Modelling and analysis of network information data for product purchasing decisions

Md Asaduzzaman[1], Uchitha Jayawickrama[2], and Samanthika Gallage[1]

[1] Staffordshire University, Stoke on Trent ST4 2DE, UK,
md.asaduzzaman@staffs.ac.uk,
WWW home page: http://www.mdasad.com/
[2] Loughborough University, Loughborough LE11 3TU, UK

Abstract. Technology has enabled consumers to gain product information from different online platforms such as social networks, online product reviews and other digital media. Large manufacturers and retailers can make use of this network information to forecast accurately, to manage the demand and thereby to improve profit margin, efficiency, etc. This paper proposes a novel framework to model and analyses consumers' purchase decision for product choices based on information obtained from two different information networks. The model has also taken into account variables such as socio-economic, and demographic characteristics. We develop a utility-based discrete choice model (DCM) to quantify the effect of consumers' attitudinal factors from two different information networks, namely, social network and product information network. The network information modelling and analysis are discussed in detail taking into account the model complexity, heterogeneity and asymmetry due to the dimension, layer and scale of information in each type of network. The likelihood function, parameter estimation and inference procedures of the full model are also derived for the model. Finally, extensive numeric investigations were carried out to establish the model framework.

Keywords: consumer purchase decision, social media analytics, discrete choice model, network information analysis

1 Introduction

Due to the prodigious advancement of technologies and increasing competition the lifecycle of products is shortened. An important source of pre-purchase information is online media for many buyers nowadays. Online reviews, feedback and comments have been playing a significant role for customers to make a purchase decision [1], [2]. In a survey by [3], it is found that 50% customers have used online media before making their purchases. Although there are multiple studies discussing the influence of online information on consumer purchase decision the literature contradicts in establishing relationships.

Marketing literature suggests that consumers usually go through five steps while making a purchase decision. In the first step, a customer recognises the problem, the second step is the search of information, in the third step the customer evaluates the alternatives, the customer takes the product purchase decision in the fourth phase and the fifth step is the post-purchase behaviour [4], [5]. Therefore, a customer's purchase decision can be defined as the purchase intention of a product and loyalty to the product in the post-purchase behaviour. However, the discipline of operational research suggests a slightly different approach. [6] introduced a decision-making model which consists of three stages: (1) formulation, (2) evaluation and (3) appraisal.

In each of the consumer decision-making process, digital media and information influence significantly to the process of purchase [7], [4], where customers start with their intention to purchase a product. In contrast, it is natural that loyal customers eliminate competitors from their consideration while making a purchase decision. "True loyalty" is a cognitive behaviour that leads to positive word of mouth (WOM) and repeats purchases [8], [9]. More diverse communication allows effective purchasing process. Recent studies show that social media helps organizations and customers to communicate dynamically and more effectively [10], [11].

Electronic Word of Mouth (eWOM), for instance, social media information, online product rating and reviews play a prominent role in customers' product choice [2], [12]. Online product reviews on the internet and product information and rating posted on social media are found to be useful sources of online WOM communication [13]. Due to the great advancement of wireless communication and internet of things (IoT) product information, reviews and rating on the internet are found to be extremely useful when gathering pre-purchase product information and making purchase intentions by the customers [14]. eWOM communication provides consumers opinion, which is easily and quickly accessible [10]. Due to higher accessibility, the eWOM has been proven to be more effective than offline WOM [15].

As there has been increased availability of information, their accessibility and higher knowledge on consumer psychology, as a consequence, the market competition has also been growing exponentially. Nonetheless, these give more opportunities than ever to be exploited by the organisations to increase product performance, profit margins, etc. More specifically, organizations can easily identify factors affecting the product choice, product features customers likes the most, etc. Hence, the main purpose of this paper is to model these factors and analyse the consumers' purchase decision for product choices and quantify the effect of consumers' attitudinal factors. Specifically, we are mainly concerned with the following questions: (i) how to model and analyse the consumers' purchase decision for product choices and (ii) how to quantify the effect of consumers' attitudinal factors, social factors, demographic factors and economic factors.

There are many model-based and algorithm-based analytical and machine learning techniques to extract social media information [16], [17], [18]. See the book [17], the review paper by [16] and references therein for the details of the methods and analysis. The influence of social media on the product purchase decision is a well-established fact [19], [20], [21], [22]. The online product reviews are also known to have a significant impact on product purchase decisions [23], [24], [25]. Social media and online product reviews are the most popular and important information sources in digital age product purchase decisions [23], [24], [25], [22]. However, consumers express their perception and views on social media and product reviews in many different forms for various products. The expressions can be in nominal, ordinal, ratio or interval scale of measurements while some of the reviews are in text formats. Although many works have been done in the past, it yet remains unclear how we can model the consumers' complex attitude expressed in the social media in different forms, scales and dimensions.

Our aim in this paper is to demonstrate the modelling and analysis of the complexities of the consumers' purchase decision and quantify the effects of many factors related to consumer behaviour. We propose a novel framework to model and analyse the consumers' purchase decision for product choices using a network model and a utility-based discrete choice model (DCM) to quantify the effect of various factors related to consumers' including attitudinal factors, which has received little attention in the literature. We consider two central information sources i.e., social networks and online product reviews also referred to as eWOM along with some socio-economic and demographics factors. Our detailed simulation experiments show that the complexities of data structure measured in many dimensions and scales can be captured through a network model and convert them into some meaningful factors through our latent class analysis.

The rest of this paper is organised as follows. In Section 2, we review the different streams of literature relevant to our research. Section 3 presents the network model formulation, discrete choice model and parameter estimation. Section 4 provides some simulation results and discussion. In Section 5 we discussed some management insights and concluding remarks of the paper.

2 Literature review

Our study builds on literature including marketing and operational research. We link these two to develop a model to build a relationship between digital information and consumer product purchase decision. According to marketing literature, decision-making is a complex cognitive process that involved multiple steps. [26] presented a recognized model of consumer purchase decision- making with five stages: (i) problem recognition, (ii) information search, (iii) evaluation of alternatives, (iv) purchase decision, and (v) post-purchase behaviour. According to operational research literature, widely used [6] model suggests three different stages such as formulation, evaluation and appraisal [27]. Information search (in marketing model) or appraisal stage (operational research) is the first step consumers actively seek for information. In these stages, consumers are motivated to activate their knowledge from the memory or acquire information from external sources [26], [28]. Memory related information is the fundamental source of decision making when a consumer is planning to purchase a product. However, when no such information is available from the memory or a customer is unable to reacquire the information from memory or if the information is not acquired previously, the customer relies upon obtaining information from the external sources [29].

In the external environment, online and offline information sources influence to a greater extend. While offline information sources are peers, family, friends, the company generated information, online sources are internet-based platforms such as social media, web sites, chat rooms, blogs etc. [30]. The motivation for external information search depends on many factors, for example, involvement, the need of cognition and the stage of decision-making process they are in, etc. Especially when it comes to high involvement products (such as electric cars), consumers would put more efforts in searching and processing information from various sources [31]. In the digital age, online information sources have become more significant and widely accessible than offline information sources [1], [2]. Although information search using internal and external sources has been properly established by researchers over the past years [32], [33], [34] majority of these studies have either focused on online or offline information sources only to a limited depth from the marketing point of view for product promotions. Thus, in this study, we focus on online information sources in-depth and to enable manufacturers to forecast their demand.

Social media information regarding products and services is perceived to be more trustworthy by consumers than corporate-sponsored communications transmitted via traditional information sources [35]. Therefore, the importance of social media marketing rapidly grew over the past few years. Social media technologies enabled both consumers and producers to create and distribute information easily. There are many of such established technologies, for example, *Wikipedia* for collaborative writing, IoT devices for content sharing (text, video, and images), *Facebook, Twitter, Delicious, LinkedIn, Youtube* for social networking, social bookmarking and syndication (e.g. ratings, tagging, RSS feeds). These are popular internet-based applications built on the Web 2.0 technological platform, and allow to generate and share user-generated contents [36]. Therefore, social media marketing includes both user-generated content and firm generated content. It is a general understanding that the user-generated contents are stored online by users themselves for relatively reliable sites other than a few exceptions. The firm generated content includes page publishing, stories, apps and advertisements by marketing organisations. These online information sources and offline sources are working as hybrid models to provide information to consumers.

The offline and online media activity can be categorized into three categories: paid media, owned media, and earned media [37]. Paid media can be referred to as a media activity that a company and its agents generate either in online or offline channels. Owned media can be referred to as the media activity that a company and its agents generate in channels it controls such as press releases, brochures, and posts, etc. Earned media are referred to the media activity that is generated by consumers as electronic word of mouth (eWOM) in online platforms. It is challenging for researchers and practitioners is how to use these factors to model consumer purchase decision-making. Yet, this is crucial in the current context.

The internet penetration in the UK is 95% and social media penetration is 66% [38]. This has enabled the eWOM as a powerful communication source for consumer decision making [39]. eWOM is defined as the statement (positive or negative) made by customers (potential, actual, or former) about a product or an organization, which is available via the internet to the mass people and institutions [10]. Further, it is also known as the informal person to person communication done through online channels [39] and has become many-to-many communications between consumers via online platforms [40]. Consumers usually join online groups to seek for advice, information and exchange ideas specific to their interest and this information has made a significant influence in their decision-making process. However, the level of interaction in these platforms depends on many factors and based on the level of interaction the speed and amount of eWOM exchange vary.

Having known the different sources of information on consumers' perception, views and attitudes toward a specific product alternative, it is a complex task to establish connections among those information and extract the actual significant factors of purchase decisions. This is also difficult given that modelling has received less attention in understanding consumer purchase decision [27]. Extensive investigations have been performed and more efforts are continuing to gather and to make the best use of those optimised factors and other extracted information. Big data modellers, data analytics and marketing analytics experts have been making consistent efforts to model and make valid conclusions from the network-based information. Some notable efforts have been made recently [41], [42], [43], [44]. [41] proposed a utility-based analytical model to model the network connections and information exchange while [42] performed a multi-dimensional network analysis for consumer preference modelling. Recently, a Bayesian social learning model has been proposed by [44] based on binary product reviews.

Although there have been some efforts on information extract from the web or social media and modelling and analysis of that information, to the best of our knowledge there is no study to capture such social media data, convert them into suitable factors with consistent dimensions, scales and layers for further modelling and analysis in the context of consumers' purchase decisions. In the next section, we propose an information network model to gather attitudinal or factor information and consumers' purchase decision analysis using a discrete choice model.

3 Methodology

3.1 The information network model formulation

Consider a network $\mathcal{G}$, where a set of individuals $\{1, 2, \ldots, n\}$ are connected by a set of direct and indirect nodes. The jth individual in the network is identified by a vector consisting of a set of p exogenous characteristics $X_j = \{X_{j1}, X_{j2}, \ldots, X_{jp}\}$. The collection of characteristics of n individual is the matrix $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$. The network $G \in \mathcal{G}$ can be represented as an $n \times n$ matrix with elements being either 0 or 1. The entry of the matrix $g_{jk} = 1$ if individual j forms a connection with individual k and $g_{jk} = 0$ if not. We assume $g_{jj} = 0$ for any j, by convention. As the network is directed, $g_{jk} = 1$ does not necessarily implies $g_{kj} = 1$. Network links among the potential customers are depicted in Figure 1 and 2. However, a potential customer gets information on his product information not necessarily from all his links.

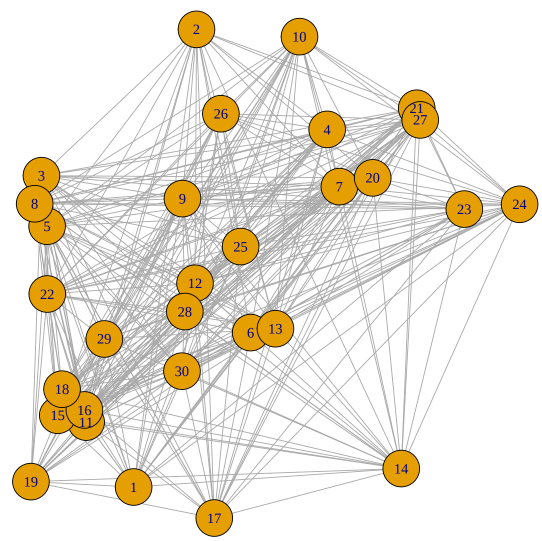

Fig. 1. A typical network with direct and indirect links

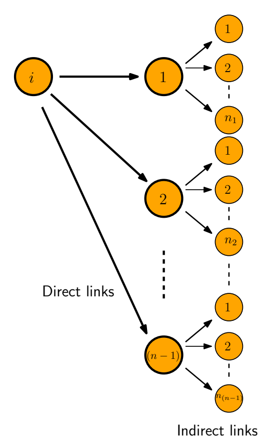

Fig. 2. Network connections of the ith person.

Suppose that the total utility obtained by an individual j for the product alternative i from the network g with the population attribute $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ can be composed into two components: (i) a social network $\mathcal{G}^{(s)}$ and (ii) a product information network $\mathcal{G}^{(p)}$. Define $g_{jk}^{(h)}$ be the connection that individual j forms with individual k for the network h, which takes the values $\{s, p\}$ corresponding to a social and product information network with parameter vectors (θ_u^s, θ_v^s) and (θ_u^p, θ_v^p), respectively. The amount of utility of the jth person for the ith product alternative obtain from a network h is denoted by $u_{ij}(g^{(h)}, \theta)$ and can be defined as

$$u_{ij}(g^{(h)};\boldsymbol{\theta}) = \begin{cases} \underbrace{\sum_{k=1}^{n} g_{jk}^{(s)} \cdot f(\theta_u^s)}_{\text{direct links}} + \underbrace{\sum_{k=1}^{n} g_{jk}^{(s)} \sum_{\substack{l=1\\ l\neq j,k}}^{n} g_{kl}^{(s)} \cdot f(\theta_v^s)}_{\text{indirect links}};\ g^{(s)} \in \mathcal{G}^{(s)} \\ \underbrace{\sum_{k=1}^{n} g_{jk}^{(p)} \cdot f(\theta_u^p)}_{\text{direct links}} + \underbrace{\sum_{k=1}^{n} g_{jk}^{(p)} \sum_{\substack{l=1\\ l\neq j,k}}^{n} g_{kl}^{(p)} \cdot f(\theta_v^p)}_{\text{indirect links}};\ g^{(p)} \in \mathcal{G}^{(p)} \end{cases} \tag{1}$$

For the product alternative i, total utility obtained from the two information networks– social network and product information network by an individual j is assumed to be additive and can be written as

$$\begin{aligned} \mathbf{u}_{ij}(g,\mathbf{Z};\boldsymbol{\theta}) &= \sum_{h\in\{s,\,p\}} u_{ij}(g^{(h)},\mathbf{Z};\boldsymbol{\theta}) \\ &= u_{ij}(g^{(s)},\mathbf{Z};\boldsymbol{\theta}) + u_{ij}(g^{(p)},\mathbf{Z};\boldsymbol{\theta}). \end{aligned} \tag{2}$$

However, the utility values from the social and product network cannot be observed directly and are, in general, latent variables. Therefore, they are described by the measurement model equations as

$$\mathbf{u}_{ij}(g^{(s)},\mathbf{Z};\boldsymbol{\theta}) = \boldsymbol{\Lambda}_s\boldsymbol{\xi}_s + \boldsymbol{\epsilon}_s \tag{3}$$

and

$$\mathbf{u}_{ij}(g^{(p)},\mathbf{Z};\boldsymbol{\theta}) = \boldsymbol{\Lambda}_p\boldsymbol{\xi}_p + \boldsymbol{\epsilon}_p, \tag{4}$$

where equations (3) and (4) are latent factor model measurement equations with parameters $\boldsymbol{\Lambda}_s$ and $\boldsymbol{\Lambda}_p$. The variables $\boldsymbol{\xi}_s$ and $\boldsymbol{\xi}_p$ are two sets of measured variables from social and product networks, respectively. These variables can be categorical and can be measured in nominal and ordinal scale and also can be interval and ratio scale.

3.2 The discrete choice model

A discrete choice model is a utility-based choice model composed of two main components, namely, the observed exogenous variables (demographic and socio-economic) and accumulated utility based on the information obtained from the social networks and product information networks. The choice of a particular product alternative i for an individual j can be generated by the utility obtained from different networks and effect of the exogenous factors. We assume that an individual's choice is a single nominal indicator and we express the choice as a function of the effect from the exogenous variables and total utility from the individual's information network. The model can be expressed as

$$v_{ijl} = \sum_{j=1}^{n}\sum_{l=0}^{q} \beta_{jl}\, x_{ijl} + \sum_{h\in\{s,\,p\}} \beta_{ij}^{(h)}\, u_{ij}(g^{(h)}, \mathbf{Z}; \boldsymbol{\theta}) + \epsilon_{ijl}, \quad \begin{cases} i\ = 1, 2, \dots m \\ j = 1, 2, \dots n \\ l\ = 1, 2, \dots q \end{cases} \tag{5}$$

where v_{ijl} is the random utility of an individual j for the product alternative i, x_{ijl} denotes the lth endogenous variable of the individual j for the product alternative i, $u_{ij}(\cdot)$ is the expected utility of an individual j for the product alternative i from the hth information network, β_{jl} denotes the effects of the lth endogenous variables for the ith product alternative and $\beta_{ij}^{(h)}$ the effect of the hth information network of the jth individual for the product alternative i and ϵ_{ijl} is the error term distributed as $\sim$ Gumbel $(0, \sigma_{ij})$. The discrete choice model is developed under the assumption of an individual's utility-maximization behavior and can be written as

$$Y_{ij} = \begin{cases} 1, & \text{if } v_{ij} = \max\limits_{i,j} \{v_{ij}\} \\ \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

In matrix notation, the model for ith product alternative is written as

$$\mathbf{P}(Y_i = 1) = \frac{\exp\left(\mathbf{X}'\boldsymbol{\beta} + \mathbf{u}'\boldsymbol{\beta}^h\right)}{1 + \exp\left(\mathbf{X}'\boldsymbol{\beta} + \mathbf{u}'\boldsymbol{\beta}^h\right)}. \tag{7}$$

The likelihood of the model can be written as

$$\mathcal{L}(\mathbf{y}_n) = \int_{\mathbf{u}_{ij}} f\left(\mathbf{y}_{ij} \,|\, \mathbf{X}_{ij}, \mathbf{u}_{ij}^{(h)}\right)\, f(\mathbf{u}_{ij} \,|\, \mathbf{Z}_{ij})\, \mathrm{d}\mathbf{u}_{ij} \tag{8}$$

where $f\left(\mathbf{y}_{ijk} \,|\, \mathbf{X}_{ij}, \mathbf{u}_{ij}^{(h)}\right)$ is the likelihood of the n sample observation under the set of covariates $\mathbf{X}_{ij}$ and expected utility $\mathbf{u}_{ij}^{(h)}$. The function $f(\mathbf{u}_{ij} \,|\, \mathbf{Z}_{ij},)$ is the joint density of the latent utility function given the observed set of covariates $\mathbf{Z}_{ij}$.

Model fitting and parameter estimation To estimate the parameters of the model, we perform a two-stage estimation procedure. In stage 1, we estimate the latent utility variables with covariates $\mathbf{Z}_{ij}$. The method analogous to factor score estimation with a mixture of categorical and continuous variables. End of stage 1, utility scores are calculated using the estimated coefficients. These utility scores are then passed to stage 2 with the assumption that these are fixed value explanatory variables for the full regression model. The rest of the parameters of the full model are estimated by the maximum likelihood method in stage 2.

4 Simulation experiment and discussion

To establish the model framework, we have taken the "electric car purchase decision" for our simulation experiment. Electric cars are becoming popular around the world despite their high price. Many countries are imposing laws and providing incentives to their citizens to popularise electric cars to reduce carbon emissions. The "electric car purchase decision" has received significant attention recently and many factors are found to be involved in electric vehicle purchase decision [45, **?**,**?**]. Among the wide range of factors involved in the purchase decisions of an electric car, some crucial factors are highly influenced by social network information and online reviews (product information). Therefore, the "electric car purchase decision" is considered as one of the most suitable examples to check the performance of our proposed model framework.

Table 1. List of variables for simulation experiment

Variable list		
Social network:	**Product network:**	**Demographic Socio-economic variables:**
Purchase price	Purchase price	Age
Cruising range	Cruising range	Gender
Charging time	Charging time	Marital Status
Max. speed	Max. speed	Educational level
Charing point distance	Charing point distance	Income level
Battery life	Battery life	Educational level
Environmental aspect		
Innovation aspect		

In our model framework, we considered that a consumer's decision is influenced by social network data, in which their online reviews along with their socio-economic and demographic variables. Previous studies [45, **?**,**?**] have reported that purchase price, cruising range, charging time, max. speed, charging point distance, battery life, environmental aspect and innovation aspect are the key influential variables for purchasing decisions. We have considered these variables for our study given in Table 1. However, the data structure is complex in types and characteristics. They can be nominal, ordinal, ratio and interval scaled variables due to the type of entries in the social media. We considered that the social media data: purchase price, cruising range, charging time, max. speed, charing point distance are of the ordinal type and battery life, environmental aspect, innovation aspect are nominal type.

In our simulation, we have used parameter values that are consistent with previous studies. In the first step, 5 ordinal variables to represent purchase price, cruising range, charging time, max. speed, charging point distance and 3 nominal variables for battery life, environmental aspect and innovation aspect for $n =$

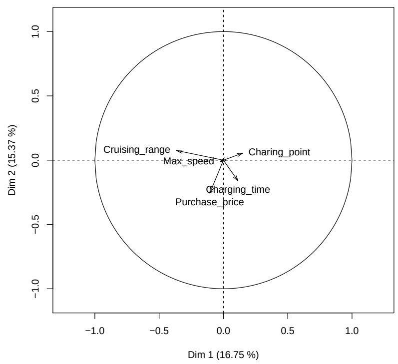

Fig. 3. Correlation circle of the numeric data (five variables) in the principal component axes for social network data.

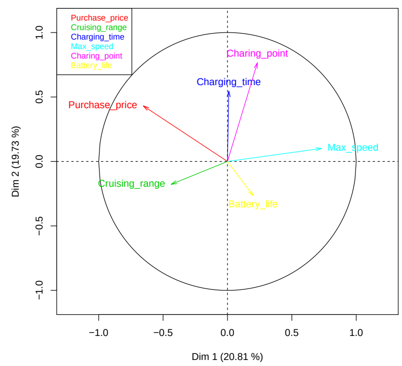

Fig. 4. Correlation circle of the numeric data in the principal component axes for product information data.

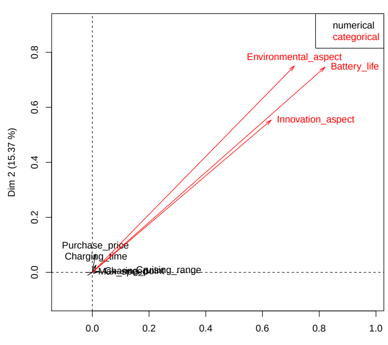

Fig. 5. Direction of variables in the principal component axes for social network data.

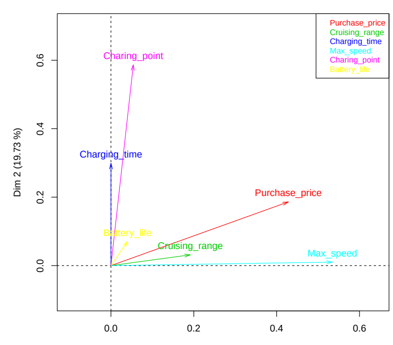

Fig. 6. Direction of variables in the principal component axes for product information data.

Table 2. Estimated parameters, std. error and p-values

Variables/Factors	Estimate	SE	p-values
Intercept	-5.651	1.252	<0.0001
Age	2.176	0.699	0.0018
Gender	1.679	0.688	0.0147
Marital tatus	2.563	0.760	0.0008
Educational level	2.520	0.759	0.0009
Income level	3.373	0.843	<0.0001
Social network	0.063	0.023	0.0073
Product network	0.110	0.027	<0.0001

100 observations are generated. These data assumed to form the latent class measurement of social influence. Another set of 6 ordinal to represent purchase price, cruising range, charging time, max. speed, charing point distance and battery life are generated to obtain the latent class factor online product network. Finally, we have generated 6 categorical variables: age (> 30 or ≤ 30), gender (M or F), marital status (married or single), educational level (lower than/equal to A level or higher than A level) and income level ($\leq 30K$ or $> 30K$).

The full set of data are generated in such a way that a more likely response (favouring to purchase an electric car) is positive or given a higher rating (in ordinal scale) for all 8 items in set 1 (social network influence) and positive for all 6 items in set 2 (online product network). The socio-economic and demographic variable set is also generated according to have a positive and consistent impact as found in the literature. The correlation structures for variables considered in social network latent factor and online product information are displayed in Figure 3 and 4, respectively. At the first stage of estimation, we calculate the scores of the latent social network factor and the score of latent online product information, which is considered as expected utility from these two factors using the equation (3) and (4), respectively. The direction of those variables on the two latent factors (social network factor and online product information) are shown in 5 and 6, respectively. In the second step, the discrete choice model in equation (5) is fitted using these two sets of utility scores along with the socio-economic and demographics factors. Higher values of the generated variables mean that the association between purchase decision and social network & online product information influence is stronger. The estimated parameter values in Table 2 indicate a strong positive impact of the considered variables in the study.

5 Conclusion

In this study, we proposed a framework to model and analyze the impact of network information on consumer product purchase decision: social media and online product information using a latent class approach. The nature of data obtained in social networks are highly complex and measured in several measurement scales. We proposed to model the expected random utility measure for each customer on their purchasing decisions based on data obtained from social networks. The unstructured product preference data are converted into structured format along with their socio-economic and demographic characteristics. The measured expected utility for two latent factors social network influence and online product information with the socio-economic and demographic are then regressed with their purchasing decisions in the discrete choice modelling framework. The likelihood function method is used to fit the full model. Using the electric car purchase decision as a case study with simulated data, we showed that the model framework can find the significant factor explicitly. The model will be useful for decision makers to forecast the demand by understanding how consumers make decisions [16, **?**]. It also provides a yardstick for marketers who make communication decisions. In future, a more generalized approach of complex data extraction will be adopted, and each separate variable will be taken as the latent factor for analysis.

References

1. Jiménez, F.R., Mendoza, N.A.: Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. Journal of Interactive Marketing **27**(3), 226–235 (2013)
2. Prasad, S., Garg, A., Prasad, S.: Purchase decision of generation Y in an online environment. Marketing Intelligence & Planning (2019)
3. PWC: They say they want a revolution, total retail 2016 (2016)
4. Kotler, P., et al.: Marketing management: The millennium edition, vol. 199. Prentice Hall Upper Saddle River, NJ (1999)
5. Zwass, V., Kendall, K.: Structure and macro-level impacts of electronic commerce. Emerging Information Technologies: Improving Decisions, Cooperation, and Infrastructure, Sage, Beverly Hills, CA pp. 289–315 (1999)
6. Holtzman, S.: Intelligent decision systems, Addison Wesley. Reading, MA (1989)
7. Zeithaml, V.A., Berry, L.L., Parasuraman, A.: The behavioral consequences of service quality. Journal of Marketing **60**(2), 31–46 (1996)
8. Shankar, V., Smith, A.K., Rangaswamy, A.: Customer satisfaction and loyalty in online and offline environments. International Journal of Research in Marketing **20**(2), 153–175 (2003)
9. Wang, L.C., Baker, J., Wagner, J.A., Wakefield, K.: Can a retail web site be social? Journal of Marketing **71**(3), 143–157 (2007)
10. Hennig-Thurau, T., Gwinner, K.P., Walsh, G., Gremler, D.D.: Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? Journal of Interactive Marketing **18**(1), 38–52 (2004)
11. Ho, C.H., Chiu, K.H., Chen, H., Papazafeiropoulou, A.: Can internet blogs be used as an effective advertising tool? the role of product blog type and brand awareness. Journal of Enterprise Information Management (2015)
12. Kulmala, M., Mesiranta, N., Tuominen, P.: Organic and amplified eWOM in consumer fashion blogs. Journal of Fashion Marketing and Management: An International Journal (2013)

13. Sen, S., Lerman, D.: Why are you telling me this? an examination into negative consumer reviews on the web. Journal of Interactive Marketing **21**(4), 76–94 (2007)
14. Zhu, F., Zhang, X.: Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. Journal of Marketing **74**(2), 133–148 (2010)
15. Chatterjee, P.: Online reviews: do consumers use them? (2001)
16. Ferrara, E., De Meo, P., Fiumara, G., Baumgartner, R.: Web data extraction, applications and techniques: A survey. Knowledge-based systems **70**, 301–323 (2014)
17. Hai-Jew, S.: Social Media Data Extraction and Content Analysis. IGI Global (2016)
18. Gupta, D., Tripathi, S., Ekbal, A., Bhattacharyya, P.: A hybrid approach for entity extraction in code-mixed social media data. Mon **25**, 66 (2016)
19. Wang, X., Yu, C., Wei, Y.: Social media peer communication and impacts on purchase intentions: A consumer socialization framework. Journal of Interactive Marketing **26**(4), 198–208 (2012)
20. Powers, T., Advincula, D., Austin, M.S., Graiko, S., Snyder, J.: Digital and social media in the purchase decision process: A special report from the advertising research foundation. Journal of Advertising Research **52**(4), 479–489 (2012)
21. Goodrich, K., De Mooij, M.: How socialare social media? a cross-cultural comparison of online and offline purchase decision influences. Journal of Marketing Communications **20**(1-2), 103–116 (2014)
22. Cheung, C.M., Xiao, B.S., Liu, I.L.: Do actions speak louder than voices? The signaling role of social information cues in influencing consumer purchase decisions. Decision Support Systems **65**, 50–58 (2014)
23. Chen, Y., Xie, J.: Third-party product review and firm marketing strategy. Marketing Science **24**(2), 218–240 (2005)
24. Pickett-Baker, J., Ozaki, R.: Pro-environmental products: marketing influence on consumer purchase decision. Journal of Consumer Marketing (2008)
25. Jang, S., Prasad, A., Ratchford, B.T.: How consumers use product reviews in the purchase decision process. Marketing Letters **23**(3), 825–838 (2012)
26. Engel, J., Blackwell, R.D., Miniard, P.: Consumer Behavior, Chicago. Dry Den **12**, 251–264 (1995)
27. Karimi, S., Papamichail, K.N., Holland, C.P.: The effect of prior knowledge and decision-making style on the online purchase decision-making process: A typology of consumer shopping behaviour. Decision Support Systems **77**, 137–147 (2015)
28. Schmidt, J.B., Spreng, R.A.: A proposed model of external consumer information search. Journal of the Academy of Marketing Science **24**(3), 246–256 (1996)
29. Gursoy, D., McCleary, K.W.: Travelers prior knowledge and its impact on their information search behavior. Journal of Hospitality & Tourism Research **28**(1), 66–94 (2004)
30. Klein, L.R., Ford, G.T.: Consumer search for information in the digital age: An empirical study of prepurchase search for automobiles. Journal of Interactive Marketing **17**(3), 29–49 (2003)
31. Zhang, T., Zhang, D.: Agent-based simulation of consumer purchase decision-making and the decoy effect. Journal of Business Research **60**(8), 912–922 (2007)
32. Furse, D.H., Punj, G.N., Stewart, D.W.: A typology of individual search strategies among purchasers of new automobiles. Journal of Consumer Research **10**(4), 417–431 (1984)
33. Beatty, S.E., Smith, S.M.: External search effort: An investigation across several product categories. Journal of Consumer Research **14**(1), 83–95 (1987)

34. Moorthy, S., Ratchford, B.T., Talukdar, D.: Consumer information search revisited: Theory and empirical analysis. Journal of Consumer Research **23**(4), 263–277 (1997)
35. Foux, G.: Consumer-generated media: Get your customers involved. Brand Strategy **8**(202), 38–39 (2006)
36. Kaplan, A.M., Haenlein, M.: Users of the world, unite! The challenges and opportunities of social media. Business Horizons **53**(1), 59–68 (2010)
37. Lovett, M.J., Staelin, R.: The role of paid, earned, and owned media in building entertainment brands: Reminding, informing, and enhancing enjoyment. Marketing Science **35**(1), 142–157 (2016)
38. Number of daily internet users in Great Britain from 2006 to 2019. https://www.statista.com/statistics/275786/daily-internet-users-in-great-britain/. Accessed: 2020-03-18
39. Chan, Y.Y., Ngai, E.W.: Conceptualising electronic word of mouth activity. Marketing Intelligence & Planning (2011)
40. Jalilvand, M.R., Samiei, N.: The effect of electronic word of mouth on brand image and purchase intention. Marketing Intelligence & Planning (2012)
41. Acemoglu, D., Bimpikis, K., Ozdaglar, A.: Dynamics of information exchange in endogenous social networks. Theoretical Economics **9**(1), 41–97 (2014)
42. Wang, M., Chen, W., Huang, Y., Contractor, N.S., Fu, Y.: Modeling customer preferences using multidimensional network analysis in engineering design. Design Science **2** (2016)
43. Besbes, O., Scarsini, M.: On information distortions in online ratings. Operations Research **66**(3), 597–610 (2018)
44. Ifrach, B., Maglaras, C., Scarsini, M., Zseleva, A.: Bayesian social learning from consumer reviews. Operations Research **67**(5), 1209–1221 (2019)
45. Sierzchula, W., Bakker, S., Maat, K., Van Wee, B.: The influence of financial incentives and other socio-economic factors on electric vehicle adoption. Energy Policy **68**, 183–194 (2014)