2016 | ISSUE 2

A first look at new ideas in marketing

insights from MSI

Marketing ROI

Why firms should invest in social media

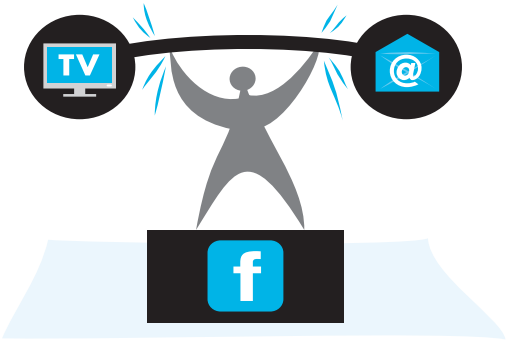
Although there's little doubt that a social media presence increases sales, marketers have had a hard time measuring the return on their investments in social media. A 2014 decision by Facebook that requires firms to pay to reach their fans through social media has only increased the pressure to justify spending. How can marketers make the case that the cost of creating and updating social media content is worth it?

New research offers critical guidance. "The Effects of Firm Generated Content in Social Media on Customer Behavior: An Empirical Examination," by Ashish Kumar, Aalto University, Ram Bezawada, SUNY Buffalo, Rishika Rishika and Ramkumar Janakiraman, University of South Carolina, and P.K. Kannan, University of Maryland, empirically links firm-generated social media to in-store purchase behavior. It also illustrates a methodology to calculate returns on social media investments by using a novel data set.

"Our study is the first to provide evidence of the impact of firm-generated content on the individual customer metrics that matter most," says Bezawada. "It also points to important ways that FGC complements and strengthens the impact of TV and email marketing."

In their study, Bezawada and colleagues used data from a large retailer of wine and spirits with multiple stores in the northeastern U.S. Data included customer transaction data, social participation data (i.e., customers who "liked" or shared the firm's social media content or commented on the firm's social media page), and attitudinal data obtained through a customer surveys.

Social media postings were assessed for valence (positive, neutral, negative), receptivity or popularity (that is, the total of comments, likes, and shares), and susceptibility (customers' predisposition to social media as reported on the customer survey). They investigated the effect of this content on three key customer metrics: spending, cross-buying, and customer profitability, as well as the synergistic effects of FGC with television advertising and email communication.



Power of synergy: Firm content on social media significantly strengthened the effect of TV and email marketing on spending and cross-buying.

inside

Strong case for growing the fan base

The authors found that the receptivity, or popularity, of a social media posting had the greatest effect on customer purchases. (Receptivity's elasticity with respect to customer spending and cross-buying was .019 and .086 respectively; these values were .013 and .029 for valence and .009 and .008 for customer susceptibility.) In other words, a neutral or even negative post with a high customer response could impact sales more than a positive post that draws no comments, likes, or shares. This was true even among customers who said their purchase decisions were not swayed by what they read on social media.

A neutral or even negative post with a high customer response could impact sales more than a positive post that draws no comments, likes, or shares.

"Since FGC receptivity involves direct customer involvement, we believe our results make a strong case for firms to grow [their] 'fan' base and monitor the level of customer engagement and post measures of popularity on their brand social media pages," the authors write.

Even more powerful were the synergistic effects of social media and traditional forms of advertising. While firm-generated content had less effect than television and email on total sales, it significantly *strengthened* the effect of TV and email marketing. Synergy with firm-generated social media content was responsible for substantial increases in both customer spending and cross-buying (1.03% and .84% respectively for TV, and 2.02% and 1.22% respectively for email).

Firm-generated content had a *greater* effect than TV and email advertising for cross-buying (elasticity .059 versus .051 and .043), which is an important indicator of customer loyalty, since customers who buy across categories have greater switching costs and longer relationships with firms.

The firm-generated content effects were greater among customers who had a longer relationship with the retailer, were tech savvy, or were more active in social media. Based on this finding, "we suggest that special product focused 'interaction forums' could be created for such customers," the authors write. Surveys can help firm identify their tech-savvy and social network-prone customers and encourage them to join the firm's social media page.

Overall, the authors write, "The clear message from our study is that social media marketing matters and that managers should embrace it to communicate and nurture relationships with customers. We find that investing in developing a social media community with a dedicated fan base (e.g., Facebook page) can significantly strengthen customer-firm relationships and can lead to a definitive impact on the firm's revenues and profits."

BY DEBORAH KREUZE

From "The Effects of Firm Generated Content in Social Media on Customer Behavior: An Empirical Examination" by Ashish Kumar, Ram Bezawada, Rishika Rishika, Ramkumar Janakiraman, and P.K. Kannan (MSI Report No. 16-111)

Issue theme

New data, new methods, new skills

This issue of *Insights from MSI* focuses on the fourth of MSI's five 2016–18 research priority topics: New data, new methods, and new skills: How to bring it all together?

Facing a flood of new data and analytic approaches, marketers need to synthesize disparate methods—old and new—to gain insight and drive action. It is a huge challenge: one MSI member described it as "the end of marketing research as we know it." Each of the working papers featured here addresses an important aspect of this brave new world of marketing analytics.

"Why firms should invest in social media" describes research by Ashish Kumar, Ram Bezawada, Rishika Rishika, Ramkumar Janakiraman, and P.K. Kannan—the first to offer evidence linking firm-generated content to individual customer purchase behavior. Not only does a brand community "fan base" boost customer spending and cross-buying, they find; firm social media content strengthens the effects of traditional marketing communications.

How can marketers extract timely intelligence from the noise of big data? Yi Zhao, Nuo Xu, and Yingge Qu develop an innovative approach, outlined in "Detecting customer milestones in cross-channel data." Adapting the control chart method, their approach enables them to detect customer life changes, signaled by small changes in behavior, *before* the window of marketing opportunity closes.

"Mobile diaries in marketing research" describes work by Mitchell Lovett and Renana Peres to evaluate the accuracy of mobile diaries to measure TV audiences. The researchers suggest that mobile diaries can complement traditional measures like the Nielsen's panel, capturing on-the-go viewing that passive measures miss.

How do customers feel about all this data collection? Our page three article offers a rare customer-centric view of marketing practices. Kelly Martin, Abhishek Borah, and Rob Palmatier find that merely *collecting* information makes consumers feel vulnerable. The bright side? Marketers can increase transparency and customer control to mitigate those effects.

If you'd like to dig deeper into any of these ideas, full working papers can be downloaded at www.msi.org.

—Susan Keane, Editorial Director

Marketing Science Institute Working Paper Series 2016
Report No. 16-111

The Effects of Firm Generated Content in Social Media on Customer Behavior: An Empirical Examination

Ashish Kumar, Ram Bezawada, Rishika Rishika, Ramkumar Janakiraman, and P.K. Kannan

Report Summary

Social media spending is expected to grow faster than any other form of online marketing, and marketers are concerned about measuring the returns on those investments.

In this study, Ashish Kumar, Ram Bezawada, Rishika Rishika, Ramkumar Janakiraman, and P.K. Kannan examine the effect of social media engagement on individual-level customer purchase behavior. Specifically, they examine the effect of firm generated content (FGC) in social media on three key customer metrics: spending, cross-buying, and customer profitability. They also examine the synergistic effects of FGC with television advertising and email communication.

Based on a large New York state retailer, their dataset comprises customers' social media participation data, customer transaction data, and attitudinal data obtained through surveys.

Their results indicate that after accounting for the effects of television advertising and email marketing, FGC has a positive and significant effect on customer spending and cross buying. Further, FGC works synergistically with both television advertising and email marketing.

They also find that the effect of FGC on spending and cross buying is greater for more experienced, tech-savvy and social-media-prone customers.

For three characteristics of FGC—valence, receptivity, and customer susceptibility—they find that all the three components have a positive impact, and the effect of FGC receptivity is the largest.

These findings suggest that managers should embrace social media and exploit the synergistic relationship between social media and TV and emails. Further, they should encourage participation by customers who have a longer tenure with the firm, and should monitor the popularity of firm generated content.

Ashish Kumar is Assistant Professor of Marketing, School of Business, Aalto University. Ram Bezawada is Associate Professor of Marketing, School of Management, State University of New York, Buffalo. Rishika Rishika is Clinical Assistant Professor of Marketing, Darla Moore School of Business, University of South Carolina. Ramkumar Janakiraman is Associate Professor of Marketing and Business Partnership Foundation Research Fellow, Darla Moore School of Business, University of South Carolina. P.K. Kannan is Ralph J. Tyser Professor of Marketing Science, Robert H. Smith School of Business, University of Maryland.

Acknowledgements
We thank Professor Arun Jain and the Research Group in Integrated Marketing at the School of Management, SUNY Buffalo for access to the data and research assistance; and the participants at the INFORMS Marketing Science and AMA Summer Educators Conferences for their suggestions. We also thank Marketing Science Institute for financial support made available through a competitive research grant (#4-1871).

1. INTRODUCTION

With the dramatic change in the media landscape in recent years, firms have embraced social media as a means to engage with their customers. Recent business reports suggest that total spending on social media advertising has increased worldwide ($17.74 billion in 2014 versus $11.36 billion in 2013 which amounts to an increase of 56.2%) and that social media engagement drives sales (*PR Newswire* 2011; *eMarketer* 2015). However, the same studies also suggest that over 80% of marketers are concerned about measuring the returns on investment from social media. Recently, the popular social networking site, Facebook, implemented policy changes to filter out unpaid promotional material in users' news feed that businesses post as status updates. This policy change will make it difficult for businesses to reach their Facebook "fans" with marketing content that is not paid for (*Wall Street Journal* 2014). While this illustrates the value of firm initiated content on firms' social media pages, it also calls into question the added value of such postings beyond traditional media marketing (e.g., television advertisements) and/or other digital media marketing communication (e.g., emails).

In this study, we examine the effect of social media engagement on (individual level) customer purchase behavior. More specifically, we study the effect of *firm generated content* (FGC henceforth), i.e., firm initiated marketing communication in its official social media pages, on two key customer metrics — customer spending and cross-buying behavior — that capture two different dimensions of customer-firm relationship, the transaction side and the relationship side, respectively. We note that while customer spending is the "customer basket size" based business performance metric that firms typically focus on, the degree of customer cross-buying captures the breadth of a customer's relationship with a firm (Kumar, George and Pancras 2008). Given the focus on customer profitability in the customer relationship management (CRM)

literature (Kumar, Venkatesan and Reinartz 2008), we also examine the effect of FGC on customer profitability.

While firms are investing more in social media, marketing communications via television advertisements and emails are also an important avenue for firms to connect with their customers. From the perspective of integrated marketing communications (Naik and Raman 2003), it is vital to understand the relative efficacy and synergy between these different media for marketing communications. Therefore, the first objective of this study is to examine the *main* effects of FGC and its *synergistic* effects with television advertising and email marketing on customer spending and cross-buying.

In the social media era, the term "social CRM" is gaining prominence whereby firms engage in managing customer relationships through social media (Malthouse et al. 2013). For *social CRM* to be effective, it is crucial for a firm to understand how customers respond to FGC and whether certain segments of customers can benefit more from the firm's social engagement efforts. Thus, our second objective is to understand how the effect of FGC varies across different customer segments. We focus on the following customer characteristics that are relevant in our context: length of customer-firm relationship, customers' technology savviness, and customers' propensity to use social networking sites regularly. These characteristics account for customers' motivation and ability to process information available via online channels such as email and social media marketing communications.

To meet our objectives, we use micro-level customer behavior data compiled from multiple sources. We combine data on customers' participation in a focal retailer's social media (hosted by a popular third party) page, individual customer level in-store transaction/purchase data available both before and after the retailer's social media engagement efforts and survey data on customer attitudes towards technology and social media. Leveraging this unique dataset, we

study the effect of customer engagement via social media on three customer metrics, spending, cross-buying and customer profitability. To account for endogeneity concerns that could arise due to (customer) self-selection and to establish the effect of FGC on customer behavior, following recent studies in marketing (Huang et al. 2012), we employ the combination of propensity score matching (PSM) and difference-in-differences (DID) analysis. A recent business report (*Loyalty One* 2012) suggests that while even short public statements on social media can spur transaction activity, more elaborate posts that elicit higher level of customer participation can have a significant impact on an individual's purchase behavior. We thus incorporate a rich formulation of FGC that captures three different components: FGC valence/sentiment, FGC receptivity, and customers' susceptibility to FGC. Following recent literature (e.g., Tirunillai and Tellis 2012; Das and Chen 2007) we rely on the *naïve Bayes Algorithm/Classifier* to classify the valence or the sentiment of the postings. FGC receptivity takes into account customers' response to social media messages and FGC susceptibility measures customers' predisposition towards using social media. This construction of FGC not only captures a firm's effort in creating meaningful content but also helps shed light on the role of customer response to FGC and the underlying mechanisms that may drive observed FGC effects.

Our results indicate that after controlling for the main effects of television advertising and email marketing and after ruling out the issue of customer self-selection, FGC has a positive and significant effect on customer spending and cross buying behavior. We further find that FGC works synergistically with both television advertising and email marketing. We find that the synergistic effect of FGC and email marketing is greater than the synergistic effect of FGC and television advertising. Our results also suggest that the effect of FGC is greater for more experienced, technologically-savvy and more social-network prone customers. More

importantly, we find that FGC is positively associated with customer profitability. By linking the effects of FGC and its interaction effects with television advertising (traditional media) and email marketing (digital media) to customers' in-store purchase behavior, the results of this study contribute to the research streams in integrated marketing communications and multi-channel marketing. From practitioners' perspective, we quantify and compare the size effects of FGC to that of television advertising and email marketing. By establishing the effect of FGC on customer profitability, we show that brand managers can use FGC not only for promoting products in social media but also for engaging with their customers and nurturing profitable relationships with them.

2. LITERATURE REVIEW

Studies in the area of social media have primarily focused on the effects of user generated content (UGC) on market outcomes in various contexts such as book sales (Chevalier and Mayzlin 2006), movie box office revenues (Chintagunta, Gopinath and Venkataraman 2010) and music album sales (Dhar and Chang 2009). Some studies have examined the motivations that underlie individuals' decisions to contribute content to social media (e.g., Toubia and Stephen 2013) while others have focused on how UGC interacts with traditional media marketing (Stephen and Galak 2012). Given that UGC serves as an effective source of word of mouth (Godes 2011) and an indicator of product quality (Tirunillai and Tellis 2014), the focus on the effects of UGC is understandable.

As firms increasingly rely on social media to engage with customers, recent studies have attempted to understand different aspects of firms' engagement via social media. For example, Schulze, Schöler, and Skiera (2014) study the reach of viral marketing campaigns shared via social media and examine the effect of different types of sharing mechanisms (unsolicited

messages, messages with incentives, direct messages from friends and broadcast messages from strangers) on the reach of high-utilitarian versus low-utilitarian Facebook apps. Some studies have also examined how firms can harness the power of social media and the impact of social media marketing efforts on firms' ROI (Kumar et al. 2013). With respect to current research, one of the studies that is relevant to ours is Danaher and Dagger (2013). The authors focus on the impact of a single promotional sale campaign that was advertised via ten different types of media including traditional and social media. They find that seven out of the ten media influence purchase outcomes and provide important insights into multi-media resource allocation.

In this study, we extend and contribute to this stream of literature by examining the effect of a firm's initiative to engage with its customers via social media (FGC) over time. Since engagement via social media may take time to impact customer purchase behavior and could become more effective as the size of the social media community increases, studying social media communications and customer behavior over time can provide more meaningful insights. Unlike other media, firms' communications through social media platforms could be part of "equity" building efforts that are particularly aimed at managing brands and nurturing customer relationships (Gensler et al. 2013). Therefore, while promotional sale campaigns such as the one analyzed in Danaher and Dagger (2013) help marketers understand the value of multi-media blitz campaigns, our focus on FGC includes both promotional and non-promotional messages that go beyond the purpose of generating short term sales and help strengthen the bond between customers and firms.

Firm generated content (FGC) is, in effect, a multi-faceted construct and its effect will depend on the message sentiment, customers' response to the message and customers' innate disposition towards social media. We take into account all three of these factors and construct a composite measure that comprises FGC valence, receptivity, and customer susceptibility (these

dimensions are further explained in the Methods section). This permits us to advance the understanding of how FGC works in creating and sustaining firms' long-term relationship building efforts and helps differentiate our study with other studies engaged in studying the effect of social media communications. Similar to Danaher and Dagger (2013), we assess the ROI of social media by linking the effect of FGC to customer profitability. Finally, we make a concerted effort to account for customer heterogeneity and rule out inherent self-selection issues to establish the link between FGC and customer behavior thus making new and significant contributions to this nascent research stream.

3. CONCEPTUAL FRAMEWORK

In this section, we develop and present a conceptual framework (see Figure 1) that helps link the effect of FGC on customer purchase behavior. Before we do so, we first describe our key dependent variables below.

3.1. Customer Behavior: Spending Versus Cross-buying

In this study, we focus on understanding the effect of FGC on two different customer behaviors —customer spending and customer cross-buying —that help us look at two different aspects of the customer-firm relationship: the transaction side and the relationship side, respectively. Customer spending (in total dollars) captures the transactional value of the customer to the firm. By focusing on customer spending, we are able to capture how FGC impacts a firm's top-line and also benchmark the effectiveness of FGC vis-à-vis television advertising and email marketing. However, from a long-term perspective, a customer's cross-buying behavior, expressed in terms of the number of *different* product categories that a customer purchases, signals the intensity of the relationship between the customer and the firm (Verhoef and Donkers 2005; Shah et al. 2012). Firms often attempt to sell additional products and services to customers

in order to engender greater customer loyalty. The reasoning behind firms' cross-selling approach is that customers who buy across several different categories have greater switching costs and longer relationship duration with the firm and contribute more towards firms' revenues and profits (Li, Sun and Wilcox 2005; Kumar, Venkatesan and Reinartz 2008). In our context, linking FGC to both the breadth and the intensity of customer-firm relationship will help better understand the role of FGC and will allow to measure the returns of investing in social media. Although our primary focus is on customer spending and cross-buying behavior, we also supplement our core findings by examining the impact of FGC on customer profitability.[1]

3.2. Firm Generated Content (FGC)

We refer to the messages posted by firms in their official social media pages as FGC and argue that through the interactive nature of the social medium, FGC can help firms develop one-on-one relationships with their customers. A recent business report (*Inc.* 2012) suggests that unlike traditional media, interaction between customers and firms via social media is perceived as being mutually beneficial. We argue that FGC will positively affect customer behavior for the following reasons. First, similar to the role of traditional advertising in informing consumers and driving sales (Vakratsas and Ambler 1999), FGC can help a firm inform customers about its current product offerings, prices, and promotions. Secondly, interaction with and virtual presence of other brand aficionados or fans can help in reinforcing favorable brand attitudes. Naylor, Lamberton, and West (2012) refer to "mere virtual presence" as the passive exposure to a brand's supporters in social media and argue that the inferred commonality between a focal user and other users in a social media community can create positive brand evaluations. Lastly, when firms post content in social media, customers can respond by 'liking' the content or commenting on it which can generate more positive brand evaluations.

[1] Note that although our conceptual framework is very general, we use customers' in-store channel transaction data to measure their purchase behavior as the channel is the dominant channel for the retailer that we study.

3.3. Interaction of FGC with Television Advertising and Email Marketing

Research in the area of integrated marketing communications (IMC) has emphasized the complex role of interactions between multiple media on the link between marketing and sales (Smith, Gopalakrishna, and Chatterjee 2006). In a recent study, Li and Kannan (2014) argue that there may be spillovers across multiple customer-firm touch points that should be accounted for when measuring cross-channel campaign effectiveness. While television advertising and social media marketing differ in several aspects, they both serve as critical communication stimuli and can influence desired outcomes such as brand awareness in a similar way by reinforcing each other (Bruhn, Schoenmueller, and Schäfer 2012). FGC can help increase the effectiveness of television advertising by providing appropriate links to products. Television advertising also often provides links to social media (e.g., '*follow us on Facebook'* message in a TV ad for a product) thus integrating the two different channels.

Similar to traditional advertising, email messages may also work in tandem with social media. Email messages provide easy access to a firms' products and can be customized to increase customer response rates. Email marketing is known to be an effective and efficient way to reach a customer base (Ansari and Mela 2003). Thus, in our context, we examine the potential positive spillovers and synergies that may exist between FGC and traditional and email marketing messages.

3.4. Interaction Effects of FGC with Customer Characteristics

For social media strategies to be effective, it is important for firms to understand if the effect of social media engagement varies across different types of customers. We focus on the following three individual characteristics that can affect customers' motivation and ability to process information: length of relationship, technological skill level, and social network proneness.

3.4.1. Length of Relationship. Studies in the branding area suggest that brand familiarity is an important component of brand equity (Aaker 1991), and that marketing messages evoke greater and more positive attitudinal response from consumers for familiar brands versus unfamiliar brands (Campbell and Keller 2003). Research also suggests that customers with a greater length of relationship also have greater levels of satisfaction with the firm (Palmatier et al. 2006). Satisfied customers in turn may feel a higher level of commitment towards the firm (Ranaweera and Prabhu 2003) and thus are more likely to exhibit a favorable response towards FGC.

3.4.2. Tech-savviness. Firms are constantly introducing new technologies to appeal to tech-savvy customers. For example, Macy's recently introduced a new "app" so that its tech-savvy customers can shop while on-the-go from catalogs, billboards and magazine ads (*RetailingToday* 2013). Tech-savvy customers are accessible via multiple digital touch points and as retailers and brands engage with such customers through FGC, we expect tech-savvy customers to reciprocate by engaging more interactively with a firm through its social media platform. Tech-savvy customers are also more likely to supplement information they receive from FGC with other online sources (Schivinski and Dabrowski 2014) and derive greater benefits from social media engagement thus leading to a greater response to FGC.

3.4.3. Social network-proneness. In the digitally connected world, consumers are spending increasing amounts of time online interacting with other consumers with whom they may share common interests and consumption experiences. We refer to such consumers as '*social network-prone*' consumers. Mere virtual presence of a brand's supporters in social media can positively influence a focal consumer's purchasing behavior (Naylor, Lamberton and West 2012). We argue that social network-prone consumers will place a greater value on the ability to connect with a firm's other customers to share their consumption experiences and thus be more receptive to social media engagement. Furthermore, customers who use social media regularly will place

more weight on the opinions of individuals with similar views (Schulze, Schöler, and Skiera 2014) and thus will likely exhibit a greater response to FGC and messages by other customers in the social media community.

4. RESEARCH METHODOLOGY

4.1. Research Setting and Data

The data set for this study comes from a large specialty retailer that sells wine and spirit products. The retailer operates multiple stores in the Northeastern United States and manages an extensive loyalty card program through which it tracks customers' transactions at the individual product level. The retailer relies on both traditional media (e.g., TV advertising) and emails to convey information about its offerings to its customer base.[2] The retailer started its foray into social media in August 2009 by creating a social media page and posting content on a popular social media networking site. It subsequently encouraged customers to sign-up to become fans and interact. The social media site is not owned by the firm but is instead operated by a third party and is a popular social networking site.[3] The firm conducted a marketing campaign via television and email media informing customers about the presence of its social media page. No incentives (either monetary or promotional such as coupons) were offered and customers signed-up and participated of their own volition. Once a customer participates in the firm's social media site (e.g., by clicking on the "like" button on the firm's social media page), FGC appears on the participating customer's social media page.[4] These "participating" customers may also receive email messages about FGC as and when they are posted by the firm on its social media site. To that end, FGC is more readily available to participating customers and thus, more likely to have an effect on the behavior of participating customers (versus "non-participating" customers). We

[2] The content used by the firm across different media (i.e., social media, television and emails) is usually distinct.
[3] We cannot disclose the name of the social media platform because of confidentiality reasons.
[4] While customers can subsequently "unlike" the firm, we do not find such instances in our data set.

note that the focal firm posts both promotional and non-promotional content on its social media page. In the Appendix, we provide a few examples of FGC from the focal firm's social media page that illustrate that the firm uses social media for both promotional and non-promotional marketing communication.

We gathered detailed information on the customers who participated in the focal firm's social media site and merged this *social media participation dataset* with the *customer transaction dataset*. For the purpose of our empirical analyses, we work solely with customers' in-store purchases as the majority of retailer's sales occur via its physical stores. This process of merging customers' social media participation with (in-store) transaction data is a multistep one that was done carefully in conjunction with the cooperation of the focal firm. An important aspect of our dataset is that we can also identify customers who do not take part in the focal firm's social media site (i.e., non-participants) and we have information on their purchases as well.

We also conducted a survey (in February 2011) on the *same* set of customers who participated in the firm's social media page. This survey was also sent to a subset of non-participant customers, i.e., those who did not participate in the focal firm's social media site. Consequently, we have attitudinal information obtained through surveys for customers who participated in the firm's social media as well as those who did not participate. We merged the survey data with customers' social media participation data and their in-store transaction data (customer in-store purchases) to create a comprehensive dataset that we subsequently employ for our empirical analysis.

4.2. Model Development

Before we present the econometric model to establish the effect of FGC on customer spending and cross-buying behavior, we discuss some pertinent issues that need to be taken into account. Customers who have a higher affinity towards the retail firm may be more likely to have a better

transactional relationship with the firm (i.e., exhibit higher levels of spending and cross-buying) and also be more responsive towards FGC through participation in the firm's social media page. In other words, customer intrinsic variables (beyond the ones we control for) may simultaneously influence customer purchase behavior and their responsiveness towards FGC. We, thus, need to account for this plausible endogeneity/self-selection and rule out the reverse causality issue to establish the effect of FGC on customer spending and cross-buying.

4.2.1. Propensity score matching. To account for the self-selection issue, we use the propensity score matching (PSM) technique which helps create two groups of customers, *participant* customers ("treatment" group customers)—those who choose to receive FGC by participating in the firm's social media—and *non-participant* customers ("control" group customers)—those who choose not to receive FGC and do not participate in the firm's social media—that resemble each other before the firm's foray into social media, thereby creating a statistical equivalence between the two groups (Rosenbaum and Rubin 1983). After employing the PSM technique to create the two groups of customers, we use the difference-in-differences (DID henceforth) modeling framework to examine the behavior of the two groups before and after the firm's venture into social media (i.e., pre-FGC and post-FGC time periods respectively). In other words, by comparing the difference in behavior between the treatment and the control group customers before and after their social media engagement (via FGC), we can estimate the impact of FGC.

Following prior literature (e.g., Huang et al. 2012; Girma and Gorg 2007), we perform the matching procedure using data from the pre-FGC time period. We model customers' (binary) decision to participate in the firm's social media (and thereby choosing to receive FGC) via a logistic regression model of customer specific explanatory variables.[5] We obtain the propensity

[5] Details are available upon request.

scores to (pair) match the control group customers that resemble the treatment group customers on the basis of propensity score similarity using the 1:1 nearest neighbor matching technique (Rosenbaum and Rubin 1985). As noted in prior research, this enables us to avoid bias that may occur when linking multiple, potentially dissimilar treatment and control group customers (Smith 1997; Huang et al. 2012).

We perform several checks to assess the validity of the PSM method. We do a visual analysis of propensity score distributions through box-plots and histograms to ensure there is a common support between the treatment and the control groups (Guo and Fraser 2010). We also perform the Kolmogorov-Smirnov test to verify that the propensity distributions of the above two groups of customers are similar and conduct sensitivity analysis to check for the presence of any hidden bias in the matching process. We also check if the logistic regression model used for matching is able to correctly predict "group membership." We find that the model is able to predict treatment versus control group membership accurately for more than 93% of the customers. Having created a matched pair of treatment and control group customers via the PSM technique, we proceed to perform our DID analysis.

4.2.2. Difference-in-differences model. For the DID analysis, the treatment group customers were selected using a multi-step process that yielded 412 customers. We used the PSM method outlined above to create a matched control group that has the same number of customers as the treatment group. For the DID model, we work with 85 weeks prior to and following the inception of the focal firm's social media page in August 2009. The pre-FGC period spans from January 2008 to July 2009 for all customers. The post-FGC period spans from August 2009 to March 2011 for the control group customers. For the treatment group customers, since they join the social media page at different points in time (see Figure 2), the post-FGC period becomes

effective only after they join the firm's social media page. Our analysis is at the weekly level and this corresponds to a total of 170 weeks.

We conduct the DID analysis utilizing techniques expounded in the current literature (e.g., Bollinger, Leslie and Sorensen 2011; Huang et al. 2012). For each matched pair (denoted by *i*), we model a focal customer's (denoted by *h*) spending and cross-buying behavior at time *t* (week) as follows:[6]

$$\begin{aligned} Spend^{*}_{iht} = {} & \alpha_{0ih} + \alpha_1 TCust_{ih} + \alpha_2 FGC_{iht} + TCust_{ih} \times FGC_{iht}(\alpha_3 + \alpha_4 TVAd_{iht} + \alpha_5 Email_{iht} \\ & + \alpha_6 CExp_{iht} + \alpha_7 TechS_{ih} + \alpha_8 SocialNet_{ih}) + \alpha_9 TVAd_{iht} + \alpha_{10} Email_{iht} + \alpha_{11} CExp_{iht} \\ & + \alpha_{12} TechS_{ih} + \alpha_{13} SocialNet_{ih} + \alpha_{14} PromD_{iht} + \alpha_{15} Dist_{ih} + \varepsilon_{1iht} \end{aligned} \quad (1)$$

$$\begin{aligned} CrossBuy^{*}_{iht} = {} & \beta_{0ih} + \beta_1 TCust_{ih} + \beta_2 FGC_{iht} + TCust_{ih} \times FGC_{iht}(\beta_3 + \beta_4 TVAd_{iht} + \beta_5 Email_{iht} \\ & + \beta_6 CExp_{iht} + \beta_7 TechS_{ih} + \beta_8 SocialNet_{ih}) + \beta_9 TVAd_{iht} + \beta_{10} Email_{iht} + \beta_{11} CExp_{iht} \\ & + \beta_{12} TechS_{ih} + \beta_{13} SocialNet_{ih} + \beta_{14} CrossP_{iht} + \beta_{15} Dist_{ih} + \varepsilon_{2iht} \end{aligned} \quad (2)$$

In Equations (1) and (2), $Spend^{*}_{iht}$ and $Crossbuy^{*}_{iht}$ refer to a customer *h*'s (in the matched pair *i*) (in-store) spending and cross buying respectively at time *t*. $TCust_{ih}$ is a dummy variable that takes the value 1 if customer *h* belongs to the treatment group and 0 otherwise. FGC_{iht} is a dummy variable that is equal to 1 if a customer *h* is a recipient of FGC at time *t* and 0 otherwise.[7] $TVAd_{iht}$ and $Email_{iht}$ refer to a customer *h*'s exposure to television advertising and email messages respectively at time *t*.[8] $CExp_{iht}$, $TechS_{ih}$ and $SocialNet_{ih}$ denote customer *h*'s length of experience with the firm, technological savviness and social network proneness respectively. The rest of the independent variables serve as customer specific control variables which include promotion depth index (*PromD*), cross category promotion (*CrossP*)—these variables proxy for

[6] We note that we included all lower order two way interaction effects in the DID model. We find that our results are robust to the inclusion of all lower order interaction effects. For the sake of brevity, we only report the relevant coefficients.

[7] FGC_{iht} =1 for a treatment group customer only after a focal customer chooses to participate in the firm's social media page.

[8] For the DID model, television advertising and email messages are coded as high (coded as 1) and low levels (coded as 0) based on median split across the panel of customers.

the customer's propensity to buy products on promotion—and distance of customers' residence from the store they shopped (*Dist*). We explain the operationalization of all these variables in the next sub-section (see also Table 1). Note that *TechS* and *SocialNet* are measured using survey data. ε_{1iht} and ε_{2iht} refer to the error terms associated with the two equations respectively. Since the error terms may be serially correlated, following Chib and Greenberg (1995), we specify the following structure:

$$\begin{bmatrix} \varepsilon_{1iht} \\ \varepsilon_{2iht} \end{bmatrix} = \begin{bmatrix} \rho_1 \varepsilon_{1iht-1} + \upsilon_{1t} \\ \rho_2 \varepsilon_{2iht-1} + \upsilon_{2t} \end{bmatrix} \quad 0 \le \left| \rho_{j=1,2} \right| < 1 \tag{3}$$

The residual error terms ($\Psi = (\upsilon_{1t}, \upsilon_{2t})$) are distributed as $\Psi \sim N(0, \sigma^2 I)$. In Equations (1) and (2), since the values of customer spending and cross-buy may be 0 for some weeks, we use a type-1 Tobit model as follows:

$$Spend_{iht} = \begin{cases} 0 & if \ \ Spend^*_{iht} \le 0 \\ Spend^*_{iht} & otherwise \end{cases} \text{ and } Crossbuy_{iht} = \begin{cases} 0 & if \ \ Crossbuy^*_{iht} \le 0 \\ Crossbuy^*_{iht} & otherwise \end{cases}.$$

To account for heterogeneity in customers' response behavior, we use a hierarchical Bayesian framework (e.g., Rossi, Allenby, and McCulloch 2006) as follows:

$$\Lambda_h = \Theta_0 + \Theta_1 G_h + \Phi_h \tag{4}$$

where Λ_h is the vector comprising of $[\alpha_{0ih}, \beta_{0ih}]'$, while G_h and Θ_1 are the matrix of customer demographic variables consisting of age, gender and race (Table 1) and the corresponding coefficients associated with them respectively. $\Phi_h \sim MVN(0, \mathrm{K})$ with *K* denoting the variance-covariance matrix.

In Equations (1) and (2) as mentioned previously, following prior literature (e.g., Angrist and Pischke 2009), we define the main variables $TCust_{ih}$ and FGC_{iht} as dummy variables. This lets us interpret the relevant parameters as causal effects. We caution that the causal interpretation is

valid assuming that except for social media participation, the treatment and the control groups are similar (within the bounds of propensity score matching that we employ). However, the operationalization of FGC as a categorical variable does not allow us to capture the various dimensions of FGC such as valence and response from customers (we discuss these salient dimensions of FGC in the next sub-section). We, thus, first utilize the DID model to establish the causal impact of FGC (within the bounds of PSM) and subsequently use another model specification – a variation of the *treatment effects* model— to capture the richness of the FGC construct.

4.2.3. Treatment effects model. The treatment effects (TE) model accounts for endogeneity issues (e.g., due to self-selection) by explicitly considering the endogenous variable and incorporating it into the modeling framework. A typical TE model set-up consists of two components: (1) a selection equation that models the endogenous variable through which customers might self-select. In our case, the endogenous variable is customers' self-selection via participation in the firm's social media page and them thereby receiving FGC, and (2) given the self-selection variable, a set of outcome equations, to model the phenomenon of interest (in our case customer spending and cross-buying). We first describe the outcome equations of our proposed TE model below:

$$Spend^{*}_{iht} = \gamma_{0ih} + FGC_{iht}(\gamma_1 + \gamma_2 TVAd_{iht} + \gamma_3 Email_{iht} + \gamma_4 CExp_{iht} + \gamma_5 TechS_{ih} + \gamma_6 SocialNet_{ih}) + \gamma_7 TVAd_{iht} + \gamma_8 Email_{iht} + \gamma_9 CExp_{iht} + \gamma_{10} TechS_{ih} + \gamma_{11} SocialNet_{ih} + \gamma_{12} PromD_{iht} + \gamma_{13} Dist_{ih} + \varepsilon_{3iht} \quad (5)$$

$$CrossBuy^{*}_{iht} = \delta_{0ih} + FGC_{iht}(\delta_1 + \delta_2 TVAd_{iht} + \delta_3 Email_{iht} + \delta_4 CExp_{iht} + \delta_5 TechS_{ih} + \delta_6 SocialNet_{ih}) + \delta_7 TVAd_{iht} + \delta_8 Email_{iht} + \delta_9 CExp_{iht} + \delta_{10} TechS_{ih} + \delta_{11} SocialNet_{ih} + \delta_{12} CrossP_{iht} + \delta_{13} Dist_{ih} + \varepsilon_{4iht} \quad (6)$$

We model the selection equation (the endogenous variable by which customers self-select) in a probit framework as:

$$CustPar^{*}_{iht} = \varphi_{0ih} + \varphi_1 PrivCon_{ih} + \varphi_2 IMov_{ih} + \varphi_3 TechS_{ih} + \varphi_4 SMov_{ih} + \varphi_5 OEnt_{ih} + \varphi_6 TimeOnSocialNet_{ih} + \varepsilon_{5iht} \quad (7)$$

where *i* indexes the matched pair of treatment or control group customer, *h* indexes customer and *t* indexes time period. In Equation (7), $CustPar^{*}_{iht}$ is the latent utility that a customer *h* derives by participating in the firm's social media site at time *t* and thereby chooses to receive FGC. We model a focal customer's propensity to participate in the firm's social media as a function of the customer's attitudes towards online privacy concerns (*PrivCon*), motivation to use the Internet for seeking information (*IMov*), technology savviness (*TechS*), motivation to socialize online (*SMov*), the proclivity to use the Internet for online entertainment (*OEnt*), and time spent in online social networking per day (*TimeOnSocialNet*).[9] Note that since we estimate the TE model on the matched sample, we do not use these variables for matching. Furthermore, whereas all these Internet, technology and social media related variables are likely to affect customers' decision to participate in social media, they are not likely to affect customer *in-store* behavior. This helps us rule out reverse causality. We link the latent utility of customer participation to the observed social media participation (and customers thus receiving FGC) as follows. Let $CustPar_{iht}$ denote a binary variable that takes on the value 1 if the focal customer *h* (of the matched pair *i*) participates in the firm's social media site at time *t* and 0 otherwise. Then, we have the following: $CustPar_{iht} = \begin{cases} 1 \text{ if } CustPar^{*}_{iht} > 0 \\ 0 \text{ otherwise} \end{cases}$.

We would like to draw the readers' attention to a few points with respect to the sample, the formulation and the estimation of the TE model. First, we note that the TE model is relevant only after the firm started engaging with its customers via FGC (August 2009 onwards). We thus use data only from the post-FGC period (that spans 85 weeks from August 2009 to March 2011). We

[9] The details on the measurement of these constructs from survey are available upon request.

select treatment group customers using the same multi-step process as mentioned earlier and we select the control group customers using the same PSM technique as outlined earlier. However, for the TE model, we end up with slightly fewer (394) treatment customers due to the restrictions imposed.[10] Thus, our sample for the TE model consists of 394 customers each from the treatment and the control groups (i.e., a total of 788 customers). Second, unlike in case of the DID model, for the TE model, FGC, television advertising and email messages are continuous variables (albeit we use the same notation as the DID model for the sake of convenience). As with the DID model, we estimate Equations 5, 6 and 7 jointly. Likewise, we also account for heterogeneity and cast the TE model in a hierarchical Bayesian framework.

For both the DID and the TE models, given that we are estimating the outcome variables (Equations 1- 2 and 5-7) simultaneously, we use exclusion restrictions to aid in the econometric identification of the model parameters. We exclude promotion depth (*PromD*) and cross category promotion (*CrossP*) from cross buying and spending equations respectively. Our arguments are as follows. Promotion depth represents the extent or the amount of savings realized by the customer from buying the products that are on sale. We expect customers who realize these higher savings to spend more because of their (increased) purchases of either planned and/or unplanned products (Stilley, Inman, and Wakefield 2010). However, there is no reason to believe that they would buy more from different product categories. In a similar vein, if more categories are on promotion, the greater is the likelihood that a customer would buy from multiple categories, thereby positively impacting cross-buying. Thus, we expect cross category promotion (*CrossP*) to be more strongly correlated with the cross-buying variable. In the robustness checks sub-section, we empirically test for the validity of the set of exclusion restrictions.

[10] For the DID model, the restriction is that customers need to make at least one purchase each in the pre-FGC period and the post-FGC period while for the TE model, the restriction is that customers have to make at least two purchases in the post-FGC period. These restrictions are imposed to reliably estimate the heterogeneity parameters.

We estimate our models using hierarchical Bayesian methods with Markov Chain Monte Carlo (MCMC) techniques. A total of 50,000 with a "burn-in" of 40,000 iterations are used. We use the last 10,000 iterations for calculating posterior means and standard errors of the model parameters upon ensuring that convergence criteria are met.[11]

4.3. Variable Operationalization

4.3.1. Spending and Cross-buying. We measure focal customer *h*'s spending (denoted by *Spend*) by the dollar amount spent on alcohol/liquor products in a given time period. Customer cross-buying behavior (*CrossBuy*) represents the breadth in customers' buying patterns and is operationalized as the number of distinct categories in which a customer purchases in a given time period *t*. (Kumar, George and Pancras 2008; Shah et al. 2012). In our context, to construct this variable, we first need to identify the relevant categories. We rely on reports from trade magazines, interviews with managers and sales information derived from the transaction data to define the categories. This process led to the categorization of wines into red, white and sparkling wines and spirits into whiskey, tequila, rum, vodka, and gin. Thus, we have a total of eight categories representing wine and spirits.

4.3.2. FGC, Television Advertising and Email Messages. For the TE model, we operationalize FGC as the number of original messages posted by the firm that are accessible to a participating customer in a given time period. We give different weights to different postings based on the three dimensions of FGC valence, FGC receptivity and (customers') susceptibility to FGC. The valence of each post reflects the sentiment conveyed by that post (*VPost*). Based on current literature (e.g., Tirunillai and Tellis 2012; Das and Chen 2007), we rely on the *naïve Bayes Algorithm/Classifier* to classify sentiment into three categories: positive, neutral or negative which we code as 1, 0, and -1 respectively. Customers may respond to FGC by liking

[11] Details of estimation process are available upon request.

the post, commenting on the post and sharing the post with their network. For each FGC posting, we sum up the total comments, likes and shares received by that posting to calculate the receptivity measure (*RPost*). Since customers may differ in their susceptibility towards social media and FGC, we include this measure as an additional dimension of FGC. For operationalizing susceptibility, we used the customer survey where we queried customers with regard to their predisposition towards social media. Based on these three dimensions, we formulate FGC (across N_t postings each denoted by k for each customer h and for time period t) as follows:

$$FGC_{ht} = \left(\left(\frac{\sum_{k=1}^{N_t} (VPost_{kt} \times RPost_{kt})}{N_t} \right) \times FGCSuscept_h \right) \quad (8)$$

This FGC formulation is then used in Equations (5) and (6) for the treatment group customers from the matched pair (note that FGC will be 0 for the control group customers) upon participation. Note that we also investigate into the efficacy of the individual components of the FGC defined above.

We operationalize television advertisements (denoted by *TVAd*) at the customer-week level. Previous literature has documented that due to memory decay and related reasons, past advertising may not be as effective as the recent ones. To accommodate this, following Tellis and Weiss (1995), we first adopt a stock formulation for television advertising (denoted by $TVAdStock_t$) as follows: $TVAdStock_t = \eta GRPTVAd_t + (1-\eta)TVAdStock_{t-1}$ where $\eta \in (0,1)$ [12] is the decay parameter and $GRPTVAd_t$ refers to average gross rating points (GRPs) across all broadcast TV advertisements at time t. To make this variable customer specific, we multiply it

[12]To reduce computational burden, we do not estimate η. Instead, we use a grid search procedure to obtain its optimum value. We find that η =0.8657 provides the lowest log marginal likelihood. More information is available upon request. Note that although we use stock formulation to capture the phenomenon of interest more accurately, we caution the reader that using the stock variable may complicate the interpretation of interaction parameters.

by the average number of hours of television watched weekly by the customer as follows: $TVAd_{ht} = TVAdStock_t \times W_h$ where W_h is the average weekly number of hours of television watched by the customer.[13] In our dataset, we observe if a customer *h* opened a specific email sent by the focal retailer. We operationalize $Email_{ht}$ as the number of emails sent by the firm that a customer *h* opens in time period *t*.

4.3.3. Customer Specific Variables. We capture customer experience (*CExp*) by the length (in weeks) from the date of the focal customer's first transaction with the retailer until time *t*. We utilize survey data to measure the attitudinal variables used in our empirical analysis (see Table 1). These include the two focal customer specific variables in the conceptual framework, i.e., customers' technology savviness (*TechS*) and social network proneness (*SocialNet*) and the explanatory variables to model customers' social media participation (see Equation 7). This set of variables includes customers' attitudes towards online privacy concerns (*PrivCon*), their motivation to use the Internet for seeking information (*IMov*), their motivation to socialize online (*SMov*), the use of the Internet for online entertainment (*OEnt*), and time spent on social network per day (*TimeOnSocialNet*). We also use the number of online social profiles of a customer and the focal customer's use of the Internet to escape reality as matching variables in PSM. All the factor loadings are significant ($p < .01$) which suggests convergent validity. Cronbach's alpha for the constructs ranges from .73 to.89 which indicates good reliability.

4.3.4. Other Control Variables. With respect to the control variables, the promotion depth index variable (*PromD*) is calculated for each customer *h* at time *t* as the weighted average of all price discounts availed by that customer across all product items (across all the eight categories) purchased. The weights are given by the volume share of each product item bought by the customer and are computed as constant weights— i.e., average share weights across the full

[13]We obtain this information from the customer survey. The average weekly number of hours of television watched for the treatment group customers is 22.38 while the corresponding number is 19.82 for the control group customers. This difference is not significant at the 5% level.

sample period (Pauwels and Srinivasan 2004). Cross category promotion (*CrossP*), which captures a focal customer's proneness towards buying products on promotion across categories, is operationalized as the proportion of total categories bought on promotion by customer *h* at time period *t*. The operationalization of customer demographic variables is straightforward as is the distance variable (average distance in miles from customers' residence to the store they shop). We refer the readers to Table 1.

Customers, in our sample, on average, spend $14.34 per week (or $49.80 per purchase visit). However, note that there are weeks during which customers do not make any purchases. Customers on an average buy from 3.23 different categories per visit. This is typical of the product category where customers tend to be variety seeking. As is evident from the table, there is significant heterogeneity across customers on both of these dependent variables. The retailer posts about three to four messages in a typical week. The high level of television advertising indicates that the retailer uses the television medium regularly to communicate with its clientele. The response of customers to the email messages sent by the retailer is quite high. On an average, in a given week, a customer opens 1.63 emails sent by the retailer. With respect to customer characteristics, the average customer experience is 94.26 weeks. Customers travel an average distance of 5.17 miles to shop at the store. The average age of consumers in our sample is 45 years and the minimum age is 23 years. In the interest of space, we summarize the operationalization of the variables and their descriptive statistics in Table 1.

5. RESULTS

5.1. Model Free Evidence

Before we report the results of the model, we present model free evidence (see Table 2) using the average difference-in-differences for the outcome variables. The "raw" numbers compare the

outcome variables across the two groups— the treatment and the control group — of customers and across the two time periods— the pre-*FGC* period and the post-*FGC* period. We perform these calculations at the weekly level accounting for the different participation times of the customers in the treatment group. While we find that there is no significant difference (at the 5% level) in weekly spending and cross-buying between the treatment versus the control group customers in the pre-FGC period ($13.52 vs. $13.35 and 3.08 vs. 2.95 for spending and cross-buying respectively), a significant difference (at the 5% level) emerges in the post-FGC period across these customer groups ($15.96 vs. $14.53 and 3.74 vs. 3.15 for spending and cross-buying respectively). Following Dagger and Danaher (2014), we find statistically significant and positive overall "difference-in-differences" values ($1.26 and .46 for spending and cross-buying respectively) indicating the positive effect of FGC on customer behavior. These set of results help rule out the issue of reverse causality and lend prima facie evidence to the (positive) effect of FGC on customer spending and cross-buying behavior. In the following, we present the formal results of the DID and the TE models.

5.2. DID and TE Model Results

We present the results of the proposed DID model (Equations 1-4) along with its (several) alternative specifications in Table 3. We rely on log marginal density (LMD) computed using the Newton-Raftery method (Rossi, Allenby, and McCulloch 2006 pg. 168) to assess the fit of the models. The first model in Table 3 (M1) is the standard DID model reported in the literature and serves as a benchmark model. This model does not contain variables related to television advertising, email marketing and control variables. It also does not incorporate the Tobit type-1 specification and does not account for serial and cross correlations. The subsequent models (M2-M5) build on the basic model sequentially by considering additional variables and model specifications. The "full" model (M5) includes the main effect of FGC, its interaction effects

with television advertising and email marketing and control variables. Moreover, the full model also incorporates customer heterogeneity (via hierarchical Bayesian methods; see Equation 4), adopts a Tobit type-1 specification and accounts for serial and cross correlations. Not surprisingly, this model has the best fit. The main parameters of interest are α_3 and β_3 that capture the effect of FGC on the spending and the cross-buying behavior respectively of participant customers vis-à-vis the non-participant customers in the post FGC period (as compared to the pre-FGC period). These parameters are positive and significant across all models attesting to the positive effect of FGC on customer spending and cross-buying.

We present the results of the TE model in Tables 4a and 4b. As before, we have the standard TE model (Model 1) which includes only the main effect of FGC. In Model 2 to Model 5, we sequentially enrich Model 1 by including additional variables and extending model specifications. Model 5 ("full" model) captures the main effect of FGC along with its interactions with television advertising, email marketing and includes control variables and customer characteristics. This model also accounts for customer heterogeneity, Tobit type-1 specification, serial and cross correlations and offers the best fit. Because the full models have the best fit for both DID and TE models, we will refer to them while discussing our results.

Our results suggest that FGC has a positive and significant effect on customer spending and cross-buying. This is true for the parameter estimates obtained from both the DID model (1.32 and .49 for spending and cross-buying respectively) and the TE model (.18 and .11 for spending and cross-buying respectively). The parameters associated with the interaction effects between FGC and television advertising (DID: .0057 and .0001; TE: .0621 and .0019 for spending and cross-buying respectively) and the interaction effects between FGC and email marketing (DID: .0991 and .0047; TE: .1755 and .1137 for spending and cross-buying respectively) are positive for customer spending and cross-buying for both the models. These results suggest that there are

synergistic effects between FGC and television advertising and between FGC and email based marketing communication.

Turning our attention to the interaction effects between FGC and customer characteristics, our results suggest that FGC has a greater effect on more experienced customers for both customer spending and cross-buying (DID: .1686 and .0797; TE: .0450 and .1228 for spending and cross-buying respectively). We find that FGC has a greater effect on customers who are more tech-savvy (DID: .2753 and .0303; TE: .0918 and .0281 for spending and cross-buying respectively) and those who are more prone to using social media (DID: .1906 and .1523; TE: .1091 and .0956 for spending and cross-buying respectively). Taken together, the results from both the DID and the TE models are consistent with each other.

With regard to other results, we find that, as expected, both television advertising and email marketing have a positive effect on customer spending and cross-buying. In Table 4b, we present the results related to customers' participation in the focal firm's social media site. We find that most of the results are in the expected direction. Customers with greater privacy concerns are less likely to participate in the firms' social media site. Likewise, customers who have a greater motivation to seek information, are more tech-savvy, have a greater motivation to socialize online, and who use the Internet for online entertainment are more likely to become part of the firm's social media site. We also find that customer social media participation, spending and cross-buying are positively correlated.

6. SUPPLEMENTARY ANALYSIS

6.1. Profitability Analysis

In order to assess the impact of FGC on the focal retailer's bottom-line, we analyze the impact of FGC on customer profitability. Leveraging our access to data on both retailer prices and costs at

the individual product level, we compute *customer specific* aggregated net total profits per period accrued to the firm as a result of purchases made by our sample customers. We use this measure in lieu of customer spending in our analysis. Raw difference in difference analysis (See Table 2) suggests that while there is no significant difference in customer profitability for the control group across the pre-FGC and the post-FGC periods, there is a significant difference (at 5% level) in customer profitability for the treatment group customers before and after social media participation. The average difference-in-differences value for customer profitability is $1.02 and is statistically significant at the 5% level. This suggests that compared to the customers who do not participate, those customers who participate in the focal firm's social media contribute $1.02 more towards the focal firm's profits (in the post-FGC period vis-à-vis the pre-FGC period).

Next, to formalize the effect of FGC on customer profitability, we first begin by re-estimating the DID model presented in Equations (1)-(4). Given that our proposed DID model is a joint model (with cross-buying), we re-estimate the model by replacing customer spending with customer profitability and leave the rest of the variables and the model specification unchanged. We find that FGC has a significant effect on customer profitability with the other results being substantively similar as earlier (see Table 5). As before, we also re-estimate the TE model (Equations 5-7) with customer profitability as the dependent variable. The results of this estimation are presented in Table 6 (note that as the results from customers' social media participation equation are similar to the earlier results, they are not presented in the paper and are available upon request). The main takeaway from these additional analyses is that FGC has a positive impact on customer profitability. These results suggest that FGC may be used not only to promote products on sale, but can also be used to influence customer purchases of high margin products thus leading to increased customer profitability.

6.2. Robustness Checks

We perform various checks to ascertain that our core results are robust to alternative operationalization of variables, alternative model specifications and presence of any outliers. With respect to variable operationalization, instead of using length of tenure, we use the number of purchase occasions as a proxy for customer experience. Following Reinartz and Kumar (2003), we also look into the effect of "profitable" customer duration; we do so by weighting customer experience by average profitability of each customer. For the operationalization of television advertising, we weight the television ad stock by customers' trust in television ads which is measured through the customer survey (instead of the number of hours of television watched). We also use the actual length of the television advertisements. We use an alternative operationalization for cross category promotion variable where we take into account the proportion of product items bought on promotion within a category. We find the results to be robust to all these alternative variable operationalizations.

Next, we re-estimated the DID model using a different post-FGC time period. In our sample, 90% of the customers in the treatment group joined the firm's social media page within one year of the firm's social media initiative (August 2009). Thus, we construct a new post-FGC time period by removing data spanning from August 2009 to August 2010 (56 weeks) from the post-FGC period. We then re-estimated the DID model with the remaining 29 weeks constituting the post-FGC period. We also estimated our TE model by selecting the control group customers randomly instead of using a matched sample. Finally, we estimated the DID and the TE models by assuming heterogeneity in all of the coefficients (as opposed to only the intercepts). We find the results to be robust to all these alternative model specifications.

We conduct the following checks to assess the sensitivity of results to the presence of potential outliers. First, we re-estimated the DID and the TE models by removing those

customers who have very high customer spending and cross-buying. We remove those customers whose spending and cross-buying are two standard deviations above the respective mean values. Second, since the retailer advertises heavily via television during certain time periods (such as Christmas, Thanksgiving etc.), the GRP of TV ads are higher during these time periods. We plotted the GRP of TV ads across the study time period and identified those periods where these values are higher than the normal. We then removed these time periods and re-estimated the DID and the TE models. We find the results to be robust to the exclusion of outliers. Details and results of all these alternative models are available from the authors upon request.

Finally, we use empirical tests to ensure that our exclusion restrictions (related to the joint models presented in Equations 1-2 and Equations 5-7 are valid). We follow Woolridge (2010) and regress the error terms of an equation on the excluded variable to see the significance of the parameter associated with the relevant excluded variable. Woolridge (2010) suggests that the restrictions criterion is met if the parameter corresponding to the excluded variable is insignificant. We find that the criterion for the exclusion restrictions is met in our empirical context.

7. DISCUSSION

7.2. Summary of Findings and Theoretical Implications

According to the latest CMO survey (2015), social media spending as a percentage of marketing spending is expected to more than double in the next five years. Yet only 13.2% of executives surveyed in the CMO survey report that they have been able to measure the impact of social media spending. Using actual in-store purchase data, we find that FGC can enhance not just the transaction and the relationship sides of customer-firm interaction (measured by spending and cross-buying respectively) but can also play a role in increasing customer profitability. We also

find that FGC works synergistically with television and email based marketing communication. Our results suggest that FGC has a greater effect on customers with a longer customer-firm relationship, who are technologically savvy and are more prone to social networking.

From a theoretical contribution perspective, whereas much of the extant studies in the area of social media help understand the impact of UGC on market outcomes, the effect of FGC on customer behavior has received less attention. Our study contributes to the social media literature by demonstrating the impact of FGC on three key customer metrics—spending, cross-buying and profitability. It is critical to note that we do so after ruling out customer self-selection, a thorny issue in the context of customers' social media participation. Second, our study contributes to the literature on integrated marketing communications and multichannel marketing which has called for developing better understanding of synergies across multiple media to build brand equity (Naik and Raman 2003; Lin, Venkatraman, and Jap 2013; Joo et al. 2014). By studying the synergistic effects of FGC on customers' in-store (offline) purchase behavior, we not only establish the synergy effect across different media (social media-television-email marketing communication), but also establish cross channel (online-offline) synergy effects. This has implications not only for optimal resource allocation across media but also for cross-channel coordination of firms' communication strategies. Our utilization of a richer specification for FGC that captures valence, receptivity and susceptibility also helps provide a deeper understanding of the theoretical underpinnings of the FGC effect. We demonstrate that the effect of FGC receptivity is the greatest which suggests that it is the ability of social media to give 'voice' to customers in the form of 'likes and comments' that make FGC more effective.

7.3. Managerial Implications

Based on the results from our study, we offer the following prescriptions for managers.

7.3.1. Embrace Social Media. The clear message from our study is that social media marketing matters and that managers should embrace it to communicate and nurture relationships with customers. We find that investing in developing a social media community with a dedicated fan base (e.g., Facebook page) can significantly strengthen customer-firm relationships and can lead to a definitive impact on the firm's revenues and profits. In our study, we note that 4.95% of our focal firm's total customer base elects to receive FGC. While this level of participation is fairly consistent with social media participation rates for other brands and retailers[14], we believe that as more of a firm's clientele participates in a firm's social media page, the resulting benefits for the firm (in terms of customer spending, cross buying and profitability) can be greater. However, firms' also need to take into consideration the costs of assembling and constantly updating FGC. Given that television advertising is measured in gross rating points, it is difficult for us to compare the effects of FGC and television advertising, but nevertheless, to quantify the impact of FGC on customer behavior, we conduct an exercise that is akin to elasticity analysis. We report the results in Table 7. We find that the elasticity of FGC with respect to customer spending is .014 which is lower than the elasticities of television advertising and email messages (.101 and .091 respectively). However, we find that the elasticity of FGC (.059) is greater than that of television advertising and email messages (.051 and .043 respectively) for customer cross-buying. This suggests that FGC can play a key role in strengthening customers' relationship with the firm by encouraging them to buy across several product categories.

To further understand if the positive effect of FGC is simply due to its attractiveness as a new media or if the effect is a long lasting one, we split the post-FGC period into two six-month time

[14] To compare how our focal firm's customer participation rate of 4.95% compares against other brands and retailers, we collected information on the number of people who signed up on the Facebook page of some well-known brands (such as Tide, Campbell's Condensed Soup, Honda CRV, Organic Valley and Cheerios) and retailers (such as Stop&Shop, Kroger, Shoprite, Safeway and HEB) and performed additional analysis. We find that the percentage of customer base that signs up and participates in the social media sites of these firms ranges from 1.00% to 7.35% with an average of 3.70%. Thus, we feel that the customer participation rate for our focal retailer is consistent with current industry trends. More details are available from the authors upon request.

periods. From Table 7, we note that the elasticities of FGC in the last six months (in the post-FGC period) for customer spending and cross-buying are actually greater than the corresponding elasticities from the first six months of the post-FGC time period. These findings suggest that the FGC effect is persistent in the post-FGC time period and that its impact is not purely attributable to the novelty of the launch of the social media page by the focal firm. Thus, while traditional media advertising is still found to be more effective in our context, we believe that FGC, though a nascent channel in our study, also yields sustained results for the firm.

7.3.2. Exploit Synergies across Media. Our study suggests a synergistic relationship between social and other media used for marketing communication—television and emails. As social media gains importance and becomes the proverbial 'talk of the town', managers must take care to not abandon traditional or other forms of advertising as these have substantial synergies between them. To illustrate these synergistic effects, we performed additional analysis. These results are reported in Table 8. We find that the percentage increases in customer spending and cross-buying due to the synergistic effect of FGC and television advertising are quite substantial (1.03% and .84% for customer spending and cross-buying respectively). The percentage increases in customer spending and cross-buying due to the synergistic effect between FGC and emails are 2.02% and 1.22% respectively. These results highlight the need for integration of marketing communication across different media and help allay managers' concerns regarding measurable returns to social media marketing. We encourage social media managers to perform simulation exercises to determine optimal allocation of resources across different media.

7.3.3. Monitor FGC popularity. The return on investment in 'social CRM' is determined not only by a firm's investment in social media but also by consumers' level of engagement with the firms' social media page (Hoffman and Fodor 2010). In our study, we incorporate a rich measure of FGC that comprises of sentiment (or valence), popularity (or receptivity) of posts and

customer susceptibility toward social media posts. While FGC valence captures a firm's effort in creating meaningful content that facilitates more positive customer-firm interactions, receptivity and susceptibility capture the extent to which customers' interest is piqued by FGC and their predisposition to using social media. Although our FGC measure (presented in Equation 8) is a composite measure, from a managerial perspective, it may be more useful to assess the differential effects of these three dimensions. Therefore, we compute the elasticities of these three dimensions with respect to customer spending and cross-buying behavior. We present the results of this elasticity analysis in Table 9.

Our results suggest that FGC receptivity has the greatest impact followed by FGC valence and susceptibility. From Table 9, we note that the elasticity of FGC receptivity with respect to customer spending and cross-buying is .019 and .086 respectively. The next most effective FGC component is valence with elasticity values of .013 and .029 with respect to customer spending and cross-buying respectively. Although we find that customer susceptibility also affects customer behavior, the elasticity of this particular FGC component is the lowest among the three components. Since FGC receptivity involves direct customer involvement, we believe our results make a strong case for firms to grow its "fan" base and monitor the level of customer engagement and post measures of popularity on their brand social media pages.

Besides providing more general prescriptions for marketing managers, we also conducted a more context specific post-hoc analysis by identifying the alcoholic products for which FGC was most effective. We worked with the treatment group consumers and simulated their shopping baskets in response to an increase in FGC level. We find that the shopping basket of these set of consumers mostly constituted of red, white and sparkling wine categories. Since the selection of wine category depends on several different features such as varietal, label (region of origin), ratings (such a Parker's rating), sweetness, acidity, tannin, fruit, and body among others,

consumers may need more information on these attributes (Sáenz-Navajas et al. 2013). We believe that this is where social media communications can increase customer access to more nuanced product related information. We note that in our context, wine products are also, in general, higher margin products (on average a 30% higher margin than other alcoholic beverage products). The take away is that FGC is able to drive sales towards higher margin categories in our study.

7.3.4. Utilize Social Media for Strengthening Brand Connections. Based on our findings that FGC has a greater impact on customers who have longer tenure with a firm and on customers who are tech-savvy and active in social media, we suggest that special product focused '*interaction forums*' could be created for such customers. By administering surveys, a firm can identify tech-savvy and social network-prone customers and encourage them to join the firm's social media page. We suggest that developing brand communities that consist of loyal, tech and social-savvy customers will aid firms' long-term financial interests.

8. CONCLUSION AND LIMITATIONS

Although our study offers key insights into FGC's impact and contributes to both theory and practice, it has several limitations. While we leverage a unique dataset that is built on customer social media participation and transaction data, we acknowledge that we analyze only one type of social media. Future research can explore the role of other kinds of social media such as blogs or tweets. Due to the lack of data, we did not consider the behavior of the same set of customers across different types of social media (e.g., Facebook and Twitter) and this could be an avenue for later research. We considered sales at the physical store for the focal retailer as this channel constitutes the majority of the sales. Other research can examine online channel sales in addition to sales at the brick and mortar stores. This study is in the context of experience goods;

extensions to other contexts (e.g., search goods, consumer packaged goods) could lend generalizability to our results. Future research could also conduct a detailed message level analysis that examines the type of messages (e.g., informative vs. persuasive) and also includes supply side analysis of FGC. We wish to note that not all the customers that participate in social media will intently read FGC. As is typical of research in the social media domain that uses observational data, we are unable to parse out this issue. Furthermore, we are analyzing the impact of FGC at a point in time when the potential customer base that opts to receive FGC is moderate; and it may be worthwhile for later research to revisit this issue when the FGC participation rates are high. We realize that some of the limitations of the study are due to lack of relevant datasets (e.g., not having TV advertising data at the individual customer level) and hope that as more data becomes available, future research can build on our study to explore other issues related to firm generated content in social media.

REFERENCES

Aaker, Davis (1991), *Managing Brand Equity: Capitalizing on the Value of a Brand Name.* New York, NY: Free Press.

Angrist, Joshua D., and Jörn Steffen Pischke (2009), "Mostly harmless econometrics: An empiricist's companion," *Princeton University Press*, N.J.

Ansari, Asim, and Carl F. Mela (2003), "E-Customization," *Journal of Marketing Research*, 40(2), 131-145.

Bollinger, Brian, Phillip Leslie, and Alan Sorensen (2011), "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 3 (1), 91-128.

Bruhn, M., Schoenmueller, V. and Schäfer, D.B. (2012), "Are social media replacing traditional media in terms of brand equity creation?," *Management Research Review*, 35(9), 770-790.

Campbell, Margaret C. and Kevin L. Keller (2003), "Brand Familiarity and Advertising Repetition Effects," *Journal of Consumer Research*, 30 (2), 292-304.

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

Chib, Siddhartha, and Edward Greenberg (1995), "Hierarchical Analysis of SUR Models with Extension to correlated serial errors and time-varying parameter models," *Journal of Econometrics*, 68(2), 339-360.

Chintagunta, Pradeep K., Jinhong Chu, and Javier Cebollada (2012), "Quantifying Transaction Costs in Online/Offline Grocery Channel Choice," *Marketing Science*, 31 (1), 96-114.

Chintagunta, Pradeep K., Shyam Gopinath, and Sriram Venkataraman (2010), "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation across Local Markets," *Marketing Science*, 29 (5), 944–957.

CMO Survey (2015), "CMO Survey Report: Highlights and Insights", [available at http://cmosurvey.org/files/2015/02/The_CMO_Survey-Highlights_and_Insights-Feb-2015_Compressed.pdf], accessed on October 3, 2015.

Dagger, T. S., and Peter J. Danaher (2014),"Comparing the effect of store remodeling on new and existing customers," *Journal of Marketing*, 78(3), 62-80.

Danaher, Peter J. and Tracey S. Dagger (2013), "Comparing the Relative Effectiveness of Advertising Channels: A Case Study of a Multimedia Blitz Campaign," *Journal of Marketing Research*, 50, (4), 517-534.

Das, Sanjiv R., and Mike Y. Chen (2007), "Yahoo! for Amazon: Sentiment extraction from small talk on the web," *Management Science*, 53 (9), 1375-1388.

Dhar, Vasant and Elaine A. Chang (2009), "Does Chatter Matter? The Impact of User-Generated Content on Music Sales," *Journal of Interactive Marketing*, 23 (4), 300-307.

eMarketer (2015), "Social Network Ad Spending to Hit $23.68 Billion Worldwide in 2015", [available at: http://www.emarketer.com/Article/Social-Network-Ad-Spending-Hit-2368-Billion-Worldwide-2015/1012357#sthash.lBpDFlob.dpuf], accessed on June 05, 2015.

Gensler, Sonja, Franziska Völckner, Yuping Liu-Thompkins, and Caroline Wiertz (2013), "Managing Brands in the Social Media Environment," *Journal of Interactive Marketing*, 27 (4), 242–256.

Girma, Sourafel, and Holger Görg (2007), "Multinationals' Productivity Advantage: Scale or Technology?", *Economic Inquiry*, 45 (2), 350-362.

Godes, David (2011), "Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science*, 30 (2), 224–229.

Guo, Shenyang and Mark W. Fraser (2010), *Propensity Score Analysis: Statistical Methods and Applications*, Thousand Oaks, CA: Sage.

Hoffman, D. L., and M. Fodor (2010), "Can you measure the ROI of your social media marketing," *MIT Sloan Management Review*, 52 (1), 41-49.

Huang, Qingyi, Vincent R. Nijs, Karsten Hansen, and Eric T. Anderson (2012), "Wal-Mart's Impact on Supplier Profits," *Journal of Marketing Research*, 49 (2), 131-143.

Inc (2012), "The New Rules of Customer Engagement", [http://www.inc.com/wendy-lea/new-rules-of-customer-engagement.html], accessed on June 9, 2015.

Joo, M., Kenneth C. Wilbur, Bo Cowgill and Yi Zhu (2014), "Television Advertising and Online Search", *Management Science*, 60 (1): 56-73.

Kumar, V., Vikram Bhaskaran, Rohan Mirchandani, and Milap Shah (2013), "Creating a Measurable Social Media Marketing Strategy: Increasing the Value and ROI of Intangibles and Tangibles for Hokey Pokey," *Marketing Science*, 32(2), 194-212.

Kumar, V., Morris George, and Joseph Pancras (2008), "Cross-buying in Retailing: Drivers and Consequences," *Journal of Retailing*, 84 (1), 15-27.

Kumar, V., Rajkumar Venkatesan, and Werner Reinartz (2008), "Performance Implications of Adopting a Customer-Focused Sales Campaign," *Journal of Marketing*, 72 (5), 50–68.

Li, Hongshuang (Alice) and P. K. Kannan (2014), "Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment," *Journal of Marketing Research*, 51 (1), 40-56.

Li, Shibo, Baohong Sun and Ronald T. Wilcox (2005), "Cross-selling Sequentially Ordered Products," *Journal of Marketing Research*, 42 (2), 233–239.

Lin, C., Sriram Venkataraman, and Sandy D. Jap (2013), "Media multiplexing behavior: Implications for targeting and media planning", *Marketing Science*, 32(2), 310-324.

Loyalty One (2012), "The Missing Link between Social Media and ROI," [available at https://loyalty.com/knowledge/articles/social-media-payoff; accessed June 1, 2015]

Malthouse Edward C., Michael Haenlein, Bernd Skiera, Egbert Wege and Michael Zhang (2013), "Managing Customer Relationships in the Social Media Era: Introducing the Social CRM House," *Journal of Interactive Marketing*, 27 (10), 270–280.

Naik, Prasad A. and Kalyan Raman (2003), "Understanding the Impact of Synergy in Multimedia Communications," *Journal of Marketing*, 40 (4), 375-388.

Naylor, Rebecca W., Cait P. Lamberton, and Patricia M. West (2012), "Beyond the "Like" Button: The Impact of Mere Virtual Presence on Brand Evaluations and Purchase Intentions in Social Media Settings," *Journal of Marketing*, 76 (6), 105-120.

Palmatier, Robert, Rajiv Dant, Dhruv Grewal, and Kenneth Evans (2006), "Factors Influencing the Effectiveness of Relationship Marketing," *Journal of Marketing*, 70 (4), 136-153.

Pauwels, Koen, and Shuba Srinivasan (2004), "Who benefits from store brand entry?," *Marketing Science*, 23 (3), 364-390.

PR Newswire (2011), New Study Suggests Integrated Social Media Drives Sales," [available at http://www.prnewswire.com/news-releases/new-study-suggests-integrated-social-media-drives-sales-123803569.html], accessed June 1, 2015.

Ranaweera, Chatura and Jaideep Prabhu (2003), "The Influence Of Satisfaction, Trust and Switching Barriers on Customer Retention In A Continuous Purchasing Setting," *International Journal of Service Industry Management*, 14 (4), 374-395.

Retailingtoday (2013), "Macy's Appeals to Tech-Savvy Consumers with New App," [available at www.retailingtoday.com/article/macy-appeals-tech-savvy-consumers-new-app].

Reinartz, W. J., and V. Kumar (2003), "The impact of customer relationship characteristics on profitable lifetime duration," *Journal of Marketing*, 67(1), 77-99.

Rosenbaum, Paul R., and Donald B. Rubin (1985), "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score", *The American Statistician*, 39 (1), 33-38.

Rossi, Peter E., Greg M. Allenby, and Rob McCulloch (2006), Bayesian Statistics and Marketing. Hoboken, NJ: John Wiley & Sons.

Sáenz-Navajas, M. P., Eva Campo, Angela Sutan, Jordi Ballester, and Dominique Valentin (2013), "Perception of wine quality according to extrinsic cues: The case of Burgundy wine consumers", *Food Quality and Preference*, 27(1), 44-53.

Schivinski, Bruno and Dariusz Dabrowski (2014), "The Effect of Social-Media Communication on Consumer Perceptions of Brands," *Journal of Marketing Communications*, 12(6), 6-10.

Schulze, Christian, Lisa Schöler, and Bernd Skiera (2014), "Not All Fun and Games: Viral Marketing for Utilitarian Products," *Journal of Marketing*, 78 (1), 1-19.

Shah, Denish, V. Kumar, Yingge Qu and Sylia Chen (2012), "Unprofitable Cross-Buying: Evidence from Consumer and Business Markets," *Journal of Marketing*, 76 (3), 78–95.

Smith, Herbert L. (1997), "Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies," *Sociological Methodology*, 27, 325:353.

Smith, Timothy M., Srinath Gopalakrishna and Rabikar Chatterjee (2006), "A Three-Stage Model of Integrated Marketing Communications at the Marketing–Sales Interface," *Journal of Marketing Research*, 43 (4), 564-579.

Stephen, Andrew T. and Jeff Galak (2012), "The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace," *Journal of Marketing Research*, 49 (5), 624-639.

Stilley, Karen M., J. Jeffrey Inman and Kirk L. Wakefield (2010), "Spending on the Fly: Mental Budgets, Promotions, and Spending Behavior," *Journal of Marketing*, May 2010, 74(3), 34-47.

Tellis, Gerard J. and Doyle L. Weiss (1995), "Does TV Advertising Really Affect Sales?," *Journal of Advertising*, 24 (3), 1-12.

Tirunillai, Seshadri and Gerard J. Tellis (2012), "Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance," *Marketing Science*, 31 (2), 198–215.

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

Toubia, Olivier and Andrew T. Stephen (2013), "Intrinsic vs. Image-Related Utility in Social Media: Why do People Contribute Content to Twitter?," *Marketing Science*, 32 (3), 368-392.

Vakratsas, Demetrios and Tim Ambler (1999), "How Advertising Works: What do We Really Know?," *Journal of Marketing*, 63 (1), 26–43.

Verhoef, Peter C. and Bas Donkers (2005), "The Effect of Acquisition Channels on Customer Loyalty and Cross-Buying," *Journal of Interactive Marketing*, 19 (2), 31–43.

Wall Street Journal (2014), "New Facebook Rules Will Sting Entrepreneurs," November 27.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Table 1: Variable Operationalization and Summary Statistics

Variable Description	Variable Notation	Variable Operationalization	Mean	SD
Social Media Participation Data and Traditional (TV Ad) & Digital (Email) Marketing Communications Data				
Firm Generated Content	*FGC*	Messages posted by the firm. (# of postings/week)	3.78	1.82
Customer Participation	*CustPar*	=1 if customer *h* participates in firm's social media page at time *t*; 0 otherwise.		
Television Advertising	*TVAd*	Television ad based on gross rating point. (GRP/week)	40.79	26.42
Email Advertising	*Email*	Total number of emails sent by the firm that are opened by customer *h* during week *t*. (# of emails opened/week/customer)	1.63	1.02
Customer Transaction Data				
Spending Behavior	*Spend*	Total dollar amount spent by customer *h* during week *t* on firm's products. ($/week/customer)	14.34	7.26
Cross-buying Behavior	*CrossBuy*	Total number of distinct categories customer *h* purchased during week *t*. (# of distinct categories bought/week/customer)	3.23	1.24
Customer Experience	*CExp*	Duration of the relationship of customer *h* with the firm until week *t*. (# of weeks of relationship with focal firm/customer)	94.26	17.89
Promotion Depth Index	*PromD*	Weighted average of price cuts availed by customer *h* for week *t* across all product items purchased. (cents/ml/week)	.38	.65
Cross Category Promotion	*CrossP*	Proportion of categories bought on promotion by customer *h* for week *t*. (proportion of categories bought on promotion/week/customer)	.46	.21
Survey Data				
Time Spent on Social Network	*TimeOnSocialNet*	Time spent on social networks (from survey). (minutes/customer)	41.54	8.50
Online Social Profile	*OProf*	Number of online social profiles (# online social profile /customer)	2.47	1.15
Age	*Age*	Age of customer *h*. (years/customer)	45.46	18.52
Distance	*Dist*	Distance between the customer *h* and the focal store. (miles/customer)	5.17	8.06
Attitudinal Variables	*TechS*		3.69	.80
	SocialNet		2.77	1.22
	PrivCon		1.55	.51
	IMov	These constructs are measured via survey.	4.08	.65
	SMov		2.88	.76
	EscMov		2.12	.80
	OEnt		2.77	.78
Gender	*Gender*	=1 if customer *h* is male; 0 otherwise.	47%	
Race	*Race*	=1 if customer *h* is white; 0 otherwise.	86%	

Note: Summary statistics for customer transaction data are unconditional on purchase occasions. The summary statistics are based on a matched pair of 412 customers from each of the treatment and the control groups.

Table 2: Average Difference in Differences: Purchase Behavior of Treatment vs. Control Groups

Outcome Variable	Groups	Post-FGC	Pre-FGC	Difference	Difference-in-Differences
Spending	Treatment Group	15.96	13.52	2.44** (.92)	1.26**
	Control Group	14.53	13.35	1.18 (1.01)	
Cross-buying	Treatment Group	3.74	3.08	.66** (.25)	.46**
	Control Group	3.15	2.95	.20 (.19)	
Profit	Treatment Group	5.91	4.79	1.12** (.43)	1.02**
	Control Group	4.78	4.68	.10 (.09)	

Note: The table presents the mean values of customer behaviors (spending, cross-buying, and customer profitability) at the weekly level. The pre-FGC and the post-FGC period comprise of 85 weeks for all the customers in the control group. Given that the treatment group customers can join the focal firm's social media page at different times, the number of weeks for the pre-FGC and post-FGC analysis vary across the treatment group customers. Whereas the "difference" measure is based on the paired-sample t-test, the "difference-in-differences" measure is derived by calculating the difference in the outcome variables between post- and pre- FGC periods and then comparing the means of these differences between the treatment and control groups. The above analysis are based on a matched pair of 412 customers from each of the treatment and the control groups.

***Difference is significantly different from zero, p≤.01. Standard errors are shown in the parentheses.*

Table 3: Results of Difference in Differences Model: Spending and Cross-buying Behavior

Variables	M1		M2		M3		M4		M5	
	Spending	Cross-buying	Spending	Cross-buying	Spending	Cross-buying	Spending	Cross-buying	Spending	Cross-buying
TCust	.1860	.1316	.1723	.1292	.1632	.1099	.1518	.1029	.0803	.0941
FGC	1.1639	.1898	1.064	.1878	1.0463	.1725	1.0166	.1558	.9851	.1291
TCust×FGC	1.2468**	.4582**	1.2766**	.4597**	1.3031**	.4693**	1.3093**	.4790**	1.3233**	.4987**
TCust×FGC×TVAd	---	---	---	---	.0085**	4.E-05**	.0026**	.0009**	.0057**	.0001**
TCust×FGC×Email	---	---	---	---	.0791**	.0020**	.0704**	.0026**	.0991**	.0047**
TCust×FGC×CExp	---	---	---	---	.1282**	.0777**	.1007**	.0788**	.1686**	.0797**
TCust×FGC×TechS	---	---	---	---	.1816**	.0090**	.1351**	.0097**	.2753**	.0303**
TCust×FGC×SocialNet	---	---	---	---	.1978**	.1370**	.0957**	.1161**	.1906**	.1523**
TVAd	---	---	.4285**	.2221**	.0737**	.1928**	.0260**	.1714**	.1373**	.1733**
Email	---	---	1.3489**	.2406**	1.1074**	.0157**	.9895**	.0215**	.9268**	.1242**
CExp	---	---	.5005**	.3206**	.6232**	.3001**	.5519**	.2897**	.7579**	.336**
TechS	---	---	-.6772**	-.0983**	-.4884**	-.2494**	-.4636**	-.2048**	-.5285**	-.2062**
SocialNet	---	---	-.3592**	-.0417**	-.8194**	-.3392**	-.7791**	-.3343**	-.5906**	-.3596**
PromoD	---	---	5.0113**	---	4.3664**	---	4.4414**	---	3.9708**	---
CrossP	---	---	---	1.1168**	---	1.1192**	---	1.0130**	---	.6949**
Dist	---	---	-.5020*	-.0628	-.4686*	-.0429	-.3949*	-.0259	-.3019*	-.0316
Gender	---	---	1.3071	.2368	1.1031	.2057	1.0071	.2310	.9660	.4133
Race	---	---	1.0456	.1251	.6228	.3838	.7753	.2758	.7214	.2093
Age	---	---	.0555	.1927	.1773	.2396	.1158	.2758	.2778	.3005
Intercept	13.2638**	3.1962**	12.3664**	3.5017**	11.0745*	3.2707*	11.9997*	3.1307**	6.2314*	2.6182**
Serial-correlation	×	×	×	×	×	×	.0200*	.0509	.0187*	.0482
Cross-correlation	×		.5321**		.5134**		.5018**		.4765**	
Heterogeneity	×		✓		✓		✓		✓	
# of Observations	140080		140080		140080		140080		140080	
Sample Size	824		824		824		824		824	
LMD	-79248.44		-73270.09		-67878.06		-66366.26		-56313.75	
Model Description	M1: Standard DID		M2: M1 + Controls		M3: M2 + Interaction		M4: M3 + Serial Correlation		M5: M4 with Tobit	

**p≤.01 (Parameter is significant at 99% level i.e., the 99% confidence interval does not contain 0)
*p≤.05 (Parameter is significant at 95% level i.e., the 95% confidence interval does not contain 0)
The sample size of 824 customers is based on a matched pair of 412 customers from each of the treatment and the control groups.

Table 4a: Results of Treatment Effect Model: Spending and Cross-Buying Behavior

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
	Spending	Cross-buying	Spending	Cross-buying	Spending	Cross-buying	Spending	Cross-buying	Spending	Cross-buying
FGC	.1153**	.0158**	.1240**	.0237**	.1394**	.0205**	.1384**	.0328**	.1835**	.1103**
FGC×TVAd	---	---	---	---	.0388**	.0019**	.0444**	.0012**	.0621**	.0019**
FGC×Email	---	---	---	---	.1653**	.0864**	.1546**	.0866**	.1755**	.1137**
FGC×CExp	---	---	---	---	.0104**	.0468**	.0206**	.0611**	.0450**	.1228**
FGC×TechS	---	---	---	---	.0799**	.0142**	.0736**	.0076**	.0918**	.0281**
FGC×SocialNet	---	---	---	---	.0780**	.0252**	.0697**	.0468**	.1091**	.0956**
TVAd	---	---	.0410**	.0332**	.0443**	.0450**	.0731**	.0615**	.1082**	.1218**
Email	---	---	.7943**	.1087**	.8627**	.1199**	.8832**	.1141**	1.0258**	.8954**
CExp	---	---	.2650**	.0163**	.3724**	.0468**	.3602**	.0441**	.5853**	.2055**
TechS	---	---	-.1845**	-.0279**	-.2174**	-.0288**	-.1849**	-.0351**	-.1537**	-.0975**
SocialNet	---	---	-.1758**	-.0207**	-.2451**	-.0782**	-.2140**	-.0772**	-.1976**	-.4158**
PromoD	---	---	5.2579**	---	4.6264**	---	3.9354**	---	2.2350**	---
CrossP	---	---	---	1.2462**	---	1.1310**	---	1.1246**	---	.5707**
Dist	---	---	-.0788*	-.0656	-.0334*	-.1471	-.0027*	-.1346	-.0394*	-.0385
Gender	---	---	.7644	0.0218	.8504	.0524	.8482	.0456	.9690	.1114
Race	---	---	.8390	0.0406	.8460	.0671	.8523	.0634	1.0673	.1102
Age	---	---	.0482	0.1082	.0666	.1487	.0850	.1399	.6973	.1844
Intercept	12.7153**	2.6589**	11.7622**	2.5514**	11.2091**	2.3329**	9.5218**	1.6543**	6.6342*	1.2867*
Serial-Correlation	×	×	×	×	×	×	.0178*	.0491	.0076*	.0128
Cross-Correlation	×		.4928*		.4726*		.4568*		.4124*	
Heterogeneity	×		✓		✓		✓		✓	
# of Observations	66980		66980		66980		66980		66980	
Sample Size	788		788		788		788		788	
LMD	-97534.30		-92754.71		-91243.21		-83385.43		-75752.36	
Model Description	Standard TE Model		Model 1 + Controls		Model 2 + Interaction		Model 3 + Serial-correlation		Model 4 with Tobit	

**p≤.01 (Parameter is significant at 99% level i.e., the 99% confidence interval does not contain 0)
*p≤.05 (Parameter is significant at 95% level i.e., the 95% confidence interval does not contain 0)
The sample size of 788 consumers is based on a matched pair of 394 customers from each of the treatment and control groups.

Table 4b: Results of Treatment Effect Model: Customer Social Media Participation

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
PrivCon	-2.2369**	-.7076**	-.7377**	-.6685**	-.1766**
IMov	1.7925**	.4170**	.1901**	.7802**	.2865**
TechS	.8375*	.0270*	.3165*	.4064*	.0521*
SMov	.7714**	.2279**	.0167**	.4032**	.0771**
OEnt	1.0763**	.3422**	.1826**	.2069**	.0509**
TimeOnSocialNet	1.8047**	1.4342**	1.4077**	1.2919**	1.1770**
Gender	---	.0114	.0160	.0287	.0095
Race	---	-.1673	-.0253	-.3322	-.0297
Age	---	-.0029	-.0061	-.0052	-.0043
Intercept	-4.2156*	-.5074*	-1.3272*	-.3041*	-.4516*
Cross-Correlation (Spending, Participation)	---	.2432**	.2134**	.1842**	.1426**
Cross-Correlation (Cross-buying, Participation)	---	.0967	.0826	.0626	.0414
Heterogeneity	×	✓	✓	✓	✓
# of Observations	66980	66980	66980	66980	66980
Sample Size	788	788	788	788	788
LMD	-97534.30	-92754.71	-91243.21	-83385.43	-75752.36

**p≤.01 (Parameter is significant at 99% level i.e., the 99% confidence interval does not contain 0)
*p≤.05 (Parameter is significant at 95% level i.e., the 95% confidence interval does not contain 0)
The sample size of 788 consumers is based on a matched pair of 394 customers from each of the treatment and control groups.

Table 5: Results of DID Model: Customer Profitability and Cross-Buying Behavior

Variables	Profit	Cross-buying
TCust	.0190	.1029
FGC	.0114	.1990*
TCust×FGC	1.1511**	.4873**
TCust×FGC×TVAd	.0034**	.0067**
TCust×FGC×Email	.0764**	.0115**
TCust×FGC×CExp	.1962**	.0507**
TCust×FGC×TechS	.6921**	.0586**
TCust×FGC×SocialNet	.7856**	.1709**
TVAd	.4310**	.2762**
Email	.6482**	.4325**
CExp	1.3341**	.2294**
TechS	-.7435**	-.4107**
SocialNet	-1.1248**	-.6669**
PromoD	-1.7639**	---
CrossP	---	.6543**
Dist	-.4610*	-.0310
Gender	.4452	.4224
Race	.9708	.2184
Age	.2844	.3019
Intercept	3.3443**	1.7202**
Serial-Correlation	.0089	.0321
Cross-Correlation (Profit, Cross-buying)	.2134*	
Heterogeneity	✓	
# of Observations	140080	
Sample Size	824	
LMD	-54233.61	

**p≤.01 (Parameter is significant at 99% level i.e., the 99% CI does not contain 0)
*p≤.05 (Parameter is significant at 95% level i.e., the 95% CI does not contain 0)
The sample size of 824 consumers is based on a matched pair of 412 customers from each of the treatment and the control groups.

Table 6: Results of Treatment Effect Model: Customer Profitability and Cross-Buying Behavior

Variables	Profit	Cross-buying
FGC	.0808**	.0826**
FGC×TVAd	.0011**	3.E-05**
FGC×Email	.0294**	.0853**
FGC×CExp	.0114**	.0756**
FGC×TechS	.0581**	.0189**
FGC×SocialNet	.0541**	.0376**
TVAd	.0157**	.0089**
Email	.1546**	.1084**
CExp	.0392**	.3093**
TechS	-.9235**	-.0356**
SocialNet	-.1407**	-.0822**
PromoD	-2.7377**	---
CrossP	---	.9564**
Dist	-.0834*	-.0004
Gender	1.2483	.0753
Race	1.1031	.0473
Age	.0589	.0059
Intercept	3.7150**	1.0353*
Serial-Correlation	.0038*	.0087
Cross-Correlation (Profit, Cross-buying)	.3164*	
Heterogeneity	✓	
# of Observations	66980	
Sample	788	
LMD	-80741.24	

**p≤.01 (Parameter is significant at 99% level i.e., the 99% CI does not contain 0)
*p≤.05 (Parameter is significant at 95% level i.e., the 95% CI does not contain 0)
The sample size of 788 customers is based on a matched pair of 394 customers from each of the treatment and control groups.

Table 7: Elasticity Analysis: Social, Traditional and Digital Media

Type of Marketing Communication	Overall		First six months		Last six months	
	Spending	*Cross-buying*	*Spending*	*Cross-buying*	*Spending*	*Cross-buying*
Social Media (FGC)	.0140	.0587	.0060	.0441	.0108	.0493
	(.0027)	(.0035)	(.0007)	(.0061)	(.0023)	(.0049)
Digital Media (Email)	.0913	.0433	.0799	.0409	.0888	.0408
	(.0215)	(.0165)	(.0137)	(.0043)	(.0358)	(.0202)
Traditional Media (TV Ad)	.1010	.0507	.0949	.0457	.0968	.0532
	(.0075)	(.0038)	(.0023)	(.0202)	(.0097)	(.0085)

Notes: Calculation of elasticity is based on the results of the TE model.
First 6 months refers to the period from August 2009 to January 2010 (first six months since inception of the firm's social media page). Last six months refers to the period from October 2010 to March 2011 (last 6 months of our data's post-FGC period). Standard errors are shown in parentheses.

Table 8: Synergistic Effects of FGC with TV Advertising and Email Marketing

Simulation	Variables in the Model	% Change in Spending	% Change in Cross-buying
Traditional Ad (TVAd) and FGC	Main effects of FGC and TV	2.08	1.06
		(.23)	(.14)
	Main effects of FGC and TV and interaction effect between FGC and TV	3.11	1.90
		(1.33)	(.85)
	Incremental change	1.03	.84
Digital Ad (Email) and FGC	Main effect of FGC and email	3.25	1.54
		(.54)	(.35)
	Main effects of FGC and email and interaction effect between FGC and email	5.27	2.76
		(.68)	(1.18)
	Incremental change	2.02	1.22

Notes: The base case is the TE model with the main effect of FGC but without TV and Email. All the models include the control variables that we presented in Equations (5)-(7).
Standard errors are shown in parentheses.

Table 9: Elasticity Analysis: Effect of the Three Dimensions of FGC on Customer Behavior

Focal Dimension of FGC	Spending	Cross-buying
Valence	.0128	.0285
	(.0068)	(.0061)
Receptivity	.0188	.0864
	(.0020)	(.0162)
Susceptibility	.0092	.0076
	(.0032)	(.0028)

Notes: Calculation of elasticity is based on the results of the TE model. Standard errors are shown in parentheses.

Figure 1: Conceptual Framework

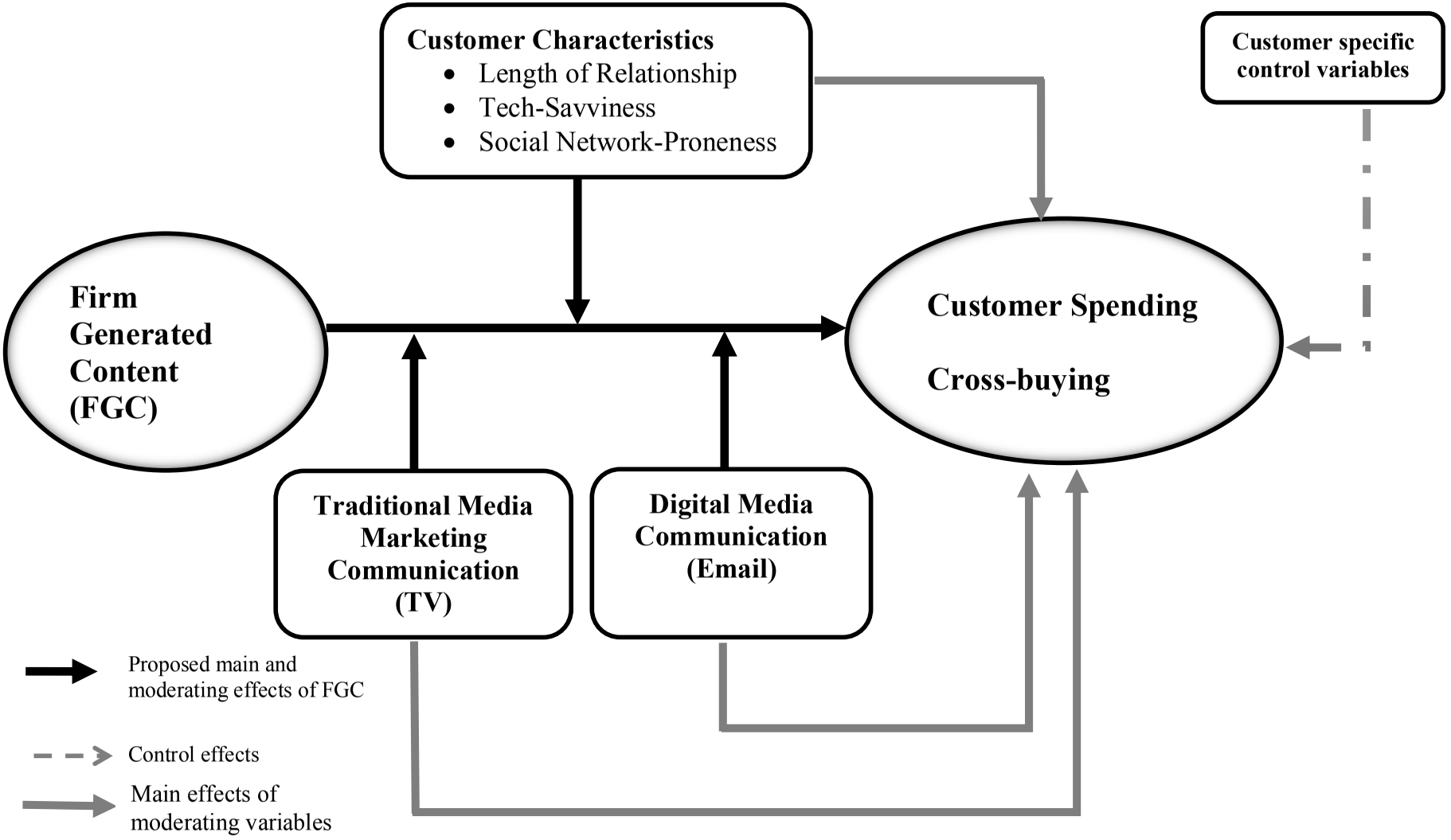
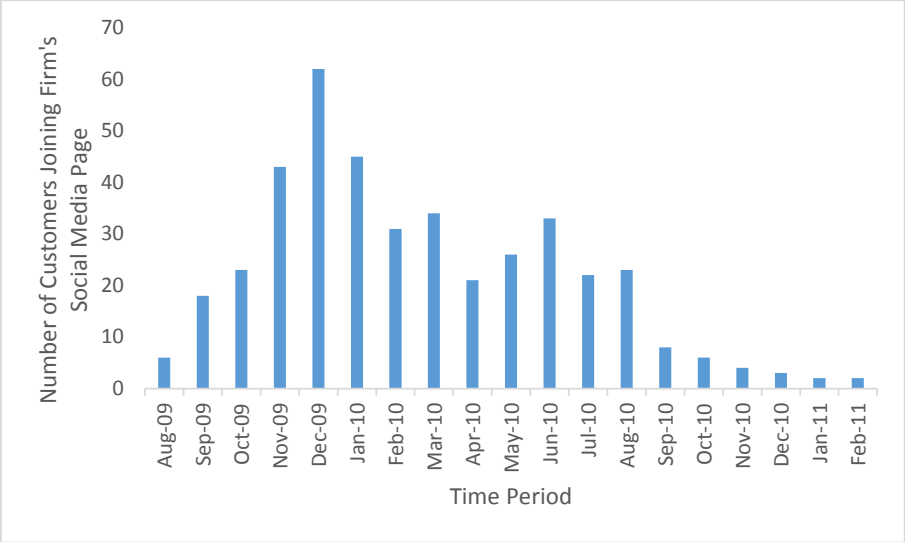


Figure 2: Histogram of Treatment Customers Joining the Firm's Social Media Page



APPENDIX

Examples of FGC

The focal retailer uses FGC to communicate several different types of information such as information about specific products, events in physical stores, specific wines in stock along with images and links to relevant content. Messages posted in social media may convey information about special events such as wine tastings, new wine/liquor arrivals and inventory information. Sometimes the postings may also relate to local events (e.g., art shows). Thus the focal firm's postings on the social media platform constitute both promotional and non-promotional messages. We present below some examples of FGC from the retailer' social media page. Please note that the postings are usually accompanied by relevant visuals such as photos and videos.

- *Join Kate in our Reserve Room today as we pour wines from famous producers, such as, Schloss Vollrads, Banfi and Fonseca. Join us 4:00 - 7:00pm!*
- *Plenty of wines will be open at the store this weekend, including these staff favorites! Join us Friday and Saturday 12:00-6:00pm.*
- *There are more amazing wines open at our tasting center than we can even fit in this status update... get to the store before 6pm and get your sample on!*
- *Our class tonight on is nearly full but there are a few spots left in our Friday class an evening with Hermann J. Wiemer estate manager Oskar Bynke.*
- *FYI: We're open Monday 9:30-5:00 for any Labor Day needs!*
- *FREE Jim Beam White Label & Red Stag samples on Saturday July 10th from 12-3pm*