Social Media Broadcasts and the Maintenance of Diverse Networks

Completed Research Paper

Yotam Shmargad
University of Arizona
yotam@email.arizona.edu

Abstract

Social media platforms like Facebook, Twitter, and LinkedIn let people broadcast messages to their entire network of contacts all at once. As the number of users and the amount of information they broadcast grow, platform managers face an increasingly pressing problem – which broadcasts should they show their users? To address this question, I study the database of a social media company that started charging its users to receive broadcasts about their contacts. By relating purchase rates to properties of users’ social networks, I identify which ties are most valuable to maintain through broadcasts. I find that strong ties increased purchase rates more than weak ties. However, purchase rates also increased with the structural diversity of users’ ties. Social media platforms should thus prioritize broadcasts from ties that are either strong or structurally diverse.

Keywords: Broadcasts, Social Media, Social Networks, Structural Diversity, Tie Strength

Introduction

Social media platforms like Facebook, Twitter, and LinkedIn let people broadcast messages to their entire network of contacts all at once. As a result, users of social media now receive more information from their contacts than ever before. However, as both the number of social media users and the amount of information they broadcast grow, platform managers face an increasingly pressing problem – which broadcasts should they show their users? For example, Facebook users, on average, could potentially see 1,500 different broadcasts each time they log into the service.¹ Most users do not have the time to see all of these broadcasts, so Facebook uses an algorithm to determine which broadcasts to display. Platform managers at Facebook are thus interested in variables that allow them to know how users evaluate different broadcasts, so that they can display the information that users want to see most.

In this paper, I argue that platform managers can leverage information about users’ online social networks to make better decisions about which broadcasts to display to them. To show that the value of receiving broadcasts depends on properties of users’ networks, I analyze the database of a social media company that started charging its users to receive broadcasts in their email address books. Before monetization, users’ address books were automatically updated each time one of their contacts broadcast a piece of information about themselves (e.g. mailing address, email address, phone number, or place of work). Once the feature was monetized, however, users needed to buy the company’s premium bundle for $60 to continue receiving broadcasts in their address books.

By comparing how premium purchase rates varied with properties of users’ networks before and after this change, I identify which ties are most valuable to maintain through broadcasts. I find that strong ties (e.g. family members, business friends) increased the likelihood of paying for broadcasts more than weak ties (e.g. ties that are only business in nature). However, purchase rates also increased with the

¹ News Feed FYI: A Window Into News Feed, Facebook, August 2013
structural diversity of users’ ties – the extent to which these ties connected users to different regions of the social network. Social media platforms should thus prioritize broadcasts from ties that are either strong or structurally diverse. Given that platforms already collect extensive data on users’ social networks, it would certainly be feasible for them to include social network measures in the algorithms they use to determine which broadcasts users see.

**Theoretical Development**

One key variable that Facebook uses to determine whether or not users see a broadcast from a particular contact is the frequency with which they interact with that contact. In the social networks literature, frequent interactions between two people increase with what is commonly called the strength of their tie (Granovetter 1973). By prioritizing broadcasts from contacts with whom users interact frequently, Facebook prioritizes broadcasts from strong ties over weak ties. In fact, there are good reasons to believe that users prefer receiving information from their strong ties. For one, strong ties are more similar to one another than weak ties (McPherson et al. 2001). Information that a user’s strong tie broadcasts may thus conform to that user’s own preferences, which may be why users are more likely to re-share information they receive from a strong tie (Bakshy et al. 2012). Moreover, strong ties are more likely to provide social support than weak ties (Wellman and Wortley 1990), something which is true online as well as offline (Burke et al. 2013). Thus, users may simply care more to hear from their strong ties than their weak ones.

Marsden and Campbell (1984) find that tie strength is associated positively with kinship and negatively with co-workership. Family ties, in particular, are more likely to provide emotional and financial support (Wellman and Wortley 1990). Thus, I assume that family ties are, on average, stronger than friend ties, and that friend ties are, on average, stronger than ties that are business in nature. If users prefer to receive broadcasts from strong ties than weak ones, then the following hypotheses should hold:

H1a: Family ties increase the likelihood of paying for broadcasts more than friend ties  
H1b: Friend ties increase the likelihood of paying for broadcasts more than business ties

The strength of a tie also depends on its multiplexity – the number of different types of relations associated with the tie (Van den Bulte and Wuyts 2007). In particular, ties that share both a friendship and a business relationship are often more motivated to help one another, because resources from one relationship can be used for the other (Uzzi 1997). For example, a favor at work can be reciprocated with a drink or a meal after work. I assume that ties that are types friend and business are, on average, stronger than ties that are either friend or business, but not both. This motivates the following hypotheses:

H1c: Business friends increase the likelihood of paying for broadcasts more than friend ties  
H1d: Business friends increase the likelihood of paying for broadcasts more than business ties

On the other hand, some of the early work surrounding social media suggests that they may be particularly useful for maintaining networks of weak, diverse ties (Donath and boyd 2004). People already have ways of communicating with their strong ties, but the impersonal nature of broadcasts seems to make them ideal for keeping tabs on ties that are harder to reach. In support of this, Ellison et al. (2007) find that Facebook usage increases more with bridging social capital than bonding social capital. While bonding social capital results from keeping strong, cohesive ties, bridging social capital has more to do with maintaining weak, diverse ties (Putnam 2000). While the previous hypotheses speak to tie strength, I appeal directly to the difficulty of reaching one’s contacts by abstracting away from the strength of users’ ties and studying their structural diversity.

---

2 I exclude multiplex family ties from my analyses because my dataset contains few instances of such ties. The results of my study are robust to including these ties.
Structurally diverse ties connect individuals to regions of the social network to which they otherwise would not be connected. While structurally diverse ties tend to be weaker, tie strength and structural diversity are distinct constructs (Granovetter 1973, Aral and Van Alstyne 2011). There are two reasons to think that broadcasts are particularly valuable for maintaining relationships with diverse ties. First, people are more likely to lose touch with these ties than other ties (Burt 2002). Second, there exist fewer alternative ways for people to obtain information about these ties. Figure 1 illustrates this logic. The leftmost diagram shows an individual, A, with ties that are structurally diverse — that is, B and C connect A to different regions of the social network. If A loses touch with B (denoted by a dashed line), then A can only obtain information about B through a broadcast. Hence, broadcasts are valuable for maintaining the relationship with that tie.

Contrast this with the value of receiving broadcasts in networks that are not structurally diverse. Burt (1992) argues that two different structures limit the diversity of one’s network. The first, redundancy by cohesion, occurs when one’s contacts are connected to each other. The second, redundancy by structural equivalence, occurs when one’s contacts are connected to the same people. The middle diagram in Figure 1 shows an individual, A’, with a network that is redundant by cohesion. If A’ loses touch with her contact, B’, then A’ can obtain information about B’ through their mutual contact, C’. The rightmost diagram shows another individual, A”, with a network that is redundant by structural equivalence. If A” loses touch with her contact, B”, then A” can obtain information about B” through C”. Even though C” is not directly connected to B”, B” and C” share a mutual contact, D”. Thus, information can still flow indirectly from B” to A”.

Figure 1: Illustration of the relative value of broadcasts in different networks.

Since individuals in networks that are redundant by cohesion or structural equivalence can obtain information about one contact from another, they should not depend on broadcasts for this information as much as individuals in structurally diverse networks. Thus, the likelihood of paying for broadcasts should decrease with the amount of redundancy present in an individual’s network:

H2a: The likelihood of paying for broadcasts decreases with redundancy by cohesion
H2b: The likelihood of paying for broadcasts decreases with redundancy by structural equivalence

Before introducing the empirical setting I use to test these hypotheses, I discuss the properties of individuals’ networks that influence their likelihood of paying for broadcasts.
Network Measures

I extend Katona et al.’s (2011) notation to accommodate for social networks with multiple relation types and denote a social network of type $r$, $r \in \{\text{Family, Friend, Business}\}$, by $G^r(V,E^r)$. Here, $V$ is a set of individuals and $E^r$ is the set of type $r$ relations among them. I define the set of individual $i$’s type $r$ contacts as $N_i^r = \{j|j \in V \text{ and } (i,j) \in E^r\}$. The following measures count the family-only, friend-only, and business-only contacts, respectively, in individual $i$’s network:

$$S_i^1 = |N_i^{\text{Family}} \setminus (N_i^{\text{Business}} \cup N_i^{\text{Friend}})|$$

$$S_i^2 = |N_i^{\text{Friend}} \setminus (N_i^{\text{Business}} \cup N_i^{\text{Family}})|$$

$$S_i^3 = |N_i^{\text{Business}} \setminus (N_i^{\text{Friend}} \cup N_i^{\text{Family}})|$$

The next measure counts the contacts in $i$’s network that are types business and friend, but not family:

$$S_i^4 = |(N_i^{\text{Business}} \cup N_i^{\text{Friend}}) \setminus N_i^{\text{Family}}|$$

My measures of diversity do not depend on the types of relations in one’s network, but rather on the patterns of relations among one’s ties. I collapse the family, friend, and business networks into one network and denote it by $G(V,E)$, where $E = E^{\text{Family}} \cup E^{\text{Friend}} \cup E^{\text{Business}}$. Similarly, I define the set of $i$’s contacts of any type to be $N_i = N_i^{\text{Family}} \cup N_i^{\text{Friend}} \cup N_i^{\text{Business}}$. In order to construct measures of diversity, it is useful to refer to single ties in the social network. For any two individuals $i,j \in V$, I let

$$e_{ij} = \begin{cases} 1 & \text{ if } (i,j) \in E \\ 0 & \text{ otherwise.} \end{cases}$$

I measure the redundancy by cohesion in an individual’s network using the clustering coefficient (Watts and Strogatz 1998), which is the proportion of the ties that could exist among one’s contacts that actually do exist:

$$D_i^1 = \frac{\sum_{j \neq q \in N_i} e_{ij}}{\sum_{j \neq q \in N_i} 1}$$

My next measure captures the redundancy by structural equivalence in an individual’s network, and is called local betweenness (Katona et al. 2011). To calculate this measure, I look at the pairs of contacts in one’s network that do not share a connection, and count how many people in the broader social network connect each pair. The contribution of a given pair to one’s local betweenness is the reciprocal of the number of people that connect them. Individuals with contacts that are connected to many of the same people will thus have lower betweenness scores. The final local betweenness score is the average of these reciprocals over all pairs of unconnected contacts in an individual’s network:

$$D_i^2 = \frac{\sum_{j \neq q \in N_i} \frac{1-e_{ij}}{\sum_{k \neq q \in N_i} e_{jk} \times e_{kj}}}{\sum_{j \neq q \in N_i} 1-e_{ij}}$$

Next, I describe the setting that lets me test how these measures influence individuals’ decisions to pay to receive broadcasts from their contacts.
A Natural Experiment in Monetization

When social media platforms first emerged, they were accessed through websites on web browsers. For this reason, early literature refers to these technologies as social network sites; “sites” being short for websites (e.g. Ellison et al. 2007, boyd and Ellison 2008). More recently, largely driven by an increase in smartphone and tablet usage, consumers have started accessing these services through software applications (or “apps”) developed solely to enable such access. The dataset that I analyze in this paper originates from a California-based company that managed a social media platform which was accessed in two ways: through an app and a website. Both the app and the site are important for understanding the empirical setting of my analyses, and so I discuss each one in turn.

Rather than being installed on smartphones or tablets, the app in my study was installed on email clients like Microsoft Outlook, Mac Mail, and Mozilla Thunderbird. When the app was first introduced in 2002, it was used to help people keep their email address books up-to-date. Users of the app could update their contact information, and these updates were automatically broadcast to any of their contacts that also had the app installed. By the time of my study, the app was used for more than broadcasting updated contact information and receiving broadcasts from others. By then, the app provided many of the same features as the company’s social network site. Figure 2 shows a screenshot of the app for Microsoft Outlook, which displays some of its more recent features.

![Figure 2: Screenshot of the app for Microsoft Outlook](image)

The company launched its social network site in 2007. The site allowed users to maintain an online profile, connect with other users, and broadcast messages, photos, videos, and other content to their connections. A unique feature of the site is that it required users to specify the types of connections
they had with other users. Connections were specified as family, friend, business, or any combination of these types, and only the types that were mutually agreed upon by both users were associated with a connection on the site. This feature made it easy for users to broadcast content to each group separately. Importantly, the site and the app stayed in sync, so that updates made on either were broadcast to both. For example, if users updated their profiles on the site, their contacts could access the new information through the app.

In July of 2009, the company announced that it would start charging users for a central feature of the app for Microsoft Outlook. Their announcement stated that non-paying users would no longer receive broadcasts from their contacts in Outlook address books. In order to continue receiving these broadcasts, users needed to buy an annual subscription to the company's premium bundle for $60. By studying which users purchased a premium subscription to receive broadcasts in Outlook, I identify for whom these broadcasts were most valuable. Next, I develop an econometric model to relate users' premium purchase decisions to properties of their social networks.

**Modeling the Natural Experiment**

To examine how properties of users' networks influenced their likelihood of paying to receive broadcasts, I construct a regression triple-difference model (Angrist and Pischke 2009, p. 242-243). While difference-in-differences models let researchers measure the average effect of nonrandom treatments, triple-difference models measure how treatment effects vary with observable characteristics off the treated population (e.g. Goldfarb and Tucker 2011). In my research setting, only users of the app for Microsoft Outlook needed to pay to receive broadcasts in their address books. Users of the app for alternative clients, like Mac Mail or Mozilla Thunderbird, continued to receive broadcasts in their address books for free. Therefore, I consider the treated population to be the subset of users that installed the app for Outlook and the control population to be the subset of users that installed the app for clients other than Outlook. I define the following indicator to distinguish between these two groups of users:

$$o_i = \begin{cases} 
0 & \text{if } i \text{ installed the app for an alternative client} \\
1 & \text{if } i \text{ installed the app for Microsoft Outlook} 
\end{cases}$$

I argue that two main factors underlie a user's likelihood of paying to receive broadcasts in Outlook: 1) her preference for Outlook, and 2) the value of receiving broadcasts about her network. Users with a low preference for Outlook could use the company's site or an alternative email client to get broadcasts for free, and so were unlikely to pay to receive broadcasts in Outlook. Similarly, users that did not value receiving broadcasts were unlikely to pay even if they had a strong preference for Outlook. Thus, once users needed to buy a premium subscription to get broadcasts in Outlook, the probability of buying premium was not just a function of these two factors alone but of their interaction as well. I start by modeling user i's probability of buying premium as a function of these two factors and their interaction:

$$\text{Prob}(p_i = 1) = \alpha_0 + B_0U_i + \alpha_1o_i + B_1o_iU_i + \epsilon_i$$  \hspace{1cm} (1)

Here, $p_i = 1$ if i purchased a premium subscription in the period after monetization and 0 otherwise, $U_i$ is a vector of network properties and control variables measuring how much i valued receiving broadcasts, and $\epsilon_i$ captures unobserved factors that influenced i's premium purchase likelihood.

An assumption implicit in this model is that the two factors described above would not have influenced premium purchase had receiving broadcasts in Outlook not been monetized. If, for example, these factors influenced the value of another tool in the premium bundle, the model would detect an effect that had little to do with the value users get from receiving broadcasts. Prior to the announced feature change, a premium subscription included tools that helped users manage their email address books, like automatic backup and recovery, removal of duplicate entries, and 24 hour customer service. These tools remained part of the premium bundle after the change. To control for effects that existed before the monetization of Outlook broadcasts, I define the following variable:
Social Media Broadcasts

I control for the effects of users’ Outlook preferences, network properties, and control variables that were unrelated to the value of receiving Outlook broadcasts by allowing these effects to vary by period.

\[
Prob(p_{it} = 1) = \alpha_0 + B_{00}U_{it} + \alpha_{10}O_{it} + B_{10}tU_{it} + \alpha_{01}t + B_{01}tU_{it} + \alpha_{11}tO_{it} + B_{11}tO_{it}U_{it} + \epsilon_{it}
\]  

Here, I define \( p_{it} \), \( U_{it} \), and \( \epsilon_{it} \) as before, but for each period \( t \in \{0,1\} \). Unlike the model specified earlier, this specification considers two observations for each user; one before monetization (\( p_{i0} \)) and one after (\( p_{i1} \)). In the final specification, I exploit having two observations per user by adding user-specific fixed effects. This final specification allows for a different baseline purchase rate for every user:

\[
Prob(p_{it} = 1) = \sum \alpha_i + B_{00}U_{it} + B_{10}tU_{it} + \alpha_{01}t + B_{01}tU_{it} + \alpha_{11}tO_{it} + B_{11}tO_{it}U_{it} + \epsilon_{it}
\]

By adding user fixed effects, I control for user-specific unobservables whose effect remained constant across the two time periods. Next, I describe the data I use to estimate the models outlined in this section.

Data

To estimate the effects of network properties on premium purchase, I compute a set of network measures for each user by analyzing her connections on the social network site. I fix the network of connections at two points in time: January 1, 2009 and July 1, 2009. I use the January network to predict users’ premium purchase decisions between January 1 – June 30, 2009 (\( t = 0 \)), and the July network to predict their purchases between July 1 – December 31, 2009 (\( t = 1 \)). I only consider the decisions of users who, at both time points, had at least two connections that were not connected to each other. This requirement ensures that all of the network measures were well-defined at both time points. Additionally, I restrict my analyses to the purchase decisions of users that installed the app for at least one email client. This ensures that both the treatment and control groups were receiving broadcasts in their email clients when broadcasts were monetized for the treatment group.

In total, 665,448 users satisfied both requirements. To calculate the networks measures for each of these users, I analyzed the connections they formed and the connections their connections formed on the site. This broader network consisted of 6,896,780 users and 38,886,367 connections. In addition to the network measures, I include several control variables when estimating my models. Table 1 describes these control variables. In the fixed effects model, these control variables cannot be identified and were dropped from the regression. However, any interactions between the controls and period dummy \( t \) are identified, and measure the incremental effects of the controls in the period when broadcasts were monetized for the treatment group.

<table>
<thead>
<tr>
<th>Variable (( X_i ))</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (( X_a ))</td>
<td>Indicator equal to 1 for females users</td>
</tr>
<tr>
<td>Email Contacts (( X_2 ))</td>
<td>Number of contacts in user’s email address book</td>
</tr>
<tr>
<td>Email Contact Users (( X_3 ))</td>
<td>Number of contacts in user’s email address book who are users</td>
</tr>
<tr>
<td>Days Since Joined (( X_4 ))</td>
<td>Number of days since user joined the service</td>
</tr>
<tr>
<td>Used Premium Before (( X_5 ))</td>
<td>Indicator equal to 1 if user purchased premium in previous year</td>
</tr>
</tbody>
</table>
Table 2 provides summary statistics of the network measures and control variables by user group. The treatment group, which needed to start paying to receive broadcasts, consisted of 481,354 users – about 72% of my sample. The remaining 184,094 users make up my control group, which did not need to pay to receive broadcasts. The network measures of these users vary somewhat across the two groups, which is to be expected because users were not randomized into them. For this reason, I estimate models that control for differences in purchase that were present before broadcasts were monetized for the treatment group.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Treatment Users of Outlook App</th>
<th>Treatment Users of Other Apps</th>
<th>Control Users of Outlook App</th>
<th>Control Users of Other Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering Coefficient (D₁)</td>
<td>.114</td>
<td>.136</td>
<td>.144</td>
<td>.154</td>
</tr>
<tr>
<td>Local Betweenness (D₂)</td>
<td>.863</td>
<td>.166</td>
<td>.831</td>
<td>.192</td>
</tr>
<tr>
<td>Family-Only Ties (S₁)</td>
<td>1.07</td>
<td>1.94</td>
<td>1.02</td>
<td>1.99</td>
</tr>
<tr>
<td>Friend-Only Ties (S₂)</td>
<td>5.96</td>
<td>10.4</td>
<td>6.53</td>
<td>14.4</td>
</tr>
<tr>
<td>Business-Only Ties (S₃)</td>
<td>25.0</td>
<td>64.1</td>
<td>22.5</td>
<td>63.8</td>
</tr>
<tr>
<td>Business Friend Ties (S₄)</td>
<td>3.91</td>
<td>8.30</td>
<td>3.78</td>
<td>10.3</td>
</tr>
<tr>
<td>Gender (X₁)</td>
<td>.265</td>
<td>.441</td>
<td>.323</td>
<td>.467</td>
</tr>
<tr>
<td>Email Contacts (X₂)</td>
<td>1190</td>
<td>2600</td>
<td>566</td>
<td>1560</td>
</tr>
<tr>
<td>Email Contact Users (X₃)</td>
<td>115</td>
<td>242</td>
<td>74.1</td>
<td>175</td>
</tr>
<tr>
<td>Days Since Joined (X₄)</td>
<td>1380</td>
<td>546</td>
<td>897</td>
<td>624</td>
</tr>
<tr>
<td>Premium User Before (X₅)</td>
<td>.030</td>
<td>.171</td>
<td>.010</td>
<td>.089</td>
</tr>
<tr>
<td>Number of Users</td>
<td>481,354</td>
<td>184,094</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To motivate my results and to illustrate the triple-difference empirical strategy, I plot the relationships between users’ network measures and premium purchase rates. Figures 3 and 4 show how users’ purchase rates varied by their clustering coefficients and local betweenness, respectively. The right panels depict purchase rates after broadcasts were monetized, and the interactions between user group and the network measures are unmistakable. For users in the treatment group, purchase rates decreased with the clustering coefficient and increased with local betweenness. For users in the control group, purchase rates hardly varied with these measures. The left panels depict purchase rates in the period before broadcasts were monetized for the treatment group. While interactions may be apparent here as well, they are not nearly as strong.

Figure 5 shows the relationship between users’ purchase rates and the number of ties they articulated through the service. The lower-right panel is the focus of the plot, and depicts the purchase rates of users in the treatment group after they needed to start paying to receive broadcasts. Family ties increased purchase rates more than friend ties, which in turn increased purchase rates more than business ties. Ties that were both friend and business increased purchase rates by more than ties that were either only friend or business. These patterns are weaker in the other panels, which either depict purchase rates before monetization (left panels), or of users in the control group (top panels). Thus, these plots strongly suggest that broadcasts are valuable for maintaining ties that are either strong or diverse. Next, I describe the results from estimating the regression models developed in this section.
Figure 3: Purchase rates by the clustering coefficient

Figure 4: Purchase rates by local betweenness

Figure 5: Purchase rates by tie type
Results

To measure how users’ network properties influenced their likelihood of paying to receive broadcasts, I estimate the regressions developed in the previous sections with a linear probability model. Linear probability models generate unbiased estimates of the marginal effects of the variables of interest and, unlike probit and logit models, are flexible enough to accommodate for user fixed effects (Chamberlain 1980). These models have heteroskedastic error terms by construction, and I use robust standard errors in my estimations to overcome this issue (Wooldridge 2002, p. 454). Table 3 displays the parameter estimates from these models. Each estimate is interpreted as the effect of a unit increase in the variable on the difference in purchase rates between users in the treatment and control groups. Models 2 and 3 control for any effects on purchase rates that were also present in the period before monetization, and Model 3 includes user fixed effects.

Estimates from Model 1 show that the clustering coefficient ($D_1$) had a significant negative effect on purchase, while local betweenness ($D_2$) had a significant positive effect. The number of family ($S_1$), friend ($S_2$), business ($S_3$), and business friend ($S_4$) ties all had significant positive effects on purchase. To test how the effects of ties varied by type, I compared the coefficients in a series of Wald tests. These tests show that an additional family tie increased purchase rates more than an additional friend tie ($p < .001$), which in turn increased purchase more than an additional business tie ($p < .001$). The effect of additional business friends was larger than that of ties that were only business ($p < .001$), though was statistically indistinguishable from the effect of ties that were only friend ($p = .55$).

Estimates from Models 2 and 3 generally support those of Model 1. The clustering coefficient had a significant negative on purchase, while local betweenness had a significant positive effect. Business friends did not have a statistically significant effect on purchase under Model 2, though they regained their significance when adding user fixed effects in Model 3. Ties of all other types had significant and positive effects on purchase under both specifications. Wald tests reveal that family ties increased purchase rates more than friend ties ($p < .001$ for both models), which increased purchase rates more than business ties ($p < .001$ for both models). Under Model 2, the effect of business friends on purchase was statistically indistinguishable from that of friend ($p = .42$) and business ties ($p = .35$). Under Model 3, business friends increased purchase rates more than business ties ($p < .001$), though their effect was indistinguishable from that of friends ($p = .26$).

I thus find strong support for the influence of tie strength and structural diversity on users’ likelihoods of paying to receive broadcasts. Hypotheses 1a, 1b, 2a, and 2b were supported by all three models. There was also substantial support for Hypothesis 1d, since business friends increased purchase rates more than business ties in Models 1 and 3. While Models 1 and 3 also show directional support for Hypothesis 1c, none of the models could reject the hypothesis that business friends and friends had equal effects on purchase. Table 3 also sheds light on how the likelihood of paying for broadcasts varied with the control variables. Across all three specifications, females had significantly lower purchase rates than males. Furthermore, users’ purchase rates increased with the number of days since they started using the service. Finally, there was some indication from Models 1 and 2 that past premium purchase predicts future premium purchase, though these effects disappeared when I included user fixed effects.

Discussion

In this study, I demonstrate that social media broadcasts are particularly valuable for maintaining relationships with ties that are strong or structurally diverse. While social media platforms like Facebook already prioritize information from strong ties by showing users broadcasts from those with whom they interact frequently, there is no indication that they take structural diversity into account. I argue that broadcasts are valuable for maintaining relationships with structurally diverse ties because users are more likely to lose touch with these ties, and, when they do lose touch, are less likely to obtain information about these ties from their other contacts.

These findings reveal an underlying tension that can arise when platforms prioritize information based on tie strength alone. While users, on average, prefer information from strong ties, these ties tend to be less structurally diverse (Granovetter 1973, Aral and Van Alstyne 2011). In this sense, information about strong ties can easily be obtained through means other than social media broadcasts. By abstracting
Table 3  Regression Models of Premium Purchase

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering Coefficient (D₁)</td>
<td>-.019***</td>
<td>-.013***</td>
<td>-.015***</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.002)</td>
<td>(.002)</td>
</tr>
<tr>
<td>Local Betweenness (D₂)</td>
<td>.013***</td>
<td>.0096***</td>
<td>.0091***</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.002)</td>
<td>(.002)</td>
</tr>
<tr>
<td>Family-Only Ties (S₁)</td>
<td>.0023***</td>
<td>.0015***</td>
<td>.0018***</td>
</tr>
<tr>
<td></td>
<td>(.0002)</td>
<td>(.0002)</td>
<td>(.0002)</td>
</tr>
<tr>
<td>Friend-Only Ties (S₂)</td>
<td>.00043***</td>
<td>.00025***</td>
<td>.00030***</td>
</tr>
<tr>
<td></td>
<td>(.00004)</td>
<td>(.00006)</td>
<td>(.00006)</td>
</tr>
<tr>
<td>Business-Only Ties (S₃)</td>
<td>.000096**</td>
<td>.000049**</td>
<td>.000064***</td>
</tr>
<tr>
<td></td>
<td>(.00001)</td>
<td>(.00001)</td>
<td>(.00001)</td>
</tr>
<tr>
<td>Business Friend Ties (S₄)</td>
<td>-.0085***</td>
<td>-.0065***</td>
<td>-.0072***</td>
</tr>
<tr>
<td></td>
<td>(.00007)</td>
<td>(.0001)</td>
<td>(.0001)</td>
</tr>
<tr>
<td>Gender (X₁)</td>
<td>-.00054</td>
<td>.00014</td>
<td>.00044***</td>
</tr>
<tr>
<td></td>
<td>(.0006)</td>
<td>(.0006)</td>
<td>(.0006)</td>
</tr>
<tr>
<td>Email Contacts (X₂)</td>
<td>3.6 x 10⁻⁷</td>
<td>2.9 x 10⁻⁷</td>
<td>4.0 x 10⁻⁷</td>
</tr>
<tr>
<td></td>
<td>(2 x 10⁻⁷)</td>
<td>(4 x 10⁻⁷)</td>
<td>(3 x 10⁻⁷)</td>
</tr>
<tr>
<td>Email Contact Users (X₃)</td>
<td>-2.3 x 10⁻⁶</td>
<td>7.1 x 10⁻⁸</td>
<td>-4.9 x 10⁻⁸</td>
</tr>
<tr>
<td></td>
<td>(3 x 10⁻⁶)</td>
<td>(4 x 10⁻⁶)</td>
<td>(4 x 10⁻⁶)</td>
</tr>
<tr>
<td>Days Since Joined (X₄)</td>
<td>7.5 x 10⁻⁶***</td>
<td>5.5 x 10⁻⁶***</td>
<td>7.0 x 10⁻⁶***</td>
</tr>
<tr>
<td></td>
<td>(5 x 10⁻⁷)</td>
<td>(6 x 10⁻⁷)</td>
<td>(6 x 10⁻⁷)</td>
</tr>
<tr>
<td>Used Premium Before (X₅)</td>
<td>.029***</td>
<td>.024***</td>
<td>.013</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.005)</td>
<td>(.01)</td>
</tr>
<tr>
<td>Before Period Controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>User Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>.02</td>
<td>.02</td>
<td>.59</td>
</tr>
<tr>
<td>Observations</td>
<td>665,448</td>
<td>1,330,896</td>
<td>1,330,896</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; *** p < 0.001. Standard errors are in parentheses.
away from tie strength and focusing on structural diversity, I show that broadcasts are valuable for maintaining relationships with ties that are harder to reach. By analyzing both the tie and network level, I bridge the gap between research which shows that social media are primarily used to interact with strong ties (e.g. Bakshy et al. 2012, Burke et al. 2013), and that which argues that social media are most valuable for maintaining weak, diverse ties (e.g. Donath and boyd 2004, Ellison et al. 2007).

If platforms do not include measures of structural diversity in the algorithms they use to determine which broadcasts to display, they could be doing their users a disservice. To see this, consider an academic researcher who wants to analyze data from a particular company. Would a broadcast stating that the researcher’s cousin started working for that company be more or less valuable than a similar broadcast from a long lost friend (or colleague)? On the one hand, the researcher’s cousin, presumably a stronger tie, may be more motivated to help the researcher obtain the data than a long lost tie. On the other hand, the researcher could come across information about her cousin though other means, like through a parent or at a family gathering. However, without the broadcast, it is likely that the researcher would not have come across this valuable piece of information about her long lost tie. Structural diversity captures whether information about one’s contacts is readily available or hard to come by.

This study has a number of limitations that are worth discussing. For one, the type of information that comprises a broadcast in my setting was of a particular nature – it was primarily contact information, such as mailing addresses, email addresses, phone numbers, and places of work. The broadcasts found on social media vary substantially more than this, and future research should investigate the extent to which the results found here generalize to other types of information. For example, the argument underlying my findings about structural diversity hinges on information about one’s contacts not being available through other means. Thus, the more private a piece of information, the more likely the argument is to hold. For information that is publicly available, such as news articles, structural diversity may not be at play.

Another limitation has to do with the type of control variables that were available in the database I analyze. For example, personal income may influence the likelihood of purchasing the premium bundle, and may be correlated to some of the network measures I use. While data on users’ personal income was not available, the natural experiment helps control for this to some extent. Since I look at the effect of the network measures on the difference in purchase probability before and after broadcasts were monetized, my identifying assumption is that income did not affect purchase in the period after monetization any more than it did in the period before. A second potential confound could also arise if some users were required by their organizations to purchase a premium subscription. Again, this would only be a confound if the requirement were present in the period after monetization but not before.

The results of this research suggest that individuals in structurally diverse networks find it more difficult to stay in touch with their contacts. This could explain recent findings, which show that people in structurally diverse networks adopt social media platforms like Facebook earlier than people in networks that are less diverse (Ugander et al. 2012). Managers should thus consider the specific benefits of a technology (e.g. receiving broadcasts) when trying to stimulate its diffusion. If individuals in structurally diverse networks find it more difficult to get information about the people around them, then network diversity is likely associated with certain psychological mechanisms – in this case, relational information deficits – that drive user behavior. Future research should explore the theoretical relationships between network structure and psychology.
References Cited


Wellman, Barry and Scot Wortley (1990), "Different Strokes from Different Folks: Community Ties and Social Support," American Journal of Sociology, 96(3), 558-588.