sales and column (2) shows the overall impact on the Google search of cameras.

Table 4A: Overall Impact of Reviewers and Critics on *BOX_OFFICE* and *SEARCH*

	Hedonic Product (Movies)	Utilitarian Product (Cameras)
	$DV: BOX\_OFFICE_{it}$ (1)	$DV: SEARCH_{it}$ (2)
	Mean [5th percentile, 95th percentile]	Mean [5th percentile, 95th percentile]
$VALENCE\_REV_{it-1}$	1.3606* [0.1349, 2.5865]	3.2409* [1.3579, 5.1238]
$VALENCE\_CRIT_{it-1}$	2.0481*[1.4305, 2.6658]	0.7510 [-2.0453, 3.5473]
$PERCENTAGE\_REV\_WOMEN_{it-1}$	5.0757*[4.7023, 5.4492]	4.6640* [4.3334, 4.9945]
$PERCENTAGE\_REV\_NONWHITE_{it-1}$	-2.1847*[-2.6435, -1.7260]	-1.2932* [-1.6542, -0.9322]
$PERCENTAGE\_CRIT\_WOMEN_{it-1}$	3.8223*[3.3790, 4.1966]	3.1615* [1.9044, 4.4185]
$PERCENTAGE\_CRIT\_NONWHITE_{it-1}$	4.0166*[3.4443, 4.5886]	3.6353* [2.4453, 4.8252]
$VOLUME\_REV_{it-1}$	3.3289* [3.1081, 3.5497]	2.5920*[2.2994, 2.8846]
$VOLUME\_CRIT_{it-1}$	3.8998*[3.7708, 5.1036]	1.4500* [0.1601, 2.7398]
AGE_{it}	-0.5339[-1.4044, 0.3367]	1.9976* [0.6798, 3.3155]
$COMPETITION_{it}$	0.1485 [-0.0445, 0.3415]	-0.0215 [-0.3281, 0.2852]
$THEATER_{it}$	0.0935*[0.0311, 0.1558]	-
$WEEKEND_{t}$	0.0129 [-0.1460, 0.1717]	-
$MONTH\ FIXED\ EFFECTS$	-	Yes
$PRODUCT\ FIXED\ EFFECTS$	Yes	Yes
Number of Observations	2,697	1,664

* The 90% confidence interval does not include zero.

The first key finding from Table 4A is that although both reviews (*VALENCE_REV*) and critics (*VALENCE_CRIT*) matter for box office sales, critics have stronger impact. In contrast, for cameras, only reviews valence has strong and positive impact on Google search. The volume of audience reviews (*VOLUME_REV*) and critics (*VOLUME_CRIT)* also have positive impacts on sales. However, similar to the valence impacts, the volume effects are weaker for the camera category.

When we look at the results for different gender groups, we find that women reviewers have a significant positive impact on movie box office sales and the Google search of cameras (5.0757 and 4.6640, respectively). Also, higher percentage of women critics increases box office sales as well as the search of cameras (3.8223 and 3.1615, respectively). The results suggest that there seems to be no gender bias towards women when it comes to online WOM in either the movie or camera category.

However, when it comes to racial groups, the result shows that higher percentage of reviews written by non-white audience has a negative impact on movie box-office sales (-2.1847), and that higher percentage of reviews by non-white reviewers have negative impact on consumers' interest in search for the camera (-1.2932), which indicates that there is potential racial bias towards non-white reviewers. However, when the reviews come from non-white critics (i.e. experts), the effect is mitigated and remain positive (4.0166 and 3.6353, respectively). This finding aligns with the fact that the identity of expert can mitigate the potential bias towards minority groups.

Table 4B and 4C shows the estimation results for *VALENCE_REV* and *VOLUME_REV*, respectively. From Table 4B we can see that overall, the product demand, volume of reviews, valence of critics and volume of critics in the last period have significant positive impact on the valence of reviews in current period. Table 4C shows similar pattern that the volume of reviews in current period is driven by product demand, valence of reviews and critics in the previous period.

Table 4B: Overall Impact on *VALENCE_REV*

	Hedonic Product (Movies)	Utilitarian Product (Cameras)
	$DV: VALENCE\_REV_{it}$ (1)	$DV: VALENCE\_REV_{it}$ (2)
	Mean [5th percentile, 95th percentile]	Mean [5th percentile, 95th percentile]
BOX_OFFICE_{it-1}	3.9222*[3.7425, 4.1019]	-
$SEARCH_{it-1}$	-	3.5888* [3.3476, 3.8300]
$VOLUME\_REV_{it-1}$	5.8553*[5.5724, 6.1381]	6.0593*[5.8554, 6.2632]
$VALENCE\_CRIT_{it-1}$	6.2714*[5.5887, 6.8964]	3.0300* [0.5485, 5.5115]
$VOLUME\_CRIT_{it-1}$	6.5527*[5.7230, 6.8197]	5.6886* [4.5952, 6.7820]
AGE_{it}	-0.0308 [-0.6671, 0.6057]	1.3194*[0.2308, 2.4080]
$COMPETITION_{it}$	-0.0470 [-0.2319, 0.1379]	0.0797 [-0.2086, 0.3682]
$WEEKEND_t$	-0.0511 [-0.2066, 0.1044]	-
MONTH FIXED EFFECTS	-	Yes
PRODUCT FIXED EFFECTS	Yes	Yes
Number of Observations	2,697	1,664

* The 90% confidence interval does not include zero.

Table 4C: Overall Impact on *VOLUME_REV*

	Hedonic Product (Movies)	Utilitarian Product (Cameras)
	$DV: VOLUME\_REV_{it}$ (1)	$DV: VOLUME\_REV_{it}$ (2)
	Mean [5th percentile, 95th percentile]	Mean [5th percentile, 95th percentile]
BOX_OFFICE_{it-1}	2.7048*[2.5635, 2.8460]	-
$SEARCH_{it-1}$	-	2.6063*[2.3499, 2.8626]
$VALENCE\_REV_{it-1}$	1.9131*[0.6158, 3.2105]	3.0033*[1.1685, 4.8381]
$VOLUME\_CRIT_{it-1}$	2.7482*[2.2007, 3.2958]	2.8548*[1.7382, 3.9713]
$VALENCE\_CRIT_{it-1}$	1.4652*[0.8268, 2.1036]	-1.5477 [-4.2684, 1.1729]
AGE_{it}	-0.0478 [-0.9441, 0.8484]	0.6356 [-0.5029, 1.7740]
$COMPETITION_{it}$	0.1396 [-0.0517, 0.3308]	0.0379 [-0.2423, 0.3180]
$THEATER_{it}$	-0.1500*[-0.2292, -0.0708]	-
$WEEKEND_t$	0.2011*[0.0342, 0.3681]	-
$MONTH\ FIXED\ EFFECTS$	-	Yes
$PRODUCT\ FIXED\ EFFECTS$	Yes	Yes
Number of Observations	2,697	1,664

* The 90% confidence interval does not include zero.

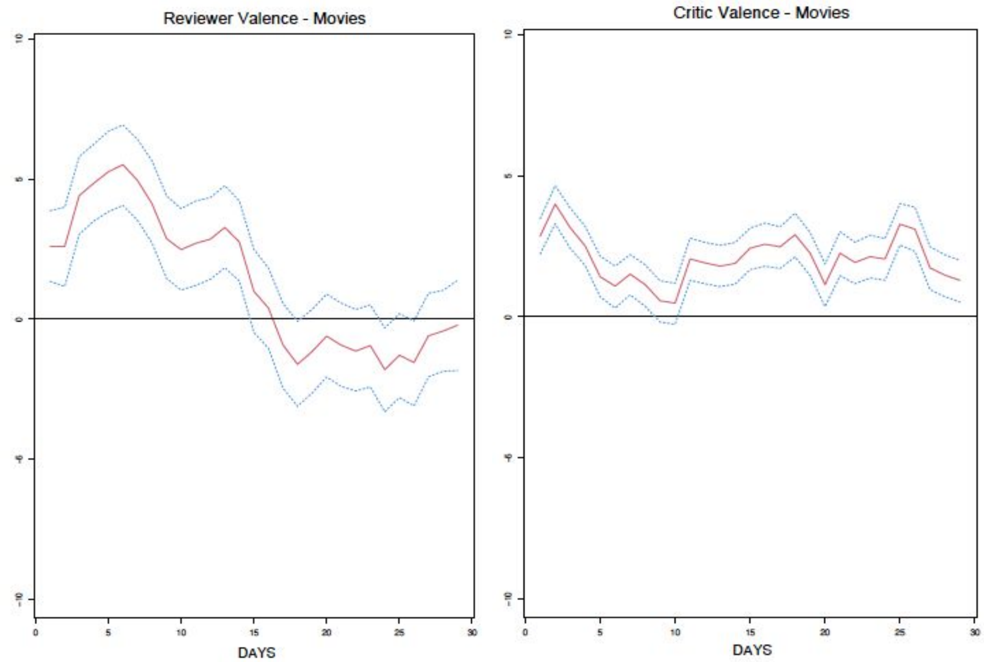
Time-Varying Impact

In addition to these overall impacts, our DLM framework allows us to arrive at the time varying impact of these different source types on consumer demand. Figures 1-6 shows these dynamic effects which provides a better understanding of how the roles of these important OWOM sources evolve over time after product introduction.

Figure 1 presents the time-varying impact of reviewers and critics valence. The first key result from the dynamic trend is: for movies, the impact of review valence on box office sales is decreasing over time and that the impact of critic valence on box office sales decreases in the early stage and then becomes more stable. The trend that we observed may be due to the fact that for new movies which are just released, there isn't much information available for potential consumers except for OWOM. Hence, consumers rely more on the review valence and critics valence in the early stage. Later, the information source becomes more diversified, so the impact of reviews and critics will decrease. For cameras, the impact of review valence is more significant compared critic valence, and the impacts are stable over time. This pattern may arise because the critics of cameras contains many technical evaluations, which might be difficult for non-expert consumers to understand. Therefore, consumers tend to refer to the reviews written by other consumers more when they make purchase decisions. Similarly, due to the complexity of technical specifications and features, consumers may find it challenging to evaluate these technical aspects themselves, which makes the valence of reviews a consistent and reliable factor in their decision-making process. Hence, the impact of review valence remains stable over time for cameras.

Figure 1: Time-Varying Impact of Reviewers and Critics Valence

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

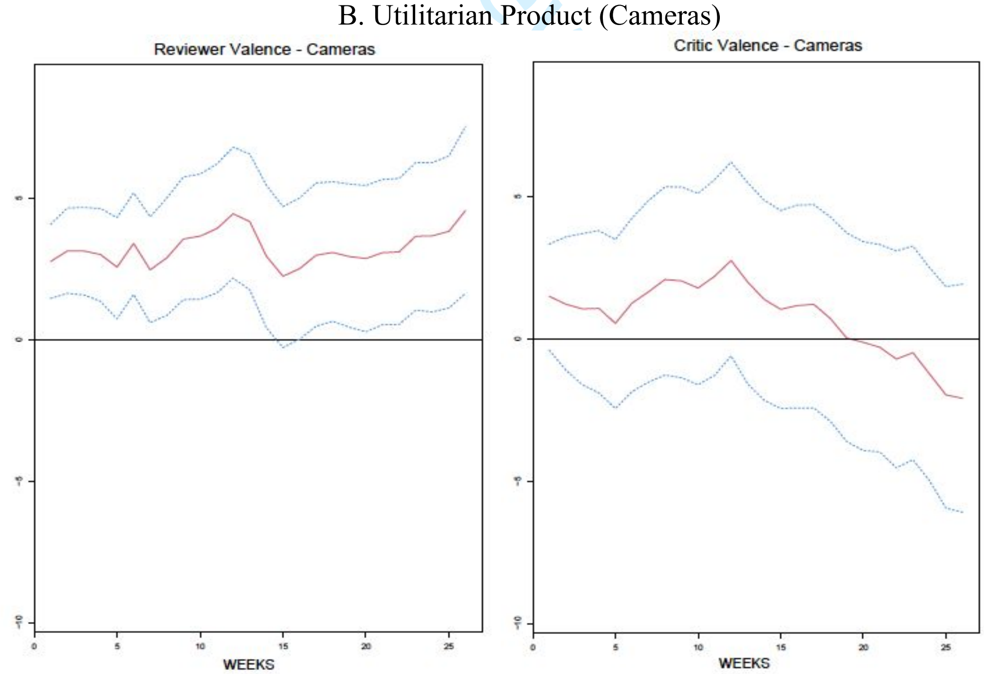
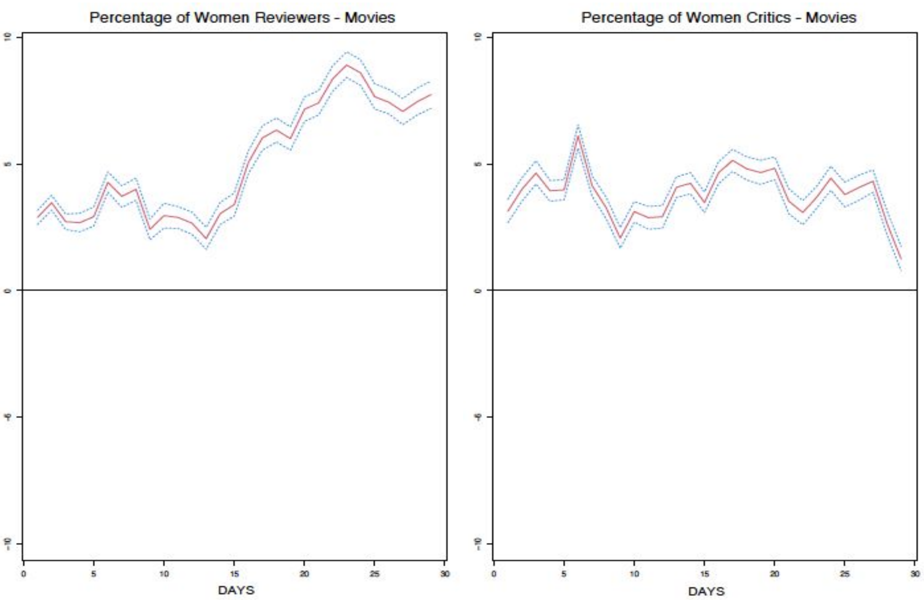


Figure 2 shows the time-varying impact of gender of reviewers and critics. From the figure we can see that the percentage of reviews written by women reviewers has a positive and increasing impact on both the movie box office sales and google search of cameras. The percentage of women critics also increases the box office sales and google search of cameras, but the impact is more stable over time. The results indicate that there's no gender discrimination against women reviewers and women critics for OWOM.

Figure 2: **Time-Varying Impact of Gender of Reviewers and Critics**

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

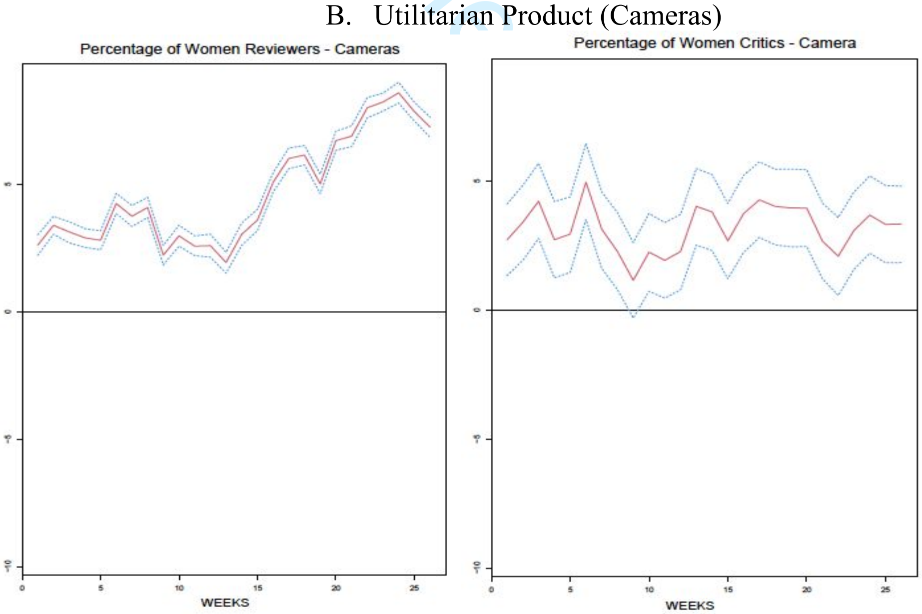
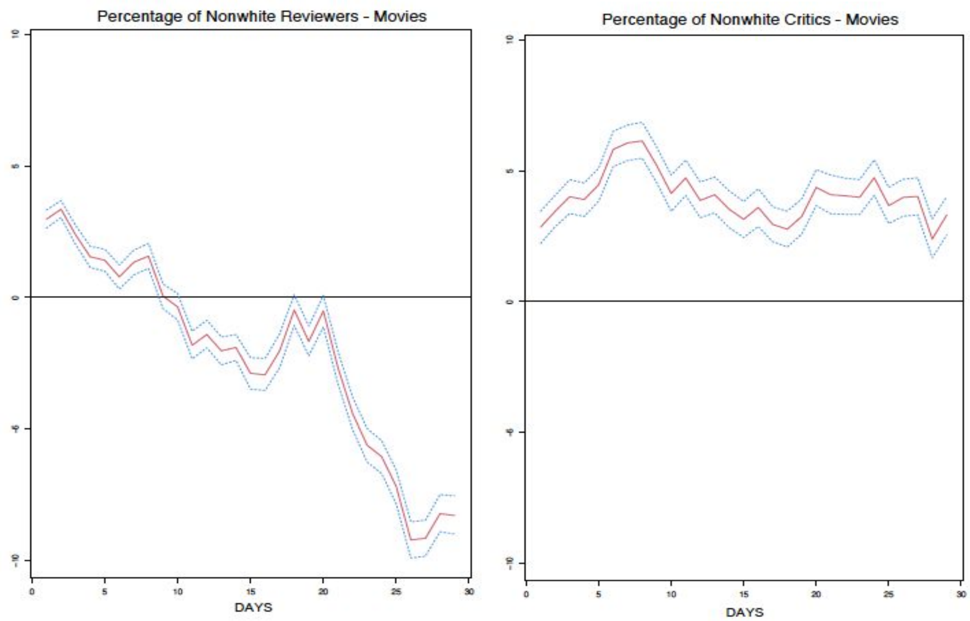


Figure 3 presents the time-varying impact of ethnicity of reviewers and critics. Different from the results from gender analysis, the analysis on ethnicity group shows that the percentage of non-white reviewers have negative impact on movie box office sales and google search of cameras. The negative impact persists and gets stronger over time. Interestingly, in the early time periods there is no negative bias for both movies and cameras. When it comes to critics, the negative impact we see for reviewers is mitigated by the identity of experts: the percentage of non-white critics have positive impact on movie box office sales and google search of cameras.

Figure 3: Time-Varying Impact of Race of Reviewers and Critics

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

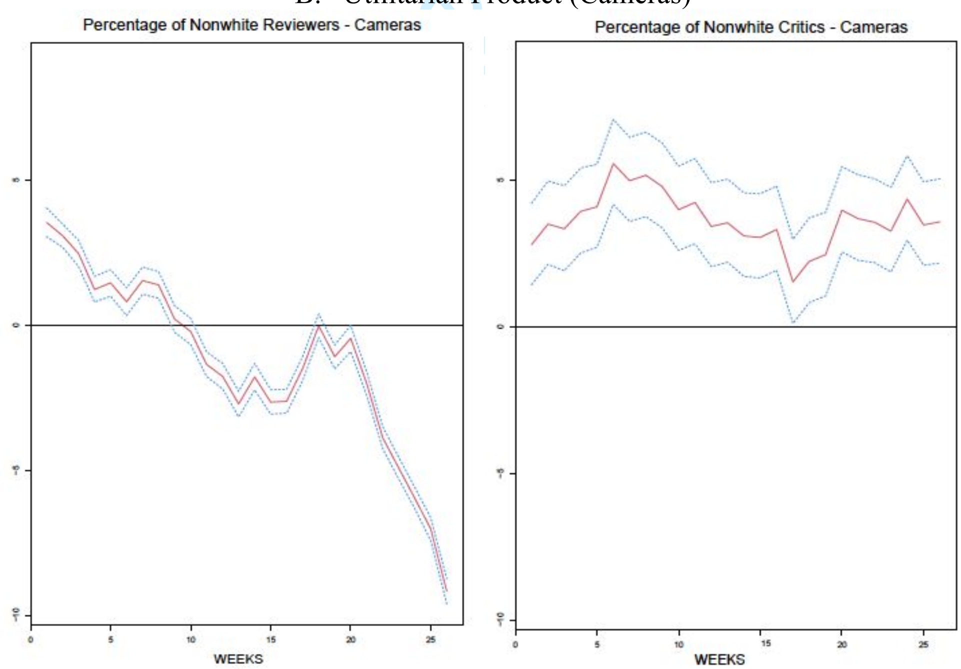
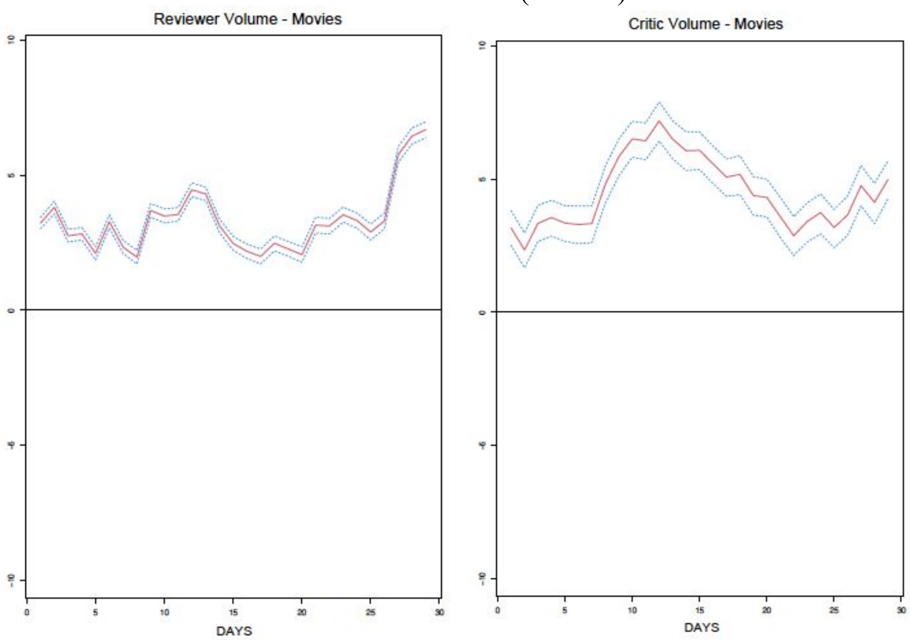


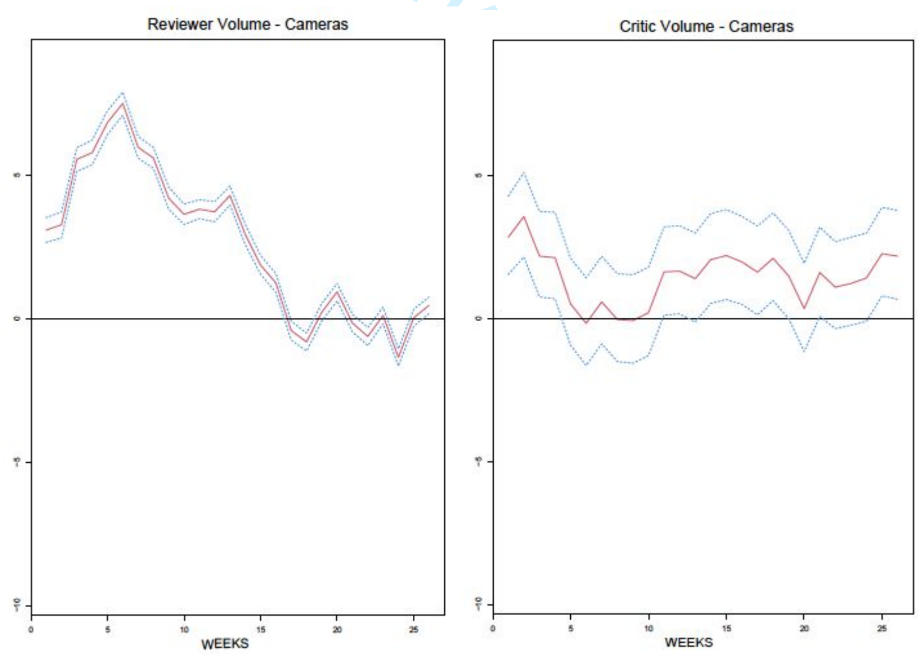
Figure 4 shows the time-varying impact of volume of reviews and critics. In general, the volume of reviews and critics have positive impact on product sales. Dynamically, the impact of critic volume on product sales are more stable in the later time periods compared with the impact of review volume. The time-varying impacts on review valence and review volume are delegated to Figure WB1 and Figure WB2 in the Web Appendix B.

Figure 4: Time-Varying Impact of Reviewers and Critics Volume

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)



5.2 Moderating Role of Reviewer Demographics

Overall Impact

In this section, we look at the interplay of the demographic groups and the valence of reviews and critics to see whether the identity of those who write reviews would affect consumers response to OWOM. For this purpose, we include relevant interactions in our main model. Table 5 shows the moderating role of gender and race of reviewers and critics. From the table we can see that increase in the percentage of women and nonwhite makes the impact of review valence and critic valence on product sales/search stronger except for non-white reviewers. This moderating effect indicates that there might be a stronger social identity among the members in minority groups (women or non-white). The results are aligned with the findings presented in Table 4A.

Table 5: Overall Moderating Role of Gender and Race of Reviewers and Critics

	Hedonic Product (Movies)	Utilitarian Product (Cameras)
	$DV: BOX\_OFFICE_{it}$ (1)	$DV: SEARCH_{it}$ (2)
	Mean [5th percentile, 95th percentile]	Mean [5th percentile, 95th percentile]
$PERCENTAGE\_REV\_WOMEN_{it-1} * VALENCE\_REV_{it-1}$	3.5160*[2.4086, 4.6233]	2.8308*[1.2610, 4.4005]
$PERCENTAGE\_REV\_NONWHITE_{it-1} * VALENCE\_REV_{it-1}$	-2.7177*[-3.5734, -1.8622]	-0.4016 [-2.0419, 1.2387]
$PERCENTAGE\_CRIT\_WOMEN_{it-1} * VALENCE\_CRIT_{it-1}$	2.6765*[2.2676, 3.0855]	4.9073*[2.7933, 7.0214]
$PERCENTAGE\_CRIT\_NONWHITE_{it-1} * VALENCE\_CRIT_{it-1}$	4.0013*[3.4094, 4.5932]	3.6333*[1.6078, 5.6589]
$VALENCE\_REV_{it-1}$	-0.3800 [-2.6049, 1.8449]	5.0712*[1.7762, 8.3663]
$VOLUME\_REV_{it-1}$	10.7222*[10.4910, 10.9535]	1.5340* [1.2344, 1.8337]
$VALENCE\_CRIT_{it-1}$	0.8664 [-0.2556, 1.9885]	2.1294 [-1.4882, 5.7470]
$VOLUME\_CRIT_{it-1}$	4.8592*[4.2371, 5.4813]	1.7331*[0.1541, 3.0685]
$PERCENTAGE\_REV\_WOMEN_{it-1}$	1.9192* [0.1390, 3.6995]	3.4943*[0.7543, 6.2343]
$PERCENTAGE\_REV\_NONWHITE_{it-1}$	5.3645*[4.0402, 6.9380]	2.1952 [-0.6631, 5.0535]
$PERCENTAGE\_CRIT\_WOMEN_{it-1}$	3.1175*[1.3512, 4.8839]	0.2914 [-2.9137, 3.4964]
$PERCENTAGE\_CRIT\_NONWHITE_{it-1}$	0.5490 [-1.9858, 3.0838]	2.6285 [-0.4001, 5.6571]
AGE_{it}	2.9020*[1.0417, 4.7623]	2.8206* [1.1103, 4.5310]
$COMPETITION_{it}$	0.0204 [-0.1786, 0.2194]	-0.1548 [-0.4650, 0.1554]
$THEATER_{it}$	-0.0191 -0.1808, 0.1427]	-
$WEEKEND_{t}$	-0.0606 [-0.1212, 0.0001]	-

MONTH FIXED EFFECTS	-	Yes
PRODUCT FIXED EFFECTS	Yes	Yes
Number of Observations	2,697	1,664

* The 90% confidence interval does not include zero.

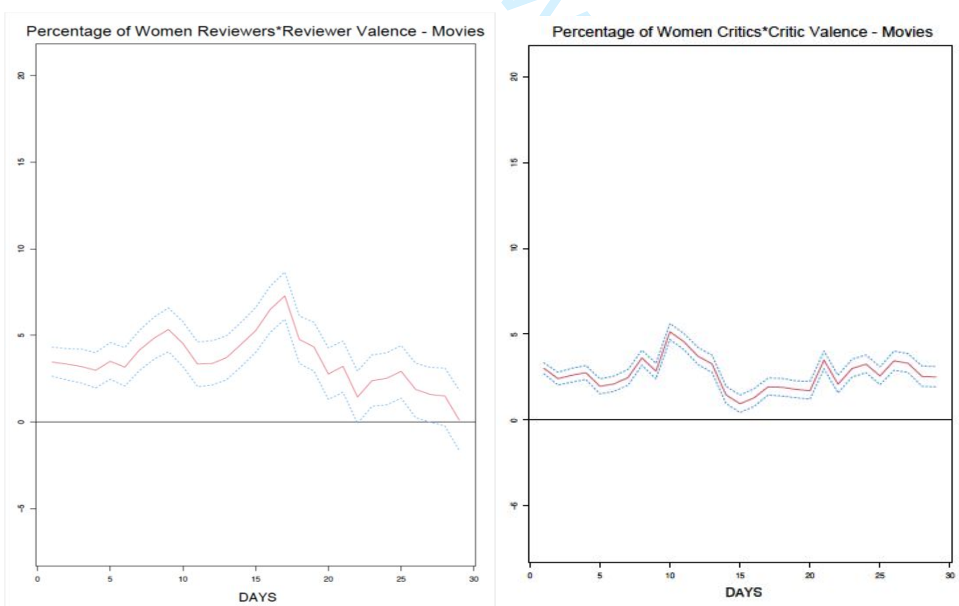
The results of overall impact on *VALENCE_REV* and *VOLUME_REV* with gender and race moderation effect are delegated to Table WA1 and Table WA2 in the Web Appendix A. Overall, the results are consistent with the results in Table 4B and Table 4C.

Time-Varying Impact

Figure 5 shows the time-varying moderating role of gender of reviewers and critics. We can see that the impact of valence (both review valence and critic valence) on product demand (box office for movies and google search for cameras) are enhanced by the percentage of women reviewers. Similarly, Figure 6 shows that the impact of critics' valence on product demand are also enhanced by the percentage of non-white critics, but the effect is different for non-white reviewers. The time-varying impact on review valence and review volume with gender and race moderation are delegated to Figure WB3 and Figure WB4 in the Web Appendix B.

Figure 5: Time-Varying Moderating Role of Gender of Reviewers and Critics

A. Hedonic Product (Movies)



B. Utilitarian Product (cameras)

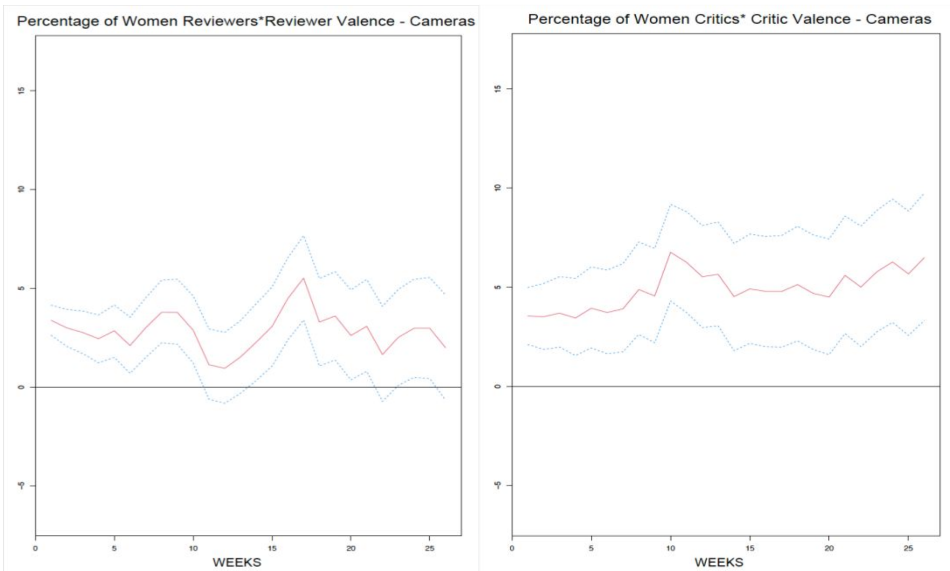
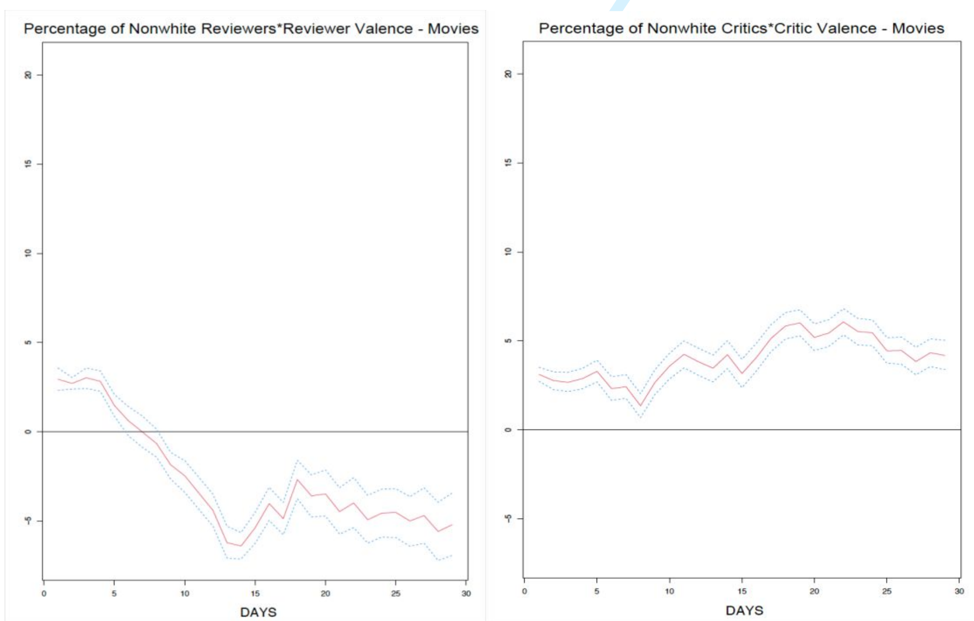
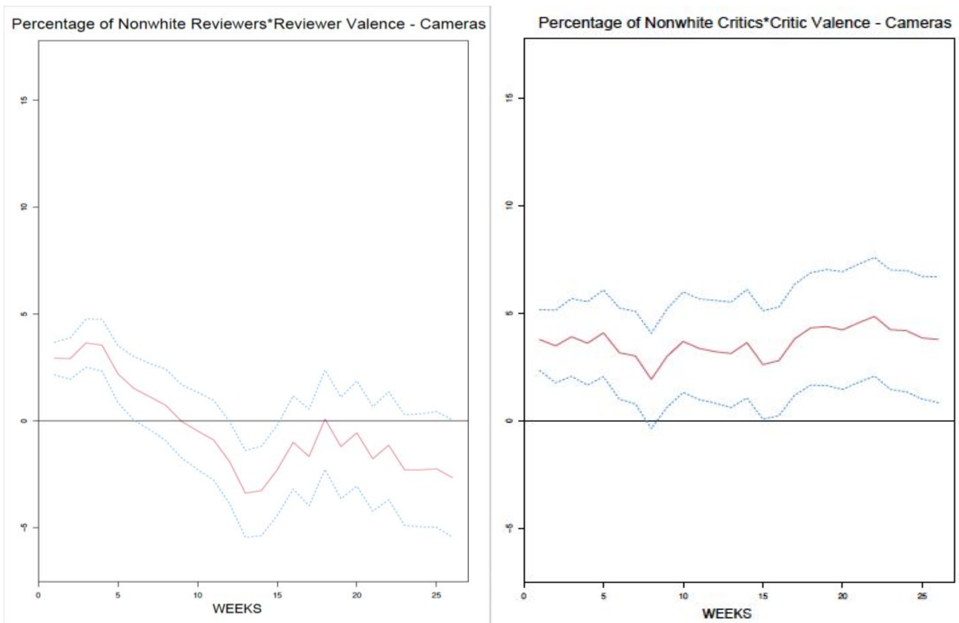


Figure 6: Time-Varying Moderating Role of Race of Reviewers and Critics

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)



6. Discussion and Managerial Implication

Active online reviewers are important players in disseminating information and driving product demand. The message senders exhibit heterogeneity within the group due to differences in identity-related factors, such as demographic traits. There is considerable heterogeneity in their ages, socioeconomic backgrounds, genders, and racial groups, which results in a diversified collection of senders with messages to share. Although there has been extensive research on a more general level of how online reviews would affect product sales, the differential impact of reviewers and critics with different demographics (e.g., gender and ethnicity) on product demand have not been fully investigated.

In this research, we first explore the overall impact of review valence and critics valence on product sales/search. Then, we look at gender and ethnicity heterogeneity of message senders' influence on product demand. More precisely, we are curious about how the identity of being a minority group member (e.g., female, non-white) interacts with subject expertise (critics vs. reviewers). We use data on movies (hedonic product category) and cameras (utilitarian product category) to study these research questions. Both the movie and the camera datasets consist of sales/search of the products, reviews, critics as well as the demographics. Moreover, we use a multivariate Dynamic Linear Modeling (DLM) framework to track the evolution of the time-varying effect on product demand for reviewers and critics. Our dataset and model specifications allow us to get a deeper understanding of the heterogenous impact of OWOM from different message sender types.

Our research has many interesting and novel findings. First, for the movie and camera categories, we find that both the valence of reviews and the valence of critics have significant positive impact on product sales/search, which is consistent with current literature. However, adding to this literature, we find that there are interesting heterogeneities across different product categories: For movies, overall, the valence of critics matters more than valence of reviewers on box office sales. In addition, the effect of review valence dies out quickly, but the impact of critics' valence stays relatively stable over time; For the camera category, we find that consumer reviews have larger impact on product sales than critics and that the impact stays positive and stable over time.

With respect to the demographics, we find very similar and consistent patterns across the two product categories: First, we didn't find the gender stereotype towards female reviewers or female experts: higher percentage of reviews and critics written by women increases the movie box office sales as well as the search for cameras. Second, there is a potential bias against nonwhite reviewers: reviews written by nonwhite reviewers have significant negative impact on product sales, and the effect becomes stronger overtime. However, being an expert can mitigate this bias, which shows that the identity of experts can act as a signal of competence.

In addition to examining the overall impact of review valence on product sales, our analysis delved into the moderating role of demographic factors, particularly gender and race, in shaping consumer responses to OWOM. Notably, we find that women and individuals from nonwhite racial backgrounds notably intensifies the influence of review valence on product sales. However, this trend does not hold true for nonwhite reviewers, where the impact on product sales did not exhibit the same enhancement. This discrepancy may stem from various factors, including differing perceptions of credibility or expertise among nonwhite reviewers, cultural nuances influencing consumer trust, or other factors within the demographic interplay of reviewers and consumers.

This study contributes to the literature by providing a deeper understanding of heterogenous OWOM impact from reviewers, critics, and different demographic groups. The novel findings also shed light on some important managerial implications: First, managers should recognize the significance of both reviews and critics valence in driving product sales. Managers need to develop strategies to leverage positive OWOM to enhance the overall perception of the products. In addition, given the dynamic nature of these impacts, the managers should track the sentiment of the reviews and critics overtime and adjust promotional efforts based on the changing influence of reviews and critics at different stages of the product cycle. For instance, our results show that the valence of critics remains more stable and influential over time compared to the valence of reviews, which underscores the importance of curating expert reviews early in the product life cycle.

Moreover, managers should recognize the importance of demographic characteristics when analyzing the impact of online reviews on product sales. Given the positive impact of women reviewers and critics on product sales, managers should create marketing strategies that resonate with female audiences and highlight female critics during promotional campaigns. In addition, they can further consider some novel strategies to encourage female audience or consumers to write their comments on the online platforms. Given the negative impact of reviews by non-white reviewers on sales but the mitigated negative impact when these reviews come from non-white experts, managers should emphasize expert opinions in marketing materials, especially from diverse backgrounds, to counteract potential biases. Also, they can think about the potential mechanisms which can be used to mitigate the negative impact of reviews written by non-white audiences or non-white consumers such as adding expertise badges to nonwhite consumers who have product category expertise to show their competence etc. Thus, this study overall reveals the importance of a diversified approach to OWOM marketing, suggesting that integrating reviews from a broad demographic spectrum can enhance product appeal and mitigate biases.

Finally, Managers should note the time-varying effects, particularly the increasing impact of reviews written by women and the mitigating effect of expert identity on racial biases over time, to tailor their promotional strategies accordingly. Looking at Figures 1-3, we recommend that managers highlight positive critic valence in marketing materials during early product lifecycle, especially from diverse experts, to build credibility and counteract potential demographic biases. During the growth phase, it is better to amplify the voices of ordinary reviewers and encourage diversity in reviewer demographics with a special emphasis on women reviewers. When it comes to maturity and decline phase in the product life cycle, managers are better off to continue monitoring the dynamic impact of reviews and adapt marketing strategies, for example leveraging longstanding positive effects of expert reviews and addressing any emerging negative trends in public perception, especially from underrepresented groups. By tailoring strategies throughout the product lifecycle with a keen eye on the nuanced impacts of reviewer demographics and expertise, managers can more effectively harness online reviews to bolster product success.

7. Limitations and Future Research

Although this study acknowledges the diversity in product types (hedonic vs. utilitarian) and how they may differently influence the impact of OWOM. The generalizability of the findings can be increased by future research by considering a broader array of products. Another limitation is the lack of detailed insights into the demographics or experiences of the message receivers, which can offer a promising avenue for future research. Understanding how consumers' backgrounds, experiences, and demographics influence their interpretation and trust in online reviews can provide deeper insights into the effectiveness of OWOM. In addition, given the global nature of online markets, exploring how cultural differences affect the reception and impact of OWOM could offer valuable insights. Cross-cultural studies could uncover how societal norms and values influence trust and reactions to online reviews. Finally, the study uses aggregate level data. Having access to individual consumer level purchase data could provide interesting additional insights. By exploring suggested future research paths, subsequent studies can build on the foundational insights provided by this research, offering richer, more nuanced understandings of the complex dynamics at play in in a continually changing digital ecosystem.

REFERENCES

Adamopoulos, P., Ghose, A., & Todri, V. (2018). The impact of user personality traits on word of mouth: Text-mining social media platforms. *Information Systems Research*, 29(3), 612-640.

Ataman, M. B., Van Heerde, H. J., & Mela, C. F. (2010). The long-term effect of marketing strategy on brand sales. *Journal of Marketing Research*, *47*(5), 866-882.

Azer, J., Anker, T., Taheri, B., & Tinsley, R. (2023). Consumer-driven racial stigmatization: The moderating role of race in online consumer-to-consumer reviews. *Journal of Business Research*, *157*, 113567.

Barasch, A., & Berger, J. (2014). Broadcasting and narrowcasting: How audience size affects what people share. *Journal of Marketing Research*, *51*(3), 286-299.

Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, *15*(5), 215-227.

Bhattacharya, U., Kumar, A., Visaria, S., & Zhao, J. (2023). Do women receive worse financial advice?. *Journal of Finance*, Forthcoming.

Bhutta, N., Hizmo, A., & Ringo, D. (2022). How much does racial bias affect mortgage lending? Evidence from human and algorithmic credit decisions. SSRN: https://doi.org/10.17016/FEDS.2022.067

Brown, J.J., P.H. Reingen. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, 14(3) 350-362.

Chakravarty, A., Liu, Y., & Mazumdar, T. (2010). The differential effects of online word-of-mouth and critics' reviews on pre-release movie evaluation. *Journal of Interactive Marketing*, 24(3), 185-197.

Chen, Y., Liu, Y., & Zhang, J. (2012). When do third-party product reviews affect firm value and what can firms do? The case of media critics and professional movie reviews. *Journal of Marketing*, *76*(2), 116-134.

Chiu, Y. L., Du, J., Sun, Y., & Wang, J. N. (2022). Do critical reviews affect box office revenues through community engagement and user reviews?. *Frontiers in Psychology*, *13*, 900360.

Cho, D., Kim, S., & Acquisti, A. (2012, January). Empirical analysis of online anonymity and user behaviors: the impact of real name policy. In *2012 45th Hawaii international conference on system sciences* (pp. 3041-3050). IEEE.

Chung, S., Jung, J., Park, J., Lee, C., & Ceran, Y. (August 23, 2023). Two-sided Impacts of Identity Disclosure in e-Customer Service Platform: Evidence from a Field Experiment KAIST College of Business Working Paper Series, Available at SSRN: https://ssrn.com/abstract=4548980

Cox, J., & Kaimann, D. (2015). How do reviews from professional critics interact with other signals of product quality? Evidence from the video game industry. *Journal of Consumer Behavior*, *14*(6), 366-377.

Dekimpe, Marnik G., & Hanssens, D. M. (1999). Sustained spending and persistent response: A new look at long-term marketing profitability. *Journal of Marketing Research,* 397-412.

Dellarocas, C., Zhang, X., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23-45.

Deng, T. (2020). Investigating the effects of textual reviews from consumers and critics on movie sales. *Online Information Review*, *44*(6), 1245-1265.

Dhar, V., & Chang, E. A. (2009). Does chatter matter? The impact of user-generated content on music sales. *Journal of Interactive Marketing*, *23*(4), 300-307.

Duan, W., B. Gu, A. B. Whinston. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing*, 84(2) 233-242.

Eagly, A. H., & Carli, L. L. (2007). *Through the labyrinth: The truth about how women become leaders*. Harvard Business Review Press.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: race, crime, and visual processing. *Journal of Personality and Social Psychology*, *87*(6), 876.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291-313.

Gelper, S., Peres, R., & Eliashberg, J. (2018). Talk bursts: The role of spikes in prerelease word-of-mouth dynamics. *Journal of Marketing Research*, *55*(6), 801-817.

Godes, D., D. Mayzlin. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, 23(4) 545-560.

Godes, D., D. Mayzlin. (2009). Firm-created word-of-mouth communication: evidence from a field test. *Marketing Science*, 28(4) 721-739.

Gopinath, S., Chintagunta, P. K., & Venkataraman, S. (2013). Blogs, advertising, and local-market movie box office performance. *Management Science*, *59*(12), 2635-2654.

Gopinath, S., J.S. Thomas, L. Krishnamurthi. (2014). Investigating the relationship between the content of online word of mouth, advertising, and brand performance. *Marketing Science*, 33(2) 241-258

Heeler, R. M., & Hustad, T. P. (1980). Problems in predicting new product growth for consumer durables. *Management Science*, *26*(10), 1007-1020.

Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: the implied communality deficit. *Journal of Applied Psychology*, *92*(1), 81.

Hoffmann, D. E., & Tarzian, A. J. (2001). The girl who cried pain: a bias against women in the treatment of pain. *Journal of Law, Medicine & Ethics*, *29*(1), 13-27.

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, *113*(16), 4296-4301.

Hoskins, J., Gopinath, S., Verhaal, J. C., & Yazdani, E. (2021). The influence of the online community, professional critics, and location similarity on review ratings for niche and mainstream brands. *Journal of the Academy of Marketing Science*, *49*, 1065-1087.

Huang, N., Hong, Y., & Burtch, G. (2017). Social network integration and user content generation. *MIS Quarterly*, *41*(4), 1035-1058.

Inci, A. C., Narayanan, M. P., & Seyhun, H. N. (2017). Gender differences in executives' access to information. *Journal of Financial and Quantitative Analysis*, 52(3), 991-1016.

Kakar, V., Voelz, J., Wu, J., & Franco, J. (2018). The visible host: Does race guide Airbnb rental rates in San Francisco?. *Journal of Housing Economics*, *40*, 25-40.

Katz, E., & Lazarsfeld, P. F. (1959). *Personal influence: The part played by people in the flow of mass communications.* Free Press.

Kim, E., Ding, M., Wang, X., & Lu, S. (2023). Does topic consistency matter? A study of critic and user reviews in the movie industry. *Journal of Marketing*, *87*(3), 428-450.

Kilner, P. G., & Hoadley, C. M. (2017). Anonymity options and professional participation in an online community of practice. *In Computer Supported Collaborative Learning 2005* (pp. 272-280). Routledge.

Knobloch-Westerwick, S., Glynn, C. J., & Huge, M. (2013). The Matilda effect in science communication: an experiment on gender bias in publication quality perceptions and collaboration interest. *Science Communication*, *35*(5), 603-625.

Leshed, G. (2009). Silencing the clatter: Removing anonymity from a corporate online community. *Online deliberation: Design, Research, and Practice*, 243-251.

Li, J., & Liang, X. (2022). Reviewers' identity cues in online product reviews and consumers' purchase intention. *Frontiers in Psychology*, *12*, 784173.

Liu, Y. (2006). Word of mouth for movies: its dynamics and impact on box office revenue. *Journal of Marketing*, 70(3) 74-89.

Lin, C. A., & Xu, X. (2017). Effectiveness of online consumer reviews: The influence of valence, reviewer ethnicity, social distance and source trustworthiness. *Internet Research*, *27*(2), 362-380.

Luo, X. (2009). Quantifying the long-term impact of negative word of mouth on cash flows and stock prices. *Marketing Science*, 28(1), 148-165.

Mahajan, V., Muller, E., & Bass, F. M. (1990). New product diffusion models in marketing: A review and directions for research. *Journal of Marketing*, *54*(1), 1-26.

Money, R. B., Gilly, M. C., & Graham, J. L. (1998). Explorations of national culture and word-of-mouth referral behavior in the purchase of industrial services in the United States and Japan. *Journal of Marketing*, 62(4), 76-87.

Moon, S., Bergey, P. K., & Iacobucci, D. (2010). Dynamic effects among movie ratings, movie revenues, and viewer satisfaction. *Journal of Marketing*, 74(1), 108-121.

Niraj, R., & Singh, J. (2015). Impact of user-generated and professional critics reviews on Bollywood movie success. *Australasian Marketing Journal*, *23*(3), 179-187.

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science*, *26*(5), 617-624.

Owens, J., & Lynch, S. M. (2012). Black and Hispanic immigrants' resilience against negative-ability racial stereotypes at selective colleges and universities in the United States. *Sociology of Education*, *85*(4), 303-325.

Paskuda, M., & Lewkowicz, M. (2017). Anonymity interacting with participation on a Q&A site. *AI & SOCIETY*, *32*, 369-381.

Petty, R. E., Cacioppo, J. T., Petty, R. E., & Cacioppo, J. T. (1986). *The elaboration likelihood model of persuasion* (pp. 1-24). Springer New York.

Pu, J., Chen, Y., Qiu, L., & Cheng, H. K. (2020). Does identity disclosure help or hurt user content generation? Social presence, inhibition, and displacement effects. *Information Systems Research*, *31*(2), 297-322.

Srinivasan, V., & Mason, C. H. (1986). Nonlinear least squares estimation of new product diffusion models. *Marketing Science*, *5*(2), 169-178.

Starr, S. (2012). *Revolt in Syria: Eye-witness to the Uprising*. Hurst.

Van Heerde, H. J., Mela, C. F., & Manchanda, P. (2004). The dynamic effect of innovation on market structure. *Journal of Marketing Research*, *41*(2), 166-183.

Wang, F., Liu, X., & Fang, E. E. (2015). User reviews variance, critic reviews variance, and product sales: An exploration of customer breadth and depth effects. *Journal of Retailing*, *91*(3), 372-389.

Zhang, L., & Hanks, L. (2018). Online reviews: The effect of cosmopolitanism, incidental similarity, and dispersion on consumer attitudes toward ethnic restaurants. *International Journal of Hospitality Management*, *68*, 115-123.

Web Appendix

Web Appendix	Contents
A	Overall Impact on *VALENCE_REV* and *VOLUME_REV*
B	Time-Varying Impact on *VALENCE_REV* and *VOLUME_REV*
C	Robustness Checks of 5 Equations

Web Appendix A: Overall Impact on *VALENCE_REV* and *VOLUME_REV*

Table WA1: Overall Impact on *VALENCE_REV* with Gender and Race Moderation

	Hedonic Product (Movies)	Utilitarian Product (Cameras)
	$DV: VALENCE\_REV_{it}$ (1)	$DV: VALENCE\_REV_{it}$ (2)
	Mean [5th percentile, 95th percentile]	Mean [5th percentile, 95th percentile]
BOX_OFFICE_{it-1}	-0.1093[-0.2911, 0.0723]	-
$SEARCH_{it-1}$	-	-0.2157 [-0.4604, 0.0290]
$VOLUME\_REV_{it-1}$	5.8126*[5.5308, 6.0943]	5.5402*[5.3331, 5.7472]
$VALENCE\_CRIT_{it-1}$	3.1906*[2.6252, 3.7560]	-0.4675 [-2.7443, 1.8094]
$VOLUME\_CRIT_{it-1}$	1.9081*[1.3907, 2.4254]	2.3759*[1.3238, 3.4280]
AGE_{it}	0.6068[-0.0399, 1.2536]	1.3302*[0.2916, 2.3688]
$COMPETITION_{it}$	0.0258[-0.1634, 0.2151]	-0.0453[-0.3281, 0.2374]
$WEEKEND_t$	0.1390[-0.0201, 0.2981]	-
$MONTH\ FIXED\ EFFECTS$	-	Yes
$PRODUCT\ FIXED\ EFFECTS$	Yes	Yes
Number of Observations	2,697	1,664

* The 90% confidence interval does not include zero.

Table WA2 : Overall Impact on *VOLUME_REV* with Gender and Race Moderation

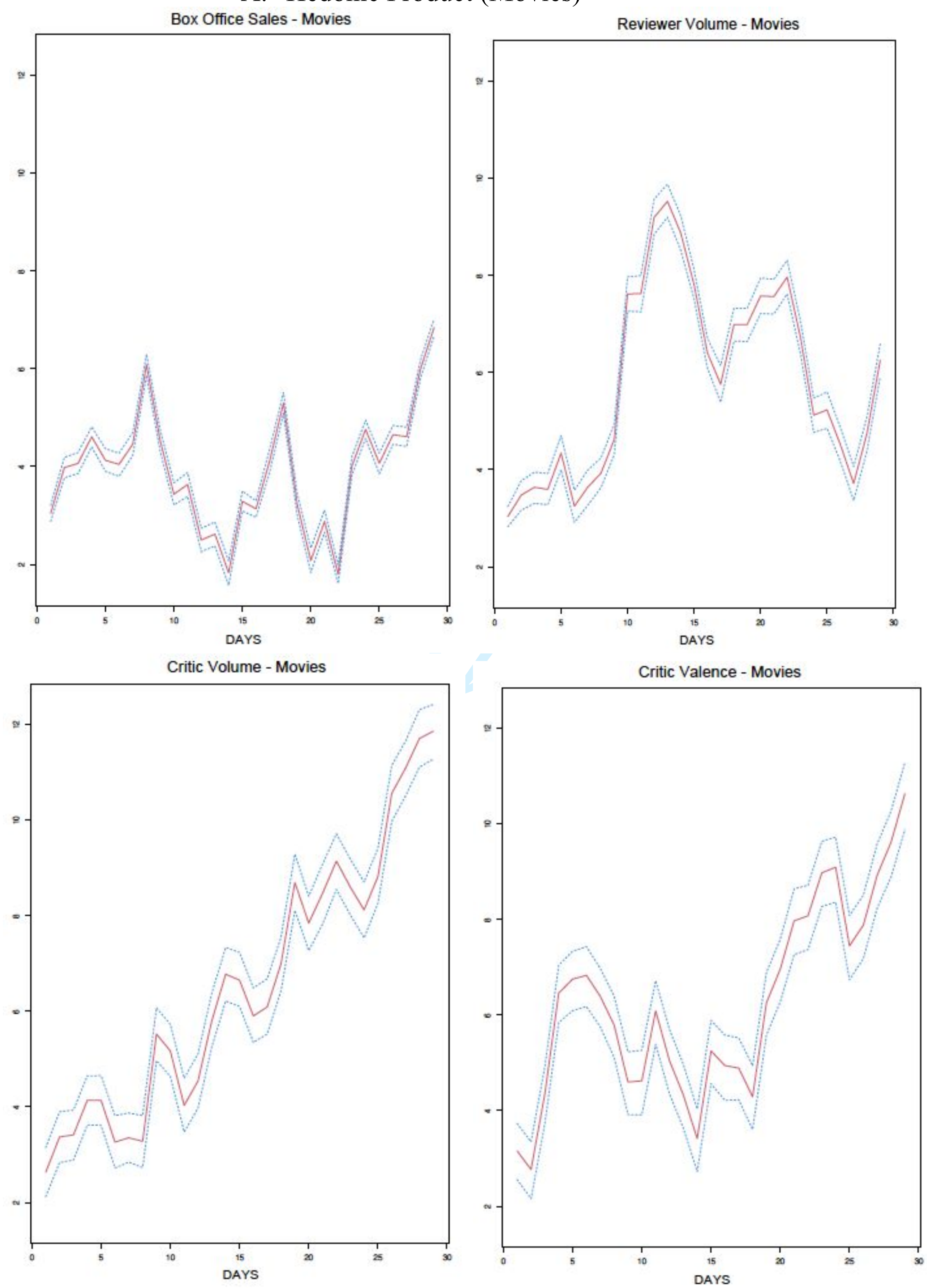
	Hedonic Product (Movies)	Utilitarian Product (Cameras)
	$DV: VOLUME\_REV_{it}$ (1)	$DV: VOLUME\_REV_{it}$ (2)
	Mean [5th percentile, 95th percentile]	Mean [5th percentile, 95th percentile]
BOX_OFFICE_{it-1}	5.4128*[5.2753, 5.5503]	-
$SEARCH_{it-1}$	-	0.0732 [-0.1873, 0.3338]
$VALENCE\_REV_{it-1}$	2.2762*[0.8817, 3.6707]	0.7201 [-1.0775, 2.5176]
$VOLUME\_CRIT_{it-1}$	0.0098 [-0.5607, 0.5803]	5.0649*[2.3816, 7.7483]
$VALENCE\_CRIT_{it-1}$	8.3080*[7.6668, 8.9492]	5.7178*[4.4992, 6.9364]
AGE_{it}	-0.0728 [-0.9986, 0.8529]	1.9287* [0.7638, 3.0936]
$COMPETITION_{it}$	0.3458* [0.1507, 0.5409]	-0.2029 [-0.4923, 0.0865]
$THEATER_{it}$	0.0199 [-0.0551, 0.0950]	-
$WEEKEND_t$	0.1620* [0.0022, 0.3218]	-
MONTH FIXED EFFECTS	-	Yes
PRODUCT FIXED EFFECTS	Yes	Yes
Number of Observations	2,697	1,664

* The 90% confidence interval does not include zero.

Web Appendix B: Time-Varying Impact on *VALENCE_REV* and *VOLUME_REV*

Figure WB1: Time-Varying Impact on *VALENCE_REV*

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

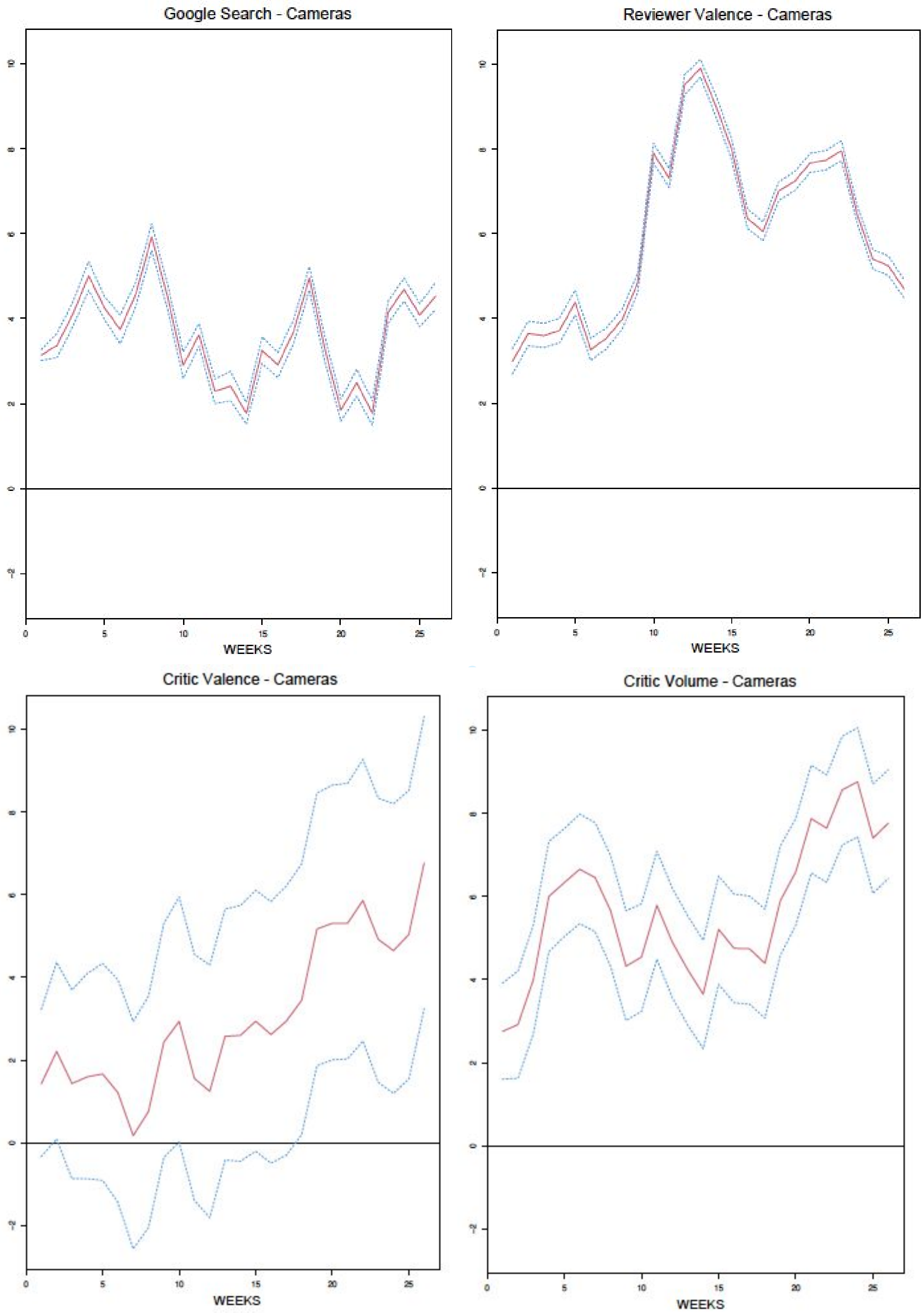
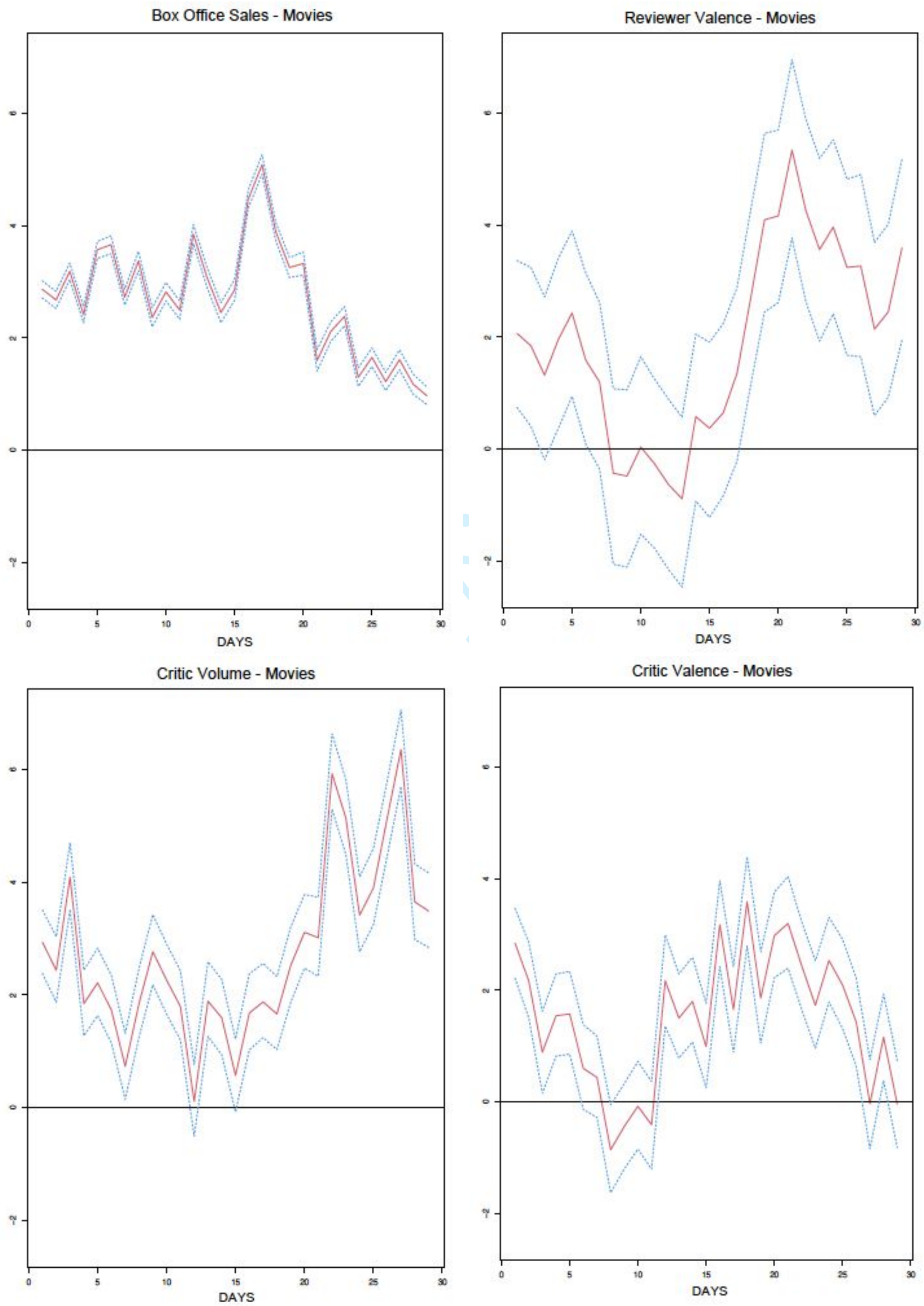


Figure WB2: Time-Varying Impact on *VOLUME_REV*

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

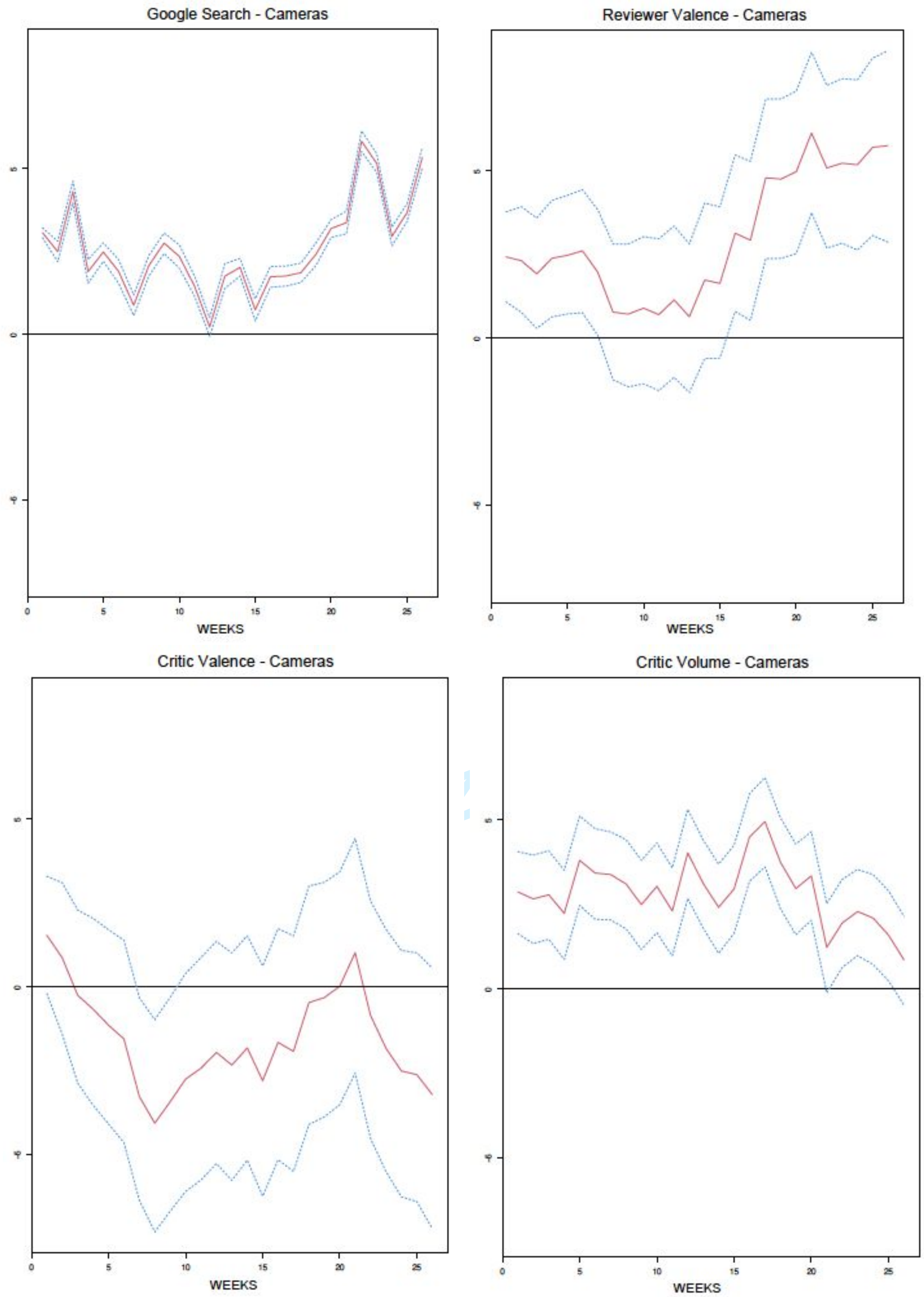
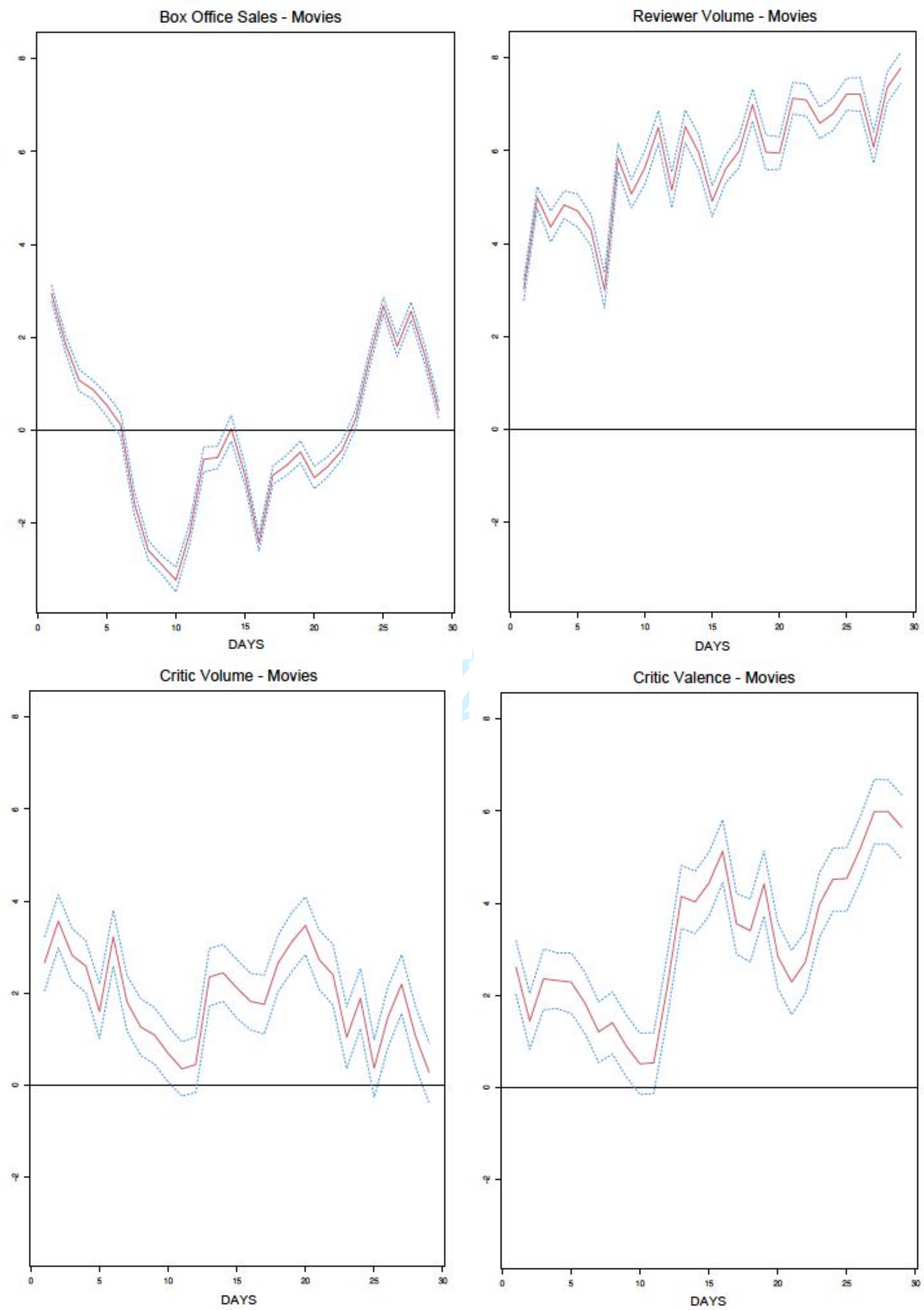


Figure WB3: Time-Varying Impact on *VALENCE_REV* with Gender and Race Moderation

B. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

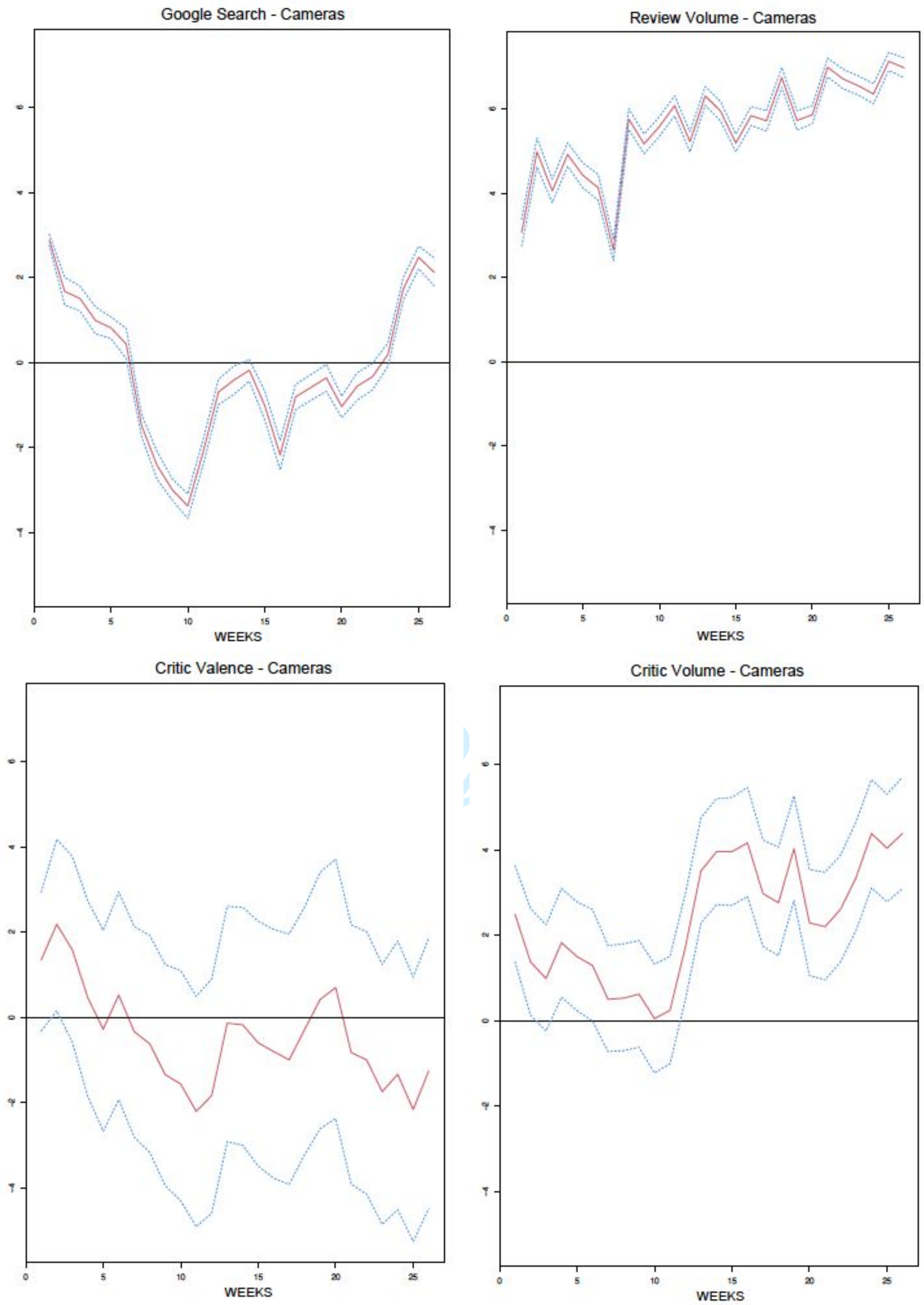
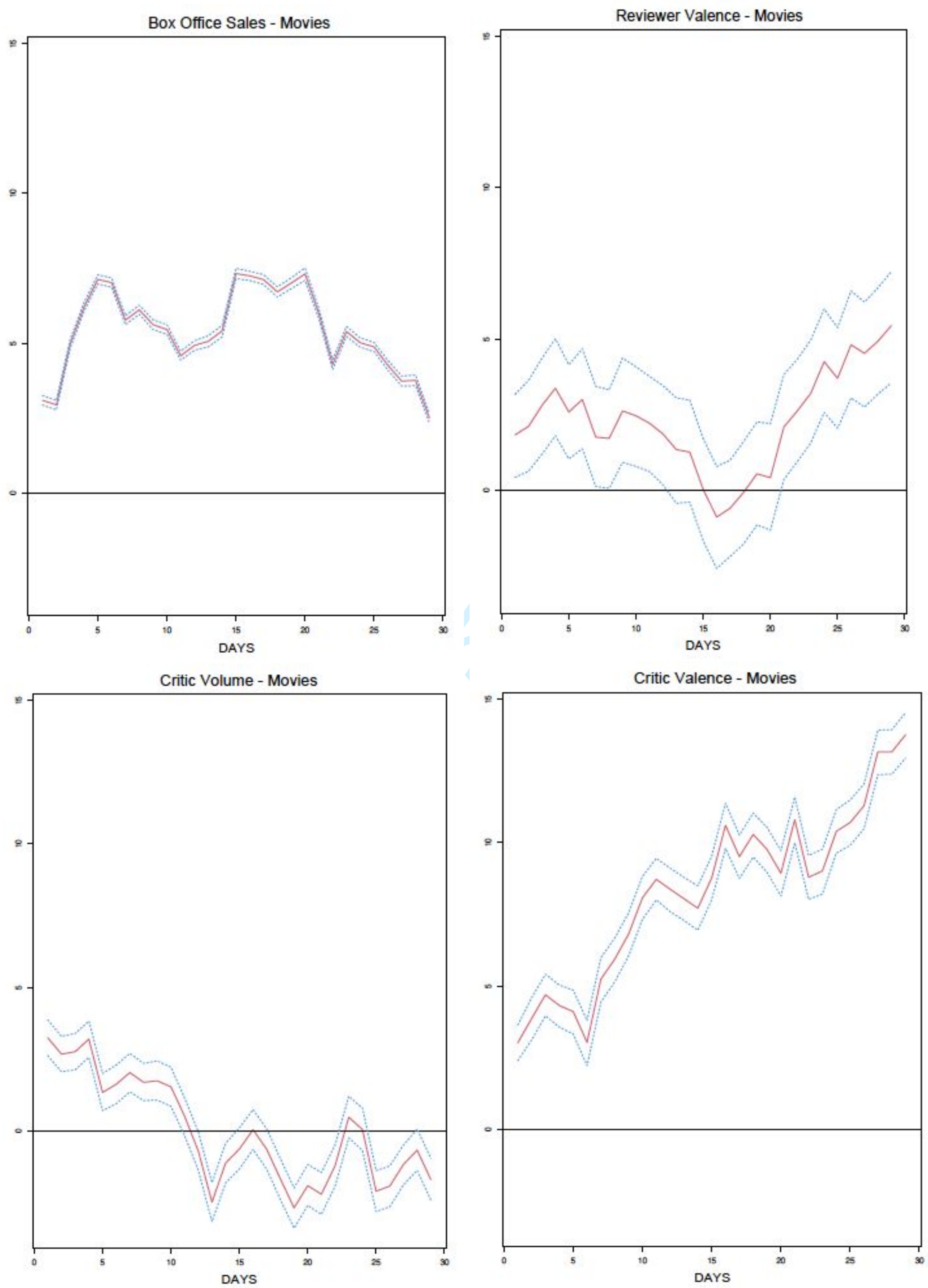
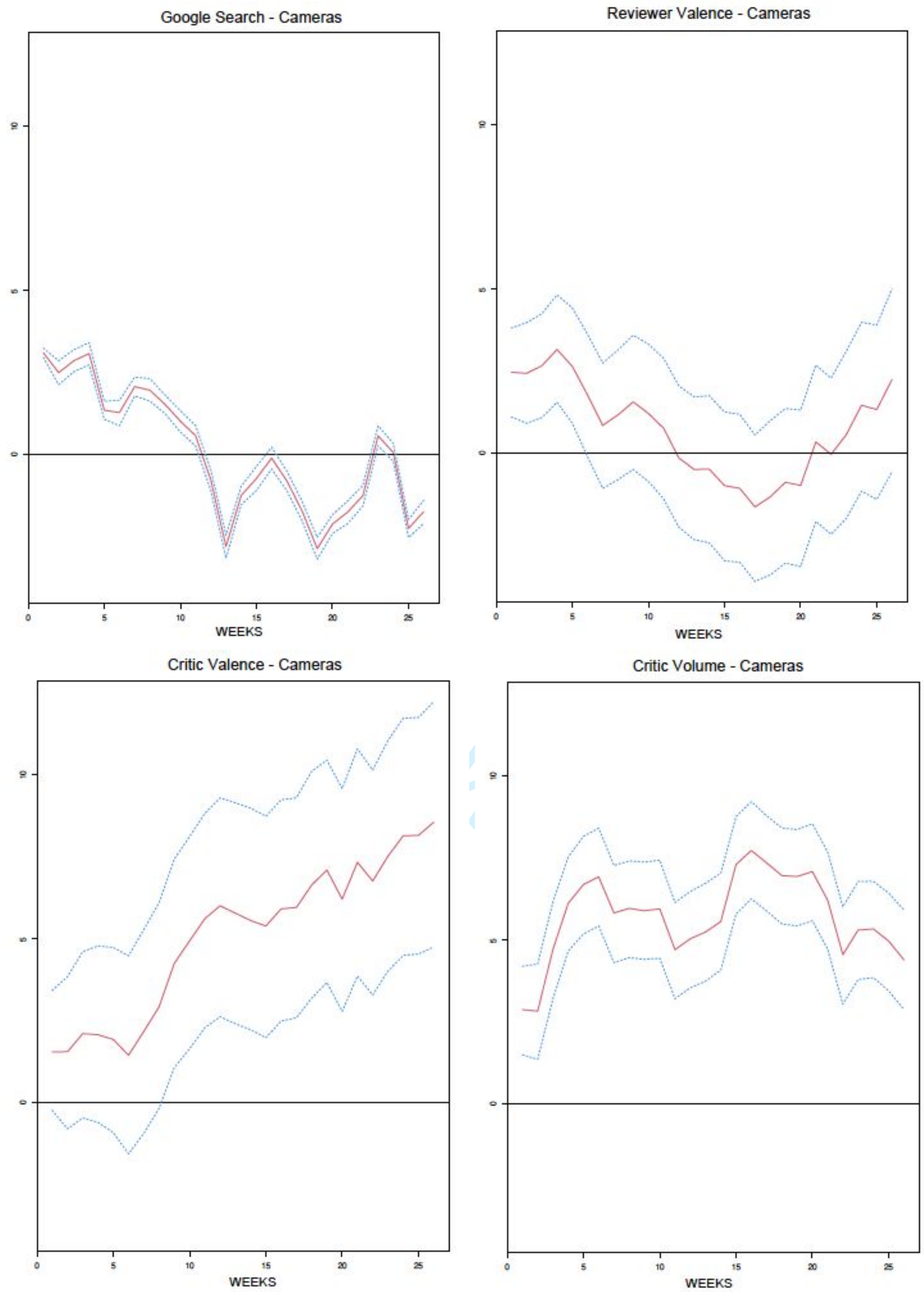


Figure WB4: Time-Varying Impact on *VOLUME_REV* with Gender and Race Moderation

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)



Web Appendix C: Robustness of 5 Equations

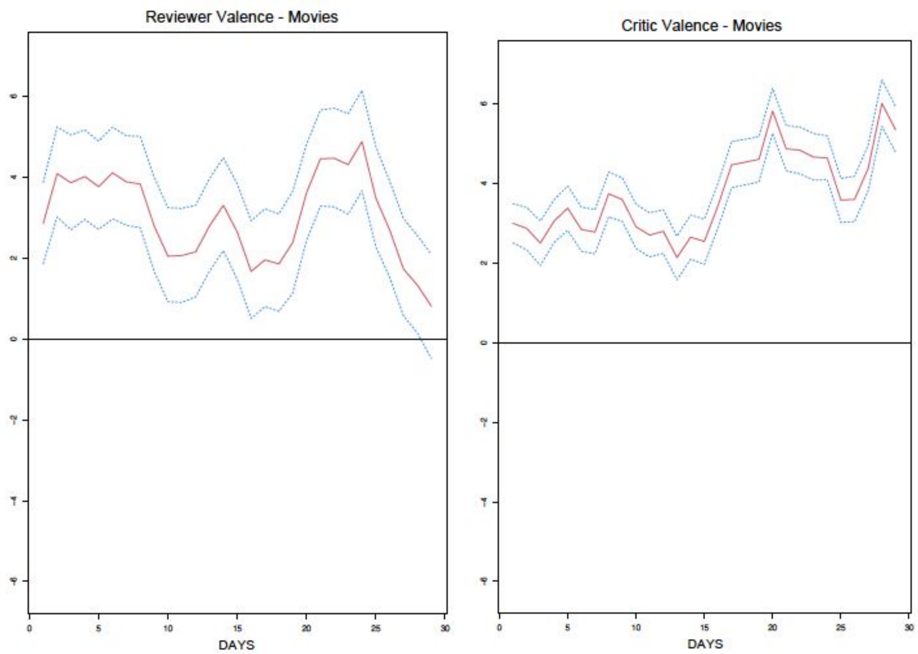
Table WC1: Overall Impact of Reviewers and Critics on *BOX_OFFICE* and *SEARCH* (5 Equations)

	Hedonic Product (Movies)	Utilitarian Product (Cameras)
	$DV: BOX\_OFFICE_{it}$ (1)	$DV: SEARCH_{it}$ (2)
	Mean [5th percentile, 95th percentile]	Mean [5th percentile, 95th percentile]
$VALENCE\_REV_{it-1}$	3.0313*[1.8825, 4.2293]	3.1144*[1.3034, 4.8980]
$VALENCE\_CRIT_{it-1}$	3.7343*[3.1777, 4.2951]	-0.0722*[-2.7616, 2.5570]
$PERCENTAGE\_REV\_WOMEN_{it-1}$	3.8082*[3.4962, 4.1193]	3.6025*[3.3293, 3.8756]
$PERCENTAGE\_REV\_NONWHITE_{it-1}$	0.2749 [-0.1185, 0.6658]	0.8391*[0.5333, 1.1443]
$PERCENTAGE\_CRIT\_WOMEN_{it-1}$	-1.3922*[-1.7048, -1.0783]	-1.1972 [-2.4625, 0.1137]
$PERCENTAGE\_CRIT\_NONWHITE_{it-1}$	2.6020*[2.1085, 3.1001]	2.8392*[1.5837, 4.1319]
$VOLUME\_REV_{it-1}$	4.2249*[4.0300, 4.4193]	3.3520*[3.0922, 3.6109]
$VOLUME\_CRIT_{it-1}$	2.3358*[1.7765, 2.8948]	3.1766*[1.8445, 4.4529]
AGE_{it}	0.4335 [-0.4328, 1.2938]	1.6370*[0.4629, 2.8623]
$COMPETITION_{it}$	-0.1081 [-0.2857, 0.06926]	-0.1012[-0.4039, 0.1864]
$THEATER_{it}$	0.0222 [-0.0320, 0.0765]	-
$WEEKEND_{t}$	0.0920 [-0.0571, 0.2485]	-
$MONTH\ FIXED\ EFFECTS$	-	-
$PRODUCT\ FIXED\ EFFECTS$	Yes	Yes
Number of Observations	2,697	1,664

* The 90% confidence interval does not include zero.

Figure WC1: Time-Varying Impact of Reviewers and Critics Valence (5 Equations)

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

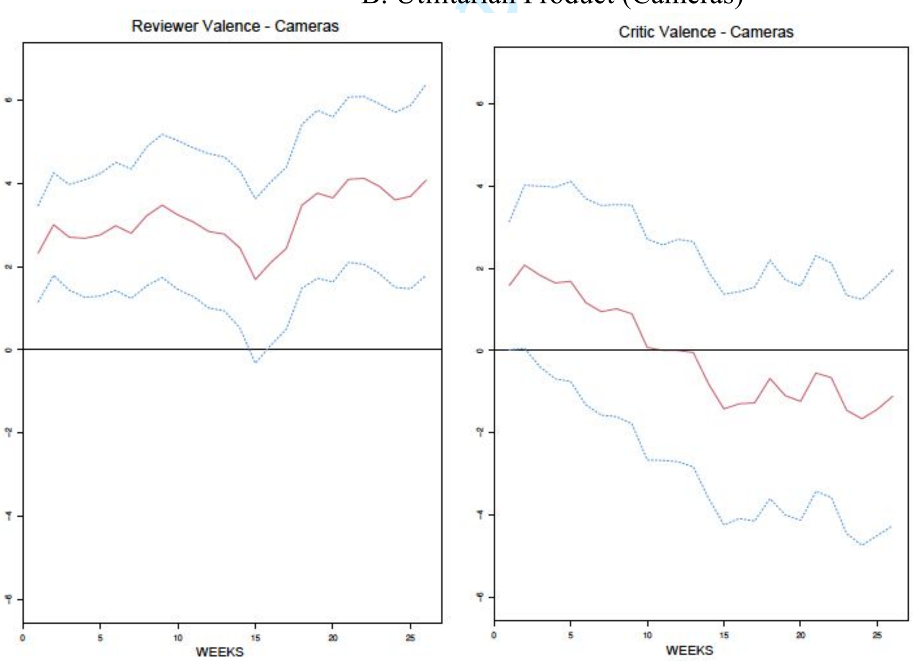
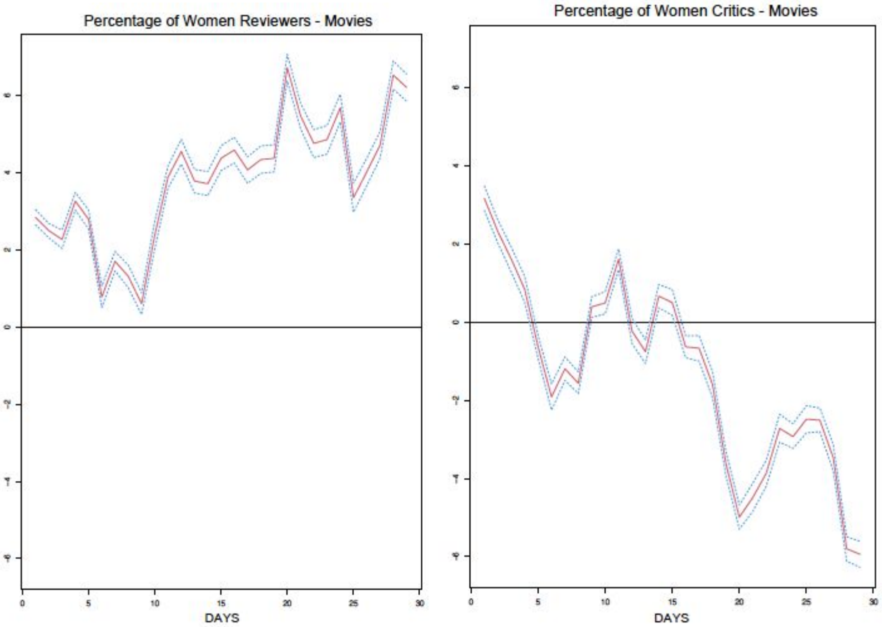


Figure WC2: **Time-Varying Impact of Gender of Reviewers and Critics (5 Equations)**

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)

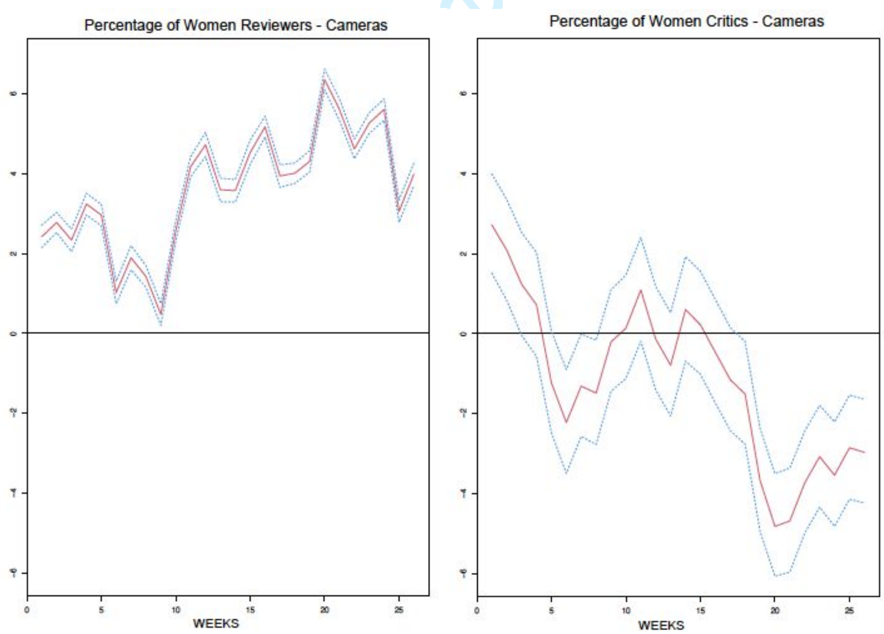
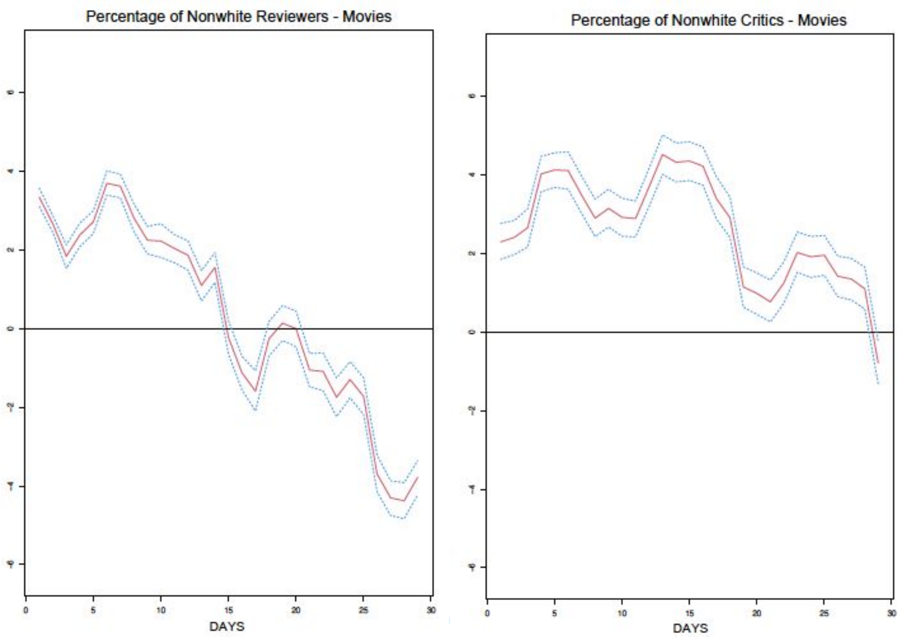


Figure WC3: Time-Varying Impact of Race of Reviewers and Critics (5 Equations)

A. Hedonic Product (Movies)



B. Utilitarian Product (cameras)

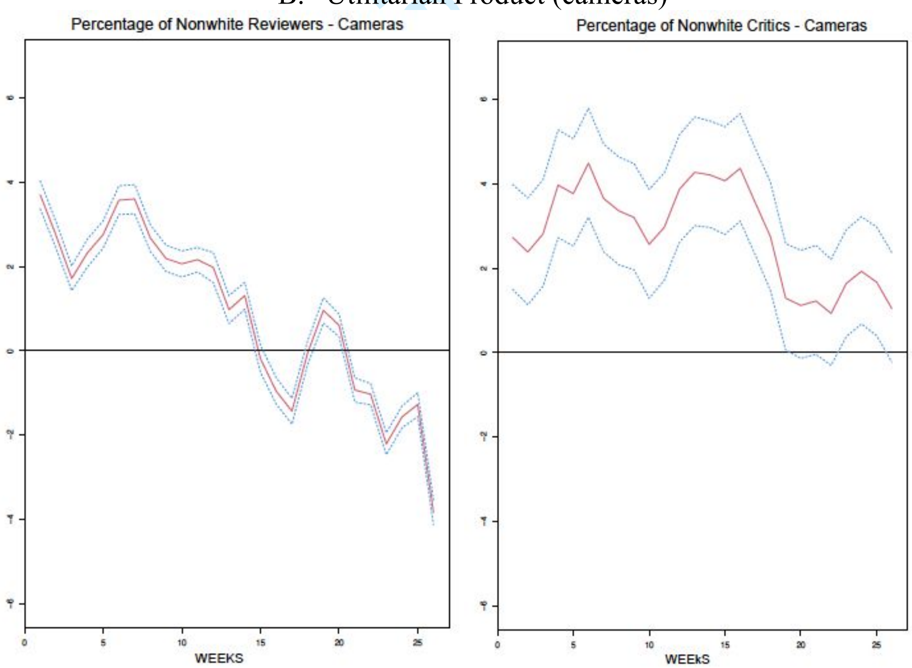
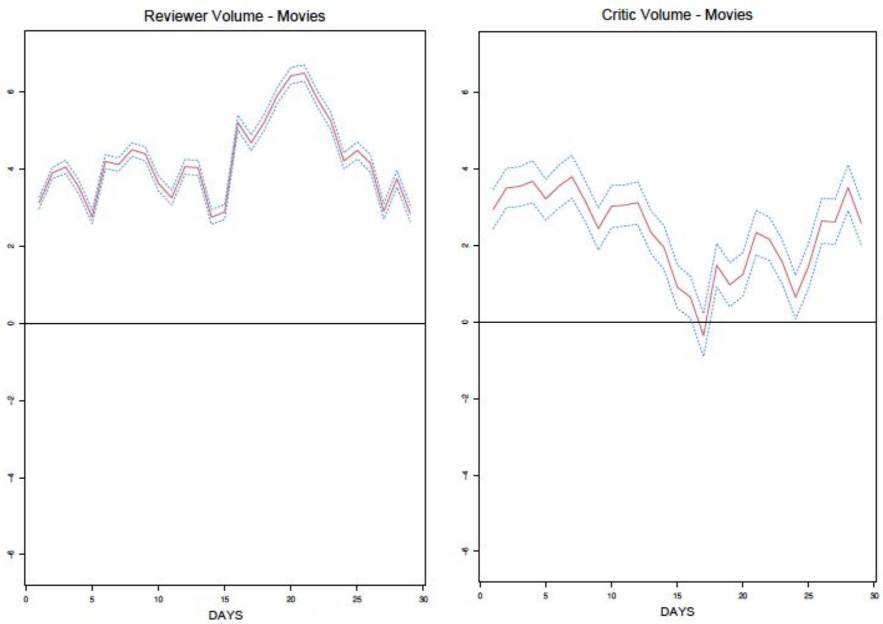
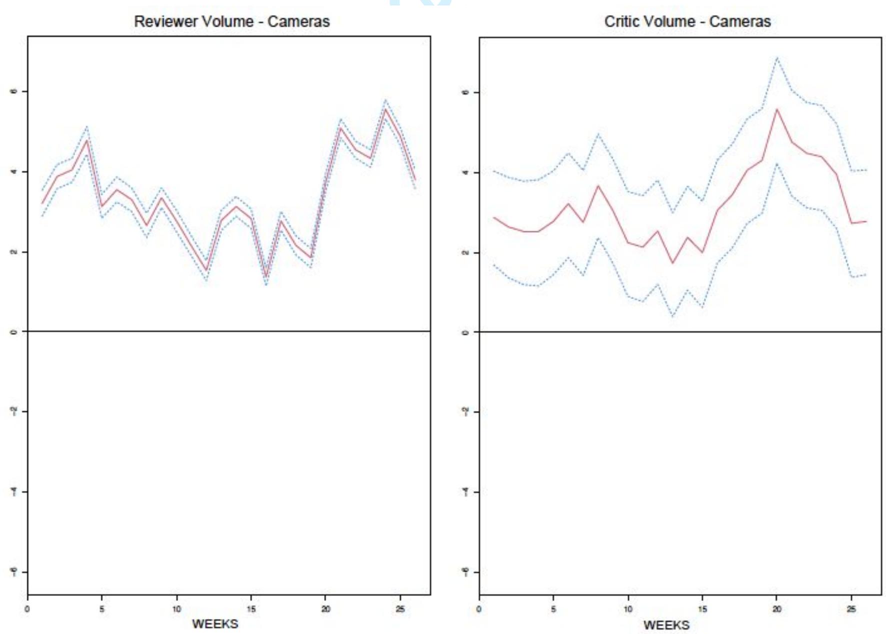


Figure WC4: Time-Varying Impact of Reviewers and Critics Volume (5 Equations)

A. Hedonic Product (Movies)



B. Utilitarian Product (Cameras)



Editor

Thank you for submitting your manuscript, POM-Mar-22-OA-0334, "Experience, a Double-Edged Sword: The Dynamic Impact of Different Message Sender Types on Different Message Receiver Types," to Production and Operations Management.

This paper adds to the growing literature on the (causal) effects of online reviews, and leverages a remarkable data set assembled from Amazon, Metacritic, and elsewhere to examine the relative / dynamic effects of different sorts of critics and reviewers. The popular DLM framework extracts a number of intriguing "crossover" effects, e.g., that novice reviewers are more impactful than experienced ones (something of a surprise), but that the opposite is true for critics. It also provides a managerially-useful method for assessing different consumer groups based on their susceptibility to various types of online opinion. The prose is cogent, covers the literature clearly and, aside from some murky graphics and nontrivial expository issues with the DLM (more on that below), makes following its conclusions fairly straightforward.

Before getting into the specifics of the reviews, let me point out that, collectively, these are among the most on-point, constructive reports – in the true sense of the term, moving the paper forward in specific ways – I've seen in my years at POM; and the SE ably summarizes the overall views in something like an action-oriented format. But there is another issue that sometimes occurs and, for the sake of full disclosure, should be mentioned as well. R1 mentions in the report that the paper has been reviewed at another journal, which is absolutely fine for POM (we do not ask for disclosure along these lines); and the authors related in their cover letter, as the journal requires, that there is overlap in the data in this paper and another in the literature. That is, there is absolutely no question of an ethical violation here (rather the opposite). However, R1 suggested that the manuscript under review at POM is rather similar to the version from another journal; so, and as part of my Editorial duties, I reached out to the other journal to verify. Although their process prevents them from supplying manuscripts, the journal did verify that many of the estimated values remained the same in the two papers, suggesting that not a great deal of updating was done between the two processes. I want to assure the authors and editorial team that none of this affected the outcome here at POM (and it's clear that R1 did not take this into account, but did rely on some of the prior report, which the authors will surely note as well).

Thank you for the opportunity to resubmit a new paper. We really appreciate the thoughtful comments from you, the SE, and the two reviewers. Based on this constructive feedback we have made substantial changes. Please see below for the key changes made to the manuscript.

<u>Summary of Key Changes</u>

1. *Two New Datasets.* Following your, SE's and reviewers' suggestions, we have collected movie dataset from hedonic category and camera dataset from the utilitarian category to make the findings more generalizable. We collected the audience reviews and critics ratings for movies from rottentomatoes.com. For cameras, we gathered consumer reviews from bhphotovideo.com, a website that sells digital products and expert critics' reviews from Google reviews.

2. *New Research Objectives.*

a. To investigate the impact of expertise (review valence and critics valence) on product sales/search.
b. To understand how identity (gender and ethnicity) of message senders influence product demand. More precisely, we are curious about how the identity of being a minority group member (e.g., female, non-white) interacts with subject expertise (critics vs. reviewers).
c. To study the time-varying impacts of different message sender types based on expertise and identity.
d. Compare the results between the hedonic and utilitarian categories.

3. *Key Findings.*
    a. *Expertise Effect.* For movies, overall, the valence of critics matters more than valence of reviewers on box office sales. In addition, the effect of review valence dies out quickly, but the impact of critics' valence stays relatively stable over time; For the camera category, we find that consumer reviews have larger impact on product sales than critics and that the impact stays positive and stable over time.
    b. *Identity Effect*. With respect to gender and ethnicity, we find very similar and consistent patterns across the two product categories.
        i. *Gender*. We didn't find the gender stereotype towards female reviewers or female experts: higher percentage of reviews and critics written by women increases the movie box office sales as well as the search for cameras.
        ii. *Ethnicity*. There is a potential bias against nonwhite reviewers: reviews written by nonwhite reviewers have significant negative impact on product sales, and the effect becomes stronger over time. However, being an expert can mitigate this bias, which shows that the identity of experts can act as a signal of competence.
    c. *Moderating role of gender and ethnicity*. We find that overall women and individuals from nonwhite racial backgrounds notably intensifies the influence of review valence on product sales. However, this trend does not hold true for nonwhite reviewers.

4. *Multivariate DLM.* Following your, SE's and reviewers' suggestions, we have significantly improved the model specification. There are five major changes regarding the dynamic linear model (DLM):
    a. We added product fixed effects in the model.
    b. We now include review valence and review volume as two additional dependent variables in the model system to account for endogeneity.
    c. We explicitly show the explanatory variables, all of the coefficients in the observation equations and state equations, the time-varying parameters, and the non-time-varying parameters.
    d. We use a multivariate t distribution for the error terms to address the potential skewness in the data.
    e. We use a random walk specification in the state equations.

    Please refer to the paper's Section 4 Model Specification for more details, pages 12–13.

5. *Streamlined Our Positioning*. We agree that there have been a lot going on in the previous version of the paper and that it made the focus of the paper unclear to readers. Following your, SE's and reviewers' suggestions, in the current version of the paper, we study the overall impact of reviews and critics on product demand, the heterogenous impact of minority message senders (women, nonwhite) as well as

the dynamic effects over time. Moreover, the discussion of experienced versus novice reviews have been removed. We also include a table to show the contributions of our research (please see table below). We hope our new positioning makes the research much clearer and more coherent.

6. *Managerial Implications*. Following your, SE's and reviewers' suggestions, we have significantly enriched the managerial implications based on the new findings of the paper. Please see the discussions in section 6 of the paper for details.

Table: Contribution Table

	Experience of Message Source - Critics and User Reviews		Demographics of Message Source				Methodology - Empirical/ Experiment	Product Category		DV
			Gender		Ethnicity					
	Overall Impact	Time Varying Impact	Overall Impact	Time Varying Impact	Overall Impact	Time Varying Impact		Utilitarian	Hedonic	
Our Study	Yes	Yes	Yes	Yes	Yes	Yes	Empirical	Yes (Cameras)	Yes (Movies)	Sales, Google Search
Azer et al. 2023	User Reviews Only	No	No	No	Yes	No	Experiment	No	Yes	Review Credibility
Chakravarty et al. 2010	Yes	No	No	No	No	No	Both	No	Yes	Sales
Chen et al. 2012	Critics Only	No	No	No	No	No	Empirical	No	Yes	Firm Value
Chiu et al. 2022	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Cox and Kaimann 2015	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Dellarocas et al. 2007	Yes	No	Entropy of gender distr. Of reviewers	No	No	No	Empirical	No	Yes	Sales
Deng 2020	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Gelper et al. 2018	User Forum Posts Only	Yes	No	No	No	No	Empirical	No	Yes	Sales
Gopinath et al. 2014	User Forum Posts Only	Yes	No	No	No	No	Empirical	Yes	No	Sales
Hoskins et al. 2021	Yes	No	No	No	No	No	Empirical	No	Yes	Reviewer Rating
Kim et al. 2023	Yes	No	No	No	No	No	Both	No	Yes	Sales
Lin and Xu 2017	User Reviews Only	No	No	No	Yes	No	Experiment	No	Yes	Purchase Intention
Niraj et al. 2015	Yes	No	No	No	No	No	Empirical	No	Yes	Sales
Wang et al. 2015	Yes	No	No	No	No	No	Empirical	Yes	Yes	Sales

Please see below for our responses to your specific comments. Your comments are shown in italics.

The paper has some compelling pros and cons. On the one side is the importance of the topic of online reviews, which has received a great deal of research attention. Balancing this is the restriction in the paper to one idiosyncratic category, sold through one online retailer (albeit by far the most prevalent one) over a short period of time; the paper is "about" message types in general… what can anyone actually say about that given results from a single category, one where experience is so critical and product "attributes" are murky? A related issue is that it's hard to see a straight connection from the topic, data, or results to Operations Management (although this is the Marketing Interface, papers do need to have a firm foot in both camps; more on this below). It's difficult to envision the paper appearing in POM without a broader base of categories, some of which would bear on production, which is often more related to durables marketing.

Thanks for the suggestions. We agree that the focus on music products in the previous version puts restrictions on the generalizability of the results and weakens the fitness of our research to the general interest of POM. We took these points seriously while revising the paper and made substantial changes to address your concerns: first, as mentioned in the previous summary, in this new version, we collect data for both hedonic product (movies) as well as utilitarian product (cameras), which we believe makes the results more generalizable across different product categories; second, based on the analysis on the two new datasets, we get some novel and interesting findings, which can have very important managerial implications for managers regarding the production and promotion of their products. For details, please see our discussion in Section 6 Discussion and Managerial Implications.

But I'm less focused on overall appropriateness than whether the model captures the correct data generating process here, and my own reading strengthens the consensus in the reviews that it doesn't fully do this. I don't want to harp on "endogeneity", so let's hold that aside and focus mainly on the dynamic elements, which are key to the entire undertaking. The paper would benefit from a DAG or some other method of precisely laying out (1) what influences what, (2) dynamically. You will see that the entire review team feels this element of the paper is opaque (e.g., R2's delicate phrasing, "The authors are a bit economical with the specifics of the model;" see R2's entire paragraph starting "First".) But the real problem concerns the dynamic effects the paper appears to assume away. R1 puts it well "DLM is now used only for sales. However, higher sales can generate more reviews and influence review valence in the future. Prices are often adjusted based on previous and expected sales."

Thanks for your comments. We agree that it is necessary to have a clear illustration of the model specification to show the inter-relationship among all the key variables and how the impact of OWOM will change over time. We have added more detailed explanations in Section 4 Model Specification. Please see pages 12-13.

Both reviewers are experts on the use of DLMs in marketing data, and it seems they cannot figure out exactly what the paper did; this means readers will find it very rough sledding. The closest I can come to understanding it is to simply focus on Equation (1), which says that log(Sales) is regressed on a set of known time-varying covariates and DMA-varying controls (let's hold aside that it's hard to know if any of these should be interacted, or have nonlinear effects, like in a GAM). This is just a regression, so the key *innovation here is the determination of the matrix, G, which will say which of the covariates has transient vs. long-term sales impact. Endogeneity is explained in Section 4.2, and involves a dichotomization (why?) of above-median precipitation; for valence, the IV is ratings for other categories (which makes sense, but is more open to question). What none of us is seeing is why the causal direction goes from all the variables in equation (1) to sales, but that sales doesn't reverse-cause some of the RHS variables. This would suggest a simultaneous equation model, and it's unclear why this was not used. Given that this came up in my own reading and those of essentially all the reviewers, especially R1 (who explains the issues very well), I'm not sure that simply using a DLM – which, though powerful, is hardly novel methodologically – is a suitable basis for any conclusions about sales drivers.*

Thanks for your helpful suggestions. We do agree with you, the senior editor and the reviewers that it's better to take endogeneity into consideration and model the key DVs and IVs using simultaneous equation model. We have revised our model specification according to your suggestions in the current paper, where sales, valence & volume of reviews and valence & volume of critics are modeled simultaneously. Please see the details of our model in Section 4 Model Specification.

So, where does this leave us? None of the review team is encouraging about the prospects for the paper, with only one opening the door, barely, to a revision. [And this is holding aside that the paper was apparently already reviewed, many of these issues raised, but were not really addressed.] POM's guideline for encouraging a revision is that there is a strong chance the paper will be close to conditional acceptance on the next round. Given that there is only one product category here and the DLM at the very least would need to include other causal possibilities, there's not much possibility of that.

Thanks for your suggestions. We understand the review teams' concerns with respect to the single product category as well as the DLM setup in the previous version of the paper. We have addressed these issues by collecting new data for movies and cameras and revising the specification of the dynamic linear model. We hope you find our results more reliable and more generalizable.

Moving forward, the paper will be categorized this way in the Manuscript Central system (note that formal rejection precludes any possibility of resubmission, even if the manuscript is radically overhauled). If the authors believe they can address the main concerns here – and I would caution that there are several distinct ones (as per the SE report), and most are nontrivial – the journal would consider a fresh review process. This serves to keep the door open. If the authors did wish to go down this road, I would endeavor to retain as much of this excellent team as practicable, unless the topic or methods changed to a degree where reviewers with alternative skill sets would be more appropriate.
We realize that this was not the desired outcome but hope that the authors will view the reports in the spirit of trying to guide the paper to external standards of excellence in the empirical marketing literature, and honestly assess whether they can move the paper in that direction for a new resubmission, as opposed to other outlets. If so, POM would be happy to consider the result.

Thanks again for giving us this opportunity to resubmit our work back to POM. We have benefited a lot from all the suggestions from the review team and we believe that there's substantial changes in the current version. We would really appreciate your consideration of this new submission.

Senior Editor

We greatly appreciate your summary of the feedback from the review team and your constructive and helpful comments. In this revision, we have closely followed the comments and suggestions from you, and the rest of the review team. We believe that the paper has greatly benefited from this process and hope that you will find the revised paper satisfactory. Please see a summary of key changes at the beginning of this document. Below we provide our point-by-point responses to your comments.

*1. **Generalizability**: I agree with Reviewer 1 (comment #1 and #2) that it is hard to generalize the results presented in the manuscript based on a single category of experience goods. In this regard, my suggestion is that the author(s) collect and analyze data from other categories, particularly utilitarian goods, in their manuscript.*

Great point. We appreciate your recommendation and agree that gathering data from a different category of utilitarian goods is important to make our findings more broadly applicable. In this version of the manuscript, we use two new datasets - a camera dataset (utilitarian good) and a movie dataset (hedonic good) for the analysis.

*2. **Model**: Reviewer 2 laments that the DLM model specification in the manuscript does not allow the reader to grasp the detailed model specification. Please also see Reviewer 1's comments (#1a, #1b, and #1c) about the usage of dynamic linear model, endogeneity, and possible skewness of the data. It appears that the author(s) have seen these comments before at another journal and have decided to do nothing about it. My suggestion is that the author(s) go back to the drawing board, examine Reviewer 1's comments very carefully, and incorporate them into their manuscript.*

Thanks for pointing this out. We completely agree that it would be helpful to make the model specification clearer. We now explicitly show the explanatory variables, all of the parameters in the observation equations and state equations, the time-varying parameters, and the non-time-varying parameters. In addition, as suggested by reviewer 1, we set up the model us a Multivariate DLM with t-distributed errors. Please refer to the paper's Section 4 for more details, pages 12–13.

*3. **Classification**: I also agree with Reviewer 1 (#1d) that the classification of novice versus experienced critics and reviewers at the corresponding 50% seems rather arbitrary. My suggestion is that the author(s) should just leave the variables as continuous (also, please see Sarang Sunder, Kihyun Hannah Kim and Eric Yorkston (2019) "What Drives Herding Behavior in Online Ratings? The Role of Rater Experience, Product Portfolio, and Diverging Opinions", Journal of Marketing, 83(6), 93-112) and examine the impact of volume directly on sales.*

We agree. In this new version of the paper, based on the review team feedback, we no longer focus on experienced versus novice reviews. Instead, our research is now centered around the roles of expertise (reviewers, critics), minority groups (female, non-white) and their interplay.

***4. Data**: The author(s) have not shown how representative and appropriate are the data from Amazon.com, metacritcs.com, and connecting them with geographically aggregated regions for the sales of music albums.*

I recommend that the author(s) collect data from a variety of platforms including Spotify (see Reviewer 1, comment #2b).

Thanks for the comment. We agree that the music category has its challenges. We now focus on two new datasets. We collected the audience review and critics ratings for movies from rottentomatoes.com. For cameras, we gathered consumer reviews from bhphotovideo.com, a popular website that sells digital products and expert critics' reviews from Google Reviews.

5. Conceptualization*: I agree with Reviewer 2 (under Positioning and Conceptualization) that the conceptualization in the paper is weak, particularly with respect to expert versus novices and online and offline sales. I suggest the author(s) come up with a solid conceptual framework and develop strong hypotheses under that umbrella framework.*

We apologize for the lack of focus in the initial submission. There were a lot of moving parts which made the conceptualization difficult. We have taken the feedback of review team to heart and have made conscientious effort to streamline our positioning. We no longer focus on experts vs novices and online vs offline sales impacts.

6. Underlying Process*: The author(s) have not established any underlying process for the results presented in the manuscript. Reviewer 2 writes, "In sum, I believe focusing on fewer, well-justified , well-motivated factors could lead to a much stronger paper." I agree with Reviewer 2's sentiment in this regard. I would suggest that the author(s) model the underlying process in their framework.*

Thanks for this comment. We agree that there have been a lot going on in the previous version of the paper and that it made the focus of the paper unclear to readers. In the current version of the paper, we focus on the overall impact of reviews and critics on product demand, the heterogenous impact of different demographic groups as well as the dynamic effect over time. We draw on literature on expertise and identity of OWOM message source to explain the findings.

7. Managerial Implications*: I also agree with Reviewer 2 that the managerial implications in the manuscript are weak. Reviewer 1 decided not to comment on the managerial implications since the results may change if the empirical model changes.*

We agree. In our current version, we have some interesting new findings, and we believe that these findings have very important managerial implications. Please see the discussions in section 6 of the paper.

8. Empirical Analysis*: Reviewer 2 (comments under "Empirical Analysis, Model Specification, and Results) raises four separate issues. One relates to album specific effects. Second is about time varying coefficients and their interpretation. The third relates to the intercept in the state equation. The final suggestion is to develop a Hierarchical Bayesian version of the model presented in the manuscript. My suggestion is that the author(s) rework their model and results and take into consideration the ideas presented by Reviewer 2.*

We really appreciate the comments and suggestions by you and Reviewer 2. First, we have added movie/camera specific effects into our model specification. Second, we have made the interpretation of time varying coefficients clearer. Third, we now use a random walk specification in the state equations. Fourth, we agree that a HB version would be interesting, but we don't have market level data for the two new datasets.

9. Other Issues Raised by the Review Team*: The reviewers have provided other valuable comments. These include (but not limited to): (i) using individual purchase data to uncover heterogeneous effects of reviews, (ii) distinguishing between different effects in markets with different market compositions versus different effects on different demographic groups, (iii) exploring heterogeneity in geographical markets, etc. The authors should carefully sort through these issues in an effort to craft a more insightful paper.*

We completely agree that the reviewers' other comments are very helpful as well. Please see our specific responses to Reviewer1-2b, Reviewer2-Positioning and Conceptualization (1), Reviewer 2-Positioning and Conceptualization (4), and Reviewer 2- Empirical Analysis, Model Specification, and Results (5).

Thank you again for your clear guidance. We really appreciate your time and effort in processing our manuscript and providing constructive feedback.

Reviewer 1

Thank you for your thoughtful comments. We found your constructive feedback very helpful in revising the paper. Please see a summary of key changes at the beginning of this document. Below we provide our point-by-point responses to your comments.

1. As I mentioned last time, the dataset used in this paper is a single category of experience goods, so the results are hard to generalize. If it is difficult to obtain data from another category of utilitarian goods, at least the modeling needs to be revised for this dataset.

Thanks for this suggestion. We agree that to make our results more generalizable, it is necessary to collect data from another category of utilitarian goods. In this version, we conduct new analysis using two new datasets - a movie dataset (hedonic good) and a camera dataset (utilitarian good).

1.a Although using a dynamic linear model (DLM) for studying the "evolutionary" effects of reviews is the right idea, the current set-up in Equations (1) and (2) is not appropriate for this application. DLM is now used only for sales. However, higher sales can generate more reviews and influence review valence in the future. Prices are often adjusted based on previous and expected sales. Therefore, PRICE, VOLUME_REV, VALENCE_REV_EXP, VALENCE_REV_NOV, VOLUME_CRITIC, VALENCE_CRITIC_EXP, VALENCE_REV_NOV on pages 12 and 13 should also be modeled as the y_t *variable in the DLM, instead of being in the* F_t *matrix. How to add the two variances VARIANCE_REV and VARIANCE_CRITIC into the DLM will require more deliberation and perhaps a stochastic volatility model can be useful. A multivariate DLM is not difficult to fit given the data you have.*

Great points. Based on your advice we now fit a multivariate DLM. Please see below for model specifics.

1) In the current version of the mansucript, we added review valence and review volume as another two dependent variables in the DLM model. To keep the model parsimonious, we focus on the review valence and review volume, and model the volume and valence of reviews simultaneously in the current model specification. However, we also did a robustness check using a model setup with 5 dependent variables (SALES/SEARCH, VALENCE_REV, VOLUME_REV, VALENCE_CRIT, VOLUME_CRIT) in the observation equation and we get similar results as the main model with three dependent variables SALES/SEARCH, VALENCE_REV and VOLUME_REV. Please see Web Appendix C for more details.

2) In the movie dataset, there exists a significant correlation between VARIANCE_REV and VALENCE_REV, which is -0.8451. Additionally, the correlation between VARIANCE_CRIT and VALENCE_CRIT is -0.5453, also significant at the 0.05 level. Similarly, in the camera dataset, the correlations between VARIANCE_REV and VALENCE_REV is -0.8190, and the correlations between VARIANCE_CRIT and VALENCE_CRIT is -0.3527, both of which are statistically significant. To avoid potential multicollinearity concerns and given that they are not central to our research, VARIANCE_REV and VARIANCE_CRITIC variables have been omitted from the model. For more detailed information, please see model specification on pages 12-13 of Section 4.

3) Thanks for your comments regarding the potential impact of prices. We agree that modeling price is important in the music category (our dataset in the first round). However, we are now focusing on two completely new datasets - movies and cameras. For the movie dataset, consistent with prior research, our DV is box office sales (in dollars), so price won't be an independent variable in the model. For cameras, we checked the listing prices for cameras over time, and found that the prices of the cameras are very stable and there's little variation during our data period (please see the table below for more information). However, there are variation in the price levels across cameras which we control for by including product fixed effects in the model. Therefore, we no longer have *PRICE* in our model specification.

Table: Summary of Six Months' Listing Price for Cameras since Release

Cameras	Mean_price	Sd_price	Min_price	Max_price
Canon EOS M50 Mark II	929.98999	0	929.98999	929.98999
Canon EOS R10	1379	0	1379	1379
Canon EOS R5	4999.8999	0	4999.8999	4999.8999
Canon EOS R5 C	5550	0	5550	5550
Canon EOS R6	3599.99	0	3599.99	3599.99
Canon EOS R6 Mark II	2799	0	2799	2799
Canon EOS R7	2099	0	2099	2099
Canon EOS Rebel T8i	899.98999	0	899.98999	899.98999
Canon EOS-1D X Mark III	6499	0	6499	6499
Canon PowerShot Zoom	299.98999	0	299.98999	299.98999
DJI Pocket 2	349	0	349	349
DJI RS 2	499	0	499	499
Fujifilm GFX 100S	5999.9502	0	5999.9502	5999.9502
Fujifilm GFX 50S II	4499.9502	0	4499.9502	4499.9502
Fujifilm X-E4	1049.95	0	1049.95	1049.95
Fujifilm X-H2	2899.95	0	2899.95	2899.95
Fujifilm X-H2S	1379	0	1379	1379
Fujifilm X-S10	999.17651	0.37051052	999	999.95001
Fujifilm X-T200	799.95001	0	799.95001	799.95001
Fujifilm X-T30 II	999.95001	0	999.95001	999.95001
Fujifilm X-T4	2099.95	0	2099.95	2099.95
Fujifilm X-T5	2199.95	0	2199.95	2199.95
Fujifilm X100V	1399.95	0	1399.95	1399.95
GoPro Hero9 Black	449	0	449	449
Hasselblad X2D 100c	14884.95	0	14884.95	14884.95
Leica M10 Monochrom	8923.415	46.471096	8895	8999
Leica M10-R	9999.9004	0	9999.9004	9999.9004
Leica M11	9999	0	9999	9999
Leica Q2 Monochrom	7999.9902	0	7999.9902	7999.9902
Leica SL2-S	8393	0	8393	8393

Nikon Coolpix P950	796.95001	0	796.95001	796.95001
Nikon D780	2796.95	0	2796.95	2796.95
Nikon Z fc	959.95001	0	959.95001	959.95001
Nikon Z30	849.95001	0	849.95001	849.95001
Nikon Z5	1399.95	0	1399.95	1399.95
Nikon Z6 II	2996.95	0	2996.95	2996.95
Nikon Z7 II	2996.95	0	2996.95	2996.95
Nikon Z9	5499.9502	0	5499.9502	5499.9502
OM System OM-1	1199.99	0	1199.99	1199.99
OM System OM-5	1199.99	0	1199.99	1199.99
Olympus OM-D E-M1 Mark III	1799.99	0	1799.99	1799.99
Olympus OM-D E-M10 IV	799.98999	0	799.98999	799.98999
Olympus PEN E-P7	1199.99	0	1199.99	1199.99
Panasonic Lumix DC-BS1H	1997.99	0	1997.99	1997.99
Panasonic Lumix DC-G100	748.39862	0.80855638	747.98999	749.98999
Panasonic Lumix DC-GH6	2799.99	0	2799.99	2799.99
Panasonic Lumix DC-S5	2297.99	0	2297.99	2297.99
Pentax K-3 Mark III	2037.7878	78.544334	1999.95	2199.95
Pentax KF	849.95001	0	849.95001	849.95001
Ricoh GR IIIx	1049.95	0	1049.95	1049.95
Ricoh WG-70	279.95001	0	279.95001	279.95001
Sigma fp L	2499	0	2499	2499
Sony FX3	3899.99	0	3899.99	3899.99
Sony FX30	2199.99	0	2199.99	2199.99
Sony ZV-1	798	0	798	798
Sony ZV-1F	499.98999	0	499.98999	499.98999
Sony ZV-E10	698	0	698	698
Sony a1	6499.9902	0	6499.9902	6499.9902
Sony a7 IV	2499.99	0	2499.99	2499.99
Sony a7C	1798	0	1798	1798
Sony a7R IIIA	2799.99	0	2799.99	2799.99
Sony a7R IVA	3499.99	0	3499.99	3499.99
Sony a7R V	3899.99	0	3899.99	3899.99
Sony a7S III	3499.99	0	3499.99	3499.99

1. b. If the DLM is set up appropriately as suggested above, some of the endogeneity issues can be resolved by discussion. Again, the discussion of the endogeneity problems in Section 4.2 is not rigorous. The mechanism of the endogeneity of those variables is unclear. The authors should demonstrate how these endogenous variables in each period are correlated with which errors in the DLM and what factors may cause such correlations. To justify the choice of the instrumental variables, it must be shown how exactly these IV's can identify the effects of the endogenous variables in this particular setting.

Thanks for the comment. We completely agree. We no longer use instrumental variables instead we set up the model as a multivariate DLM as you suggested.

1.c. Although I have not seen the exact data in this paper, my own experience with sales and online review data shows the data can be highly skewed or heavy-tailed, even after log transformation. Using normal errors in the DLM may not give you the best model. Implementing t-distributed errors in this model is not difficult, especially using Bayesian estimation (normal and inverse chi-squared mixture). This model tends to be more robust, so it should be compared with the normally distributed model.

This is an excellent point. We agree that multivariate t distribution would be a better option to address the potential skewness in the data. We have implemented the multivariate t distribution errors in the current model. Please see model specification in section 4.

1.d. The classification of experienced and novice reviews are not handled rigorously. It is not clear why the top 50% of reviewers and critics based on their review volume should be considered experienced whereas the bottom 50% are novice. As I mentioned, you can try to incorporate heterogeneous effects and hierarchical modeling in the DLM based on reviewers' historical data. If such a hierarchical model is hard to fit, at least you should try some clustering methods to segment reviewers.

We agree. In this new version of the paper, based on the feedback from the review team we no longer focus on experienced versus novice reviews. Instead, our research is now centered around the roles of expertise (reviewers, critics), minority groups (female, non-white) and their interplay.

2. For more generalizable results, I still suggest the authors assemble another dataset for a category of utilitarian goods. It is well known that reviews and critics have different effects on experience and utilitarian goods. For example, the effects of technical and expert critic reviews on electronics should be very different from those of music albums and fictions. Finding data from such as category should not be that difficult.
2.a From my own experience, computer games and music albums tend to have most sales concentrated in the first few weeks after their release. Utilitarian goods such as electronics can have very different sales patterns over time. The volume, valence and variance of the reviews generated by the purchasers of utilitarian goods will also have different dynamic patterns. The short and long-term influence of reviews can also be different. It should be interesting to compare the effects of reviews for experience and utilitarian goods.

Based on your advice, we collected movie dataset from hedonic category and camera dataset from the utilitarian category. We conduct similar analysis for these two product categories and compare the effects of reviews for these two types of goods. Please see the detailed results in section 5 of the paper.

2.b. The paper needs to justify why Amazon.com and metacritic.com reviews are representative of consumer and expert reviews for music albums. Note that music albums may include singles which are sold and reviewed on other websites (e.g., Spotify). Is it appropriate to use only the reviews from these websites for both online and offline sales?

Good points. We agree that the music category has its challenges. Hence, in this round we move to two completely new datasets. Specifically, we collect audience reviews and critics ratings for movies from rottentomatoes.com. For the cameras, we gathered consumer reviews from bhphotovideo.com, a website that sells digital products and expert critics' reviews from Google reviews.

The paper uses aggregated sales data at a certain level instead of individual purchase data Can such aggregate data identify the heterogeneous effects of reviews? Rigorous theoretical and statistical discussions are needed to justify this approach. It is not wrong to use such data, but different effects in the markets of different demographic compositions are not equivalent to different effects on different demographic groups.

Thanks for your suggestions. Our research questions and datasets have changed. We no longer focus on the demographics of the consumers. Instead, we look at the expertise and identity of the message senders (experts, critics, women, non-white). However, we agree that one could get interesting insights if we had access to individual level purchase data. We now mention this in future research section. Please see page 27.

Thank you again for your guidance. We hope that we have adequately addressed your concerns in this round of revision.

Reviewer 2

Thank you for your constructive feedback and suggestions. We apologize for not being very clear in the writing in the previous round. We have taken your feedback to heart and completely rewritten the paper. We hope you find the revised version much easier to follow. Please see a summary of key changes at the beginning of this document. Below we provide our point-by-point responses to your comments.

Positioning and Conceptualization

1. I believe that the authors are trying to do too much in a single paper. Whereas I find the idea of exploring the dynamic effect of ratings on early sales of products interesting, I am not sure whether the additional factors explored contribute much given the setting and the data that is available.

Variation across reviewer types and, to a lesser extent, geographical markets are potentially interesting sources of variation to explore. However, I cannot say the same for reviewer activity levels or sales channels.

Thanks for your great points. We agree that there was a lot going on in the previous version of the paper and that it made the focus of the paper unclear. In the current version of the paper, we focus on the overall impact of reviews and critics on product demand, the heterogenous impact of different demographic groups as well as the dynamic effect over time. Hope you like the current positioning of the paper.

2. Especially for online versus offline sales distinction, I cannot clearly see why we should expect different effects. The motivation (p.3), the intuition behind the findings (p.4), and the expectations (p.8) are not very clear. Should the authors decide to pursue this line of inquiry, stronger justification may be needed.

We completely agree. In this new version, we use two new datasets and have modified our research questions based on the feedback received. We no longer have online versus offline analysis.

3. As for reviewer activity levels, I believe the potential differences rest on a strong assumption. That is, the decision-maker consuming the user- or expert-generated content is readily able to distinguish the content generators' review activity levels. This is not the case. While it is certainly possible to access this information, it is not clear how heavily the decision-makers rely on this. Consequently, I wonder whether the reviewing activity level is confounded with other characteristics of the reviews. In sum, I believe focusing on fewer, well-justified, well-motivated factors could lead to a much stronger paper.

Great point. As we mentioned earlier, there have been substantial changes in the current version. In current manuscript, we don't focus on the reviewer activity levels (experienced versus novice) anymore.

4. As mentioned above, exploring the heterogeneity in geographical markets seems interesting. However, I am not sure age, gender, race, and income are the most appropriate variables to characterize the geographical markets in this context. Yet, I am unable to propose alternatives at this point.

Thanks for pointing this out. We agree. Based on the feedback in the previous round we no longer focus on consumer demographics in different geographical markets in the current version.

5. Finally, I was slightly puzzled by the authors' decision to position this work in the influencer marketing literature. User-generated reviews on e-commerce platforms, such as Amazon, cannot be considered influencer-generated content. Therefore, the reach vs. relevance discussion or managerial challenge does not apply in this context. Related, the managerial implications presented on p.22 cannot be drawn. Though I am less familiar with the Metacritic platform, I gather that the critics' reviews cannot be considered influencer-generated content either. Therefore, I urge the authors to position this work in the broader eWoM literature.

Thanks for this comment. Based on your advice we now position our research around the literature on eWOM and user-generated content. We have now removed the influencer marketing discussion.

Empirical Analysis, Model Specification, and Results

1. First, the authors specify the model on p.12. However, this compact DLM form does not allow the reader to grasp the details of the specification. Please present the observation and the state equations in as much detail as needed. Please explicitly show all the explanatory variables in the observation equation and all the coefficients in both equations with the appropriate indices. It is unclear in this version which variation identifies which coefficients at what level.

Thanks for your great suggestions. We totally agree that it would be beneficial for the readers to further clarify all the details in the model. Based on your advice, in the current version, we explicitly show all the explanatory variables, all the parameters in both the observation equations and state equations, the time-varying parameters and the non-time-varying parameters. Please see more details on Pages 12-13 in Section 4 of the paper.

2. Second, I am not sure whether there is an album-specific fixed-effect somewhere in the model. If not, the (critics') ratings might be capturing quality differences across the albums. However, the authors state that they estimate the model separately for each music album (p.12 and p.15). If that is the case, album-specific fixed-effects may not be needed. But this raises another question. At what point in the estimation process these album-specific time-varying effects are pooled, and how, to arrive at the results presented in Tables 2-4 and Figure 2? Can the authors provide more information on this? Please also show the 95% CIs of the sampled states in Figure 2.

Thanks for the excellent points.
1) In the current version of the paper, we have a movie or camera fixed effect in the model.
2) We did the parameter estimation for each period and then calculated the average across all the time periods to arrive at the results presented in the tables.
3) Based on your suggestion, we now present the confidence intervals for the figures as well as the tables.

3. Third, the specification presented on p.12 suggests that all covariates have time-varying effects. First, do we need all of them to be time-varying (e.g., age or time since launch)? Second, if they all have time-varying coefficients, then what do the coefficient estimates presented in Tables 2-4 correspond to? The last period? The first period? I don't understand where these estimates are coming from. Could it be that the authors estimate and report results from a time-invariant version of the model at times?

Thanks for your great suggestion. This was very helpful. We now have two sets of variables – one with time varying parameters and one with time invariant parameters. In our current model specifications, we are interested in the dynamic impact of OWOM and reviewer demographics on movie box office revenues, therefore they are included in the time-varying variable matrix. Variables with non-time-varying parameters are the control variables. Please see our detailed model specification on Pages 12-13.

4. Fourth, as far as I can see, the state equation does not have an intercept. Such AR(1) specifications produce the parameter evolution patterns presented in Figure 2. Introducing an album-specific intercept in the state equation might be an option. However, without exogenous variables in the state equation, separately identifying the intercepts and the carryover coefficients might be a challenge. As the authors are primarily interested in the temporal variation in coefficient estimates, a random walk specification for the state equations might be the better option (i.e., setting G to an identity matrix). This will produce the most flexible coefficient evolution pattern.

We agree. As per your suggestion, we have added a movie/camera fixed effect in our current model specification. We also use a random walk specification in the state equations.

5. Finally, to explore the variation across "receiver types", the authors add interaction terms to the model (p.19). Is it the case that the authors are now estimating separate time-varying coefficients for, say, the variables "valence of expert reviews" and "valence of expert reviews x % female population"? What do the temporal patterns of the coefficients of these interaction terms look like? Interpreting such time-varying coefficients may not be very straightforward. A more intuitive approach could be a Hierarchical Bayesian version of the model. For instance, the authors can specify market-specific state equations with random drift and, if they can justify the choice, random carryover coefficients. The means of these random-effects distributions can then be expressed as functions of market-specific covariates. An alternative approach might be a Hierarchical DLM a la Gamerman and Migon (1993, JRSS Series B).

Thanks for this comment. We no longer focus on receiver types given the review team feedback from the previous round. Instead, we now focus on the message source. Specifically, we study the overall impact of reviews and critics on product demand, the heterogenous impact of minority message senders (women, nonwhite) as well as the dynamic effects over time.

Thank you again for your time and effort in reviewing our manuscript. We really appreciate your guidance and hope that we have adequately addressed your concerns in this revised manuscript.

References:

Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing*, 84(2), 233-242.