Proceedings of the ASME 2020 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2020
August 16 - 19, 2020, St.Louis, MO, USA

DETC2020-22591

A WEIGHTED NETWORK MODELING APPROACH FOR ANALYZING PRODUCT COMPETITION

Yaxin Cui
Dept. of Mechanical Engineering
Northwestern University
Evanston, IL
Email: yaxincui2023@u.northwestern.edu

Faez Ahmed
Dept. of Mechanical Engineering
Northwestern University
Evanston, IL
Email: faez@northwestern.edu

Zhenghui Sha
Dept. of Mechanical Engineering
University of Arkansas
Fayetteville, AR
Email: zsha@uark.edu

Lijun Wang
Insight and Analytics
Ford Motor Company
Dearborn, MI
Email: lwang149@ford.com

Yan Fu
Insight and Analytics
Ford Motor Company
Dearborn, MI
Email: yfu4@ford.com

Wei Chen*
Dept. of Mechanical Engineering
Northwestern University
Evanston, IL
Email: weichen@northwestern.edu

ABSTRACT

Statistical network models allow us to study the co-evolution between the products and the social aspects of a market system, by modeling these components and their interactions as graphs. In this paper, we study competition between different car models using network theory, with a focus on how product attributes (like fuel economy and price) affect which cars are considered together and which cars are finally bought by customers. Unlike past work, where most systems have been studied with the assumption that relationships between competitors are binary (i.e., whether a relationship exists or not), we allow relationships to take strengths (i.e., how strong a relationship is). Specifically, we use valued Exponential Random Graph Models and show that our approach provides a significant improvement over the baselines in predicting product co-considerations as well as in the validation of market share. This is also the first attempt to study aggregated purchase preference and car competition using valued directed networks.

*Address all correspondence to this author.

KEYWORDS

Product competition, Customer preference modeling, Exponential Random Graph Model, Weighted networks

1 INTRODUCTION

To make rational product design decisions, there is an increasing need for a better understanding of the dynamic interactions among different stakeholders in a market system. Among all types of interactions, competitive relations between similar products from multiple producers is one of the most important ones. For example, in the auto market, electric vehicles (EVs) are emerging as competitors of traditional vehicles and other alternative fuel vehicles. In 2017, the Chinese market accounted for more than half of all global EV sales, a 53% increase compared to 2016 [1]. However, only 4% of the EVs sold in China originated from the U.S. [2]. To improve U.S. global competitiveness, U.S. automakers must design EVs that are tailored to the preferences of Chinese customers. Not only does the success of a new EV design depend on its engineering performance, but it must also consider the dynamic competition among multiple competitors.

TABLE 1: Comparison of our work with past work on using network approaches for customer preference and car competition modeling

Paper	Network Type	Undirected	Directed	Weighted	Goal
Wang *et al.* (2015) [3]	Multidimensional	✓	✓	×	Introducing multidimensional network representation
Wang *et al.* (2016) [4]	Unidimensional	✓	×	×	Forecasting the existence of cars' co-consideration relationship
Fu *et al.* (2017) [5]	Bipartite	✓	✓	×	Modeling consider-and-then-choice behaviors
Sha *et al.* (2018) [6]	Unidimensional	✓	×	×	Comparing ERGM model with Dyiad model
Wang *et al.* (2018) [7]	Unidimensional	✓	×	×	Predicting product co-consideration relationships
Bi *et al.* (2018) [8]	Bipartite	✓	×	×	Studying spatiotemporal heterogeneity of customer preferences
Xie *et al.* (2019) [9]	Unidimensional	✓	×	×	Dynamic ERGM model
Sha *et al.* (2019) [10]	Bipartite	✓	×	×	Comparing network models with Discrete Choice Analysis
This Paper	Unidimensional	✓	✓	✓	Modeling product competition using weighted networks

Competition has been traditionally studied using game-theoretic models [11, 12] to support engineering designs that involve multiple decision-makers. In terms of enterprises' decisions, the studies generally focus on three types of strategies: (1) pricing strategy [13], (2) design configuration decisions in either single product design [14] or product line design [15], (3) strategic decisions on product innovation [16]. However, existing work does not explicitly take into account customer preferences and product design features in the formulation of payoff or utility function. Moreover, the competition analyzed in existing literature typically involves two or a handful of players instead of the entire market. As a result, the impact of complex market relations among entities beyond the competitors being investigated cannot be assessed. To address these issues, an approach that can model both customer preferences and complex competitive relations are needed. In the following, we briefly examine the work done in network-based customer preference modeling that aims to quantitatively model complex relations in a market system.

Customer preference modeling emphasizes the understanding of how customers make trade-offs among multiple attributes when making purchase decisions [17, 18]. While customer preferences can be modeled using many approaches [19], the network analysis approach has proven effective in facilitating understanding of customer-product relations [20, 21]. Compared to traditionally used utility-based approach [22], network approaches exhibit advantages in considering decision dependency, attributes collinearity, customer irrationality and missing information of choice set. Recent studies of network approach explored the capability of statistical network models (Exponential Random Graph Models - ERGMs) in studying customers' consideration behaviors [23], forecasting the impact of technological changes on market competitions [4], modeling customers' consideration-then-choice behaviors [5], and predicting products' co-consideration relations [6, 7].

Despite various attempts of using network models and theories in better understanding the driving factors in customers' consideration and choice behaviors, existing studies have several limitations. First, the networks are simplified as binary networks meaning that the weights/strength of links is neglected. Part of the reasons for such treatment was due to the authors' initial focus on firstly understanding the feasibility of using networks in modeling customer preferences. However, the link strength is an important aspect for understanding product competitions as well as customer preferences. This is because the prediction of link strength will enable designers to more precisely evaluate the effects of potential designs on the market demand compared to merely predicting the existence of links. For example, if a threshold is set to binarize the network links, then any links with a strength lower than the pre-defined threshold will not be considered. To probe into the question of *how much* a competition relation between two products could be changed because of the change of designs or customer preferences, the link strength must be explicitly modeled. Second, most, if not all, past research on network models on car competition analysis do not use directed networks for modeling the final choice decision of a product, but instead only focus on the first stage of decision making, where customers consider a set of products. Our work attempts to address these limitations. Unlike past approaches, our work is the first to use weighted networks as well as study both choice (directed network) and consideration (undirected network) using it. Table 1 summarizes the existing works of using network analysis in customer preference modeling and how our work differs from them.

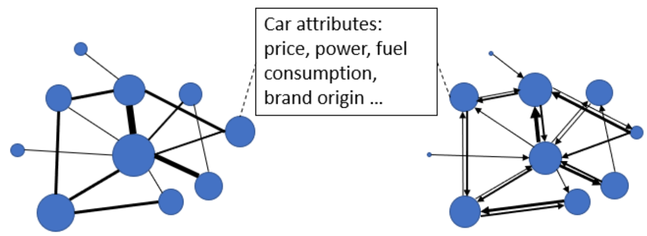

FIGURE 1: We use valued-ERGM network models to study car competition for consideration stage (left network) and choice stage (right network). The nodes represent cars in these network illustrations and links are competition strength.

The new approach we propose in this study is based on valued-ERGM models that allow a link between nodes to carry weights and such a link can be either directed or undirected. In a unidimensional car competition network, we study both customers' consideration and choice behaviors by establishing two types of networks as illustrated in Figure 1 – an undirected network, in which links represent co-consideration relationship and a directed network, in which directed links between cars which are co-considered indicates customers' aggregated preference towards the final choice (or purchase) decisions. The relaxation of the binary link assumption enables us to better understand production competitions. We also demonstrate a way to represent choices between different products using a directed network. As a summary, the **objectives of this research** are two-fold: a) to extend valued-ERGM to both weighted co-consideration and choice networks for modeling product competitions; and b) to evaluate the benefits of valued-ERGM for studying the attribute and network effects, and for network prediction when node attributes (product design) change.

2 TECHNICAL BACKGROUND

Exponential Random Graph Model (ERGM), a statistical analysis technique which serves as a formal representation of the network formation process [24], has been a popular choice in social network research. ERGMs provide a probability for every possible network that can be formed from a fixed number of nodes. This leads to a probability distribution on the set of all possible networks with the same number of nodes [24]. Mathematically, ERGMs can be expressed as a function of a set of input parameters (which can be node properties, edge properties, network structure attributes *etc.*) [25], as shown below:

$$Pr(\mathbf{Y} = \mathbf{y}) = \frac{exp(\theta^T \cdot g(\mathbf{y}))}{\kappa(\theta, \mathbf{y})} \tag{1}$$

In this equation, network structure $\mathbf{Y}$ is treated as a random variable and an observed network $\mathbf{y}$ is the network data the researcher has collected and regarded as one realization from a set of possible networks. The probability of the observed network structure is determined by network statistics $g(\mathbf{y})$, which can include attributes of nodes, attributes of links as well as network structure attributes, and corresponding model parameters θ (a vector). $\kappa(\theta, \mathbf{y})$ is a normalizing constant, to make the function a realistic probability value. Eq. 1 suggests that the probability of observing a specific network structure is proportional to the exponent of a weighted combination of network statistics [9]. To estimate the parameters (or learn the model from existing data), a Markov chain Monte Carlo (MCMC) procedure using maximum likelihood estimation [26] is typically used. The estimated parameters θ indicate the importance of different network statistics in the formation of links in a network. By analyzing the magnitude and statistical relevance of θ, one can interpret the factors which may be important in the network formation process.

ERGMs has several advantages over traditional statistical models. For instance, unlike traditional logit models [27], they allow the interdependence among network edges, which is more realistic in many network formation processes. ERGMs also provide a flexible statistical inference framework that can model the influence of both exogenous effects (*e.g.* nodal attributes) and endogenous effects (*e.g.* network structures and the relationship between nodes, such as 3-way product competition) on the probability of forming a connection between nodes.

A limitation of traditional binary ERGM is that it cannot model real-world networks with weighted links (*e.g.* flight connection intensity between two airports). If one wishes to model a weighted network with a traditional ERGM, they have to first binarize the network with a link weight threshold. This process converts each edge to a binary 0 or 1 link so that the ERGM can take the resultant network as input. This dichotomization step may lead to biases and information loss, which can eventually affect network prediction. Valued-ERGM [28], a technique recently developed in social network modeling addresses this limitation by modeling the strength of links rather than merely their presence or absence. For a given set of discrete variables, a valued-ERGM is expressed as:

$$Pr(\mathbf{Y} = \mathbf{y}) = \frac{h(\mathbf{y})exp(\theta^T \cdot g(\mathbf{y}))}{\kappa(\theta, \mathbf{y})}, \ \mathbf{y} \in \mathbb{Y} \tag{2}$$

where $\mathbb{Y}$ is called the support and $\kappa(\theta, \mathbf{y})$ is a normalizing constant, to make the function a feasible probability value. By comparing Eq. 1 with Eq. 2, one may notice two major differences — the presence of a support $\mathbb{Y}$ term and a reference distribution $h(\mathbf{y})$ term.

Different from binary ERGMs, the support of a valued-ERGM is over a set of valued graphs, which is often infinite or uncountable [29]. One cannot enumerate all possible weighted networks with real-valued link strengths. Thus in a weighted network case, we need to consider what the strength of connections are and how they are distributed. This brings in the need of specifying a reference distribution, which determines the sample space and baseline distribution of link values. The sample space is the set of possible networks given the size and density of the observed network. A reference distribution simply answers the question of what might the link distribution looks like in the absence of any model terms. The ability to model valued links has greatly advanced network research as it enables researchers to conduct more nuanced examinations of network patterns. Moreover, similar to traditional ERGMs, valued-ERGMs are capable of modeling networks with both undirected links and directed links.

Valued-ERGMs have been employed in various applications ranging from policy studies [30], organizational communication [31] to disease transmissions [32] and global migration [33]. An important step of using valued-ERGM is to first specify meaningful links between nodes and define a way to measure the link strength. The definition of link strength often depends on the domain, and in the past, researchers have determined it based on factors ranging from the level of interactions between two nodes [30], the strength of friendship [31], or the total duration of human contact [32]. These links, although valued, are typically discrete in a small range such as $\{0,1,2,3\}$. Existing methods in the social science area cannot be directly used in our study to model the valued product association networks due to: a) the link strength in the product competition networks has a substantially large range. This infinite sample space increases the complexity of the task; and b) existing studies mainly concentrate on interpreting the models, whereas we focus on both interpretation and prediction. Therefore, a new approach and validation methods for a rigorous evaluation of valued-ERGM models are needed.

3 METHODOLOGY

In a product market, the number of customers considering a pair of products (u and v) or choosing one product over the other reflects the in-between competitive strength. To capture the product competition strength based on customers' considerations and choices, we build weighted product competition networks and model them with valued-ERGMs. In this section, we outline the typical steps required for statistical modeling of a weighted competition network, which contains three main steps: 1. Create a weighted network, 2. Train a valued-ERGM model and interpret the effects of the parameters on competition, and 3. Make predictions about competition among products in the future.

While the key contribution of this work is in the choice of the modeling method, we will describe the step-by-step process of building a weighted network and analyzing it in this section. We will use cars as an example to illustrate the method, but the concepts can be generalized to many other product designs.

3.1 Weighted Product Competition Network Construction

Networks are capable of mathematically and graphically representing the product competition structure based on survey data. To capture different stages of a customer's decision-making process, we build two different unidimensional networks, which we call the "co-consideration network" and the "choice network". The first is an undirected network that represents the alternatives in the consideration stage and the second is a directed network, which represents the customer's aggregated choice behavior.

Defining link strengths for a co-consideration network A unidimensional product network can reveal product market competition by describing products' co-consideration or preference relationship. In both networks, a product (in this case, a car) corresponds to a node. Each node is associated with a set of attributes like price, fuel consumption and engine power. In the co-consideration network, we define an undirected link between node u and node v, if there exists at least one customer who considers both cars u and v together. The number of customers who consider the two cars together is set as the weight of the link ($w_{u,v}$) between nodes u and v.

Defining link strengths for a choice network In the choice network, a directed link from node u to node v is established if there exist customers who considered car u and v together but finally bought v instead of u. The total number of customers who bought car v despite considering car u denote the link strength from u to v and vice versa.

We denote both networks as $G = (V, \varepsilon, W)$, where V, ε and W represent nodes, edges and weights respectively. Figure 1 provides a simplified illustration for both the unidimensional consideration and choice networks that we investigate. The thickness of the link between two nodes is proportional to its strength (*i.e.* the number of customers who consider or choose the product), and the size of the node is proportional to the popularity of the product.

Descriptive network analysis Descriptive network analysis helps designers explore some major characteristics of the network, like which products are popular, how dense the network is, without going into the statistical relationship of nodes and links, It requires the computation of topological measures to assess the position of nodes and the implication of structural advantages [3].

A few descriptive metrics for analyzing a unidimensional weighted car competition network are *network weight distribution*, *centrality* and *clustering coefficient*. The values of weights $w_{u,v}$, which measures the competition strength between pairs of cars (u and v), can be considered as a fundamental element in the weighted network analysis. The *probability distribution of weights* $P(w)$ indicates the overall competition strength, i.e., the frequency of a pair of cars being co-considered in an undirected network. We can also calculate the *centrality* of a node, which may mean different things for different networks. For an undirected consideration network, the centrality is measured by the strength of a node, which is defined as $s(i) = \sum_{j \in V(i)} w_{ij}$ ($V(i)$ is the node-set of node i's neighbourhood). It is a measure of how popular the car is. Note that as this is a measure of popularity in consideration network, it is possible that a model is popular (considered by many people) in the consideration stage but still has a low market share. In the directed choice network, the in-strength of a node $s_{in}(i)$ equals to the sum of weights of all directed inward links, which is a measure of popularity in the final purchase decisions. Further, the fraction of a node's in-strength to the total in-strength of all the nodes ($\frac{s_{in}(i)}{\sum_{j=1}^{n} s_{in}(j)}$) for directed networks represent the market share of the car. Finally, we are also interested in observing if there are cliques of cars, which have high competition among them. To measure this, we use the *weighted global* *clustering coefficient*, which measures the overall network interconnected triplets [34]. A cluster is defined as a group of nodes with high weight links between each other and with low weight links to other nodes in the network. Therefore, a high clustering coefficient indicates interconnected communities (car competitions within market segments) are more common in the network. While descriptive analysis provides broad insights about the network structure, it does not throw light on how different attributes affect link formation, which we will discuss next.

3.2 Network Modeling and Interpretation

In this paper, we use the valued-ERGM to model both undirected consideration networks and directed choice networks, such that the link strength can be captured within the network structures. As described in Eq. 2, the input of valued-ERGMs are a reference distribution $h(\mathbf{y})$ and a vector of selected input terms $g(\mathbf{y})$ (such as car price, fuel efficiency) and a few characteristics for weighted network structures (such as network density).

Defining a reference distribution The reference distribution $h(\mathbf{y})$ acts as our prior belief about the network based on the known distribution of link weights, which refers to the distribution of the co-consideration strength and pairwise preference strength among car competition networks. While Binomial distribution is typically used for binary networks, other choices like the Poisson, Geometric, Bernoulli, Uniform and Standard Normal distribution are possible for a weighted network. The exact choice of prior belief depends on the application domain and the data distribution.

Defining input variables Many of the variables used as input in a valued-ERGM model are similar to the ones used in binary ERGM models, and can be classified into three categories: network configurations, main effects, and homophily effects [6]. Network configurations measure the network structure effects, main effects correspond to nodal effects of the product attributes, and homophily effects are the similarity or difference between the attributes of two nodes. Unlike any dyad-independent binary ERGM statistic expressed as $g_k = \sum_{(i,j)\in Y} x_{k,i,j} y_{i,j}$, where y_{ij} is allowed to have values either 0 or 1, in the valued-ERGM, y_{ij} has a larger range of choices. As for the network configuration terms, valued-ERGM can handle network sparsity, mutuality, individual heterogeneity and triadic closure via various input model terms [35].

Interpreting valued-ERGM parameters The result of the valued-ERGM is a set of estimated coefficients and associated p-value (significance level) for all variables. Network configuration effects indicate the tie independence, *i.e.* formation of ties due to the presence of other ties [36]. The estimation of those effects can be seen as evidence of the prevalence or absence of certain structures (such as edge density, transitivity and star effects) in a network. For example, a negative significant estimation of "edges" shows that the competition network has a low density. The impact of main attributes refers to how an attribute might influence a car's propensity to form a link in a co-consideration network and to receive or send a link in a choice network. We examine selected car attributes, and the result will enable designers to determine whether cars with a higher price, lower fuel consumption are more likely to be considered by customers and win a competition. The homophily effects test the hypothesis that cars with more similarities on different attributes are likely to be co-considered, which is a common explanation established in social relations and further extended to our car competition networks.

3.3 Market Competition Prediction

While statistical network models are typically used to interpret what factors lead to link formation or dissolution, predicting what the network will look like in the future is useful for manufacturers to make strategic decisions. In practice, if manufacturers can predict how the competition between car models will change when certain node attributes (product design) is changed, they can use this information to forecast the effect of changes in their models and to position their products strategically among competitors. Using the estimated parameters of input variables of observed car competition networks, we can predict unseen networks in the future, with new car attributes as input.

Based on valued-ERGM equation Eq. 2, the distribution of network models is determined by a base network structure, estimated parameters, input variables, and a reference distribution. If we want to simulate a trained network, we can substitute new car attributes to the trained model and derive the distribution of predicted network structures (with valued link strength) and draw samples from it, where each sample is a simulated network. In our validation study, we draw such simulated network samples based on the trained valued-ERGM model. By aggregating a large number of network samples, we create a single network, which represents the central tendency (highest probable network) of all simulated networks. We use this aggregated network as our prediction and compare it with the known network in the future to show our model's accuracy.

Future predictions using aggregated simulations can be made for either the co-consideration network or the choice network. In the predicted co-consideration networks, the number of competitors with corresponding competition strength is forecasted, which helps a manufacturer in understanding their overall market position. For the predicted choice networks, along with the market position, the manufacturers also get an understanding of which car models are their main competitors. In the next section, we show how the methods and the process discussed so far can be applied to two real-world datasets of vehicles.

4 CASE STUDIES

Cars are an expensive good, and customers often consider many alternatives before deciding which car to buy. This decision can be influenced by many factors, like your budget, what are your driving needs, what must-have and good-to-have features you want in a car, which cars are currently popular in your vicinity, which cars are owned or suggested by your friends and family, what past experience you had with particular brands *etc.* From a manufacturer's perspective, it is important to understand their competition and develop strategies to improve market share. Consider two brands with similar number of units sold per year. Car A has low sales figures, and is rarely even considered by any customer while making their purchase decision. In contrast, car B is considered by most customers looking to buy a car, but most customers eventually buy a competitor car C, instead of car B. Despite similar sales figures, car A's strategy of market share improvement may be very different from car B's strategy (which can be focused on the competitor). How can one throw light on such complex interactions between products?

We argue that many, if not all, of these factors can be captured by network models. Specifically, in this section, we demonstrate the use of the valued-ERGM approach to study the Chinese car market. We use a car survey data provided to us by the Ford company as a test example. Network modeling can be applied to different stages of decision making of a customer, and different types of network models. We show two case studies, covering different aspects of graph structures and decision making. The first case study focuses on the initial stage of customer decision making and uses an undirected co-consideration network model. In this case, cars compete with each other to be simultaneously considered by a customer. The second case study focuses on the final stage of decision making using a directed final choice network model. In this case, cars that are co-considered by a customer compete with each other to "win" (or be purchased) from their competitors.

4.1 Data Description

Our dataset contains customer survey data from 2012 to 2016 in the China market. In the survey, there were more than 40,000 respondents each year, who specified which cars they purchased and which cars they considered, before making their final car purchase decision. Each customer indicated at least one, and up to three cars which they considered. The dataset, resulting from the survey, also contains many attributes for each car (*e.g.* price, power, brand origin and fuel consumption) and many attributes for each customer (*e.g.* gender and age).

4.2 Case study 1: Car Co-Consideration Network

In this case study, we use valued-ERGM models to throw some light on competition between cars in the consideration stage of a customer's decision making process.

Step 1: Network construction and Characterization To study car co-consideration, we start by creating a car co-consideration network based on customers' survey responses in the 2013 survey data. The network consists of 296 unique car models as network nodes. The link between a pair of nodes (denoting cars) is allocated the weight equal to the number of customers who considered both the car models together in their consideration set. Figure 2 shows an example of a small part of the co-consideration network. In this example, cars "Great Wall Hover" and "Honda Dongfeng CRV" appear together in the consideration set of 18 customers in 2013 and 30 customers in 2014, showing that their competition has increased in one year. In contrast, cars "VW SVW Tiguan" and "Honda Dongfeng CRV" appear together in the consideration set of 201 customers in 2013 and 192 customers in 2014. This shows that their competition has decreased in one year, although both cars are still more popular (sum of all link strengths connected to a node) than the "Great Wall Hover".

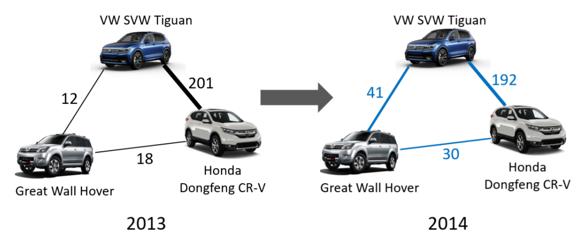

FIGURE 2: An example of co-consideration network between three cars changing from one year to another.

Table 2 presents a summary of our network's descriptive characteristics. Network density shows that among all possibly connected car models, 15.7% of them are being co-considered, and an average of 5.511 customers consider any connected car models (average link strength). The average degree means that on average, each car competes with 23.2 other cars. The average weighted degree indicates a car is considered by 127.9 customers on average. The global clustering coefficients of 0.623 gives us an overall indication of clustering and suggests that car models tend to engage in a multi-way competition. We observe from the distribution of link strengths (including non-existing links denoted by zero strength), that 98.17% values are lower than ten, and the distribution has a long-tail with a few large values. When comparing it to existing statistical distributions that can be used to initialize a valued-ERGM, its shape matches the most with the positive part of a standard normal distribution. We further verify this observation by our experiments, where using a standard normal distribution provided the best results in model fit. We use it to report all our results.

To create the ERGM network model, we select the set of most important car attributes based on the selection criterion used in a previous study [6] to associate with network nodes, includ-

TABLE 2: Summary of Co-Consideration Network Descriptive Characteristics

No. Nodes	Network Density	Ave. Strength	Ave. Degree	Ave. Weighted Degree	Global Clustering Coefficient
296	0.157	5.511	23.2	127.9	0.623

ing price, engine power, fuel consumption, market segment, import, and car make origin. This allows us to compare our results with past work too. We apply log transformation to price (in Chinese Yuan RMB) and engine power (in brake horsepower BHP) to normalize the range of attribute values and reduce the large outlier effects. Fuel consumption is calculated by the ratio of consumed gasoline (in liters) to driving distance (in 100 km), and a smaller fuel consumption value speaks for higher fuel efficiency. The market segment is a categorical variable that contains 17 car segments. These segments are provided by Ford. Import and make origin are related to car's brand information, and 35.1% of cars are imported from Europe, the United States, Japan and South Korea and 64.9% of cars are domestically produced in China.

Step 2: Network modeling and Interpretation In the implementation of the valued-ERGM model, we assign the selected car attributes to network nodes and the occurrences of co-considerations to the link strengths. Based on the shape of link strength distribution, we select the standard normal distribution as the reference distribution specified in valued-ERGM models The input variables can be divided into two categories: main effects and homophily effects. The whole set of input variables can be found in Table 3. We use the statistical network analysis package "Statnet" in R programming, in which the valued-ERGM is integrated [37].

Table 3 shows the estimated coefficients from fitting the valued-ERGM models. The sum/intercept variable serves as a constant term in valued-ERGM and it estimates the likelihood of two cars' co-consideration strength without any knowledge about the cars' attributes. All the input variables, except the main effect of power, are statistically significant with p values lower than 0.05. As all the variables are normalized to a similar order of magnitude, the differences in coefficients denote their relative importance in the model fit. We observe that the coefficients corresponding to the homophily effects are larger than those of the variables within the main effects. This indicates that the homophily effects may play a more important role in two cars being co-considered. Among the main effects, the coefficient of import effect is negative but the coefficients of brand origins from different countries are positive, which implies customers tend to consider domestically made cars with foreign brands, such as Ford Changan Focus, Honda Dongfeng Civic. Variables like price, power and fuel consumption are not as significant as the other main effects. Among the homophily effects, the market segment matching and brand origin matching are significant. This may

TABLE 3: Estimated Coefficients of the Co-consideration Network

Input Variables	Est. coef.
Network structural effect	
Sum/Intercept	- 1.64***
Main effect (nodal attributes)	
Import	- 0.30***
Price (log2)	0.07***
Power (log2)	- 0.006
Fuel consumption (per 100 BHP)	- 0.01***
Brand origin (the US)	0.29***
Brand origin (Europe)	0.19***
Brand origin (Japan)	0.09***
Brand origin (Korean)	0.15***
Homophily effect (dyadic attributes)	
Market segment matching	0.44***
Brand origin matching	0.22***
Price difference (log2)	- 0.20***
Power difference (log2)	0.08***
Fuel consumption difference	- 0.04***

Note: $*p<0.05$; $**p<0.01$; $***p<0.001$

reveal that car models within the same market segment and the same brand origin tend to be co-considered by customers. Further, a statistically significant large negative coefficient of price difference shows customers prefer to consider cars in a similar price range. This observation aligns with our intuition, as a customer may consider cars within his/her budget.

Step 3: Validation and Prediction We perform three different types of validation to examine both the model fit and the prediction power, as elaborated below:

Trained model prediction is similar to the link strengths in the training data. We start model validation by simulations on the current network structure based on the estimated coefficients of selected model terms. More concretely, we create 100 simulated networks with 2013 car co-consideration network structures and estimate parameters in Table 3, then take the average of link strength values for each pair of nodes and denote it as the aggregated simulated car co-consideration strength. As a result, the link strength comparison of the simulated network and the original network reveals the goodness of the model fit. Figure 3 (Top) plots the link strengths of the true network with the aggregated simulated network obtained from our trained ERGM model. We observe that two sets of link strengths are positively correlated, where a perfect $y = x$ line may indicate a perfect fit. We observe a Pearson coefficient of 0.995 and the coefficient of determination (R^2) of 0.990, which is strong evidence of good model fit.

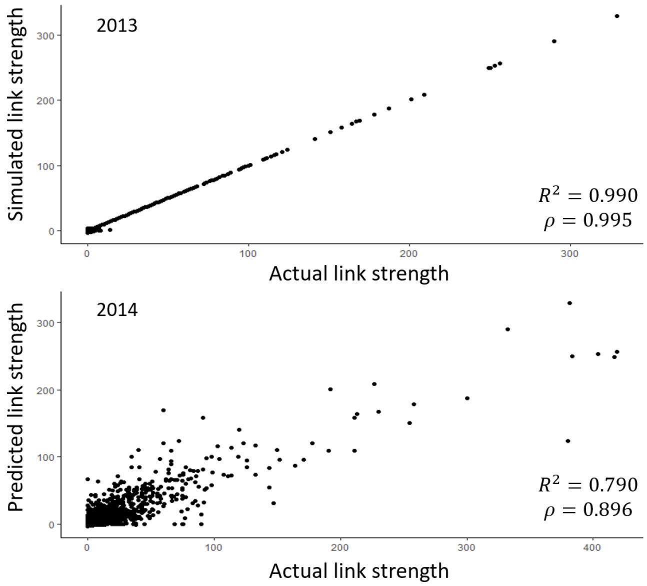

FIGURE 3: Goodness of fit using link strength comparison between the trained network and simulated network. Top: Link strengths of the trained network with the aggregated simulated network for 2013. Bottom: Link strengths of the true network with the aggregated simulated network for 2014 (unseen future data).

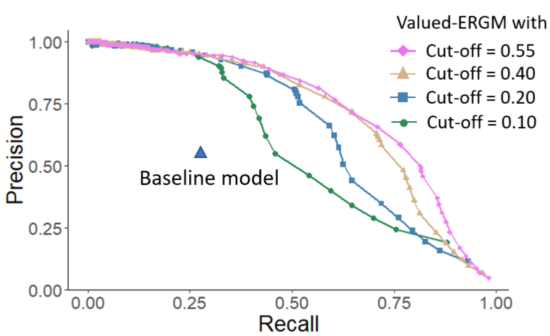

FIGURE 4: Precision and Recall Curve of Network Prediction. We observe that irrespective of different cut-off values, valued-ERGM models have a higher precision and recall than binary network used as a baseline from [6]

Trained model predicts link strengths of future unseen data reasonably well. In practice, the benefit of training a statistical model is to predict the future behavior of networks and not the network that has already been realized. While the market competition between different car models varies yearly, we test whether our fitted co-consideration model can be utilized to predict the co-consideration relationship in the future market. Figure 2 illustrates an example of the real market evolution, it can be observed that in 2014 Great Wall Hover gains more customers' consideration, and the strong co-consideration relationship between VW Tiguan and Honda CR-V decreases slightly. Our examination of the model's predicting power uses a similar method of network aggregation as used in the validation on the training data, but with the input of 2014 car attributes as the updated network nodes. With a similar simulation process, we derive the aggregated predicted co-consideration network for 2014 market data, and compare it with the actual co-consideration network in 2014. The scatter plot of actual link strength and predicted link strength is reported in Figure 3 (bottom), with a high R^2 of 0.790 and Pearson coefficient as 0.896. More importantly, we observe that though there exist some deviations between prediction and true link strength when the link strength values are low in magnitude, the prediction is better for large link strengths. In practice, large link strengths are more important to predict correctly, as they indicate the competition between major players in a market. In addition, the weighted network model is also beneficial to designers who eager to know the outcomes of design variable changes. When we conduct the model validation, we have been using the weighted network model to predict the competition relation evolution when design variables changes from 2013 to 2014. Likewise, with its predictive ability, one can simulate the weighted network model with changed design variables (design intervention), and foresee the resulted competition market. Thus, the model can provide some foresight for car manufacturers to make strategic decisions.

Valued-ERGM has higher precision and recall compared to baseline binary models We want to further compare the prediction result with the previous binary non-weighted network baseline [6]. However, for comparison, we have to convert a simulated weighted network to a binary network using some cut-off value for link strengths. We use four different cut-off values and after creating a binary network, we compare the predicted co-consideration network and the actual one. This comparison allows us to measure the precision (the fraction of true positive predictions among all positive predictions) and recall (the fraction of true positive predictions among all positive observations) as the metrics to evaluate the performance of our method. A higher value of precision and recall indicates a better predictive model.

To find a few cutoff values for comparison, we use the quantiles from the actual network link strength distribution, which are 0.10, 0.30 and 0.40 and the mean value (0.55). For the predicted link values, due to the unbalanced nature of the data (most of the link strength values are zero), we set the cutoff value of the link strength by different thresholds and draw the precision and recall curve for each cutoff values. The baseline model is reported in the past work [6], which used a binary ERGM has a precision value of 0.543 and a recall value of 0.311. As can be seen from Figure 4, the precision-recall curves for all cut-off values using the valued-ERGM model is higher than the baseline model. This demonstrates that the valued-ERGM network modeling technique has a significantly better precision and recall for predicting binary co-consideration networks too, which provides evidence that it can be utilized for better understanding of market competition.

4.3 Case Study 2: Crossover SUV Choice Network

In this case study, we use valued-ERGM models to throw some light on competition between cars in the choice stage of a customer's decision making process. We specifically focus on crossover SUV car market segment and describe the steps next.

Step 1: Network construction In the second case study, we focus on market competition among crossover SUVs, such as Ford Edge and Mazda CX-7, which are designed with the body and space of an SUV but the platform of a sedan and have received increasing attention in the car market. There are 14 crossover SUV models in 2013 survey data, and we have collected all the survey data of which customers have either considered or chosen a crossover SUV model in that year (1975 customers in total). The directed choice network is established based on the customers' purchase behavior as described in the previous section and all competitors in the network are divided into four segmentation groups: Sedan, SUV, Luxury or Sport and Crossover SUV. The visualization of the choice network is plotted in Figure 5, where the node sizes of crossover SUVs reflect the number of customers who have purchased it.

Overall, there are 217 car models in the crossover SUV choice network. All the links are directed and point to the "winner" in a competition of customers' decision making. The average link strength is 2.431 corresponding to the average number of customers' purchases among all co-considered cars. A distinct characterization of the choice network is that the in-strength of a node is correlated to its market share.

Step 2: Network modeling and Interpretation The procedure of network modeling of a choice network shares many similarities with that of a co-consideration network in the valued-ERGM package. However, as the choice behavior is not symmetric between pairs of nodes, the model terms are further specified for inward nodes or outward nodes. Specifically, the main effects in Table 4 refer to the nodal attributes of the inward nodes, hence we can learn the important attributes of the "winners" and find possible reasons behind the popularity of a car model. Besides,

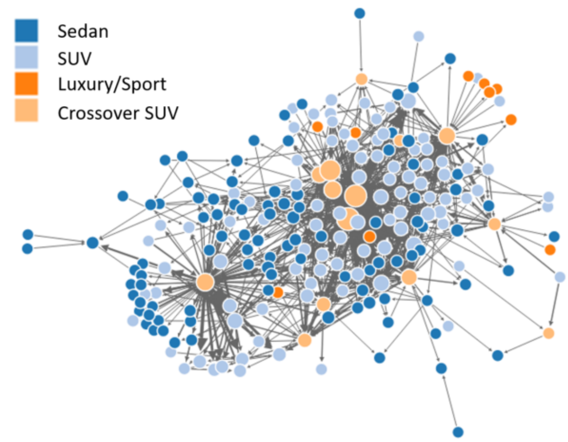

FIGURE 5: A force directed graph visualization of the choice network for Crossover SUVs. We observe that most crossover SUVs compete with Sedans and SUVs.

TABLE 4: Estimated Coefficients of the Choice Network

Input Variables	Est. coef.
Network structural effect	
Sum/Intercept	0.19
in-2-star (Popularity)	0.002***
Main effect (inward node attributes)	
Import	0.01
Price (log2)	-0.02.
Power (log2)	0.02
Fuel consumption (per 100 BHP)	0.01*
Brand origin (the US)	0.005
Brand origin (Europe)	0.002
Brand origin (Japan)	0.003
Brand origin (Korean)	0.01
Homophily effect (dyadic attributes)	
Market segment matching	0.04**
Brand origin matching	0.04***

[1] Note: .p <0.1; *p<0.05; **p<0.01; ***p<0.001

we have added a network structural effect - "in-2-star", which measures the uncentered covariance of in-dyad values incident on each node. More precisely, it accounts for the in-strength heterogeneity and is used to represent the popularity of a car model.

Table 4 shows the estimated coefficients from fitting the directed valued-ERGM models. The network structural effects and homophily effects play an important role in a directed network. A significant and positive estimation of in-star effects indicates a high heterogeneity of the node in-strength distribution and there are some car models more popular than the others (because the more edges on a node, the more two stars an additional edge will create). An evidence of this interpretation can be found in Figure 6 that in crossover SUV market, Chevrolet Captiva and Toyota Highlander account for more than 40% of the market share. As a choice network also implicitly contains customers' consideration information as they link two co-considered cars together, the estimation of homophily effects implies the positive relationship between car similarities and the tendency of they are co-considered. Moreover, many of the main effects don't have a significant p-value, but their signs are consistent with common sense as people prefer a car model with a lower price (negative) and a stronger power (positive). What is interesting about the data in the main effect category is that higher fuel consumption has a positive impact on customers' choice. One possible explanation for this observation is that the customers who consider or purchase crossover SUVs seek a better off-road performance instead of purchasing the car with lesser fuel consumption.

Step 3: Validation with pairwise competition comparison We validate our model using two methods — predicting pairwise competition and estimating the market share of each car. We first evaluate the model fit at the pairwise competition level. Given the original network structure, one can recognize the "winner" in each pairwise competition by detecting the customers' choice prevalence. For example, among 25 customers who have considered both car A and car B, there are 15 customer prefers car A and 10 customers prefer car B, and then A is denoted as the "winner" in "A-B" competition. After simulations of the choice network structure based on the fitted model, the aggregated (*i.e.* averaged) simulated link strength can be used as the judgment of pairwise competitions. The result shows that the simulated choice network can estimate 72.06% of the pairwise competitions correctly.

Validation with market share comparison In a directed choice network, the in-strength of node $s_{in}(i)$ is related to its market share as we explained before. Hence, we can further validate the choice network by comparing the simulated market share for each crossover SUV with the true market share in the network data. Specifically, the in-strength fraction ($\frac{s_{in}(i)}{\sum_{j=1}^{n} s_{in}(j)}$) is calculated for the crossover SUVs in an observed choice network and serves as the actual market share, then the simulated market share is derived from the average in-strength of the nodes within 100 simulations. The comparison of actual market share, simulated market share and uniform market share (which assumes all crossover SUVs have the same market share and serves as a baseline) is plotted in Figure 6. Compared to the baseline of uniform market share, the simulated market share has a R^2 value equals to 0.939, which indicates that approximately 94% of the observed variation can be explained by the fitted choice network model.

While valued-ERGM shows a reasonably good fit at the pairwise relative competition and the market share, we found that it did not predict well the absolute value of weights in the choice network, both for the current and future markets. We suspect that this is due to the limited choice of reference distributions in standard software packages which were not representative of our true distribution. In future work, we will explore ways to improve valued-ERGM for better prediction of both unseen data and edge strengths.

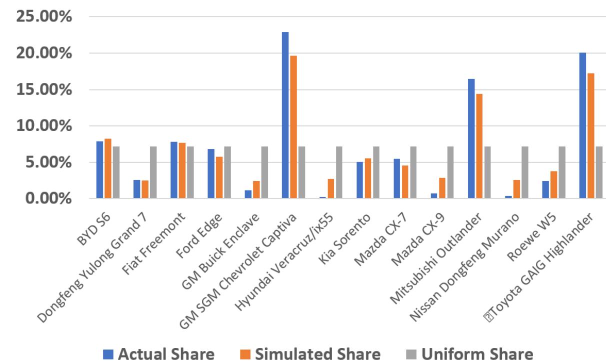

FIGURE 6: Valued-ERGM prediction of crossover SUVs market share aligns with the true market share.

5 DISCUSSION

While the valued-ERGM model provides many advantages over existing statistical models, it still has many theoretical and practical challenges to be widely adopted by car manufacturers. We discuss a few of these limitations next and discuss how they pave the path to future research directions.

Helping engineering design decisions using valued-ERGM One of the goals of using valued-ERGM model is to demonstrate how the approach can help in understanding the important factors that influence product competition. These factors can help a decision-maker in making strategic decisions, which include changes to the product. However, it is important to note that while the theoretical model allows one to estimate the importance of any attribute, the analysis for specific case studies also depend on what product data is available and whether there exist any true relationships between product attributes and customer-purchase decisions. To understand this, let us consider three hypothetical situations. In the first situation, customers make a purchase decision based only on the size of the car engine. Using an valued-ERGM model, the analysis results show that the size of the engine (or power – which is correlated with it) have a significant coefficient. In such cases, the network model will directly inform a car manufacturer that increasing the engine size will help them gain more market share. The same applies if the customers consider multiple attributes. In the second situation, we assume that the customers make a purchase decision based only on the quality of the automotive air-conditioning system. If the data we analyze does not include air conditioning as a design attribute, the results will be meaningless due to the lack of data. The only remedy for this situation is to collect the data which can capture the relevant attributes, and then use them to assist car manufacturers in making decisions. In the third situation, we assume that customers base their purchase decisions on factors that are not design-related, such as social or cultural influence. In such cases, the coefficients of all design attributes may not have statistical significance, which indicates that the improved design factors will not help automakers gain more market share. Hence, the guidance provided to a manufacturer is to not waste resources on improving factors which do not have an impact.

In our case studies, we believe that all the three scenarios discussed above might be playing a role — we find some design attributes which impact design decisions and have statistically significant values, we also discover that the dataset is limited and lacks information about designs of parts within a car and finally, many design attributes which we studied did not have statistically significant values, indicating that those factors may not play a role in customer decisions.

From our current results for both case studies, we uncover a few factors which impact engineering design decisions for product consideration. Specifically, in the co-consideration network of case study 1 (Table 3), we observe that a car designer seeking to improve the chances of their car models to be considered by customers must reduce fuel consumption (which relates to engine efficiency). Although factors like price, power and fuel consumption have significant values, they do not directly provide clear design guidance for a car manufacturer. In the choice network of case study 2 (Table 4), the model results help decision-makers with strategic planning as well. For example, in the crossover SUV market, the improvement of fuel consumption is not important in customers' purchase decision. Instead, it is helpful to reduce price to improve market share. We notice, that our dataset lacked design-based attributes about individual components in the cars, which may have played a huge role in the customer decisions. In future work, we aim to gather product attributes for a new customer product using crowdsourcing.

Trade-off between feature engineering and model interpretability In valued-ERGM models, we start with a large collection of features. These features can be node-specific (*e.g.*car fuel efficiency, price), edge-specific (*e.g.*difference in price, similar car models) or network-specific (*e.g.*popularity, density). The choice of what features to use has a large impact on the goodness of fit of the model, coefficients for each attribute as well as the statistical relevance. While we use automated methods for feature selection (which largely select features that are uncorrelated), the process is largely manual. In contrast, one can use modern deep learning models, which learn hierarchical feature representations for each item to minimize a loss function. The deep learning models are largely a black-box and are hard to interpret, which is one of the key reasons of our using ERGMs, the statistical network models, for assisting car manufacturers. In the future, we aim to find the middle ground of reducing dependence on selecting features and explore the interpretability of deep learning models with newly developed analysis, such as SHapley Additive exPlanations (SHAP) [38].

Training large scale models Valued-ERGM models are typically appropriate for small to medium-sized graphs, with a few attributes. From our experience, one cannot train such a model for tens of thousands of edges and hundreds of node attributes. For large datasets, the MCMC approach to estimate parameters does not converge. This leads to an important limitation for car manufacturers, who now want to make the most of huge datasets but still want statistical models that can help them understand what is happening under the hood. While there is some recent work on developing scalable ERGMs [39], extending them to valued-ERGMs can help alleviate the scalability problem for large datasets.

Another alternalte option is to use Graph Neural Networks (GNNs) to model large graphs. However, using GNNs has two problems. First, most GNNs are inherently transductive and can only generate embeddings for a single fixed graph. These transductive approaches do not efficiently generalize to unseen new products. Secondly, it is difficult to isolate the importance of design attributes in link formation using GNNs as the network embeddings also encode some of the attribute information. Hence, ERGM models are more suitable for studies requiring inductive models. In future studies, we will explore a few inductive GNN methods like GraphSAGE and compare them to valued-ERGM models.

Temporal domain of network analysis The temporal domain of the network evolution is an important topic in the product competition area, where the network structure evolves with time as new products or design features are introduced. Researchers have used binary STERGM models to study the competition between cars in dynamic networks [9]. In our preliminary literature review, we identified that the weighted STERGM models have not been proposed in the network literature yet. This may be because of the difficulty in modeling the network prior. In the current work, we predict the future network, by assuming that no new car is introduced, although the value of car attributes may change. While weighted dynamic networks were not the focus of this paper, we agree with the reviewer's advice and will further explore it to study the dynamic effect of network in future work.

6 CONCLUSION

In this paper, we exhibit how valued-ERGM models can be used to model directed and undirected car competition networks with non-binary link strengths. The method enables designers to estimate the major factors that affect customers' consideration and choice behavior, and which can help in predicting future market competition when a manufacturer changes some product attributes. The scientific contributions and novelty of this work are as follows:

1. We extend the newly developed valued-ERGM in social network modeling into product competition network modeling, which enriches the knowledge base of product design modeling techniques.
2. By developing a procedure of weighted network construction, interpretation and validation, we demonstrate that valued-ERGM models provide a better model than binary-ERGM, as measured by model fit and prediction accuracy for car competition.
3. As a first attempt of unidimensional network modeling for analyzing aggregated choice behavior, we demonstrate a way to represent choices between different cars using a directed unidimensional network, which enables us to capture both consideration and choice preferences in a unidimensional network.

The case studies in this paper show how network models are used to systematically analyze large real-world networks. For the first case study, which analyzes the co-consideration competitions between 296 cars, we show that homophily effects, affecting the differences between two cars, are more important than the main effects in predicting link strengths. Cars are generally found to compete more with other cars from the same market segment, same brand origin and similar price range. Moreover, the model has exhibited the predictive ability to provide some foresight for car manufacturers to make strategic decisions. From the second case study, which focuses on the crossover SUV market, we analyze a network of 217 cars and find that cars which are considered by more people are also purchased more often. In future work, we aim to analyze how valued-ERGM can help study new domains and further investigate ways to integrate feature learning methods like deep learning with valued-ERGM models while retaining their interpretability. Improving the scalability of these models to larger datasets and using them for dynamically evolving car competition is another interesting avenue of research.

ACKNOWLEDGMENT

We are grateful to the support from the National Science Foundation under Grant No. CMMI-2005661 and No. CMMI-2005665, Ford-Northwestern Alliance Project and Design Cluster Fellowship at Northwestern University.

REFERENCES

[1] Hirtenstein, A., 2017. Global Electric Car Sales Jump 63 Percent, Nov.

[2] Pontes, J. 2017 China Electric Car Sales Blow World Out Of The Water — BAIC EC-Series Is A Superstar | CleanTechnica.

[3] Wang, M., Chen, W., Huang, Y., Contractor, N. S., and Fu, Y., 2015. "A multidimensional network approach for modeling customer-product relations in engineering design". In ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers Digital Collection.

[4] Wang, M., Sha, Z., Huang, Y., Contractor, N., Fu, Y., and Chen, W., 2016. "Forecasting Technological Impacts on Customers' Co-Consideration Behaviors: A Data-Driven Network Analysis Approach". In ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pp. V02AT03A040–V02AT03A040.

[5] Fu, J. S., Sha, Z., Huang, Y., Wang, M., Fu, Y., , and Chen, W., 2017. "Modeling Customer Choice Preferences in Engineering Design using Bipartite Network Analysis". In Proceedings of the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.

[6] Sha, Z., Huang, Y., Fu, J. S., Wang, M., Fu, Y., Contractor, N., and Chen, W., 2018. "A network-based approach to modeling and predicting product coconsideration relations". *Complexity,* ***2018***.

[7] Wang, M., Sha, Z., Huang, Y., Contractor, N., Fu, Y., and Chen, W., 2018. "Predicting product co-consideration and market competitions for technology-driven product design: a network-based approach". *Design Science,* ***4***.

[8] Bi, Y., Xie, J., Sha, Z., Wang, M., Fu, Y., and Chen, W., 2018. "Modeling Spatiotemporal Heterogeneity of Customer Preferences in Engineering Design". In ASME 2018 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.

[9] Xie, J., Bi, Y., Sha, Z., Fu, Y., Contractor, N., Gong, L., and Chen, W., 2019. "Data-Driven Dynamic Network Modeling for Analyzing the Evolution of Product Competitions". In ASME International Design Engineering Technical Conferences, ASME.

[10] Sha, Z., Bi, Y., Wang, M., Stathopoulos, A., Contractor, N., Fu, Y., and Chen, W., 2019. "Comparing utility-based and network-based approaches in modeling customer preferences for engineering design". In Proceedings of the Design Society: International Conference on Engineering Design, Vol. 1, Cambridge University Press, pp. 3831–3840.

[11] Vincent, T., 1983. "Game theory as a design tool". *Journal of Mechanical Design,* ***105***(2), pp. 165–170.

[12] Kato, T., Nishida, A., Koshijima, I., and Umeda, T., 2013. "Engineering innovation methodology using evolutionary game theory". In Engineering, Technology and Innovation

(ICE) & IEEE International Technology Management Conference, 2013 International Conference on, pp. 1–9.
[13] Shiau, C.-S. N., and Michalek, J. J., 2009. "Optimal product design under price competition". *Journal of Mechanical Design,* ***131***(7), p. 71003.
[14] Kaul, A., and Rao, V. R., 1995. "Research for product positioning and design decisions: An integrative review". *International Journal of research in Marketing,* ***12***(4), pp. 293–320.
[15] Liu, X., Du, G., Jiao, R. J., and Xia, Y., 2017. "Product Line Design Considering Competition by Bilevel Optimization of a Stackelberg-Nash Game". *IISE Transactions*(just-accepted).
[16] Kaiser, U., and Licht, G., 1998. R&D cooperation and R&D intensity: theory and micro-econometric evidence for German manufacturing industries. Tech. rep.
[17] Wassenaar, H. J., and Chen, W., 2003. "An approach to decision-based design with discrete choice analysis for demand modeling". *J. Mech. Des.,* ***125***(3), pp. 490–497.
[18] Du, P., and MacDonald, E. F., 2015. "Products' shared visual features do not cancel in consumer decisions". *Journal of Mechanical Design,* ***137***(7).
[19] Donndelinger, J., and Ferguson, S. M., 2017. "Design for Marketing Mix: The Past, Present, and Future of Market-Driven Product Design". In ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pp. V02AT03A040–V02AT03A040.
[20] Wang, M., and Chen, W., 2015. "A data-driven network analysis approach to predicting customer choice sets for choice modeling in engineering design". *Journal of Mechanical Design,* ***137***(7).
[21] Wang, M., Chen, W., Huang, Y., Contractor, N. S., and Fu, Y., 2016. "Modeling customer preferences using multidimensional network analysis in engineering design". *Design Science,* ***2***.
[22] Yip, A. H., Michalek, J. J., and Whitefoot, K. S., 2018. "On the implications of using composite vehicles in choice model prediction". *Transportation Research Part B: Methodological,* ***116***, pp. 163–188.
[23] Sha, Z., Saeger, V., Wang, M., Fu, Y., and Chen, W., 2017. "Analyzing Customer Preference to Product Optional Features in Supporting Product Configuration". *SAE International Journal of Materials and Manufacturing,* ***10***(2017-01-0243).
[24] Robins, G., Pattison, P., Kalish, Y., and Lusher, D., 2007. "An introduction to exponential random graph (p*) models for social networks". *Social networks,* ***29***(2), pp. 173–191.
[25] Wasserman, S., and Pattison, P., 1996. "Logit models and logistic regressions for social networks: I. an introduction to markov graphs andp". *Psychometrika,* ***61***(3), pp. 401–425.
[26] Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., and Morris, M., 2008. "statnet: Software tools for the representation, visualization, analysis and simulation of network data". *Journal of statistical software,* ***24***(1), p. 1548.
[27] Van Der Pol, J., 2017. Introduction to network modeling using Exponential Random Graph models (ERGM). working paper or preprint, Oct.
[28] Krivitsky, P. N., 2012. "Exponential-family random graph models for valued networks". *Electronic journal of statistics,* ***6***, p. 1100.
[29] Krivitsky, P. N., and Butts, C. T., 2013. "Modeling valued networks with statnet". *The Statnet Development Team*, p. 2013.
[30] Scott, T. A., 2016. "Analyzing policy networks using valued exponential random graph models: Do government-sponsored collaborative groups enhance organizational networks?". *Policy Studies Journal,* ***44***(2), pp. 215–244.
[31] Pilny, A., and Atouba, Y., 2018. "Modeling valued organizational communication networks using exponential random graph models". *Management Communication Quarterly,* ***32***(2), pp. 250–264.
[32] Silk, M. J., Weber, N. L., Steward, L. C., Hodgson, D. J., Boots, M., Croft, D. P., Delahay, R. J., and McDonald, R. A., 2018. "Contact networks structured by sex underpin sex-specific epidemiology of infection". *Ecology letters,* ***21***(2), pp. 309–318.
[33] Windzio, M., 2018. "The network of global migration 1990–2013: Using ergms to test theories of migration between countries". *Social Networks,* ***53***, pp. 20–29.
[34] De Montis, A., Barthélemy, M., Chessa, A., and Vespignani, A., 2007. "The structure of interurban traffic: a weighted network analysis". *Environment and Planning B: Planning and Design,* ***34***(5), pp. 905–924.
[35] Krivitsky, P. N., 2012. "Exponential-family random graph models for valued networks". *Electronic journal of statistics,* ***6***, p. 1100.
[36] Robins, G., 2013. "A tutorial on methods for the modeling and analysis of social network data". *Journal of Mathematical Psychology,* ***57***(6), pp. 261–274.
[37] Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., Bender-deMoll, S., and Morris, M., 2019. statnet: Software Tools for the Statistical Analysis of Network Data, June.
[38] Lundberg, S. M., and Lee, S.-I., 2017. "A unified approach to interpreting model predictions". In *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds. Curran Associates, Inc., pp. 4765–4774.
[39] An, W., 2016. "Fitting ergms on big networks". *Social science research,* ***59***, pp. 107–119.