Distilling Brand Alliance Opportunities from Information Networks

Pankhuri Malhotra[1] Daniel Ringel [2] Keran Zhao[3]

pmal@ou.edu dmr@unc.edu krzhao@psu.edu

Yaxin Cui[4]

yaxincui2023@u.northwestern.edu

[1]University of Oklahoma
[2]University of North Carolina at Chapel Hill
[3]Pennsylvania State University
[4]Northwestern University

Abstract

Brand alliances are strategic partnerships between brands to create value and mutual benefit. Identifying brands to partner with is difficult. A promising avenue is social networks, a type of information network where consumers' perspectives could guide the discovery of brand alliance opportunities. Because consumer-generated content is noisy, it can suggest opportunities that lack substance. To mitigate the risk of being misled, the authors propose a four-phase market research approach called BANE–Brand Alliance Network Exploitation–that distills promising opportunities from social networks. The four phases of BANE identify, validate, qualify, and explain brand alliance opportunities. Applied to 450 brands, BANE considers co-mentions by both consumers and brands to identify an opportunity space. The space is then validated against theoretically-driven indicators of brand compatibility using Exponential Random Graph Models. Systematic analysis of consumer engagement qualifies opportunities further and offers a differentiated perspective on their strategic value. By examining opportunities' underlying social media content, BANE's final phase provides guidance for their implementation. Empirical analysis reveals that 92% of the opportunities suggested by the information network are either incompatible or lack substance. Findings are validated internally and externally from theory, firm, and consumer perspectives.

Keywords: Brand Alliances, Social Network Analysis, Exponential Random Graph Models.

INTRODUCTION

Brand alliances are strategic partnerships between brands to create value and mutual benefit (Swaminathan and Moorman 2009). Finding an alliance partner is a challenging task for brand managers due to the varied and dynamic nature of consumer needs and perceptions. A promising avenue for finding brand alliance partners are information networks. Formed by consumer interactions on social media and in online forums, these networks can reveal meaningful structures among brands, products, and consumers (Sundararajan et al. 2013; Oestreicher-Singer et al. 2013). Previous marketing research has successfully used brands' co-occurrence in information networks to uncover market structure and competition (Netzer et al. 2012; Kim, Albuquerque, and Bronnenberg 2011; Ringel 2023b; Yang, Zhang, and Kannan 2022).

Information networks can be formed through implicit or explicit consumer activity (Zhang, Bhattacharyya, and Ram 2016). Implicit networks arise when a user independently likes or comments on different brands, thereby creating a latent connection between those brands through common consumer activity. Explicit networks, on the other hand, arise from deliberate mentions of two brands in the same context at the same time, indicating a more conscious association. While implicit information networks, based on liking and commenting, are effective for discovering market structures as demonstrated by Yang, Zhang, and Kannan (2022), their application in the context of brand alliances is limited. Identifying and explaining brand alliances requires understanding both the structure and content of interactions.

Structure identifies alliance candidates and serves as the starting point for all further analysis. To make informed decisions, brand managers need deeper insights into the nature of a brand alliance opportunity, including what it entails and how it is perceived by consumers. This calls for the analysis of network content, not just network structure. The aim of this research is to identify and explain the most promising brand alliance opportunities from explicit information networks by considering both structure and content. For the purpose of this study, opportunities are considered promising if they align with theoretically-driven indicators of favorable brand alliances and also demonstrate sufficient salience, engagement, and positive sentiment from consumers.

Such undertaking is, however, challenging for two reasons. First, information networks are created from unstructured user-generated content (UGC) that needs to be collected, processed, and distilled to separate signals from noise. Second, while there has been a growing interest in using information networks to solve marketing problems, little is known about the validity of their structural elements for brand analysis. Lovett, Peres, and Shachar (2014) emphasize the significance of examining the nature of information networks and establishing a connection to theoretically-founded brand attributes such as mutual brand values and shared brand personalities (Aaker 1997). Yet, previous work on brand alliance opportunities from information network falls short of establishing structural validity.

We propose a new market research approach called BANE (Brand Alliance Network Exploitation), which not only identifies and explains promising brand alliance opportunities from an information network, but also empirically validates the information network's structure for brand analysis. BANE comprises four consecutive phases. Phase 1 identifies a brand network from brand co-mentions on social media. Phase 2 validates the brand network's structure using exponential random graph modeling. Phase 3 qualifies brand alliance opportunities suggested by the network using salience, valence, and support. Phase 4 explains the most promising opportunities based on social media content (i.e., posts).

The advantages of BANE over previous research approaches for identifying brand alliance opportunities from information networks (Malhotra and Bhattacharyya 2022) are fourfold. First, we identify the opportunity space for brand alliances by constructing a brand network from brands' co-mentions in consumers' posts. Previous research approaches rely on co-following or co-commenting on social media. In contrast to co-following or co-commenting, where brands may be followed or commented on for independent reasons at different times, co-mentions capture deliberate brand connections made by consumers at the same time.

Furthermore, co-mentions capture observable content, which is crucial for gaining deeper insights into how to execute the opportunity.

Second, we are the first to validate the identified brand network using theoretically driven indicators of favorable brand alliance partners. Previous work on brand alliance opportunities from information networks falls short of establishing structural validity. Using an exponential random graph model (ERGM), we examine the statistical relationships between our brand network's structure and brand personality attributes (Aaker 1997; van der Lans, Van den Bergh, and Dieleman 2014) to validate its appropriateness and relevance for identifying promising brand alliance opportunities. We then use the estimated ERGM model to predict the most probable brand alliances in the network based on the compatibility of brand attributes, thereby refining the opportunity space further.

Third, we qualify the brand alliance candidates from the opportunity space on three dimensions to distill the most promising ones from them. Previous approaches relied solely on salience (i.e., co-occurrence frequency). BANE additionally considers valence (the polarity of the posts co-mentioning brands), and support (consumer engagement with posts co-mentioning brands). By considering three dimensions (salience, valence, and support), we provide brand managers with a more nuanced perspective on the opportunity space and enable them to consciously balance these dimensions when selecting brand alliance partners.

Fourth, in contrast to previous approaches, BANE examines the social media content underlying the opportunities for substance to guide their implementation along the marketing mix. While one opportunity may be about a joint product, another opportunity might be about a joint promotion. For instance, bundling with *price* advantage could be seen in McDonald's and Coca-Cola's value meals, where various products are combined at a discounted price. Collaborating for joint *promotion*, Red Bull and GoPro created sports content featuring athletes using GoPro cameras while consuming Red Bull products. Selling one brand in the *place* of another is demonstrated by store-in-store concepts, such as CVS pharmacies within Target stores. Lastly, creating a joint *product* was highlighted when Nike launched a limited-edition Air Jordan Dior sneaker.

We apply BANE in an empirical study of 450 brands that were co-mentioned by 236,710 unique consumers in 477,933 posts on Twitter from January 2020 – December 2020. BANE identifies the top 2% most promising brand alliance opportunities from a candidate pool of 23,385 opportunities co-mentioned by social media users. Notably, over 92% of brand alliance opportunities suggested by the information network were eliminated across BANE's four phases. These findings emphasize the significance of leveraging not only the network structure but also the consumer content to uncover opportunities with the highest potential. We validate BANE's insights in the context of our empirical study both internally and externally from marketing theory, firm, and consumer perspectives.

RELATED LITERATURE

Simonin and Ruth (1998) define a brand alliance as a short-term or long-term collaboration between two or more distinct brands, products, or other unique proprietary assets. Such collaborative arrangements can take various forms, including joint sales promotions, bundled products, composite brand extensions, true product combinations, or collaborative marketing campaigns. Selecting the right partner for a brand alliance continues to remain a complex marketing problem.

Only few marketing studies examined the characteristics of individual partners (whether similar or dissimilar) to assess consumer evaluations towards a brand alliance (Simonin and Ruth 1998; Batra, Lenk, and Wedel 2010; Cao and Sorescu 2013; van der Lans, Van den Bergh, and Dieleman 2014). Those that examined partner characteristics primarily using traditional survey approaches. While surveys provide valuable insights into consumer brand perceptions, they are also subject to high recruiting costs, low response rates, extended time requirements, and self-reporting inaccuracies (Culotta and Cutler 2016). In response to the limitations of surveys, marketing researchers turned to information networks to study brand alliance opportunities (Kupfer et al. 2018; Malhotra and Bhattacharyya 2022).

Information networks can manifest implicitly or explicitly in consumer-brand engagement patterns on social media. Implicit brand information networks emerge from inferred relationships among brands. Here, relationships are inferred from, for example, aggregated preferences of a vast number of social media users. Malhotra and Bhattacharyya (2022) and Culotta and Cutler (2016) identify an implicit brand information network from brands' shared followers on Twitter. Culotta and Cutler (2016) use the implicit brand information network for mining brand attribute perceptions while Malhotra and Bhattacharyya (2022) use it for identifying brand alliance opportunities. Similarly, Yang, Zhang, and Kannan (2022) employ an implicit brand information network, derived from co-commenting and likes, to uncover market structures.

Explicit brand information networks, on the other hand, capture deliberate brand connections made by consumers within the specific context of their posts at the same time. These networks serve as a proxy for consumers' perceptual space and top-of-mind associations (Netzer et al. 2012). In explicit brand information networks, nodes represent brands, and edges between nodes capture brand co-mentions made by users in their posts at the same time. This temporal aspect distinguishes them from implicit networks, where a consumer can like brands at separate points in time. Therefore, connections in implicit networks may not provide adequate information for explaining brand alliances, which is a primary goal of this study—to go beyond prediction and explain how to execute opportunities.

This research falls into the literature stream on explicit brand information networks. As such, it differs to (Malhotra and Bhattacharyya 2022), who rely on co-followership data, where brands might be followed for independent reasons such that identified implicit network structure is subject to considerable noise. Our work also differs in objective and approach from that of Yang, Zhang, and Kannan (2022), who identify brand similarity implicitly by training an auto-encoder (neural network) on consumers' independent commenting and liking of brands to discover market structure.

More closely related to this study is Kupfer et al. (2018), who study actor-movie alliances in which the movies function as composite products and actors represent partner brands. They examine how a partner brand's social media presence (i.e., actors, in their case) relates to the success of the composite product (i.e., movies). Unlike Kupfer et al. (2018) whose focus is on preexisting alliances in cinema, our work is forward-looking, that is, we aim to discover and explain new brand alliances opportunities across a broad spectrum of B2C categories.

More importantly, the identification of brand alliance opportunities is only the starting point of our market research approach. We develop BANE, a four-phase market research approach that helps marketers distill the most promising brand alliance opportunities from consumers' social media posts. In contrast to previous work on brand alliance opportunities that leverage information networks, our BANE considers both network structure and network content to deliver additional insights to brand managers. Notably, past studies on brand alliance opportunities are silent on how identified opportunities might be executed. BANE addresses this gap by revealing "what" opportunities are about from themes (or topics) in the underlying content, and "how" they might be executed along the marketing mix.

Finally, we validate the identified brand alliance opportunity space using Exponential Random Graph Models (ERGMs). There has been a growing interest in using information networks to solve marketing problems. However, little is known about the validity of their structural elements for brand analysis. Lovett, Peres, and Shachar (2014) emphasize the significance of examining the nature of information networks and establishing a connection to theoretically-founded brand attributes such as mutual brand values and shared brand personalities. In this study, ERGMs relate the core network components of the opportunity space to established indicators of a compatible brand alliance opportunity. ERGMs have been adopted by a growing number of studies in related research fields, such as information systems to study business proximity in inter-firm networks (Shi, Lee, and Whinston 2020), healthcare disparity in tele-consultation networks (Hwang et al. 2022), and value creation in social networks (Goh, Gao, and Agarwal 2016). Table 1 provides an overview of research related to our work and highlights our contribution to the extant literature.

Table 1: Overview of Related Marketing Studies

Study	Objective	Data	Implicit vs Explicit Network	Network information used: Structure vs Content	Validation	Explanation
Netzer et al. (2012)	Market structures insights using NLP	Posts on online forums	Explicit	Structure and Content	Purchase data, survey	-
Nam, Joshi, and Kannan (2017)	Analyze user-generated tags for marketing research	Brand social tags	Explicit	Content	Brand Concept Map (Survey)	-
Gabel, Guhl, and Klapper (2019)	Market structure insights using NLP and ML	Shopping basket	Explicit	Structure	Simulation	-
Yang, Zhang, and Kannan (2022)	Market structures insights using deep learning	Facebook fan pages	Implicit	Structure	Survey, Google search trends	-
Malhotra and Bhattacharyya (2022)	Exploring Brand alliances	Consumer followership data on Twitter	Implicit	Structure	Survey	-
Ringel (2023b)	Multimarket membership of products in market structure maps	Clickstream data	Implicit	Structure	Simulation	-
This study	BANE's 4 Phases for distilling and explaining brand alliance opportunities	Posts by consumers and brands	Explicit	Structure and Content	ERGM, Survey, Past data on real-world brand alliances	Marketing-Mix

BRAND ALLIANCE NETWORK EXPLOITATION

BANE (Brand Alliance Network Exploitation) comprises four sequential phases: Identification, Validation, Qualification, and Explanation (see Figure 1). BANE casts a wide net to capture a comprehensive set of potential brand alliances from explicit co-mentions of brands on social media. We apply BANE to a large set of brands to demonstrate its implementation and validate the insights it generates. This validation allows brand managers to confidently apply BANE to explore brand alliance opportunities around a single brand using an egocentric network, providing tailored insights specific to that brand. The following subsections describe the specific mechanics and rationale that underpin each of BANE's four phases.

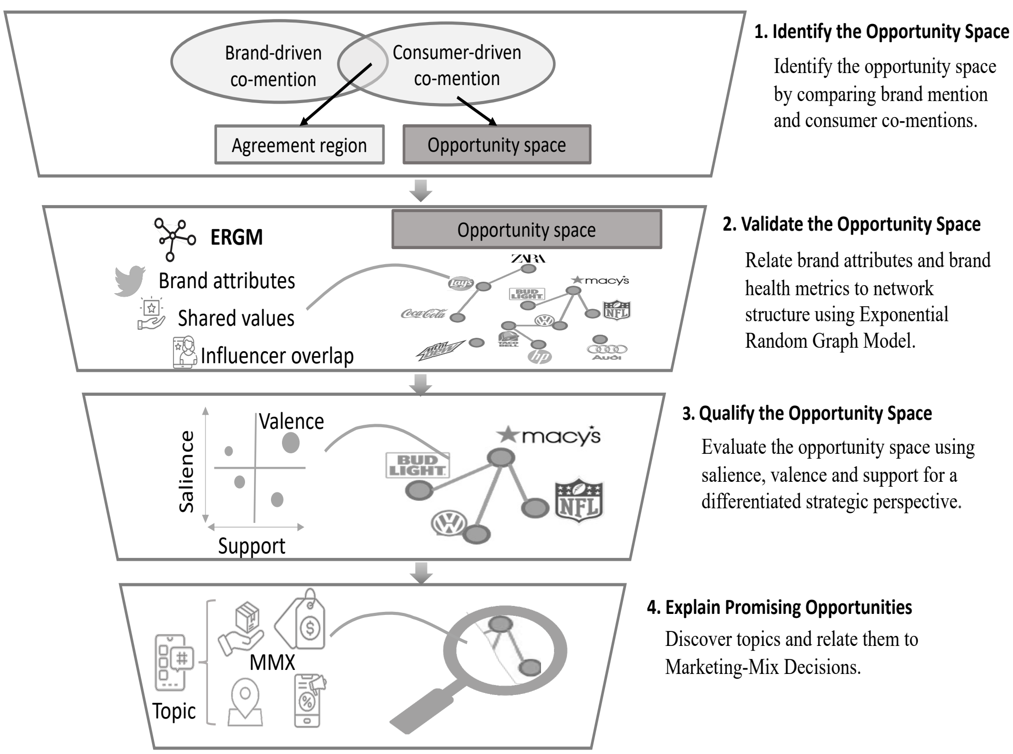


Figure 1: BANE's Four Phases

Phase 1: Identify the Opportunity Space

BANE utilizes brand co-mentions on social media to discover promising brand alliance opportunities. The objective of BANE's first phase is to identify an opportunity space for brand alliances from consumers' perspectives using network analysis techniques. This opportunity space is then passed to Phase 2 for validation and further refinement.

We begin by constructing an explicit brand information network, where nodes represent brands and edges denote explicit brand co-mentions on social media. Notably, a brand information network can be constructed from both the consumer perspective (based on explicit co-mentions by consumers) and the brand perspective (based on explicit co-mentions by brands). These two perspectives may not coincide and are shown in Phase 1 of Figure 1. Consumers might co-mention brands that the brands themselves do not, and vice versa. Because our aim is to discover brand alliance opportunities, we focus on the consumer perspective net of the brand perspective. That is, we only consider brand co-mentions that brands don't already promote in their own posts, labeled as the "opportunity space" in Phase 1 of Figure 1. Additionally, the area labeled "agreement region" identifies brand connections where consumer and brand perspectives align.

We define the "consumer-driven network" as an explicit information network based on brand co-mentions by consumers. The entailing information network provides us with a lens on consumers' perceptual brand space. The more frequently consumers mention two brands together, the closer those brands are in their perceptual space (Netzer et al. 2012), and the stronger the edge connecting the brands in our information network. The concept of using co-occurrence (here, co-mentions) has its roots in associative memory literature (Anderson and Bower 1974). Netzer et al. (2012) note that when a consumer consciously mentions two brands together, it highlights a sense of proximity or relationship in their mind. In the empirical application of our new market research approach, we also build implicit brand networks from co-followership and co-commenting data and discuss how their application is limited in the context of identifying and explaining brand alliances.

Mathematically, the consumer-driven network is defined as:

$$G_C = (V_C, E_C, W_C) \tag{1}$$

Where: G_C is the consumer-driven graph, V_C is the set of nodes, where each node in V_C represents a brand in our analysis, E_C is the set of edges, where an edge (u_C, v_C) exists if a consumer mentions both brands u_C and v_C concurrently in a single post and $W_C : E_C \to \mathbb{R}^+$ is a weight function that assigns a positive real number to each edge in E_C, indicating the number of consumers who mentioned both brands together. It's important to note that popular brands may receive more mentions from consumers, leading to higher weights in the network. To address this popularity bias, we normalize these edge weights using the Jaccard coefficient, as described in (Culotta and Cutler 2016). Edges are thus calculated as the total number of consumer-driven co-mentions divided by the sum of their total co-mentions with other brands.

Next, we define the "brand-driven network" from posts by brands that mention other brands. Again, nodes represent brands. In contrast to the consumer-driven network, edges in the brand-driven network are formed when brands are mentioned by other brands, not through co-mentions by consumers. Mathematically, the brand-driven network is defined as -

$$G_B = (V_B, E_B, W_B) \tag{2}$$

where G_B is the brand-driven graph, V_B is the set of nodes, where each node in V_B represents a brand, E_B is the set of edges, where an edge (u_B, v_B) exists if brand, u_B directly references another brand v_B in its official social media post and $W_B : E_B \to \mathbb{R}^+$ is a weight function that assigns a positive real number to each edge in E_B, indicating the number of times a brand referenced another brand.

To obtain a consumer perspective net of the brand perspective, we remove all edges from the consumer-driven network that also exist in the brand-driven network. What remains is a "pruned" consumer-driven network where only those edges (i.e., brand connections) remain that are unique to the consumer-driven network. The process is highlighted in Figure 2. All edges promoted by brands themselves in G_B were removed from G_C. This pruned consumer-driven network is the opportunity space, G_O, that serves as the foundation for all entailing analyses.

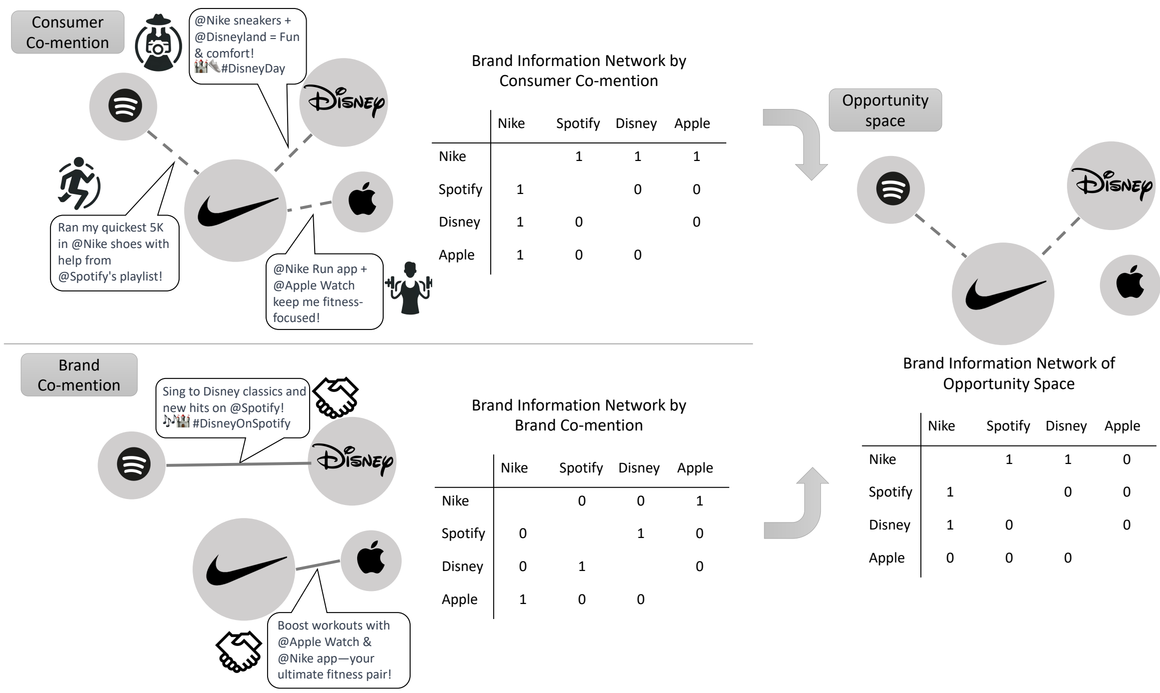


Figure 2: Construction of Opportunity Space

Finally, we further refine the identified opportunity space by removing edges between brands in the same category from the information network. Consumers may, for instance, compare the products of two brands such as McDonald's and Burger King in their posts and comments. Arguably, such product comparison is not indicative of a brand alliance opportunity. By removing same-category edges, we mitigate the risk of misleading signals into our analysis. We then pass the pruned and refined brand information network—the opportunity space—on to Phase 2.

Phase 2: Validate the Opportunity Space

To ensure that the opportunity space yields meaningful brand alliances, its structure should align with marketing theory for favorable brand alliances. The validation phase ensures that the identified brand alliances are compatible in terms of brand personality and theoretically motivated marketing variables, as previously studied in (Simonin and Ruth 1998; van der Lans, Van den Bergh, and Dieleman 2014). We validate the network structure of the opportunity space using Exponential Random Graph Models (ERGMs). The output of BANE's second phase is the most likely brand-brand connections based on variables—brand personality metrics, identity statements, and additional social media brand attributes. The most probable brand-brand connections are then passed to the third phase for further qualification.

ERGMs are a stochastic network modeling approach that allows researchers to statistically identify the underlying social features that explain the structure of an observed network (Krivitsky 2012; Lusher, Koskinen, and Robins 2013). They offer several modeling advantages over alternative statistical models, such as regression. One key advantage is their ability to incorporate interdependence among network ties, that is, edges between nodes (Hwang et al. 2022). In our network, nodes represent brands and edges represent co-mentions in consumers' posts. The edges in network data are not independent of each other as one brand may be a candidate for multiple dyads (here, alliance opportunities); this may violate the independent and identical distribution (IID) rule assumed in traditional regression models. However, ERGMs are explicitly designed to model such interdependence among edges (Hwang et al. 2022), making them a suitable choice for us to statistically analyze the identified opportunity space.

It is important to acknowledge that the role of Exponential Random Graph Models (ERGMs) is to validate the observed brand information network with respect to theoretically motivated variables for compatible brand alliances. They are not intended to draw any causal conclusions about user motivations for co-mentioning brands. Examining user motivations for co-mentioning brands on social media is beyond the scope of this research. Instead, ERGMs serve as a validation tool to ensure the network's appropriateness and relevance for identifying brand alliances.

Formally, an ERGM is expressed as follows:

$$P(\mathbf{Y} = \mathbf{y}) = \frac{\exp(\boldsymbol{\theta}^T g(\mathbf{y}))}{c(\boldsymbol{\theta})} \tag{3}$$

where $\mathbf{y}$ is our observed network and $\mathbf{Y}$ denotes all possible network realizations. The term $g(\mathbf{y})$ is a vector of network statistics responsible for the realization of $\mathbf{y}$. This includes model features like individual brand characteristics to explain the observed network structure. Here θ denotes the vector of unknown coefficients corresponding to $g(\mathbf{y})$, and is estimated using Markov Chain Monte Carlo maximum likelihood estimation (MCMC-MLE) procedures (Hwang et al. 2022). The denominator $c(\boldsymbol{\theta})$ is the normalizing constant calculated by summing up $\exp(\boldsymbol{\theta}^T g(\mathbf{y}))$ over all possible network realizations. We organize and define all variables included in our analysis in Table 2.

Table 2: Summary of ERGM Variables

Variables of Interest	**Description**
Brand Personality Metrics (Aaker 1997; van der Lans, Van den Bergh, and Dieleman 2014)	
Sincerity (Absdiff)	Measures the tendency of tie formation between any two brands subject to their absolute difference in brand *sincerity* attribute. A negative estimate would mean that brands with smaller difference in their sincerity scores tend to be more connected in the network.
Competence (Absdiff)	Measures the tendency of tie formation between any two brands subject to their absolute difference in brand *competence* attribute. A negative estimate would mean that brands with a smaller difference in their competence scores tend to be more connected in the network.
Ruggedness (Absdiff)	Measures the tendency of tie formation between any two brands subject to their absolute difference in brand *ruggedness* attribute. A negative estimate would mean that brands with a smaller difference in their ruggedness scores tend to be more connected in the network.
Sophistication (Absdiff)	Measures the tendency of tie formation between any two brands subject to their absolute difference in brand *sophistication* attribute. A negative estimate would mean that brands with a smaller difference in their sophistication scores tend to be more connected in the network.
Brand Health Metrics (Simonin and Ruth 1998)	
Familiarity (Absdiff)	Measures the tendency of tie formation between any two brands subject to their absolute difference in brand *familiarity* attribute.
Favorability (Absdiff)	Measures the tendency of tie formation between any two brands subject to their absolute difference in brand *favorability* attribute.
Brand Identity(Aaker 2012; Moreau et al. 2020)	
Brand Identity (Edgecov)	Measures the semantic similarity between the identity statements of two brands. A positive estimate would mean that brands with similar identity statements tend to be more connected in the network.

Summary of ERGM Variables

Control Variables	**Description**
Platform-related Controls	
Brand Popularity (Nodecov)	Measures the brand's popularity on social media, indicating the tendency of brands with varying levels of social media popularity to form connections in the network.
Number of Posts (Nodecov)	Measures the total number of posts by a brand, indicating the tendency of brands with varying levels of social media activity to form connections in the network.
Average Hashtags (Nodecov)	Measures the average number of hashtags used in a brand's posts, indicating the tendency of brands with varying levels of hashtag usage to form connections in the network.
Average Caption Length (Nodecov)	Measures the average number of words in a brand's posts, indicating the tendency of brands with varying content lengths to form connections in the network.
Average Engagement (Nodecov)	Measures the average level of engagement (e.g., likes, shares, comments) with a brand's posts, indicating the tendency of brands with varying levels of engagement to form connections in the network.
Influencer Overlap (Edgecov)	Measures the count of influencers who mention both brands together in their posts. We define influencers as individuals with a high number of followers on social media. In our empirical study, we established the threshold for influencers at 100K, based on the definitions of different tiers of influencers presented in Beichert et al. (2024).

Explanatory variables for ERGM

Our choice of explanatory variables for the ERGM is theoretically motivated by previous research on brand alliances and social networks. These variables include aspects related to brand personality and identity, health metrics like familiarity and favorability, as well as other online brand attributes. Table 2 provides a summary of the variables included in our ERGM. Few marketing studies previously explored the characteristics of individual partner brands to explain consumer attitudes towards an alliance (Simonin and Ruth 1998; Batra, Lenk, and Wedel 2010; van der Lans, Van den Bergh, and Dieleman 2014). These studies, primarily built on survey data, found that conceptual coherence in brand personality profiles is an important predictor of consumer attitude towards the brand alliance.

We include brand personality measures based on the framework established by Aaker (1997) and van der Lans, Van den Bergh, and Dieleman (2014) on partner selection in brand alliances. These dimensions include (1) Sincerity, characterized by attributes such as down-to-earth, real, and sincere; (2) Competence, characterized by attributes such as intelligence, reliability, and confidence; (3) Sophistication characterized by such as glamorous, upper-class and charming; and (4) Ruggedness characterized by attributes such as masculine, tough and outdoorsy. We use these brand personality dimensions in our analysis to validate the structure of the opportunity space.

Related marketing research on branding notes the importance of brand identity for building equity (Aaker 2012; Moreau et al. 2020). The brand identity reflects what a brand is (Lam et al. 2010). We use brand identity as a key variable in the ERGM model to determine how brands with similar values, as reflected in their brand identity statements, are connected in the opportunity space. Although brand identity has not been studied in the context of brand alliances previously, we posit that brands with similar values would create promising brand alliance opportunities that are beneficial for both parties involved.

When examining brand alliances, other factors considered include the familiarity and favorability ratings of brands (Simonin and Ruth 1998). We employ these variables in ERGM analysis to validate the overall network structure of the opportunity space. Finally, platform-related variables may explain connections between brands in the opportunity space. We control for brand popularity (measured using the number of followers), account age (the year when the brand account was established on social media), and content-based metrics (posting frequency, use of hashtags/caption length, and consumer engagement). Another factor driving the connections between brands could be their explicit co-mentions by influencers and/or celebrities. Generally, posts by influencers have a greater reach (Hughes, Swaminathan, and

Brooks 2019). This implies that when influencers co-mention brands in their posts, there is a greater potential to reach many consumers and shape their perceived relationships with those brands. Therefore, we include influencer overlap as an edge covariate in the model. In our analysis, we define influencers as users with more than 100K followers on social media also known as macro-influencers (Beichert et al. 2024).

It is important to note that ERGMs consider the global brand network structure to validate the opportunity space constructed from user co-mentions on social media. If brand managers choose to focus on an egocentric network and identify alliances around a single brand using user co-mentions, this phase can be skipped, as this research has already tested the relevance of the opportunity space for identifying and explaining brand alliances.

Phase 3: Qualify the Reduced Opportunity Space

BANE's qualification phase evaluates the brand alliance opportunities along three engagement metrics: salience, support, and valence. Its objective is threefold. First, to provide brand managers with a quantitative assessment of each opportunity using established social media metrics. Second, to systematically rank and compare all remaining opportunities. Third, to make brand managers aware of potential trade-offs that need to be considered when deciding for (or against) pursuing a particular brand alliance opportunity.

Salience, support, and valence provide multifaceted insights into each brand alliance opportunity. Salience is measured by the frequency of brand co-mentions, which serves as an indicator of a brand pair's prominence in consumers' minds. The marketing literature documents its positive relationship with sales (Dhar and Chang 2009) and stock market performance (Tirunillai and Tellis 2012). Salience for brand pair (A, B) is measured as:

$$Salience(A, B) = \text{Co-mentions}(A, B), \tag{4}$$

where $\text{Co-mentions}(A, B)$ = The total number of posts mentioning both brands A and B.

Support is quantified by the total number of likes and replies, normalized by the reposts received. Previous marketing studies highlight the importance of likes and replies in capturing consumer engagement (Valsesia, Proserpio, and Nunes 2020; Li and Xie 2020). Support for brand pair (A, B) is calculated by summing over the likes and replies for all co-mentions of A and B and normalizing it by the total number of reposts:

$$Support(A, B) = \frac{\sum_{t \in Co-mentions(A,B)} (Likes(t) + Replies(t))}{\sum_{t \in Co-mentions(A,B)} Reposts(t) + 1}, \tag{5}$$

where $Likes(t)$ = The number of likes for post t; $Replies(t)$ = The number of replies for post t; $Reposts(t)$ = The number of reposts for post t.

Finally, valence is measured through the polarity (or sentiment) of consumer posts. This aspect is crucial, as several studies have demonstrated the impact of sentiment on brand perception and customer decision-making (Chevalier and Mayzlin 2006; Baker, Donthu, and Kumar 2016). We estimate valence for each post using Vader (Hutto and Gilbert 2014), a widely used sentiment analysis tool in marketing research. Vader returns a polarity score between -1 (most extreme negative) and +1 (most extreme positive) for each post, indicating the sentiment orientation of the text. We take the average polarity score across all posts mentioning a given brand pair to get the valence value for that brand alliance opportunity. This valence score captures the overall sentiment towards the potential brand alliance opportunity, ranging from negative (-1) to positive (+1), with 0 being neutral.

We introduce the BANE Matrix (illustrated in Figure 3) to provide brand managers with a comprehensive perspective across all three dimensions (salience, support, and valence). It segments the reduced opportunity space from the previous phase into four quadrants, bounded at the median for salience (y-axis) and support (x-axis). Brand alliance opportunities (i.e., brand pairs) appear as bubbles inside the BANE Matrix that are ranked relative to one another within its dimensions. Brand alliance opportunities within each quadrant are further qualified by their valence (denoted by bubble size).

Despite the obvious appeal of opportunities that score high on all three dimensions of the BANE Matrix (i.e., large bubbles in its top right-hand corner), we contend that brand managers should also consider the varied potential in other quadrants. In what follows, we

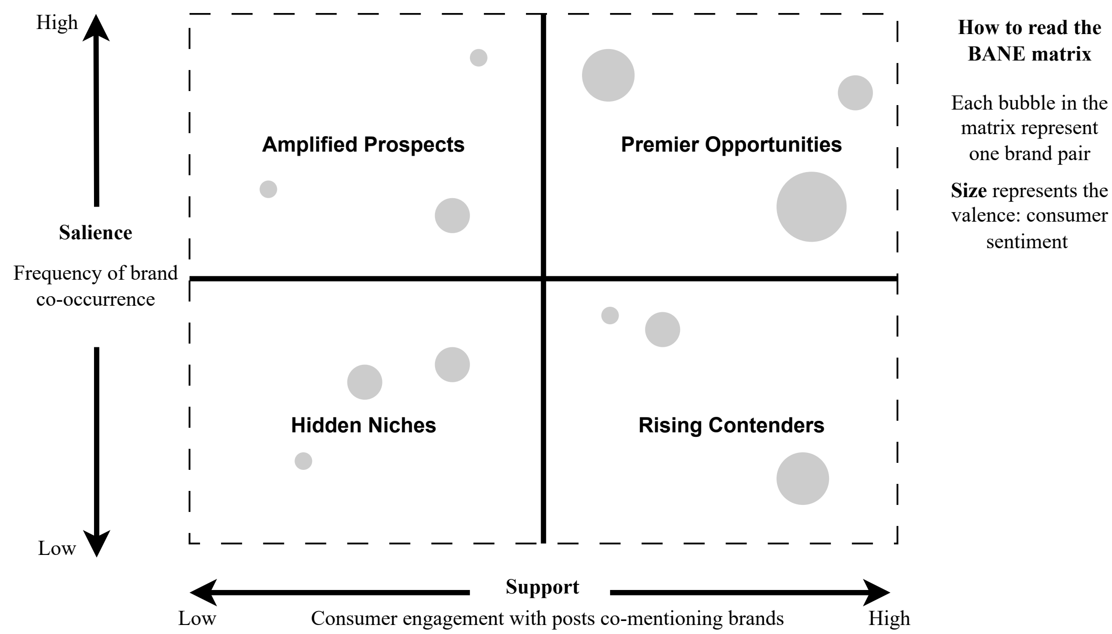


Figure 3: The BANE Matrix

name each quadrant based on its characteristics and discuss its potential for a brand alliance.

1. **High Salience, High Support: Premier Opportunities** These are the most promising brand alliance opportunities, characterized by high visibility and strong engagement. Premier opportunities represent well-established connections with significant potential for successful brand collaborations. Brand managers should focus on premier opportunities with high valence (positive sentiment) to co-create unique offerings that resonate with an already engaged and favorable consumer base.

2. **High Salience, Low Support: Amplified Prospects** While these opportunities are highly visible, they lack strong engagement. Brand managers can leverage existing awareness to actively prompt greater consumer engagement. Doing so amplifies partnering brands' consumer interactions and boosts their market presence. However, if valence is mixed or neutral, brand managers must incorporate additional actions designed to positively shift consumer sentiment.

3. **Low Salience, High Support: Rising Contenders** The high support despite low

salience in combination with high valance points to deep, authentic connections in consumers' minds. When this is the case, a brand alliance can offer unique engagement and growth possibilities despite being less obvious. Brand managers should capitalize on dedicated consumer bases to expand reach and establish a stronger presence in specialized markets.

4. **Low Salience, Low Support: Hidden Niches** Hidden niches can point to emerging trends or specialized markets. Although they require more effort to develop, they reveal unique, untapped market segments or innovative collaboration avenues. Importantly, brand alliances in this quadrant are largely exploratory, which makes their outcomes highly uncertain. It is important to understand whether these niches harbor positive consumer sentiment or not.

By exploring all four quadrants, brand managers can formulate a comprehensive strategy that leverages different market potentials. The BANE Matrix facilitates this process and leads brand managers to a reduced set of candidate opportunities that are examined in greater detail in BANE's final phase. Note that the BANE Matrix, while used in this research for a comprehensive analysis among brands, is also applicable to brand alliance opportunities around a single brand.

Phase 4: Explain Promising Opportunities

BANE's final phase examines the content of consumer posts that co-mention brands. Its objective is to explain what the most promising opportunities are thematically about (i.e., topics) and to which marketing mix component—product, price, place, or promotion—they pertain. Such detailed exploration is crucial as it informs brand managers not just about the existence of a promising opportunity, but also about its nature and the consumer interests it taps into. By understanding the specific topics within consumer discussions and their associated marketing mix component, brand managers gain valuable insights into how to construct and execute the brand alliance. Effectively, this phase lays the groundwork for strategic decisions, ensuring that any brand alliance resonates with the target audience's expectations.

Phase 4 comprises two steps. The first step determines which component of the marketing mix each post that co-mentions a focal brand pair pertains to. For example, a post like *"Love how my [Brand A lipstick] matches perfectly with my [Brand B handbag] #FashionMeetsBeauty #StyleCombo"* pertains to the product. In contrast, a post like "*Small steps for a big change: switched to [Brand A] for their tree-planting with each buy & loving [Brand B] for slashing CO2 footprints! #GoGreen #SustainableChoices*" pertains to promotion. The second step zooms in on the most prevalent marketing mix component(s) to discover what topics their associated posts are about. In the above examples, the first post is about enhancing one's overall style, while the second post is about brands' ESG efforts.

By concurrently examining the prevalent marketing mix components and the topics of associated posts, brand managers gain insights into both the 'how' and the 'what' of potential brand alliances. Such insights can support brand mangers in deciding which marketing mix component(s) to executing the alliance on (how), as well in creating a thematic focus that is likely to resonate with consumers (what). Importantly, such a deeper understanding of brand alliance opportunities also facilitates identifying brand alliance opportunities that align best the overall brand strategy.

Step one uses supervised machine learning to determine which marketing mix component each post of a brand alliance opportunity pertains to. In this study, we use a pre-trained marketing mix (MMX) classifier by Ringel (2023a) that was trained on Twitter data using generative artificial intelligence and exhibits high predictive accuracy. Given a post, the MMX classifier[1] returns all marketing mix components (if any) that the post pertains to. We then examine all posts per marketing mix component in more detail in step two.

Step two uses a natural language processing approach to discover and describe topics that

[1]This method is a text-based multiclass classifier that returns all possible MMX. For example, one post can be classified as 'Product' and 'Promotion' at the same time. Details of the MMX classifier are available in https://huggingface.co/dmr76/mmx_classifier_microblog_ENv02.

manifest in consumers' posts. We employ BERTopic (Grootendorst 2022), a topic modeling approach that embeds posts into latent feature representations and clusters these into topics based on their semantic similarity[2]. BERTopic's strength lies in its ability to handle varied topic densities and sizes, making it well-suited for discovering topics from social media posts. It filters out noise and learns the number of topics from the data.

To facilitate interpretability, we further use generative AI (ChatGPT) to summarize posts' content within each topic cluster, using the prompt[3] *"Summarize what the following social media posts are about within 50 words:"*. While BERTopic excels at quantitatively identifying keywords and topics using term frequency-inverse document frequency (TF-IDF) as an objective measure of word significance, it may not fully capture the context and narratives presented in posts' topics. Generative AI contributes additional insight by considering context and translating the tone and vocabulary of social media content into a succinct summary of what the opportunity is about. By leveraging both, we marry the quantitative aspect of topic modeling with the qualitative depth of language understanding.

In sum, Phase 4 offers a nuanced exploration of consumer discussions surrounding co-mentioned brands. It reveals prevailing marketing mix-related conversations that brand managers can use to explain brand alliance opportunities, test them for substance, and inform their realization.

EMPIRICAL APPLICATION OF BANE TO TWITTER

We use BANE to discover and explain promising brand alliance opportunities from posts on X, previously known as Twitter[4], that co-mention brands. The value of leveraging large-scale, consumer-generated content from Twitter has previously been documented both by academia and industry research (Culotta and Cutler 2016; Lambrecht, Tucker, and Wiertz

[2]We use Sentence BERT (Reimers and Gurevych 2019) with its fine-tuned large language model "all-MiniLM-L6-v2" for the embedding step in BERTopic because Sentence BERT is specifically designed for tasks involving sentence similarity.

[3]We later change "social media posts" to "tweets" in our empirical study.

[4]For simplicity, we refer to X as Twitter throughout this paper.

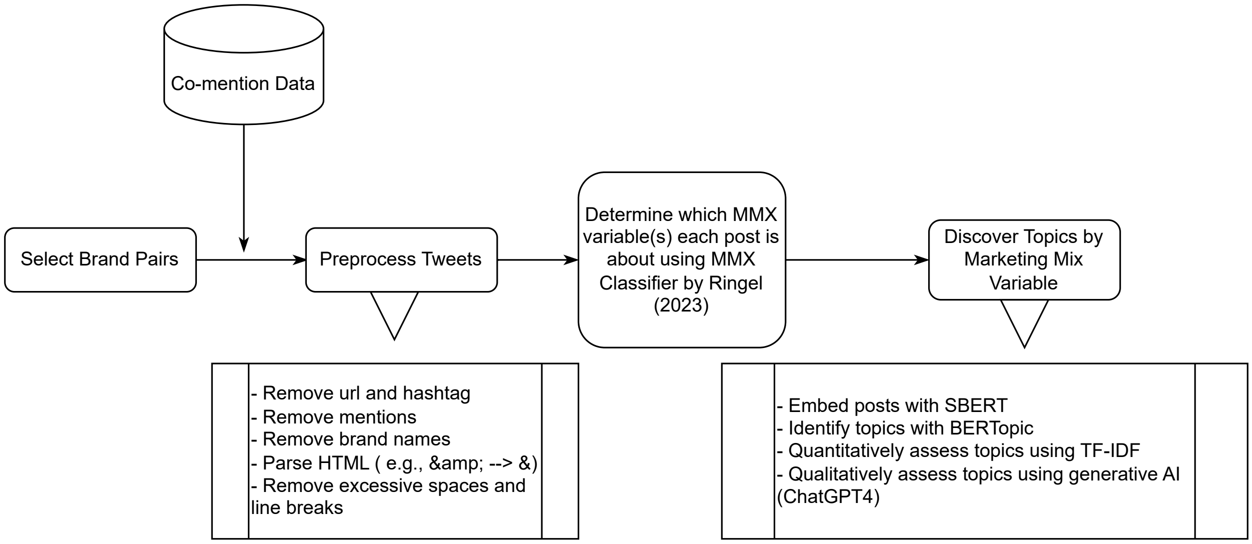


Figure 4: Discovering Marketing Mix Topics in Co-Mentions

2018; Liaukonytė, Tuchman, and Zhu 2023). Our objective is to empirically test and validate BANE's practical application for brand managers. Using BANE, we distill the 553 most promising brand alliance opportunities from an initial candidate pool of over 20,000. By zooming-in and explaining individual opportunities, we showcase how brand managers can use BANE for actionable guidance on executing these opportunities through specific components of their marketing mix.

Our analysis includes 450 B2C brands across 122 categories, including restaurants, airlines, banking, skincare and cosmetics, luxury fashion, video games, healthcare services, and retail categories such as grocery and pharmacy, among others. The choice of brands for our analysis follows the recent work of Dew, Ansari, and Toubia (2022) on data-driven branding. Their study included brands featured in Interbrand's list of the top 100 global brands of 2016, those ranked in the top 500 by brand valuation consultancy firm Brand Finance, and those listed in Forbes' 500 in 2016, with significant overlap among these lists observed. Of these 450 brands, 448 have active consumer co-mentions with other brands on Twitter. We collected all user posts mentioning these brands in 2020 using Twitter's API, totaling 477,933 posts by 236,710 unique consumers. Additionally, we also collected all social media posts from the official handles of the 448 brands.

All other relevant metrics associated with the collected posts, including author information (followers, followees, profile descriptions), tweet content, and engagement metrics (likes, replies, and retweets), were collected. To ensure data quality and relevance, we discarded posts mentioning more than five brands [5]. This approach helps us avoid the ambiguity of brand pairings, as posts containing a laundry list of brands do not contain meaningful insights on brand alliance opportunities. This data cleaning process removed approximately 30% of the initial posts, leaving us with 333,222 posts for analysis containing 443 brands. In addition to social media data, we also use external survey ratings on brand personality dimensions, brand health metrics, and brand identity descriptions from (Dew, Ansari, and Toubia 2022).

In what follows, we (1) identify the opportunity space from Twitter co-mentions by constructing an explicit brand information network, (2) validate the network structure of the opportunity space using external brand attributes, including personality and other online brand metrics, (3) qualify the reduced opportunity space using consumer engagement metrics, and (4) explain promising opportunities by marketing mix variables and topics.

Phase 1: Identify the Opportunity Space

We construct the consumer-driven graph, G_C, and the brand-driven graph, G_B, using explicit brand co-mentions on Twitter. All descriptive metrics for the consumer-driven and the brand-driven graph are included in Table 3. We observe that the brand-driven graph is much sparser than the consumer-driven graph, with respective densities of 2% and 14%, and an edge count of 788 and 14,167 brand pairs, respectively. This sparsity is likely due to brands rarely mentioning each other on social media unless they have already formed some partnership, pursued common causes, or sought to explicitly differentiate themselves from another brand.

[5] We investigated the distribution of the number of brands mentioned in users' posts over the entire year and used the median cutoff of 5 to discard any posts mentioning more than 5 brands.

Table 3: Descriptive Statistics of Graphs

Key Statistics	**Consumer-Driven Graph,** G_C	**Brand-Driven Graph,** G_B	**Opportunity Space,** G_0
Number of nodes	443	312	417
Number of edges	14,167	788	7,149
Same-Category Edges	Yes	Yes	No
Density	14%	2%	8%
Clustering Coefficient	.51	.08	.45
Maximum Degree	316 (Amazon)	38 (Salesforce)	259 (Amazon)

Next we identify the opportunity space, G_O, by removing the edges present in G_B from G_C. This leaves us with 13,633 edges and 442 brands. We further refine the opportunity space by excluding edges among brand pairs of the same category to exclude competing brand pairs from our analysis. Because typical industry classification schemes like NAICS and SIC codes are too coarse for our purposes, we classify brands into more granular categories using industry reports and domain expertise (refer to Web Appendix A for details). Unlike NAICS and SIC categories, our more detailed categorization scheme allows us to preserve edges among brands like Staples (Retail - Office Supplies) and Michaels (Retail - Arts and Crafts), where a brand alliance cannot easily be dismissed on the grounds of competition (substitution). This final distillation process removes nearly 50% of the edges between same-category brands, resulting in a more focused dataset comprising 7,149 cross-category alliance opportunities and 417 brands in the opportunity space, as shown in Table 3. Notably, 25 brands with connections only within the same category are also removed in the distillation process.

Phase 2: Validating the Opportunity Space using ERGM

We now validate that the network structure of our opportunity space relates to the prerequisites of a favorable brand alliance opportunity, as outlined in previous marketing theory (van der Lans, Van den Bergh, and Dieleman 2014). To do this, we fit an Exponential Random Graph Model (ERGM) to the opportunity space. We present descriptive statistics on all included variables in table 4. The variables of interest in the model are those indicative of favorable brand alliances: brand personality (competence, ruggedness, sincerity and sophistication), brand health indicators (familiarity and favorability), and brand similarity based on brand identity statements. The remaining variables serve as controls.

We use mean survey scores for the brand personality and brand health metrics. To quantify brand similarity, we first embed brands' identify statements into lower-dimensional feature vectors using SBERT, a large language model that was fine-tuned on sentence similarity (Reimers and Gurevych 2019). We then calculate the cosine-similarity between the latent feature vectors of each connected brand pair and use it as an edge covariate for brand similarity. The platform-level controls include the age of the brand's Twitter account, number of followers, post-level metrics (i.e., total brand posts in the given year, average caption (word) length per post, average hashtags per post, average retweets per post), and influencer overlap. Influencer overlap measures the number of times two brands have been co-mentioned in a post by a Twitter influencer.

Table 4: Descriptive Statistics for ERGM Variables

Variables	Mean	Std. Dev.	Top Brands (pairs)
Key Variables of Interest			
Brand Personality			
Competence	2.46	.93	CME Group
Ruggedness	1.16	.98	Tractor Supply
Sincerity	1.82	.94	Tyson Foods
Sophistication	1.41	.96	Burberry
Brand Health			
Familiarity	1.36	.83	Adobe
Favorability	2.62	.63	Cartier
Brand Identity			
Brand Similarity	.21	.12	(Food 4 Less, Kroger)
Platform-level Control Variables			
Age of Account	12.02	2.91	Starbucks
Number of followers	1,622,257	630.58	Twitter
Number of posts	630	2,939	ABC
Average hashtags	.79	.67	Panasonic
Average caption length	21.81	7.62	Kimberly-Clark Corp
Average retweets	134	1,564	Twitter
Influencer overlap	.02	.69	(Twitter, Facebook)

Table 5 presents the results of the ERGM analyses. Our baseline is Model 1, which includes only platform-level control variables. Model 2, the main model, introduces brand personality and health node covariates, as well as an edge covariate for brand similarity (derived from brands' identity statements).

Table 5: ERGM Model Results

	Model 1 - Baseline		Model 2 - Main	
	Estimate		Estimate	
Brand Personality				
Competence (Absdiff)			.32***	(.07)
Ruggedness (Absdiff)			-.35***	(.07)
Sincerity (Absdiff)			.04	(.07)
Sophistication (Absdiff)			-.12*	(.07)
Brand Health				
Familiarity (Absdiff)			-.60***	(.07)
Favorability (Absdiff)			.25	(.09)
Brand Identity				
Identity (Edgecov)			3.72***	(.11)
Controls				
Edge	-4.30***	(.12)	-7.47***	(.18)
Brand Account Age (Nodecov)	1.85***	(.06)	1.27***	(.08)
Num of Followers (Nodecov)	1.10***	(.09)	1.11***	(.11)
Num of Brand Posts (Nodecov)	1.01***	(.09)	.69***	(.10)
Average hashtags (Nodecov)	-1.15***	(.08)	-1.16***	(.09)
Average caption length (Nodecov)	-.65***	(.05)	.19**	(.06)
Average retweet (Nodecov)	-.13	(.14)	-.53***	(.15)
Influencer Overlap (Edgecov)	.39***	(.03)	.16**	(.03)
Akaike Information Criterion (AIC)	47,162		42,255	

Note: Significance codes: '***' .001 '**' .01 '*' .05 '.' .1. Standard errors are included in parenthesis.

We first discuss the results for platform-related controls in our main Model 2, which includes all variables (detailed in Table 5). We find a significant negative coefficient for the edges parameter, indicating that the formation of an edge between two brands is not random, unlike in a random graph model like the Erdos-Renyi model, where edges form independently with uniform probability (Robins et al. 2007). This suggests that user co-mentioning ties are concentrated among specific pairs of brands, aligning with particular brand attributes.

The significant positive estimate for Brand Account Age (NodeCov) indicates that brands with a longer social media presence are more likely to be co-mentioned than their younger counterparts. Other platform-related control variables, such as the Number of Followers (NodeCov), Average Caption Length (NodeCov), and Number of Brand Posts (NodeCov), are also significant and positive. This indicates that brands with a large number of followers, high posting frequency, and long average post caption length are more likely to be co-mentioned in the opportunity space. To account for the virality trends driven by influencers, we incorporate Influencer Overlap (Edgecov) into our set of control variables. The significant positive estimate for Influencer Overlap (Edgecov) suggests that brands co-mentioned by influencers or celebrities are more likely to be co-mentioned in the opportunity space. Taken together, our findings appear face valid and support the objective of controlling for platform-related variables in explaining brand co-mentions.

Next, we examine whether coherence in brand personality metrics, operationalized as an absolute difference (Abdsdiff), is related to the network structure of our opportunity space. A positive coefficient indicates that more different personality scores among brand pairs increase their likelihood of being connected in our opportunity space. Conversely, a negative coefficient indicates that more similar personality scores among brand pairs increase their likelihood of being connected. We find significant negative coefficients for Ruggedness (Absdiff) and Sophistication (Absdiff), which is in line with van der Lans, Van den Bergh, and Dieleman (2014), who found that similarity in ruggedness and sophistication lead to more favorable brand alliance evaluations by survey participants. Consistent with van der Lans, Van den Bergh, and Dieleman (2014), who noted that dissimilarity in competence leads to favorable brand alliance evaluations, we observe that dissimilarity in competence among brand pairs is associated with a higher likelihood of being connected in the opportunity space. We do not find a significant relationship between our opportunity space and Sincerity (Absdiff).

We further consider two external brand health characteristics in model 2: brand familiarity and favorability. Marketing theory, as outlined in Simonin and Ruth (1998), indicates that brand familiarity is important for understanding consumer attitudes towards a brand alliance. When two highly familiar brands form an alliance, they experience equal spillover effects

(Simonin and Ruth 1998). Although we cannot measure these spillover effects in our analysis, we find a significant negative coefficient for the absolute difference in familiarity (Absdiff). Our finding suggests that brands with similar levels of brand familiarity are more closely connected in the opportunity space. The coefficient for the absolute difference in favorability (Absdiff) is not significant, and thus unrelated to our opportunity space.

The seventh brand metric we examine is Identity (EdgeCov). We observe a significant and positive coefficient, indicating that brands with more similar values, as represented in their brand identity statements, are more likely to be connected in the opportunity space. These findings suggest that users tend to discuss brands together that share similarities in their value propositions, service offerings, and overarching missions. Related marketing research on mergers and acquisitions (M&A) highlights the importance of similarity in values of companies for successful outcomes (Swaminathan, Murshed, and Hulland 2008). Our analysis also reveals that alliance opportunities in our graph tend to be aligned in terms of similar values, which, according to marketing theory, plays an important role in future success.

Overall, our findings align with established marketing theory and expectations on favorable brand alliance opportunities (van der Lans, Van den Bergh, and Dieleman 2014; Simonin and Ruth 1998). Five out of seven theoretically motivated brand metrics are significant and their direction (sign) matches the expectation. We believe that this consistency with marketing theory supports the validity of the identified opportunity space, allowing us to use it to discover promising brand alliance opportunities. We provide goodness-of-fit tests in Web Appendix B, to document that all model parameters converged, indicating that the proposed ERGM appropriately fits the observed opportunity space. Finally, as observed in Table 5, Model 2 (main model) exhibits a lower Akaike Information Criterion (AIC) compared to Model 1 (baseline). This reduction in AIC suggests that the inclusion of covariates related to brand personality, brand health, and brand identity significantly enhances the model's ability to explain the opportunity space, indicating their importance as predictors.

Next, we use the fitted model to predict a probability score for each brand alliance opportunity. The probability score captures the likelihood of two brands being connected, given their compatibility based on brand personality, health, and identity metrics, as well as platform-related control variables. We then remove opportunities with probability scores below the median to retain only those opportunities that align best with marketing theory. The remaining opportunity space, now comprising 3,575 opportunities, is then passed to Phase 3 for further qualification using consumer engagement metrics.

Phase 3: Qualifying the Reduced Opportunity Space

We qualify the remaining 3,575 brand alliance opportunities using three consumer engagement metrics: salience, support, and valence. Salience is the frequency with which a brand pair is mentioned, support is the engagement with tweets mentioning a brand pair (sum of likes and replies), and valence is the mean sentiment of tweets that mention a brand pair. Because we are interested in favorable brand alliance opportunities, we only include opportunities with positive valence in our analysis, reducing our opportunity space from 3,575 to 1,793 brand pairs. We present the BANE Matrix of our empirical study in Figure 5 (for better readability, we show five randomly sampled opportunities per quadrant).

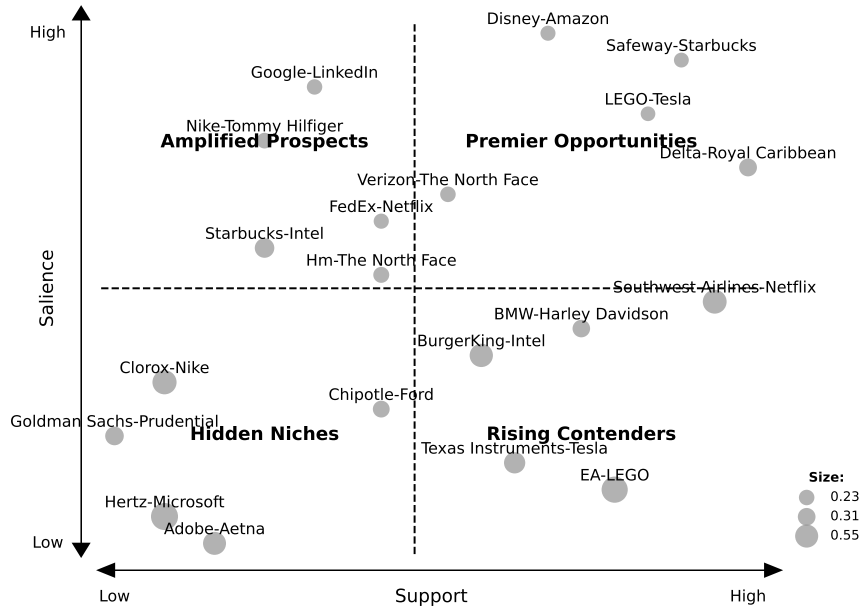


Figure 5: BANE Matrix from Twitter Brand Co-mentions. We show five randomly sampled opportunities per quadrant here for better readability.*

*Bubble size corresponds to valence (a continuous variable), with .23, .31, and .55 denoting the minimum, mean, and maximum values respectively.

Table 6 characterizes the BANE Matrix's four quadrants in the context of this study. We find that the most common opportunities are *Premier* and *Hidden Niches* (553 and 602, respectively). The top categories in these two quadrants are restaurants, department stores, automotive, electronics and e-commerce. *Amplified Prospects* (296) and *Rising Contenders* (342) are less common and are primarily centered around software, airlines, retail (Grocery), automotive, and restaurants.

Although each quadrant features different top 5 categories, no quadrant is heavily concentrated to subsets of specific industry categories, as indicated by the low Herfindahl-Hirschman Index (HHI) values. The low category concentration across quadrants demonstrates BANE's potential to identify opportunities across a broad range of categories and industries without bias toward any particular category.

Table 6: Summary Statistics of Key Attributes by Quadrant

Attributes	**Definition**	**Premier Opportunities**	**Amplified Prospects**	**Rising Contenders**	**Hidden Niches**
Opportunities	Number of opportunities	553	296	342	602
Avg. Popularity	Average number of followers of brands.	4,665,403	3,762,914	3,745,406	3,661,782
Top Categories	Top 5 categories by total count	Restaurants, Department Stores, E-commerce, Technology (Electronics), Telecommunications	Technology (Software), Financial Services, E-commerce, Restaurants, Automotive	Retail (Grocery), Department Stores, Automotive, Restaurants, Airlines	Restaurants, Automotive, Department Stores, Financial Services, Food Production
Brand Identity Agreement	Average similarity score. Higher values indicate higher agreement.	.31	.30	.28	.27
Brand Health Agreement	Mean score of difference in familiarity and favorability. Smaller values indicate higher agreement.	1.73	1.59	2.08	1.86
Brand Personality Agreement	Mean score of difference in competence, ruggedness, sincerity, and sophistication. Smaller values indicate higher agreement.	.89	.92	.90	.94
Influencer Overlap	Count of influencers mentioning the alliance	41	14	28	11
Category Concentration (HHI)	Herfindahl-Hirschman Index measuring category concentration	.03	.03	.02	.02

Among all quadrants, brand alliance opportunities exhibit consistently high agreement in brand identity, with the *Premier Opportunities* quadrant demonstrating the highest agreement. Previous marketing research in M&A (Swaminathan, Murshed, and Hulland 2008) shows that aligned values are crucial for successful collaboration. Along with high values for support and salience, strong alignment in the core values of brand pairs adds to the already high appeal of the *Premier Opportunities*. This suggests that these opportunities are not only strategically advantageous in terms of visibility and engagement but are also fundamentally aligned at the core identity level, which bodes well for a successful partnership. Such insights point toward a deeper strategic alignment that goes beyond mere market visibility or engagement metrics, anchoring on the intrinsic compatibility of core brand values.

On the other hand, *Rising Contenders* and *Hidden Niche*s exhibit the highest level of agreement in health metrics (familiarity and favorability) across brand pairs. The high health metrics suggest that these quadrants, though possibly less prominent or emerging, hold strong potential for successful alliances. Phase 3's differentiated analysis reveals that brand managers must not solely rely on the most popular or visible brand pairs on social media when searching for alliance opportunities, but might also consider those with solid underlying brand health. Our findings demonstrate that diving deeper into social media metrics and combining them with traditional brand metrics can uncover hidden gems—brand pairs that may not top popularity charts but have cultivated strong favorably and familiarity among niche audiences.

We further find that influencer overlap varies substantially across the quadrants, reaching a high of 41 in *Premier Opportunities* and a low of 11 in *Hidden Niches*. The high level of influencer co-mentions for *Premier Opportunities* reinforces the quadrant's strengths in terms of high visibility, strong engagement, and an already engaged consumer base. In contrast, *Amplified Prospects*, characterized by a lower influencer overlap of 14, are fueled more organically by regular users. While brand managers may immediately capitalize on *Premier Opportunities*, our findings also suggest that increasing influencer activity (e.g., through sponsoring) for could *Amplified Prospects* could be an avenue to increase support as the opportunity is further pursued.

Our findings from Phase 3 highlight the BANE Matrix's usefulness at a macroscopic level, documenting its capacity to offer a differentiated perspective on the strategic fit between brand alliances and overarching goals. For brand managers, a micro-level analysis tailored to their specific brands becomes the next step, enabling a focused exploration of opportunities—across various quadrants—that are most pertinent to their strategic goals.

Phase 4: Explain Promising Opportunities

In Phase 4, we zoom-in on the social media content of individual candidate opportunities to explain what they are about, offering insights to brand managers on how they might be executed. We focus on two *Premier Opportunities* in this study: One that occurred after 2020 (post our analysis) to demonstrate face validity of our approach, and one that has yet to be realized by the involved brands to demonstrate our approach's potential.

We start by discussing the Nike and Lego example, identified by the BANE matrix as a premier opportunity. In Step 1, we utilize a pre-trained marketing mix (MMX) classifier by Ringel (2023a) to determine the marketing mix component to which the Nike-Lego brand alliance opportunity pertains. We find that the majority of posts co-mentioning these two brands are related to 'Product'. In the second step, we delve deeper into the most prevalent component of the marketing mix, in this case, 'Product', to uncover the specific topics associated with these posts using BERTopic. Examples of Twitter posts related to this opportunity and identified topics are shown in Table 7.

Indeed, three years later, a formal partnership[6] between Nike and Lego was announced. According to Nike, the partnership could see a series of co-branded products, content, and experiences that combine the imaginative power of Lego bricks with the 'Just Do It' spirit of Nike to invite all kids into play and sport. Our analysis reveals that the main topics centered

[6]Details of the partnership: Link - `https://about.nike.com/en/newsroom/releases/nike-lego-partnership-announcement`

around this opportunity also involve co-branded products like sneakers and special Nike LEGO sets. The generative AI summary aligns with our own interpretation. Taken together, the close alignment of the opportunity found by BANE with the brand alliance later realized in 2024by the involved brands demonstrates BANE's forward-looking potential and the face validity of its insights.

Table 7: Discovery of Topics related to the Alliance Opportunity (Lego_Group - Nike)

No.	Topic (Keywords)	MMX	Example Tweets	ChatGPT Summary
1	Design (pitch, design, set, project)	Product	*- I think @NWSL and @Nike should team up with @LEGO_Group for a women soccer LEGO set. @budweiserusa could sponsor the drinks in fans hands...* ...	*"The tweets propose a potential collaboration between Nike and LEGO for a women's soccer LEGO set and highlight a student design project featuring LEGO minifigures and Nike Air Max trainers. They also reference past ideas linking Nike and LEGO in creative concepts."*
2	Marketing (brand, consumers, report, social,reputationscore, messaging)	Promotion, Product	*- Nowadays, more and more, "consumers want to co-create the product". Think customised sneakers (@Nike), special sets (@LEGO_Group), etc...* ...	*"The tweets emphasize the trend of consumer-driven co-creation and customization, referencing potential products like special LEGO sets and customized Nike sneakers. They also touch on the innovative use of Augmented Reality and the influence of brand reputation and social messaging on consumer engagement."*

Next, we discuss an example of a premier alliance opportunity yet to materialize, Tesla and Lego. We observe that the majority of posts co-mentioning these two brands are related to 'Product'. The results highlight consumer suggestions for a potential collaboration, with tweets uncovering topics such as creating Lego models of Tesla cars and SpaceX rockets. This indicates strong consumer interest in tangible, branded products resulting from a Tesla-Lego partnership, an insight that can be strategically valuable for both brands. The analysis thus reveals the potential for a unique product line that resonates with the consumers of both brands, tapping into a niche yet enthusiastic market segment. In addition, we use generative

AI (ChatGPT) to validate our interpretation further, following the process discussed in phase 4. The AI summary is highly consistent with our takeaways, which confirms our observations that consumers are enthusiastic about the potential collaboration between Tesla and Lego while it was not available as of the posting date. We present the results in Table 8.

Table 8: Discovery of Topics related to the Alliance Opportunity (Tesla - Lego_Group)

No.	Topic (Keywords)	MMX	Example Tweets	ChatGPT Summary
1	Rocket (lego, create, rocket, people, collab)	Product	*- What just came into my mind: How cool would it be if @LEGO_Group and @elonmusk would team up and we could build all @Tesla models and all @SpaceX rockets with #Lego bricks? I would buy everything. #Lego #Tesla #SpaceX...* ...	*"Users are excited about the potential collaboration between LEGO, Tesla, and SpaceX, suggesting building Tesla models and SpaceX rockets with LEGO bricks would be educational and fun."*
2	Cybertruck (cybertruck, fan, votes, gen, legos)	Product	*- My 8 year old just built this #cybertruck with random Legos. The love for @Tesla runs deep in our family. It looks like he threw in @elonmusk and some astronauts too. Perfect. #marscybertruck @spacex @LEGO_Group...* ...	*"Fans and children express their creativity by building LEGO Tesla Cybertrucks, showcasing familial affection for Tesla. A fan-made LEGO Cybertruck model gains significant support for potential creation."*

Validation

Distilling brand alliance opportunities from information networks inherently involves an exploratory process. Given the absence of a known ground truth to directly assess the effectiveness the proposed BANE market research approach, we validate it from three distinct perspectives: marketing theory, firms, and consumers. We aim to establish BANE's internal and external validity through a multi-perspective approach that serves as a proxy for direct ground truth comparison.

First, from the marketing-theory perspective, we reiterate the findings of ERGM analysis that tested the relationship between the network structure of the opportunity space and brand metrics from the marketing literature. We found that five out of seven theoretically motivated brand metrics are significant and match the expectations for favorable brand alliance opportunities as studied in previous research (van der Lans, Van den Bergh, and Dieleman 2014; Swaminathan, Murshed, and Hulland 2008; Simonin and Ruth 1998). The fitted ERGM also revealed that the network structure strongly relates to the similarity in brand values (Swaminathan, Murshed, and Hulland 2008), further supporting the internal validity of the proposed approach.

Second, from the firm perspective, we collected data on real-world brand alliances that materialized after our analysis (post-2020). Of the 30 real-world alliances identified, 28 were suggested by the studied information network. Of these, over 80% are included in the BANE Matrix (Phase 3), with the majority falling into the premier opportunities quadrant. For instance, alliances such as the one between Google and Ford, as well as the Starbucks and Amazon partnership, occurred after our analysis in 2020 and were successfully identified by BANE. Three real-world brand alliance opportunities were eliminated through ERGM analysis (Phase 2). A closer examination of the three eliminated brand pairs reveals that they had significantly lower agreement in brand personality scores compared to those that progressed to Phase 3. Overall, the strong alignment between the brand alliances identified by BANE and subsequent real-world alliances supports its internal validity. These findings demonstrate BANE's effectiveness in a practical business context.

Third, we externally validate brand alliance opportunities distilled by BANE through an online survey from the consumer perspective. We recruited consumers via Prolific [7] to assess whether the premier opportunities identified by BANE are systematically perceived as more appealing than those eliminated. To this end, we sampled 25 of the most promising brand alliance pairs from the premier opportunities quadrant of the BANE matrix and 25 pairs that were eliminated during Phase 2 due to low ERGM likelihood scores. We recruited 500 U.S. participants, stratified by demographic factors, who rated each pair using a 5-point Likert scale.

[7] Prolific is a survey management and distribution platform known for its ability to rapidly recruit a diverse set of nationally representative participants for research. More details on the platform can be found here: https://www.prolific.com/.

Using a Partial Block Design, we constructed 25 sets comprising six brand pairs each, with three pairs sampled randomly from both the premier and eliminated groups. Participants were randomly assigned three sets of questions, resulting in each participant rating a total of 18 brand alliance opportunities. On average, each opportunity was evaluated by 160 survey participants who passed attention checks. Further details on the survey methodology and descriptive statistics can be found in Web Appendix C. The Wilcoxon–Mann–Whitney test confirms that the premier opportunities identified in the BANE matrix were systematically rated higher than the eliminated ones ($W\ statistics = 6019374$, $Rank\ difference = 3956670$, $p < .001$), supporting the external validity of BANE insights. To control for demographic biases that might influence ratings, we employed a mixed-effects regression model:

$$y_{ij} = \beta_0 + \beta_1 \text{Treated}_{ij} + \beta_2 \text{Age}_{ij} + \beta_3 \text{Gender}_{ij} + \beta_4 \text{Income}_{ij} + u_{0j} + u_{1j} \text{Treated}_{ij} + \epsilon_{ij} \quad (6)$$

where y_{ij} is the appeal rating by respondent j on brand pair i, and other terms adjust for individual and treatment variability. We find a significant positive effect for premier opportunities ($.603, p < .001$), indicating that consumers generally find these opportunities more appealing than those eliminated earlier by BANE. We contend that this nuanced result demonstrates BANE's ability to effectively discern brand alliances that are more likely to resonate with consumer expectations, further validating its utility and external applicability. Further details and the full regression results are available in Web Appendix C.

Comparing Explicit and Implicit Networks for Identifying Brand Alliances

Implicit networks have been shown to be useful for inferring market structures and competition (Yang, Zhang, and Kannan 2022; Ringel and Skiera 2016). However, their use in identifying and explaining brand alliances is limited for two reasons. First, people can follow or comment on two brands for independent reasons, with no common preference. Second, user comments made months apart may be unrelated, and consumers may not even have the previous brand in mind when commenting on another brand later.

To investigate how our findings differ to implicit approaches (as in Malhotra and Bhattacharyya (2022); Yang, Zhang, and Kannan (2022)), we collect co-followership and co-commenting data for a set of 160 brands. We construct three opportunity spaces from these data: co-follow as intersect of followers (Malhotra and Bhattacharyya 2022), co-comment with a deep neural autoencoder (Yang, Zhang, and Kannan 2022), and explicit brand co-mentions (this study). See Web Appendix D for details on the construction of the opportunity spaces from the implicit networks. We use ERGMs to study how the observed network structures of all three opportunity spaces align with prior marketing theory on favorable brand alliances. We present the model results for the key variables of interest (Brand Personality, Brand Health, and Brand Identity) in Table 9.

Table 9: ERGM Model Results

	Co-Mentions	**Co-Followership**	**Co-Comments**
	Estimate	Estimate	Estimate
Brand Personality			
Competence (Absdiff)	.04	-.08	-.03
Ruggedness (Absdiff)	-.24*	-.30**	.03
Sincerity (Absdiff)	.45***	.16	-.12
Sophistication (Absdiff)	-.41***	-.18.	-.35***
Brand Health			
Familiarity (Absdiff)	-.86***	-1.62***	1.03***
Favorability (Absdiff)	.31*	.20	-.20
Brand Identity			
Identity (Edgecov)	2.12***	.95***	-1.08***
Akaike Information Criterion (AIC)	15704	16046	16234

Note: Significance codes: '***' .001 '**' .01 '*' .05 '.' .1.

We find that the opportunity space constructed from the explicit co-mentions network aligns with prior marketing theory on brand compatibility. We observe similar ERGM estimates with a larger set of brands in Phase 2, suggesting our findings are robust across different sample sizes. Similarity in ruggedness and sophistication among brand pairs is associated with a higher likelihood of being connected in the opportunity space, while dissimilarity in competence and sincerity also leads to a higher likelihood of connection (van der Lans, Van den Bergh, and Dieleman 2014). Our analysis further reveals that alliance opportunities in the explicit co-mentions network are most aligned in terms of similar values, which, according to marketing theory (Swaminathan, Murshed, and Hulland 2008), plays an important role in alliance success.

For the implicit networks, including the co-followership and co-commenting graphs, we observe that brand metrics related to personality are not significant, and their directions (signs) do not align with expectations. Additionally, alliance opportunities in the co-commenting graph are not aligned in terms of similar values. However, it is important to note that previous research using implicit relationships pursued different objectives than this study and demonstrated the value of implicit relationships in different contexts, such as market structure and competition.

Overall, we believe that the alignment of the co-mentions network structure with marketing theory supports the validity of the identified opportunity space, further strengthening the case for using explicit networks to discover promising brand alliance opportunities.

MANAGERIAL IMPLICATIONS

We introduce a new market research approach called BANE to identify and explain brand alliance opportunities. By synthesizing two critical streams of inquiry—the identification of potential brand alliance partners in Phase 1 ("Who") and the evaluation of brand compatibility in Phase 2 ("Why")—BANE offers a structured, comprehensive tool for navigating the complexities of brand partnerships. BANE systematically mitigates the risk of premature alliance selections by providing a rigorous validation process, ensuring that only alliances with a solid theoretical foundation are considered. This step is critical in avoiding ill-informed decisions that could dilute brand equity or misallocate resources.

The BANE Matrix's introduction in Phase 3 refines the opportunity space by offering a differentiated evaluation of opportunities across four quadrants—*Premier Opportunities*, *Amplified Prospects*, *Rising Contenders*, and *Hidden Niches*. This differentiation helps brand managers focus on those opportunities that best align with their strategic and tactical goals. For instance, brand managers might pursue *Premier Opportunities* for immediate engagement and market impact, leverage *Amplified Prospects* to enhance consumer interaction, capitalize on the dedicated consumer bases of *Rising Contenders* for niche market expansion, or explore *Hidden Niches* for innovative collaborations. This strategic flexibility ensures that brand alliances are not only chosen for their prominence but are also tailored towards specific objectives.

Phase 4's emphasis on operationalization—determining the "What" (thematic focus) and "How" (marketing mix execution)—equips brand managers with a deeper understanding on implementing alliances that are not only strategically sound but also highly aligned with consumer expectations. This phase addresses a crucial gap in previous approaches by providing a structured pathway from theoretical compatibility to practical, consumer centric execution. Importantly, by systematically examining the social media conversation around a candidate opportunity, brand managers can assess its substance and leverage consumers' ideas as crowd-sourced starting points for further ideation.

Note that, for the purpose of this research, we apply BANE to a large set of brands to demonstrate its application and assess its validity. Using ERGMs, we find that the opportunity space constructed from the explicit co-mentions network aligns with prior marketing theory on brand compatibility. Given our findings, brand managers could choose to omit Phase 2 and focus on an egocentric network to identify alliances around a single (e.g., their) brand from consumers' co-mentions.

LIMITATIONS AND FUTURE DIRECTIONS OF RESEARCH

While BANE adds deeper analysis to extant market research approaches to provide richer insights (e.g., (Malhotra and Bhattacharyya 2022)), its application is not without limitations. Although BANE is a novel approach to identifying and explaining potential brand alliance partners through the lens of the consumer, managers must also consider the strategic alignment among partners, the risk of brand dilution, and other operational and financial factors. Therefore, a comprehensive evaluation that goes beyond the consumer insights provided by BANE is essential. As such, BANE can be used as an important first step in identifying and explaining how to execute brand alliances. This step should be followed by operational and financial assessments that are beyond the scope of this study.

Another consideration is the dynamic nature of social media where the landscape of potential brand alliances is constantly evolving. Although BANE offers a structured approach to navigating this space, its insights are static and can't be extrapolated indefinitely into the future. We also find that the majority of the opportunities revealed through BANE pertain to the product dimensions of the marketing mix, offering limited insight into opportunities around place, price, and promotion. To address this limitation, future applications of BANE might augment social media data with online customer reviews and forum discussions.

It is also important to acknowledge that not all opportunities identified in Phase 4 will be meaningful, and some may lack substance upon closer examination of the underlying content. Hence, during the topic discovery phase, it is crucial to leverage domain knowledge and managerial insight when interpreting the topics to identify the most promising opportunities. Further, while Phase 2 validates identified opportunities using ERGMs, there are various external factors—ranging from market dynamics to cultural trends and external events—that may influence brand co-mentions beyond what is captured in the current approach. We use ERGMs for validation and to establish the appropriateness of the opportunity space for identifying brand alliances. However, ERGMs are not meant to draw causal conclusions about user motivations for co-mentioning brands on social media, which is a future topic of study.

Lastly, the qualification phase of BANE creates a more focused set of alliances that have sufficient salience, support, and positive valence. Utilizing positive valence ensures that only opportunities mentioned in a positive context by customers are retained in the BANE matrix. Future research could extend this analysis to look into opportunities mentioned in a negative context to identify any brand opportunities that contain customer queries or pain points–something that this study did not do. Such analysis could be used to streamline or address joint customer complaints for those brands. Integrating this approach could be a part of a brand's social media CRM (Customer Relationship Management) efforts to address customer issues in a timely and relevant manner.

CONCLUSION

In this paper, we propose, test, and validate a new market research approach called BANE, to guide brand managers in discovering and evaluating brand alliance opportunities. Unlike previous research that predominantly focused on identification (the "Who" question) (Malhotra and Bhattacharyya 2022) or assessing the compatibility of brand personality for favorable alliances (the "Why" question) (van der Lans, Van den Bergh, and Dieleman 2014), BANE connects these critical inquiries. We further address the "What" and "How" questions of brand alliance opportunities. This extension not only bridges the gap between identifying potential brand partners and understanding their compatibility, but also guides brand managers on what specific alliance opportunities to pursue and how to implement them such that resonate with consumers' expectations.

Substantively, we show that relying solely on the co-occurrence of brands on social media can be misleading because co-occurrence is not necessarily indicative of a promising brand alliance opportunity. It is important to validate, distill, and explain promising alliance opportunities to make them practically useful for brand managers. We discover that over 92% of opportunities suggested by the social network lack substance. Our findings emphasize the importance of leveraging both the network structure and content to fully understand the attractiveness and potential of these opportunities.

We further show how Exponential Random Graph Models (ERGMs) can be used to validate the structure of explicit information networks for identifying and explaining brand alliance opportunities. Additionally, the comparison of implicit versus explicit networks using ERGMs shows that the opportunity space derived from explicit consumer co-mentions is more aligned with marketing theory on compatible brand alliances. ERGMs are a powerful tool in the social sciences to analyze network data (Hwang et al. 2022) but have rarely been used in marketing research. Prior marketing research posed the question of how large-scale brand networks relate to brand attributes (Lovett, Peres, and Shachar 2014). This study addresses the question in the context of distilling brand alliance opportunities from social media. Our empirical study reveals significant relationships between the explicit brand information network constructed from consumer co-mentions data and brand personality, health, and identity attributes.

To test the validity and practical utility of BANE's empirical insights, we conduct internal and external validations through marketing theory analysis, real-world firm assessments, and consumer surveys. Our findings collectively support BANE's validity and highlight the utility of this new market research approach to brand managers. In sum, BANE marks a step forward in the discovery and evaluation of brand alliance opportunities. As such, it is part of the ongoing journey towards more sophisticated and nuanced market research methods in the digital age. We hope that our work paves the path for further research in the domain of brand alliances and social networks.

REFERENCES

Aaker, David A (2012), *Building Strong Brands.* Simon and Schuster.

Aaker, Jennifer L (1997), "Dimensions of Brand Personality," *Journal of Marketing Research*, 34 (3), 347–356.

Anderson, John R and Gordon H Bower (1974), "A Propositional Theory of Recognition Memory," *Memory & Cognition*, 2 (3), 406–412.

Baker, Andrew M., Naveen Donthu, and V. Kumar (2016), "Investigating How Word-of-Mouth Conversations About Brands Influence Purchase and Retransmission Intentions," *Journal of Marketing Research*, 53 (2), 225–239.

Batra, Rajeev, Peter Lenk, and Michel Wedel (2010), "Brand Extension Strategy Planning: Empirical Estimation of Brand–category Personality Fit and Atypicality," *Journal of Marketing Research*, 47 (2), 335–347.

Beichert, Maximilian, Andreas Bayerl, Jacob Goldenberg, and Andreas Lanz (2024), "EXPRESS: Revenue Generation Through Influencer Marketing," *Journal of Marketing* (published online 2024), https://doi.org/10.1177/00222429231217471.

Cao, Zixia and Alina Sorescu (2013), "Wedded Bliss or Tainted Love? Stock Market Reactions to the Introduction of Cobranded Products," *Marketing Science*, 32 (6), 939–959.

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.

Culotta, Aron and Jennifer Cutler (2016), "Mining Brand Perceptions From Twitter Social Networks," *Marketing Science*, 35 (3), 343–362.

Dew, Ryan, Asim Ansari, and Olivier Toubia (2022), "Letting Logos Speak: Leveraging Multiview Representation Learning for Data-Driven Branding and Logo Design," *Marketing Science*, 41 (2), 401–425.

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.

Gabel, Sebastian, Daniel Guhl, and Daniel Klapper (2019), "P2v-Map: Mapping Market Structures for Large Retail Assortments," *Journal of Marketing Research*, 56 (4), 557–580.

Goh, Jie Mein, Guodong Gao, and Ritu Agarwal (2016), "The Creation of Social Value," *MIS Quarterly*, 40 (1), 247–264.

Grootendorst, Maarten (2022), "BERTopic: Neural Topic Modeling With a Class-Based TF-IDF Procedure," arXiv (March 11), https://arxiv.org/abs/2203.05794.

Hughes, Christian, Vanitha Swaminathan, and Gillian Brooks (2019), "Driving Brand Engagement Through Online Social Influencers: An Empirical Investigation of Sponsored Blogging Campaigns," *Journal of Marketing*, 83 (5), 78–96.

Hutto, Clayton and Eric Gilbert "Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," "Proceedings of the International AAAI Conference on Web and Social Media," Vol. 8. (2014).

Hwang, Elina H, Xitong Guo, Yong Tan, and Yuanyuan Dang (2022), "Delivering Healthcare Through Teleconsultations: Implications for Offline Healthcare Disparity," *Information Systems Research*, 33 (2), 515–539.

Kim, Jun B, Paulo Albuquerque, and Bart J Bronnenberg (2011), "Mapping Online Consumer Search," *Journal of Marketing Research*, 48 (1), 13–27.

Krivitsky, Pavel N (2012), "Exponential-Family Random Graph Models for Valued Networks," *Electronic Journal of Statistics*, 6, 1100–1128.

Kupfer, Ann Kristin, Nora Pähler Vor Der Holte, Raoul V. Kübler, and Thorsten Hennig-Thurau (2018), "The Role of the Partner Brand's Social Media Power in Brand Alliances," *Journal of Marketing*, 82 (3), 25–44.

Lam, Son K., Michael Ahearne, Ye Hu, and Niels Schillewaert (2010), "Resistance to Brand Switching When a Radically New Brand Is Introduced: A Social Identity Theory Perspective," *Journal of Marketing*, 74 (6), 128–146.

Lambrecht, Anja, Catherine Tucker, and Caroline Wiertz (2018), "Advertising to Early Trend Propagators: Evidence From Twitter," *Marketing Science*, 37 (2), 177–199.

Li, Yiyi and Ying Xie (2020), “Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement,” *Journal of Marketing Research*, 57 (1), 1–19.

Liaukonytė, Jūra, Anna Tuchman, and Xinrong Zhu (2023), “Frontiers: Spilling the Beans on Political Consumerism: Do Social Media Boycotts and Buycotts Translate to Real Sales Impact?,” *Marketing Science*, 42 (1), 11–25.

Lovett, Mitchell, Renana Peres, and Ron Shachar (2014), “A Data Set of Brands and Their Characteristics,” *Marketing Science*, 33 (4), 609–617.

Lusher, Dean, Johan Koskinen, and Garry Robins (2013), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications* Cambridge University Press.

Malhotra, Pankhuri and Siddhartha Bhattacharyya (2022), “Leveraging Cofollowership Patterns on Social Media to Identify Brand Alliance Opportunities,” *Journal of Marketing*, 86 (4), 17–36.

Moreau, C. Page, Emanuela Prandelli, Martin Schreier, and Silke Hieke (2020), “Customization in Luxury Brands: Can Valentino Get Personal?,” *Journal of Marketing Research*, 57 (5), 937–947.

Nam, Hyoryung, Yogesh V Joshi, and P K Kannan (2017), “Harvesting Brand Information From Social Tags,” *Journal of Marketing*, 81 (4), 88–108.

Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Moshe Fresko (2012), “Mine Your Own Business: Market-Structure Surveillance Through Text Mining,” *Marketing Science*, 31 (3), 521–543.

Oestreicher-Singer, Gal, Barak Libai, Liron Sivan, Eyal Carmi, and Ohad Yassin (2013), “The Network Value of Products,” *Journal of Marketing*, 77 (3), 1–14.

Reimers, Nils and Iryna Gurevych “Sentence-Bert: Sentence Embeddings Using Siamese BERT-networks,” “EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference,” (2019).

Ringel, Daniel (2023a), “Creating Synthetic Experts With Generative Artificial Intelligence,” SSRN (Dec 11), http://dx.doi.org/10.2139/ssrn.4542949.

Ringel, Daniel M (2023b), “Multimarket Membership Mapping,” *Journal of Marketing Research*, 60 (2), 237–262.

Ringel, Daniel M and Bernd Skiera (2016), “Visualizing Asymmetric Competition Among More Than 1,000 Products Using Big Search Data,” *Marketing Science*, 35 (3), 511–534.

Robins, Garry, Pip Pattison, Yuval Kalish, and Dean Lusher (2007), “An Introduction to Exponential Random Graph (P*) Models for Social Networks,” *Social networks*, 29 (2), 173–191.

Shi, Zhan Michael, Gene Moo Lee, and Andrew B. Whinston (2020), “Toward a Better Measure of Business Proximity: Topic Modeling for Industry Intelligence,” *MIS Quarterly*, 40 (4), 1035–1056.

Simonin, Bernard L and Julie A Ruth (1998), “Is a Company Known by the Company It Keeps? Assessing the Spillover Effects of Brand Alliances on Consumer Brand Attitudes,” *Journal of Marketing Research*, 35 (1), 30–42.

Sundararajan, Arun, Foster Provost, Gal Oestreicher-Singer, and Sinan Aral (2013), “Research Commentary—information in Digital, Economic, and Social Networks,” *Information Systems Research*, 24 (4), 883–905.

Swaminathan, Vanitha and Christine Moorman (2009), “Marketing Alliances, Firm Networks, and Firm Value Creation,” *Journal of Marketing*, 73 (5), 52–69.

Swaminathan, Vanitha, Feisal Murshed, and John Hulland (2008), “Value Creation Following Merger and Acquisition Announcements: The Role of Strategic Emphasis Alignment,” *Journal of Marketing Research*, 45 (1), 33–47.

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.

Valsesia, Francesca, Davide Proserpio, and Joseph C Nunes (2020), “The Positive Effect of Not Following Others on Social Media,” *Journal of Marketing Research*, 57 (6), 1152–1168.

van der Lans, Ralf, Bram Van den Bergh, and Evelien Dieleman (2014), “Partner Selection in Brand Alliances: An Empirical Investigation of the Drivers of Brand Fit,” *Marketing Science*, 33 (4), 551–566.

Yang, Yi, Kunpeng Zhang, and P K Kannan (2022), "Identifying Market Structure: A Deep Network Representation Learning of Social Engagement," *Journal of Marketing*, 86 (4), 37–56.

Zhang, Kunpeng, Siddhartha Bhattacharyya, and Sudha Ram (2016), "Large-Scale Network Analysis for Online Social Brand Advertising.," *Mis Quarterly*, 40 (4).

WEB APPENDIX A: LIST OF BRANDS

All the brands used in our analysis, along with their respective categories, are included in Table 10. The choice of brands for our analysis follows the recent work of Dew, Ansari, and Toubia (2022) on data-driven branding. Please note that the assignment of categories was manually performed by three coders, using external industry reports and domain knowledge.

Table 10: List of the Brands and Categories

Brand	**Category**	**Brand**	**Category**
LabCorp	Medical Laboratories/Diagnostics	supervaluPR	Retail - Grocery
CMEGroup	Financial Services	QuestDX	Healthcare Services
CharterNewsroom	Telecommunications	GE	Multinational Conglomerate
3M	Manufacturing and Technology	FTI_Global	Business Advisory
Carrier	HVAC	mccormickspices	Food Production
Mattel	Toys	bmsnews	Pharmaceutical
PolarisORV	Automotive - Vehicles Off-Road	Thrivent	Financial Services
johnniewalker	Beverages - Alcoholic	Burlington	Discount Retail
FirstAm	Insurance	Coach	Luxury Fashion
cbrands	Beverages - Alcoholic	ABC	Broadcasting and Media
Nike	Sportswear	warnerbros	Film and Television Production
Lowes	Retail - Home Improvement	Discovery	Mass Media and Entertainment
CarMax	Retail - Automotive	PSEGNews	Energy
Harman	Audio and Infotainment	UnitedHealthGrp	Healthcare Insurance
SamsClub	Retail - Grocery	MonsterEnergy	Beverages - NonAlcoholic
MichaelsStores	Retail - Arts and Crafts	TruistNews	Banking
MurphyUSA	Gas Station and Convenience Store	Amgen	Biotechnology
riteaid	Retail - Pharmacy and Health	KCCorp	Personal Care Products
Walgreens	Retail - Pharmacy and Health	Gillette	Personal Care Products
TreeHouseFoods	Food Production	GameStop	Retail - Video Games
dhlexpressuk	Logistics and Courier	Cabelas	Retail - Outdoor Equipment
goodyear	Tires and Rubber	ATT	Telecommunications
HomeDepot	Retail - Home Improvement	tjxco	Discount Retail
MinuteMaid	Beverages - NonAlcoholic	oreillyauto	Automotive Parts
drpepper	Beverages - NonAlcoholic	SKECHERSUSA	Footwear & Outdoor Apparel
xcelenergy	Energy	Prudential	Financial Services
CenturyLink	Telecommunications	CarnivalCruise	Cruise Line
Activision	Video Games	HanesBrands	Fashion & Apparel
Ameriprise	Financial Services	Dell	Technology - Electronics
Hyundai	Automotive - Vehicles	Chase	Banking
EversourceCorp	Energy	Google	Technology - Internet Services
HCAhealthcare	Healthcare Services	StateFarm	Insurance

Continued on next page

Table 10 – continued from previous page

Brand	Category	Brand	Category
harleydavidson	Automotive - Vehicles Motorcycles	MTV	Entertainment - Television Network
FlyFrontier	Airline	molinahealth	Healthcare Insurance
JetBlue	Airline	Dior	Luxury Fashion
TysonFoods	Food Production	DominionEnergy	Energy
SUBWAY	Restaurants	Loews_Hotels	Hotel
IKEAUSA	Retail - Furniture & Home Goods	QurateRetailGrp	Retail - Multi-platform
MarathonPetroCo	Oil & Gas	Cinemark	Entertainment - Movie Theaters
GetSpectrum	Telecommunications	TXInstruments	Technology - Semiconductors
comcast	Telecommunications	CruiseNorwegian	Cruise Line
pvh	Fashion & Apparel	LEGO_Group	Toy Manufacturing
Buick	Automotive - Vehicles	KKR_Co	Investment Management
BMW	Automotive - Vehicles	ParamountPics	Film and Television Production
Hermes_Paris	Luxury Fashion	LOrealUSA	Skincare and Cosmetics
Costco	Retail - Wholesale	WellsFargo	Banking
iHeartMedia	Broadcasting and Media	CalvinKlein	Fashion & Apparel
WyndhamHotels	Hotel	NewsCorp	News and Media
DICKS	Retail - Sporting Goods	fanta	Beverages - NonAlcoholic
bookingcom	Online Travel Agency	Verizon	Telecommunications
insideFPL	Energy	OralB	Oral Care
LouisVuitton	Luxury Fashion	TheAESCorp	Energy
JNJNews	Healthcare Products	InvescoUS	Investment Management
LEVIS	Denim Apparel	BigLots	Discount Retail
TMobile	Telecommunications	BNBuzz	Retail - Books
FannieMae	Financial Services	PayPal	Financial Services
Toyota	Automotive - Vehicles	MINI	Automotive - Vehicles
EA	Video Games	CharlesSchwab	Investment Management
NextEraEnergy	Energy	DaVita	Healthcare Services
Airbnb	Online Marketplace & Hospitality	sheratonhotels	Hotel
tacobell	Restaurants	Twitter	Social Media
PPLCorp	Energy	Enterprise	Vehicle Rental
RalphLauren	Luxury Fashion	LithiaMotors	Retail - Automotive
ProcterGamble	Consumer Goods	CMSenergy	Energy
Clinique	Skincare and Cosmetics	Danone	Food Production
AsburyAutoGroup	Retail - Automotive	Nordstrom	Department Store
AltriaNews	Tobacco and Wine	MaxFactorUK	Skincare and Cosmetics
TIAA	Financial Services	RamadaWorldwide	Hotel
AvonInsider	Skincare and Cosmetics	PNCBank	Banking
LillyPad	Pharmaceutical	Hilton	Hotel
OshkoshDefense	Defense and Heavy Equipment Manufacturing	TommyHilfiger	Fashion & Apparel
NewYorkLife	Insurance	AEPnews	Energy
Expedia	Online Travel Agency	Staples	Retail - Office Supplies
VFCorp	Footwear & Outdoor Apparel	RaymondJames	Financial Services
autozone	Retail - Automotive	AIGinsurance	Insurance

Continued on next page

Table 10 – continued from previous page

Brand	Category	Brand	Category
CocaCola	Beverages - NonAlcoholic	KraftHeinzCo	Food Production
salesforce	Technology - Software	Allstate	Insurance
panerabread	Restaurants	Lenovo	Technology - Electronics
Disney	Mass Media and Entertainment	LeeJeans	Denim Apparel
PGE4Me	Energy	netflix	Streaming Services
Edison_Electric	Energy	dominos	Restaurants
MaytagBrand	Home Appliances	RegionsBank	Banking
FirstEnergyCorp	Energy	Safeway	Retail - Grocery
MichaelKors	Luxury Fashion	TheHartford	Insurance
CanonUSA	Imaging and Optical Products	LifePointHealth	Healthcare Services
MarriottIntl	Hotel	KitchenAidUSA	Kitchen Appliances
SamsungUS	Technology - Electronics	Xfinity	Telecommunications
Budweiser	Beverages - Alcoholic	21CF	Mass Media and Entertainment
westerndigital	Technology - Data Storage	WellCare_Health	Healthcare Insurance
Mastercard	Financial Services	Neutrogena	Skincare and Cosmetics
VW	Automotive - Vehicles	WynnLasVegas	Resort & Casino
StanleyBlkDeckr	Tools and Storage	UHS_Inc	Healthcare Services
FidelityNhw	Investment Management	HyattRegency	Hotel
Xbox	Video Games	amfam	Insurance
LincolnMotorCo	Automotive - Vehicles	InsidePMI	Tobacco
espn	Sports Broadcasting	Campbells	Food Production
intel	Technology - Semiconductors	SanDisk	Technology - Data Storage
eBay	E-commerce	Merck	Pharmaceutical
LandOLakesKtchn	Food Production	BACARDI	Beverages - Alcoholic
Food4Less	Retail - Grocery	TenetHealth	Healthcare Services
caseysgenstore	Retail - Grocery	GileadSciences	Biotechnology
WECEnergyGroup	Energy	GuardianLife	Insurance
Centene	Healthcare Services	DollarTree	Discount Retail
Principal	Financial Services	GreyGoose	Beverages - Alcoholic
Group1Auto	Retail - Automotive	AmericanAir	Airline
Timberland	Footwear & Outdoor Apparel	Gatorade	Beverages - NonAlcoholic
yumbrands	Restaurants	SherwinWilliams	Paints and Coatings
Folgers	Beverages - NonAlcoholic	HBO	Entertainment - Television Network
JackBox	Restaurants	pfizer	Pharmaceutical
PFGC	Food Distribution	LibertyMutual	Insurance
NVHomes1979	Real Estate Services	kroger	Retail - Grocery
Aetna	Healthcare Insurance	BurgerKing	Restaurants
RoyalCaribbean	Cruise Line	Adobe	Technology - Software
Dillards	Department Store	SunocoRacing	Oil & Gas
TracfoneCalls	Telecommunications	MandT_Bank	Banking
Travelers	Insurance	Microsoft	Technology - Software
Sprite	Beverages - NonAlcoholic	CBRE	Real Estate Services
Apple	Technology - Electronics	AssurantNews	Insurance

Continued on next page

Table 10 – continued from previous page

Brand	Category	Brand	Category
Voya	Financial Services	Allianz	Financial Services
AmerenCorp	Energy	Honeywell	Diversified Technology and Manufacturing
Calpine	Energy	blackstone	Investment Management
LasVegasSands	Resort & Casino	GEICO	Insurance
officedepot	Retail - Office Supplies	erie_insurance	Insurance
DRHorton	Home Construction	thenorthface	Retail - Outdoor Apparel & Equipment
AutoOwnersIns	Insurance	Heineken_US	Beverages - Alcoholic
MetLife	Insurance	Citi	Banking
SouthwestAir	Airline	Philips	Technology - Electronics
AmericanExpress	Financial Services	JohnHancockUSA	Insurance
DIRECTV	Satellite Television	FoxNews	News Broadcasting
NiSourceInc	Energy	RAI_News	Tobacco
LinkedIn	Professional Networking	Seagate	Technology - Data Storage
TractorSupply	Retail - Farm and Ranch Supplies	HolidayInn	Hotel
priceline	Online Travel Agency	MountainDew	Beverages - NonAlcoholic
Windstream	Telecommunications	Delta	Airline
GoldmanSachs	Investment Banking	QVC	Retail - Television and Online
OldNavy	Fashion & Apparel	CoorsLight	Beverages - Alcoholic
JLL	Real Estate Services	Progressive	Insurance
Target	Department Store	MACcosmetics	Skincare and Cosmetics
tridentgum	Confectionery	Purina	Retail - Pet Supplies
healthnet	Healthcare Insurance	AbbottNews	Pharmaceutical
Exelon	Energy	massmutual	Insurance
FreddieMac	Financial Services	blackrock	Investment Management
DelekUSHoldings	Oil & Gas	Sony	Technology - Electronics
usbank	Banking	WesternUnion	Financial Services
ToysRUs	Retail - Toys	united	Airline
Humana	Healthcare Insurance	marshalls	Discount Retail
LKQCorp	Automotive Parts	familydollar	Discount Retail
panasonic	Technology - Electronics	KelloggsUS	Food Production
CBS	Broadcasting and Media	johnsonsbaby	Healthcare Products
Facebook	Social Media	GM	Automotive - Vehicles
CourtyardHotels	Hotel	CadburyWorld	Food Production
TripAdvisor	Online Travel Agency	ultabeauty	Retail - Beauty and Cosmetics
newell_brands	Consumer Goods	jcpenney	Department Store
Prada	Luxury Fashion	Publix	Retail - Grocery
MDLZ	Food Production	AlaskaAir	Airline
Ford	Automotive - Vehicles	RepublicService	Waste Management
TiffanyAndCo	Luxury Jewelry	Doritos	Food Production
DTE_Energy	Energy	SingaporeAir	Airline
MarsGlobal	Confectionery	Chubb	Insurance
chevrolet	Automotive - Vehicles	Burberry	Luxury Fashion
Navient	Financial Services	biogen	Biotechnology

Continued on next page

Table 10 – continued from previous page

Brand	Category	Brand	Category
GeneralMills	Food Production	Nestle	Food Production
McDonalds	Restaurants	exxonmobil	Oil & Gas
UnderArmour	Sportswear	MercedesBenz	Automotive - Vehicles
adidasUS	Sportswear	HSBC	Banking
MGMResortsIntl	Resort & Casino	flowersfoods	Food Production
MarriottBonvoy	Hotel	AllyFinancial	Financial Services
Quaker	Food Production	OptimumHelp	Telecommunications
PetSmart	Retail - Pet Supplies	WeAreFarmers	Insurance
Westin	Hotel	cvspharmacy	Retail - Pharmacy and Health
Moet_UK	Beverages - Alcoholic	EsteeLauder	Skincare and Cosmetics
BestBuy	Retail - Electronics	Entergy	Energy
jpmorgan	Investment Banking	HPE_News	Technology - Software/Hardware
Listerine	Oral Care	AutoNation	Retail - Automotive
Yahoo	Technology - Internet Services	bathbodyworks	Retail - Personal Care
Realogy	Real Estate Services	Cadillac	Automotive - Vehicles
VictoriasSecret	Lingerie & Women's Apparel	amazon	E-commerce
pizzahut	Restaurants	SIRIUSXM	Satellite Radio
mohawkind	Flooring Manufacturing	Kia	Automotive - Vehicles
Ross_Stores	Discount Retail	Tropicana	Beverages - NonAlcoholic
Allergan	Pharmaceutical	hm	Fashion & Apparel
ChipotleTweets	Restaurants	Walmart	Department Store
Alphabet	Technology - Internet Services	pepsi	Beverages - NonAlcoholic
nrgenergy	Energy	nvidia	Technology - Semiconductors
Xerox	Technology - Printing and Digital Services	ConagraBrands	Food Production
Lennar	Home Construction	Hyatt	Hotel
AdvanceAuto	Retail - Automotive	Chevron	Oil & Gas
Uber	Ride-Hailing	NMFinancial	Financial Services
Macys	Department Store	JackDaniels_US	Beverages - Alcoholic
DukeEnergy	Energy	SouthernCompany	Energy
NBC	Broadcasting and Media	gucci	Luxury Fashion
PedigreeUS	Pet Care	StJude	Healthcare Services
Avis	Vehicle Rental	Sears	Department Store
GMC	Automotive - Vehicles	FedEx	Logistics and Courier
mutualofomaha	Insurance	ZARA	Fashion & Apparel
Cigna	Healthcare Insurance	smuckers	Food Production
Hertz	Vehicle Rental	PPG	Coatings and Specialty Materials
IBM	Technology - Software/Hardware	FifthThird	Banking
Fossil	Fashion & Apparel	Porsche	Automotive - Vehicles
BedBathBeyond	Retail - Home Improvement	CitizensBank	Banking
UMG	Music Entertainment	Clorox	Cleaning and Household Products
AceHardware	Retail - Hardware	Garmin	GPS Technology and Wearables
TheASF	Technology - Software	Visa	Financial Services
CapitalOne	Banking	ConEdison	Energy

Continued on next page

Table 10 – continued from previous page

Brand	Category	Brand	Category
LandRoverUSA	Automotive - Vehicles	Nationwide	Insurance
WasteManagement	Waste Management	WhirlpoolCorp	Home Appliances
ValeroEnergy	Oil & Gas	MoodysInvSvc	Financial Services
DollarGeneral	Discount Retail	BankofAmerica	Banking
Yoplait	Food Production	DeanFoods	Food Production
PenskeCars	Retail - Automotive	Starbucks	Restaurants
Aflac	Insurance	pacificlife	Insurance
Hersheys	Confectionery	AnthemBCBS	Healthcare Insurance
WholeFoods	Retail - Grocery	DISH	Satellite Television
washingtonpost	News and Media	Oracle	Technology - Software
BNYMellon	Banking	HormelFoods	Food Production
TATravelCenters	Travel Centers/Service Stations	UPS	Logistics and Courier
ExpressScripts	Healthcare Services	CNPalerts	Energy
Tesla	Automotive - Vehicles	sonicautomotive	Retail - Automotive
Colgate	Oral Care	Audi	Automotive - Vehicles
kfc	Restaurants	footlocker	Retail - Footwear
NissanUSA	Automotive - Vehicles	Cartier	Luxury Jewelry and Watches
YouTube	Video Sharing Platform	MorganStanley	Investment Banking
Kohls	Department Store	SantanderBankUS	Banking
Cisco	Technology - Networking	Optum	Healthcare Services
Honda	Automotive - Vehicles	DoubleTree	Hotel

WEB APPENDIX B: GOODNESS OF FIT AND MCMC DIAGNOSTICS FOR ERGM MODEL

In this appendix, we present a detailed analysis of the Goodness of Fit and MCMC diagnostic outcomes for the ERGM model discussed in the main text of the paper.

Goodness of Fit Plots In our ERGM analysis, Goodness of Fit plots were generated to evaluate how well the model captures various network characteristics. These characteristics include model statistics, degree distribution, edgewise shared partners, and minimum geodesic distance in Figure 6. These plots are crucial as they compare the observed network data with data simulated from our ERGM, providing a visual and quantitative measure of the model's accuracy.

MCMC Diagnostic Plots As shown in Figure 7, the MCMC diagnostic process incorporates both trace and density plots to assess the convergence and stability of the Markov Chain Monte Carlo simulations used for ERGM parameter estimation. Trace plots are important in evaluating convergence, as they illustrate the sampled values over iterations. An effectively converged trace plot demonstrates a stable, fluctuating pattern, indicating a thorough exploration of the parameter space. Density plots complement this by depicting the distribution of the sampled values, shedding light on the stability and dispersion of the parameter estimates. Collectively, these diagnostic plots are essential for verifying the reliability and accuracy of our model's parameter estimates, thereby strengthening the model's overall validity.

Figure 6: Goodness of fit of ERGM model

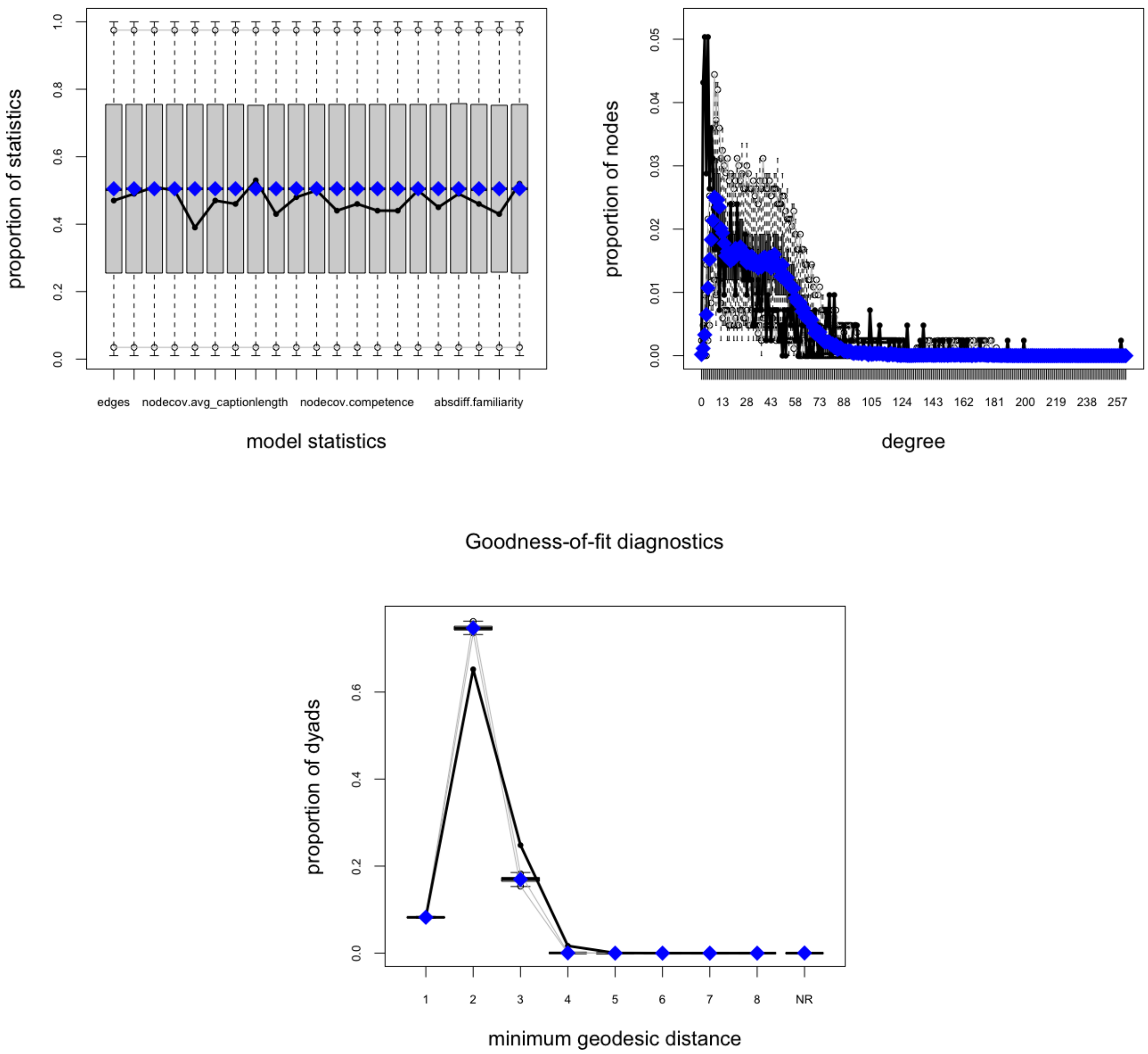
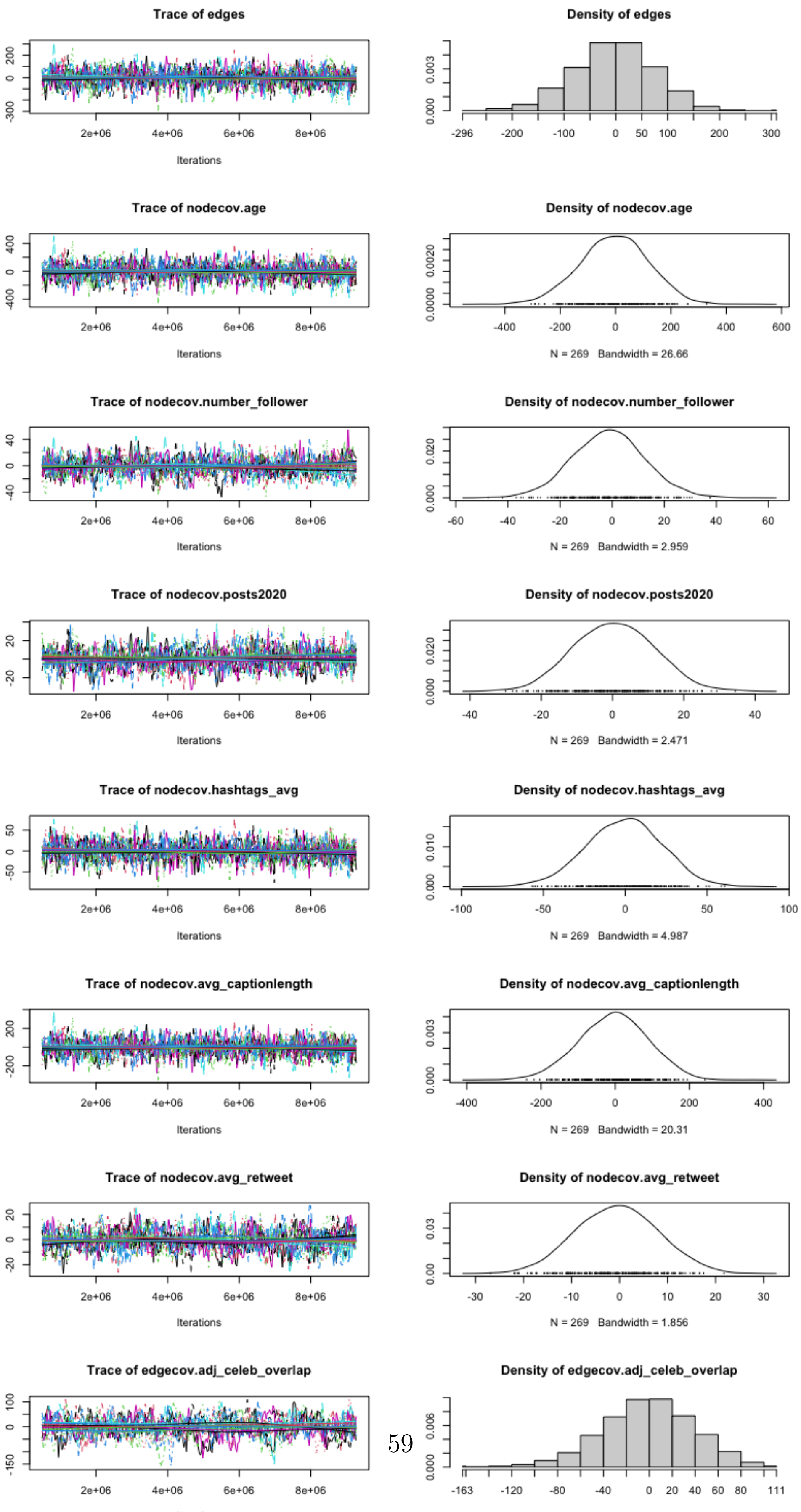
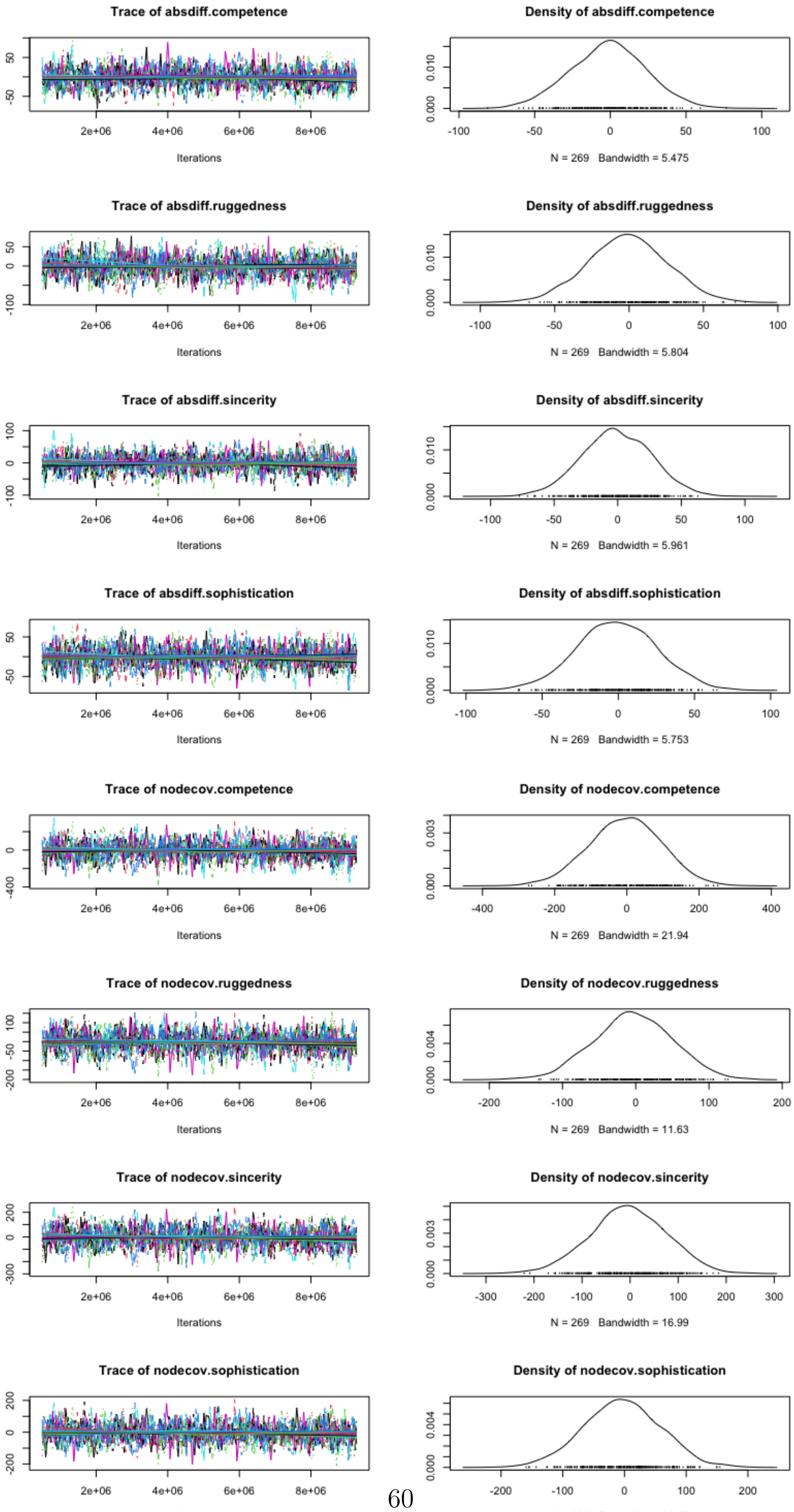
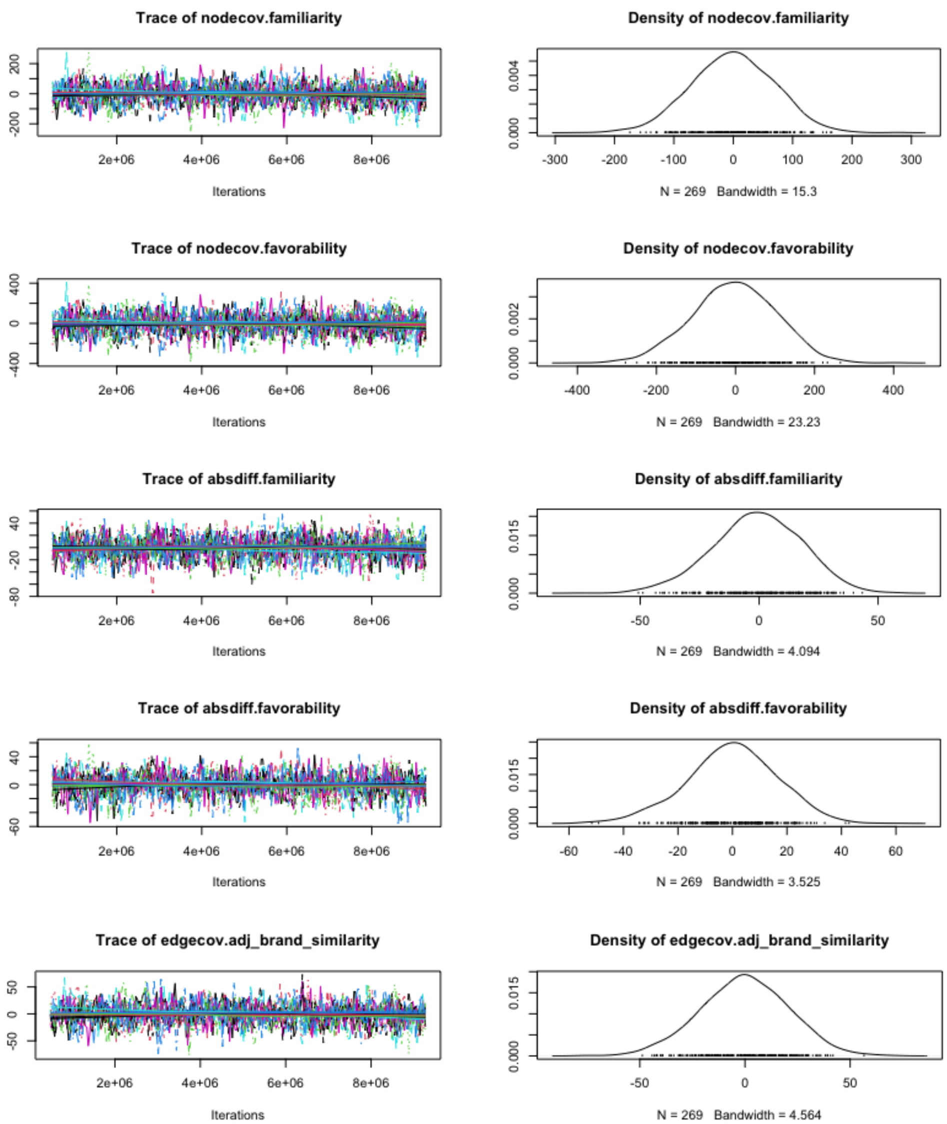


Figure 7: MCMC diagnostic of ERGM model





Trace of absdiff.competence
Iterations
Density of absdiff.competence
N = 269 Bandwidth = 5.475
Trace of absdiff.ruggedness
Iterations
Density of absdiff.ruggedness
N = 269 Bandwidth = 5.804
Trace of absdiff.sincerity
Iterations
Density of absdiff.sincerity
N = 269 Bandwidth = 5.961
Trace of absdiff.sophistication
Iterations
Density of absdiff.sophistication
N = 269 Bandwidth = 5.753
Trace of nodecov.competence
Iterations
Density of nodecov.competence
N = 269 Bandwidth = 21.94
Trace of nodecov.ruggedness
Iterations
Density of nodecov.ruggedness
N = 269 Bandwidth = 11.63
Trace of nodecov.sincerity
Iterations
Density of nodecov.sincerity
N = 269 Bandwidth = 16.99
Trace of nodecov.sophistication
Iterations
Density of nodecov.sophistication
N = 269 Bandwidth = 13.47


Trace of nodecov.familiarity
Iterations
Density of nodecov.familiarity
N = 269 Bandwidth = 15.3
Trace of nodecov.favorability
Iterations
Density of nodecov.favorability
N = 269 Bandwidth = 23.23
Trace of absdiff.familiarity
Iterations
Density of absdiff.familiarity
N = 269 Bandwidth = 4.094
Trace of absdiff.favorability
Iterations
Density of absdiff.favorability
N = 269 Bandwidth = 3.525
Trace of edgecov.adj_brand_similarity
Iterations
Density of edgecov.adj_brand_similarity
N = 269 Bandwidth = 4.564

WEB APPENDIX C: SURVEY INSTRUMENT FOR VALIDATING BANE

We recruited 500 survey participants through Prolific to examine whether opportunities identified by BANE were rated as more appealing than those eliminated. A nationally representative sample of participants was used on Prolific to ensure all demographics were equally represented. Participants first read a brief description of the meaning of brand alliances. They were then randomly assigned three sets of questions, resulting in each participant rating a total of 18 brand alliance opportunities using a 5-point Likert scale. The format of the questions in the survey is included in Table 11 below.

Table 11: Survey Measures and Corresponding Questions

	Measure	**Format in Survey**
Brand Alliance Evaluation	Appeal of Brand Alliances	Consider the following brand pairs that have formed alliances to offer combined products, services, or promotions. Rate these alliances on a scale from 1 to 5, where 1 means 'Not at all appealing' and 5 means 'Very appealing.' Your rating should be based on your personal preferences and the perceived value these alliances could provide you as a consumer.
Demographics Questions	Gender	Please indicate your gender.
	Age	Please indicate your age.
	Income	Please indicate your individual income.

The Wilcoxon–Mann–Whitney test confirms that participants systematically rated the premier opportunities identified in the BANE matrix higher than the eliminated ones (W *statistics* $= 6019374$, *Rank difference* $= 3956670$, $p < .001$). The findings support the external validity of BANE, demonstrating that it accurately identifies premier opportunities from information networks.

Next, to ensure that the observed ratings for premier opportunities are not confounded by demographic variables such as age, gender, and income, we conducted a mixed-effects regression analysis. This approach allows us to adjust for these potential biases by incorporating both fixed effects to assess the impact of these demographic factors and random effects to account for individual differences in how participants respond. The model is specified as follows:

$$y_{ij} = \beta_0 + \beta_1 \text{Treated}_{ij} + \beta_2 \text{Age}_{ij} + \beta_3 \text{Gender}_{ij} + \beta_4 \text{Income}_{ij} + u_{0j} + u_{1j} \text{Treated}_{ij} + \epsilon_{ij} \quad (7)$$

where:

- y_{ij} is the rating for brand alliance appeal given by respondent j on brand pair i.
- β_0 is the overall intercept.
- $\beta_1, \beta_2, \beta_3, \beta_4$ are the fixed effects of treatment (is a premiere opportunity), age, gender, and income respectively.
- u_{0j} is the random intercept for each respondent, capturing unobserved heterogeneity at the individual level.
- u_{1j} is the random slope for the treatment effect, allowing the impact of treatment to vary across respondents.
- ϵ_{ij} is the residual error.

The regression results, presented in Table 12, reveal a significant positive effect for premier opportunities (.603, $p < .001$), indicating that consumers generally find these opportunities more appealing than those eliminated earlier. The corresponding Cohen's d value of .435, while classified as a small effect size, approaches the threshold of a medium effect (.5), highlighting its practical significance in this context. This proximity to a medium effect size is notable, especially considering the diversity and complexity of consumer preferences captured in our nationally representative sample.

Table 12: Mixed-Effects Regression Analysis for Appeal of Brand Alliance Opportunities

Variable	**Coefficient (Std. Err.)**
Treated	.603*** (.031)
Age [T.35-54]	.114 (.089)
Age [T.55-74]	-.039 (.086)
Age [T.75 or older]	.048 (.250)
Gender [T.Male]	.041 (.070)
Gender [T.Non-binary / third gender]	-.513* (.293)
Income [T.$100,000 - $149,999]	.026 (.149)
Income [T.$20,000 - $29,999]	.075 (.154)
Income [T.$30,000 - $39,999]	.136 (.147)
Income [T.$40,000 - $49,999]	.039 (.165)
Income [T.$50,000 - $59,999]	.433** (.160)
Income [T.$60,000 - $69,999]	.209 (.158)
Income [T.$70,000 - $79,999]	.319* (.161)
Income [T.$80,000 - $89,999]	.197 (.215)
Income [T.$90,000 - $99,999]	.038 (.192)
Income [T.Less than $10,000]	-.002 (.142)
Income [T.More than $150,000]	.044 (.170)
Intercept	2.333*** (.122)
Number of Observations	8020

Significance level: ** p < .05, ** p < .01, *** p < .001*

WEB APPENDIX D: CONSTRUCTING IMPLICIT INFORMATION NETWORKS FROM CO-COMMENTS AND CO-FOLLOWING DATA

Our objective is to compare implicit and explicit information networks in terms of how their structures align with brand attributes that are indicative of favorable brand alliance partnerships. We collected co-commenting, co-following, and co-mentioning data on 160 brands. We then construct implicit information networks using the approaches proposed by (Malhotra and Bhattacharyya 2022) and Yang, Zhang, and Kannan (2022), as well as the approach outlined in BANE's first Phase. This web appendix details the construction of the implicit information networks.

DATA COLLECTION AND PRUNING

Co-commenting and co-following data was collected for 160 brands from Twitter in 2021, resulting in bipartite networks where nodes are users and brands, and weighted edges representing comments (following) of brands by users. Neither user nodes nor brand nodes are directly connected by edges. Following Yang, Zhang, and Kannan (2022), we pruned the dataset to eliminate users that only co-comment or co-follow very few brands which is not informative for the purpose of identifying latent connections among brands. The resulting dataset comprises 118,648 unique users, 159 unique brands, and 484,113 weighted edges.

AUTOENCODER MODEL

We configure an autoencoder as described by Yang, Zhang, and Kannan (2022) to jointly learn low-dimensional embeddings $(\mathbf{z})$ of dimension 128 for both brands and users. The architecture of the neural network is as follows:

Encoder

$$\mathbf{z} = f(\mathbf{x}) = \text{ReLU}(W_2 \cdot \text{ReLU}(W_1 \cdot \mathbf{x} + b_1) + b_2)$$

where:

- $\mathbf{x}$ is the input (a one-hot encoded vector),
- W_1 (2048 dimensions) and W_2 (512 dimensions) are weight matrices.

Decoder

$$\hat{\mathbf{x}} = g(\mathbf{z}) = \sigma(W_4 \cdot \text{ReLU}(W_3 \cdot \mathbf{z} + b_3) + b_4)$$

where W_3 (512 dimensions) and W_4 (2048 dimensions) are weight matrices, b are bias terms, ReLU is an activation function that introduces non-linearity to the model (ReLU(x)=max(0,x)), and σ is the sigmoid activation function used in the final layer to ensure that the output values are in the range [0, 1], making them suitable for representing probabilities.

Custom Loss Function

The configured autoencoder was trained using the custom loss function described by y Yang, Zhang, and Kannan (2022) that combines two components:

First-Order Proximity

$$\mathcal{L}_{\text{first-order}} = \sum_{(i,j)\in \text{edges}} A_{ij} \|z_i - z_j\|^2$$

This component ensures that directly connected nodes (users and brands) in the co-commenting network are embedded close to each other in the latent space. The purpose is to preserve the local neighborhood structure of the original network, so brands that are frequently co-commented by users are closer in the embedding space.

Second-Order Proximity

$$\mathcal{L}_{\text{second-order}} = \text{MSE}(\mathbf{x}, \hat{\mathbf{x}})$$

This component measures the mean squared error (MSE) between the input vector $\mathbf{x}$ and its reconstruction $\hat{\mathbf{x}}$. This term captures the global structure of the network, ensuring that the autoencoder accurately reconstructs the input data, reflecting the overall pattern of brand and user relationships.

Total Loss

$$\mathcal{L} = \mathcal{L}_{\text{first-order}} + \lambda \cdot \mathcal{L}_{\text{second-order}}$$

Following Yang, Zhang, and Kannan (2022), we set $\lambda = 1$ to balance the first-order and second-order proximity components.

Training Configuration

As in Yang, Zhang, and Kannan (2022), the autoencoder was trained using the Adam optimizer[8]. To mitigate the risk of overfitting, we added early stopping during training, halting the process when the validation loss did not decrease further for three consecutive epochs.

CO-FOLLOWING

The co-followership network is constructed following (Malhotra and Bhattacharyya 2022). The network is defined as:

$$G_F = (V_F, E_F, W_F) \tag{8}$$

Where – G_F is the co-follow graph, V_F is the set of nodes, where each node in V_F represents a brand in our analysis; E_F is the set of edges, where an edge (u_F, v_F) exists if a consumer co-follows both brands u_F and v_F; and $W_F : E_F \to \mathbb{R}^+$ is a weight function that assigns a positive real number to each edge in E_F, indicating the number of consumers who follow both brands. Popular brands may have a large number of followers, which can inflate their edge weight in the network. To mitigate this popularity bias, we normalize the edge weight by the total number of followers for each brand (Culotta and Cutler 2016).

[8]Hyperparameters: learning rate = 2e-4, dropout rate = 0.1, batch size = 1048, epochs = 50, lambda = 1.