MASTER

Cross-validation of opinion leader and market maven constructs

Tielen, M.A.J.

Award date:
2014
Cross-validation of opinion leader and market maven constructs

By
Marco Tielen

BSc Mechanical Engineering
Student identity number 0727077

In partial fulfilment of the requirements for the degree of

Master of Science
In Innovation Management

Supervisors:
dr. Ing. J.P.M. Wouters, TU/e, ITEM
prof. dr. C.C.P. Snijders, TU/e, HTI
Subject headings: social structures, network theory, graph theory, market mavens, opinion leaders, online influence
Abstract
The increased attention of businesses on online influencers and social media along with the availability of data gave rise to this study. A quantitative study was performed to validate the identification of opinion leaders and market mavens by use of sociometrics. By combining network data from an online social medium (Twitter) with the results of a scientifically grounded online questionnaire, validation of sociometrics was performed.

The results show that opinion leadership and market mavenism are positively associated with a low to medium relationship to a majority of sociometrics. Interestingly, general descriptive metrics on the intensity and type of communications (e.g. number of retweets, number of times mentioned) perform equal or better than sociometrics in covariation to opinion leadership and market mavenism.

Most important implication based upon this research is that current popular sociometrics do not suffice in identifying opinion leaders and market mavens, and as such should not be used to identify either of them for use in influencer marketing.
Preface
I would like to start with thanking my supervisor, Joost Wouters, for his patience on my progress with the master thesis. He continuously kept a positive attitude and didn’t stop losing faith in me finalizing my master thesis. Also, I would like to address a special thank you to Chris Snijders for his insightful help on the quantitative analysis that constitutes a major part of this master thesis.

Furthermore, I’d like to express my appreciation for my colleague Erik Tromp for his guidance, talks and the effort that he put into facilitating the technical environment needed to collect and process the data. Without his help, this study wouldn’t have been conducted. Another colleague I’d like to thank is Mark Graus for his interest and the help with reviewing the questionnaire. Also, I’d like to acknowledge all the Twitterers that took the time and effort to complete the questionnaire. Without their responses, all the effort would have been in vain.

From a social perspective, a warm thank you to my parents for giving me the opportunity to obtain a degree at the Technical University of Eindhoven and for their support during my entire “career” as a student. Finally, my enormous gratitude to my future wife, Joyce, for her support and motivation and patience during my study, but moreover during finalizing this master thesis.

Marco Tielen
Executive summary
Alongside with the rise of social media, the interest in influencer marketing as well as the interest in social network analysis has rapidly increased. Network data has become available in abundance, whilst influencers have gained the possibility of spreading their word of mouth faster and to a larger audience. The interest of companies is to address these influencers for aid in their marketing efforts, because influencers are “disproportionally capable of influencing others” (Rogers and Cartano, 1962).

The prerequisite to addressing influencers is to be able to identify them, which is investigated in both marketing research and research on social networks. To identify influencers, the research field on marketing has created self-report surveys that are well-tested and based on grounded theory. The research field on network theory has created “sociometrics” that are based on the structural position of people within a network. This study aims to perform validation of these sociometrics to identify influencers in relation to influencer constructs created in marketing theory.

Theory and hypothesis
Influencers are people that have “a greater than average reach or impact through word of mouth in a relevant marketplace” (Worth of Mouth Marketing Association) and can trigger word-of-mouth with possible higher diffusions of information, innovation and adoption (of products) (Rogers and Cartano, 1962; Feick and Price, 1987; Lazarsfeld, Berelson, and Gaudet, 1944). Within this definition of influencers, researches have been able to identify multiple typologies of influencers including opinion leaders (Katz & Lazarsfeld, 1957), early adopters (Rogers, 1962), market mavens (Feick and Price, 1987) and lead-users (Hippel, 1986). The literature review performed in this study focuses on opinion leaders and market mavens, because they are associated with enhanced information diffusion and distinct structural positions within a social network (Katz, 1957; Burt, 1999).

Opinion leaders are "individuals who exert an unequal amount of influence on the decisions of others" (1962, p. 435) and play a crucial role in information diffusion between mass media and people within their social network (Katz and Lazarsfeld, 1957), whereas market mavens are consumers who have “information about many kinds of products, places to shop, and other facets of markets, and initiate discussions with consumers and respond to requests from consumers for market information” (Feick & Price, 1987, p. 85). To identify opinion leaders and market mavens, scientific marketing research has developed self-reported surveys (Flynn et al., 1996; Feick and Price 1987) that provide a scale for the amount of influence.

The field of network theory has developed multiple sociometrics, that describe different characteristics of actors within a social network. Commonly used sociometrics include PageRank (Page et al.,1999), HITS (Kleinberg, 1999) and Structural Holes (Burt, 1992). By combining the properties of these sociometrics with the characteristics of the influencers described in marketing theory the following hypothesis were created and tested:

- **HO1:** Opinion leaders are positively associated with having strong ties in the social network
- **HO2:** Opinion leaders are positively associated with betweenness centrality
- **HO3:** Opinion leaders are positively associated with closeness centrality
- **HO4:** Opinion leaders are positively associated with having structural holes in the social network
- **HM1:** Market Mavens are positively associated with authority (HITS algorithm)
- HM2: Market Mavens are positively associated with PageRank
- HM3: Market Mavens are positively associated with in-degree

Additionally, explorative research was performed on the relationships between the opinion leadership and market mavenism constructs and general descriptive metrics on the intensity and type of communications from and towards actors within the network.

**Method**

Social network data was collected from Twitter based on the keywords related to the elections of the Dutch House of Commons. Using the tool SNAP (Stanford Network Analysis Project) created by Stanford University sociometrics were calculated. Secondly, based on the research by Flynn et al. (1996) and Feick and Price (1987), questionnaires on opinion leadership and market mavenism were created. These questionnaires were sent to actors randomly subtracted from the dataset collected from Twitter. Using covariance analysis the relationships between sociometrics and both the opinion leadership and market mavenism construct were analyzed.

**Results**

Results largely confirm the hypothesis, but with low to weak correlations between the sociometrics and both constructs. Moreover, general descriptive metrics on the intensity and type of communications (such as number of retweets) perform equal or better in covariance to both opinion leadership and market mavenism. Another interesting finding is the importance of active membership of a political party in being opinion leader or market maven. Although data only stems from tweets about politics, being involved with the subject is an important variable for both constructs.

Only two sociometrics have a notable difference in covariance between opinion leadership and market mavenism. PageRank and Betweenness have a stronger relationship to opinion leadership than towards market mavenism. All other sociometrics are somewhat alike in covariance towards both constructs, making PageRank and Betweenness the only variables that are somewhat capable of making a distinction between both constructs.

In general, one can state that the identification of either opinion leaders or market mavens is not valid through use of discussed sociometrics (alone). However, increasing the likelihood of targeting opinion leaders and/or market mavens is possible, because some sociometrics exclude a majority of non-opinion leaders and market mavens. Thereby potentially reducing costs of influencer marketing activities.

**Discussion**

One main limitation to the research is the incorporation of only one communication channel into the research. The social network structure which is used for the calculation of sociometrics is solely based on communications on Twitter. Therefore, it might be possible that regarding the subject of politics, the constructed network misses important nodes that can potentially have a significant impact on the network and values of the used sociometrics. Nonetheless, important implications can be derived from this research.

Most important implication is regarding the use of sociometrics to identify opinion leaders or market mavens. Current sociometrics do not suffice in identifying both influencers, and as such should not
be used to identify either of them for use in influencer marketing. Moreover, current use of the construct of opinion leadership in relation to sociometrics that indicate the importance of a node within a network in regard to the spread of information is not correct. In return, marketing should be careful in adopting techniques from network theory to identify influencers for future research. Additional research should be necessary to better identify influencers and the different types of influencers based on data available in online social networks.

Based on this research, one might conclude that research on network theory has to redefine their constructs for important nodes in the spread of information. The focus within network theory is on the spread of information rather than on the actual influence a person has on others, which isn’t in line with marketing theory from which the construct of opinion leadership stems. Future research to identify opinion leaders and/or market mavens based on network information could therefore include dynamic (time based) qualitative data, such as the messages being send. For example, relatively new text analysis methods such as sentiment and emotional analysis could be used to detect opinion changes over time of actors or groups/cliques towards certain subjects, after which potential sources of influence could be determined.
# Table of Contents

Abstract ................................................................................................................................................... 3
Preface..................................................................................................................................................... 4
Executive summary .................................................................................................................................... 5
Theory and hypothesis ........................................................................................................................ 5
Method................................................................................................................................................ 6
Results ................................................................................................................................................. 6
Discussion ............................................................................................................................................ 6
1. Introduction ................................................................................................................................... 10
2. Literature Review .......................................................................................................................... 11
   2.1. The Value of Identifying Influencers ..................................................................................... 11
   2.2. Opinion Leaders in Marketing Research ............................................................................... 13
   2.3. Market Mavens in Marketing Research ................................................................................ 14
   2.4. Influencers in Network Theory .............................................................................................. 16
   2.5. Identification of influencers based on sociometrics ............................................................. 20
   2.1. Validity of Influencer Identification methods in Network Theory ........................................ 21
   2.2. Learnings outcomes .............................................................................................................. 23
3. Method .......................................................................................................................................... 23
   3.1. Network data collection ........................................................................................................ 24
   3.2. Questionnaire data collection ............................................................................................... 25
4. Results ........................................................................................................................................... 26
   4.1. Opinion Leadership ............................................................................................................... 26
   4.2. Market Mavenism .................................................................................................................. 32
5. Discussion ...................................................................................................................................... 38
   5.1. Conclusion ............................................................................................................................. 38
   5.2. Managerial Implications ........................................................................................................ 39
   5.3. Academic Implications........................................................................................................... 39
   5.4. Limitations and recommendations for further research ....................................................... 40
6. References ..................................................................................................................................... 41
7. Appendices .................................................................................................................................... 49
   7.1. Appendix I: PageRank Calculations ....................................................................................... 49
   7.1. Appendix II: Opinion Leadership Scale Items ........................................................................ 50
   7.1. Appendix III: Market Maven Scale Items ............................................................................... 51
   7.2. Appendix V: Literature Quality Overview .............................................................................. 52
1. Introduction

Identifying influencers has been in the interest of companies and academic research for years. Ever since the discovery of “special” individuals that are disproportionally capable of influencing others (Rogers and Cartano, 1962), initiatives to identify and address them have been in place. The reasoning behind it is simple: if there are individuals capable of influencing others, why not use them for a specific cause? A good example is the use of influencers for the launch of new products. By persuading influencers to advice others about a product, one could potentially increase the adoption of the product and consecutively increase sales.

The first step of activating influencers is identification, which has been investigated in both marketing research and research on social networks. The research field on marketing has created self-report surveys, whereas research on social networks has focused on the structural position of people within a network to identify influencers. Validation of the latter technique has been done by simulation, which is a usable but also questionable technique since simulation models are based on assumptions. For example, almost all values assigned to the actors that are used to predict behaviour are randomly distributed, whereas in a social network this is typically not the case. A certain structural position implies certain behaviour.

Another way to perform the validation for the identification of influencers through their structural position would be to perform a cross-validation with self-report surveys from marketing that are based on grounded theory. The results can clarify whether current metrics on structural network position are valid to be used to identify opinion leaders and whether marketing and network theory are searching for the same constructs.

An important argument for performing this type of research is the improvement in effort and cost-efficiency of identifying influencers based on their network position in relation to self-report surveys. With the rise of web 2.0 on the internet providing a multitude of online social platforms (e.g. Facebook, Twitter, LinkedIn and privately owned forums), all data regarding social networks is readily available. Questionnaires on the other hand are more cumbersome and time-intensive and have to deal with response-rates that can prevent one from finding important influencers.

This literature study will investigate on both the marketing and social network theory solutions of identifying influencers and try to combine both theories to come to a theoretical idea for cross-validation. First, the importance of identifying influencers will be addressed. Secondly, instruments and theories from the field of marketing will be discussed, followed by instruments and theories from the field of social network theory. Afterwards, a critical comparison between both fields of study will be made, discussing both the good and bad aspects. Additionally, a solution will be designed to address influencers in a political context that can be used for cross-validation with social network theory.
2. Literature Review

The goal of this study was to validate scales developed in the scientific field of network theory with grounded theory from the scientific field of marketing. Therefore, the literature review comprehends both scientific domains, that are discussed separately, but combined in the last stage of the review. A purposive sample of the extensive literature review is selected and incorporated. In order to assess the quality of the literature the number of citations and the list for journal quality by Harzing forty-fifth edition (1 april 2012) were used.

2.1. The Value of Identifying Influencers

In the past 60 years, influencers have gotten the interest of many researchers. Influencers are thought to be special individuals that have “a greater than average reach or impact through word of mouth in a relevant marketplace” (Word of Mouth Marketing Association). The main reason to this interest is that influencers, because of their impact in the marketplace, can trigger word-of-mouth with possible higher diffusions of information, innovation and adoption (of products) (Rogers and Cartano, 1962; Feick and Price, 1987; Lazarsfeld, Berelson, and Gaudet, 1944).

Although the existence of influencers is proven by many research studies (Feick and Price, 1987; Lazarsfeld, Berelson, and Gaudet, 1944), there is still some controversy on this matter. The most recent discussion is between Watts (2007) and Gladwell (2000), in which Watts attacks Gladwell’s idea that influencers are the main drivers behind viral effects. Watts claims that if this would be the case, influencers could create a trend each time they wanted to. According to Watts the society needs to be susceptible to a certain idea for it to get adopted, whereas Gladwell puts a large emphasis on the power of influencers. However, Clark et al. (2001) have proven that role models can have a higher degree of influence on certain individuals in relation to others in their direct social environment. Secondly, Marsden et al. (2003) performed a test, which showed that brand preference was strongly influenced by the visible use of that brand by an individual that fit a certain influencer profile (ACTIVE profile; Keller and Berry, 2003). To be sure of the effect, the experiment was repeated with a non-influencer, which showed no significant influence on brand preference.

Scientific research on the possible business returns of targeting influencers hasn’t been performed. One major reason for the lack of research on this matter is the difficulty in addressing the returns of influencer’s activities. However, there are some compelling examples available in the book “Connected Marketing” by Kirby and Marsden (2006). For example, the launch of Windows 95 by Microsoft was preceded by a seeding trial to 450,000 opinion-leading PC-users. This generated an enormous goodwill and an “army of product advocates”. The launch became a great success with the first one million products sold within 4 days in respect to the previous record of a million copies sold within 40 days. It’s very difficult to prove the relationship of this return with the advocate activities of the opinion leaders that received a free copy, but the results are far above average taking into account the historical timing of this event.

Another important note is that influencers are likely to be most powerful in cases of uncertainty (Pfeffer et al. 1976). In case of easy decisions, people are less dependent on others for decision-making. However, when matters become complicated, people are more likely to refer to others for advice (Ryan and Gross, 1943; Valente, 1995; Rogers, 1995). Buying complex products is a good example of a situation in which there can be much uncertainty for the potential buyer. The potential buyer will very likely turn to people within their social network for advice. Most likely to persons that
are somewhat knowledgeable on the specific product at hand. As it turns out, these are often the same persons (Lazarsfeld, Berelson, and Gaudet, 1944; Valente, 1996) for the same subject (Tang, 2009).

The main mechanism behind their influence is word of mouth (WOM or interpersonal communication), which has an impact on sales as proven by empirical evidence (Basuroy, Chatterjee, and Ravid 2003; Eliashberg and Shugan 1997; Godes and Mayzlin, 2004; Liu, 2006). Influencers are strongly associated with WOM (Martilla, 1971), with an emphasis on advice-giving (Richins and Root-Shaffer, 1988). Moreover, some types of influencers are associated with direct influence on the decisions of others, meaning they could influence buyer’s decisions (Rogers and Cartano, 1962).

To be able to identify influencers, it is important to note the differences between typologies of influencers. In the past 60 years of research a wide variety of typologies of influencers have appeared in the literature. Most notable typologies are opinion leaders (Katz & Lazarsfeld, 1957), early adopters (Rogers, 1962), market mavens (Feick and Price, 1987) and lead-users (Hippel, 1986). All of them with different and partial overlapping characteristics and functions to the marketplace. In reference to the descriptions of these types of influencers, it is possible that individuals can have a fit with more than one typology (Morrison et al., 2000).

The focus in this study will be on the validation of the identification of influencers based on their structural position within a network. The structural position within the network is determined based upon their communications regarding a specific topic. Therefore a strong association of an influencer type with information diffusion and social relations within the marketplace is crucial in their identification within a social network. Based on the earlier stated typologies market mavens and opinion leaders are the best fit for such a study. Because, according to their theoretical founders, opinion leaders are the information filters in communication between mass media and the general public (Katz & Lazarsfeld, 1957) and have as such a great influence on information flows and are capable of adding their personal sentiment to this information. Additionally, market mavens are market experts who have “information about many kinds of products, places to shop, and other facets of markets, and initiate discussions with consumers and respond to requests from consumers for market information” (Feick & Price, 1987, p. 85). Again communication is a core component in their typology and as such relevant for this study.

In turn, lead-users are users of a product or service that currently experience needs still unknown to the public and who also benefit greatly if they obtain a solution to these needs (Hippel, 1986), whereas early adopters are typical first-buyers of products and are often inclined to help in the development of a product (Rogers, 1962). Both characteristics do not necessarily imply a certain structural social role with a particular type of communication, thus making them difficult to identify based on a social network based on communications.
2.2. Opinion Leaders in Marketing Research

The main emphasis of this study is on the validation of the identification of opinion leaders rather than market mavens, since theory on opinion leaders and scales to identify opinion leaders are well-defined and based on grounded theory. First, a (historical) overview from marketing research is given on opinion leaders to define the construct. Secondly, methods to identify opinion leaders from marketing research are discussed.

2.2.1. Opinion Leaders in marketing research

In 1955 Katz and Lazarsfeld introduced a breakthrough theory to the scientific world about a two-step flow of communication. The theory encompasses the diffusion process of information from an initial sender to the large majority. Katz and Lazarsfeld state that this process is a two-step flow, because the message of the initial sender will reach the large majority of society through special individuals called “opinion leaders”. The two-step flow model has received criticism and eventually developed to the multistep flow by Robinson (1976; Figure 1). Most important adjustment to the model is that information flows are not as unidirectional as proposed in theory (Troldahli and Van Dam, 1965).

Katz (1957) elaborated in more detail on the definition of opinion leaders. Katz states that the influence of opinion leaders is related to:

- The personification of certain values (who one is);
- Competence (what one knows); and
- Strategic social location (whom one knows)

Who one is, is especially important to the person that is influenced. For example, as shown in other studies (Lazarsfeld and Merton, 1954; Marsden, 1988), homophily can be an important determinant in the possibility that someone influences another. People tend to be “attracted” to someone when there is some form of identification in the other.

The competence aspect is important, because the impact of a fashion leader saying something about clothing is different from that same person telling about politics (Katz, 1957). Tang (2009) has also proven that in large scale networks, topics are important in determining the influence of a person.

![Figure 1: Multistep Flow Model by Robinson (1976)](image-url)
Strategic social location is an aspect that is elaborated on later in this study. Katz (1957), in his article, gives the example of a study on the spread of the farmer-innovators responsible for the diffusion of hybrid seed corn in Iowa in which was concluded that these leaders also could be characterized in terms of the relative frequency of their trips out of town. In other words, they probably had access to people (and thus information) that their local competitors hadn’t.

Next to the detailed information in the influence of opinion leaders, Rogers and Cartano give a definition about opinion leaders themselves: opinion leaders are "individuals who exert an unequal amount of influence on the decisions of others" (1962, p. 435).

2.2.2. Identifying opinion leaders from a marketing theory perspective

In marketing research 2 different methods are present to identify opinion leaders. The first is interviewing (Katz, 1957; Feick and Price, 1987), which heavily relies on research of the second method: surveying. Grounded theory on these surveys comes from King and Summers who consecutively based their design on the first design available in marketing research by Rogers and Cartano (1962). The item scale from King and Summers has gotten small revisions from Childers (1986), Reynolds & Darden (1971) and Flynn et al. (1996). The latter research by Flynn et al. (1996) has been tested it in 5 different contexts and across these 5 contexts proven itself with having “high internal consistency and test-retest reliability, yield normally distributed scores, and free from acquiescence response bias” (Flynn et al. 1996, p.145).

2.3. Market Mavens in Marketing Research

Although, the emphasis of this study is on the validation of the identification of opinion leaders, market mavens are also an interesting type of influencers. The theory on market mavens is less well-grounded (compared to opinion leadership), but has proven that market mavens have a distinct, well-defined and important role in the marketplace that provides ammunition for future research. Therefore, the inclusion of market mavens is used for exploratory research. The following chapters include an overview on the construct of market mavens and methods from marketing research to identify them are discussed.

2.3.1. Market Mavens in marketing research

The word “maven” stems from Yiddish and means the one that understands. A maven typically is an expert, therefore market maven means “market expert”. Market mavens are consumers who have “information about many kinds of products, places to shop, and other facets of markets, and initiate discussions with consumers and respond to requests from consumers for market information” (Feick & Price, 1987, p. 85). They are the experts in the market place and the diffusers of information within it. In relation to the opinion leaders, market mavens have a more general knowledge of markets. An opinion leader is more focussed towards specific products and ideas, whereas to market maven is thought to span many product categories. Market mavens have similarities with opinion leaders and early purchasers/adopters, but are nonetheless distinct from both typologies (Feick and Price 1987).

Next to being market experts, they are also active diffusers of information (Feick & Price, 1987). Feick states that market mavens are “individuals who have information about many kinds of products, places to shop, and other facets of markets and initiate discussions with consumers and respond to requests from consumers for market information” (pp. 354, Price et al., 1988). They are much more likely to read advertising and local direct mail (Price et al., 1988) than others and like to give feedback and enter discussions on market activities, products and do’s or dont’s whenever someone asks.
Compared to opinion leaders, market mavens are not as often the initial sender of information, but like to respond based on questions. They have a desire to be a competent helper to other consumers.

2.3.2. Identifying Market Mavens from a marketing theory perspective
Identifying market mavens has solely been discussed by Feick & Price (1987), because they had to prove their mere existence. Feick and Price (1987) used data gathered by phone surveys with questions about opinion leadership from the survey by King and Summers’ (1970), with additional questions about purchasing innovativeness and questions regarding market mavens. By use of statistical analysis, Feick and Price concluded that opinion leaders, early purchasers/adopters and market mavens are distinct concepts (although there certainly was some overlap). The self-report questionnaire was created based on a set of 40 questions, that was diminished to 19 by use of experts and tested in a pilot study on 265 MBA students. By use of factor analysis and computation of Cronbach’s alpha 6 items were selected (Appendix III), that also proved to give good results in the final survey.
2.4. Influencers in Network Theory
Identification of influencers from a network theory perspective is more cost-effective than traditional surveying. An additional advantage is that there can be no non-response, because network data is readily available from a diversity of online platforms, ranging from forums and guest books to complete social networking tools like LinkedIn and Twitter. The current scales developed in Network Theory to identify influencers are discussed after a small introduction to the field of network theory in general. Finally, the application of sociometrics to the constructs of opinion leader and market maven are examined.

2.4.1. A small introduction to Network Theory
Network theory is a study about the relations between objects. It has become quite popular in recent years and is used in a variety of (scientific) research fields. Computer sciences, math, biology and economics are all examples of research areas in which network theories and tools are used. One important variant of networks is the \textit{social} network, which has a special focus on the relationships between social entities. These social entities are often persons, but they also include groups of people, business units or organizations (as long as the social aspect is present). The popularity of the subject becomes clear by gathering the amount of indexed articles with “social network” in the title from Google Scholar on a yearly basis (Figure 2; Borgatti and Halgin, 2011).

![Figure 2: Share of All Articles Indexed in Google Scholar with “Social Network” in the Title, by Year](image)

The most common way of analysing social networks is with the aid of social graphs. Invented in around 1930 by Jacob Moreno, a social graph contains a set of nodes (also called objects, vertices, actors) and edges (also called ties, relations, connections). Together they form the social graph, which becomes a representation of the network of the people involved (\textit{Fout! Verwijzingsbron niet gevonden}). Attributes can be assigned to nodes as well as to relationships to be able to do research on a particular matter and even be able to control for variables like sex and age.

![Figure 3: Social Graph of an ego centred network on Facebook](image)
There are 2 basic types of network analysis. First, there is ego-centred network analysis. As implied by the name, one analyses data from the network around single persons (ego’s). This data can be gathered with the help of traditional surveys in which the respondents are asked about the people they interact with and about the relationships among those people. This type of research is typically done to evaluate a person’s network (size, available resources, diversity etc.) or to find attributes that are comparable to the alters (homophily). The advantage of ego networks data is that it can be gathered with random sampling which makes it possible to use classical statistical analysis.

The second type of analysis is complete network analysis. The associated data to this analysis contains all the relations between a set of respondents (all the relations between the actors in the set are known, not only the ones a certain person knows). This type of data provides insight into the structure of a particular network, which could possibly help to predict certain effects (e.g. contagion, actor's behaviour etc.). However, in contrast to ego-centred network analysis random sampling can’t be used with complete network analysis, because one must gather the data from all the actors in a particular network.

2.4.2. Influencers in Network Theory

In social network theory the idea of influencers was adopted quickly, but in general a deviation from the original definition can be found in the literature. Network theory literature places a large emphasis on the characteristic of spreading and diffusing information in a network, thereby overlooking the aspect of actually influencing others. For example, the commonly used definition of opinion leaders in network theory is: “special individuals who disproportionately impact the likelihood that information will spread broadly” (Weimann, 1994; Gladwell, 2000; Keller and Berry, 2003; Goldenberg, Han, Lehmann, and Hong, 2009). Although this is different from the definition in marketing research, a reference to the original concept from Katz and Lazarsfeld is always present.

In current research from network theory a large diversity of scales and models have been constructed to measure social relationships. These scales and models are called sociometrics and include basic network metrics like degree centrality and closeness centrality and more complex and theory embedded metrics like structural holes. Below these scales will be explained and discussed.

2.4.3. Sociometrics

Finding influencers online based on network theory is a popular theme in research on social networks. It started with basic network metrics for nodes in a network that were claimed to have a relation to influencers and are used in a diversity of research (Nolker, 2005; Wuyts & Bulte, 2007; Iyengar et al., 2011). They are well documented in an early network theory book from Wasserman and Faust (1994) and in a book by Hanneman and Riddle (2005). With the rise of the World Wide Web, more sophisticated types of network scales were developed to be able to find the most relevant pages for users of the World Wide Web. Although these scales are used to find pages, they can also be used for social network analysis.
2.4.3.1. **Centrality Measures**

Centrality measures were one of the earliest types of measurements to find users that are in the centre of the network and thus in the position to control information flows and gather large amounts of information and resources (others can’t). Most used centrality measures include degree centrality, closeness centrality, betweenness centrality and eigenvector centrality. Extensive information on these sociometrics are published by Wasserman and Faust (1994) and by Hanneman and Riddle (2005):

*Degree Centrality (number of alternative trading partners)* represents the number of ties that a node has. The more ties an actor has, the more power they (may) have. In the star network, Actor A has degree six, all other actors have degree one. As a result A has more opportunities and alternatives (and thus choices) compared to the other actors. This autonomy makes them less dependent on any specific other actor, and hence more powerful (Hanneman and Riddle, 2005).

Degree centrality can be divided into 2 types: *in-degree* and *out-degree* centrality. A high in-degree means that a person is receiving many ties, whereas a high out-degree means that the actor “sends” many ties. The direction of the tie (receiving versus sending) is of importance to the interpretation of the scale. Actors with a high in-degree are sought by many people and can be seen as an authority or a prominent actor. Actors with a high out-degree seek many people themselves and could be hubs, influencers or highly gregarious people (Hanneman and Riddle, 2005).

Note however that degree centrality is a relative local measurement, because a person can have a high centrality although not connected to the entire network. Degree centrality only includes direct ties to an actor (Hanneman and Riddle, 2005).

*Closeness Centrality* represents the distance of a person to alters in the network. More literally, it is the inverse of farness, which is the sum off the distances to all nodes in the network. The sum of distances can be calculated both on shortest paths and/or based on all (possible) paths (Hanneman and Riddle, 2005).

*Betweenness Centrality* can be calculated by counting the number of times that ego lies on the shortest path between two other nodes. This is an interesting scale, because it defines the possibility that ego can control a (information) flow between two nodes which gives a form of power to ego (Hanneman and Riddle, 2005).

*Eigenvector Centrality* is more sophisticated than the other centralities in the sense that it also includes the relative importance of the actors ego is connected to. Eigenvector centrality acknowledges that not all actors are equal and bases the centrality value of ego on the centrality values of the alters ego is connected to. A good example of an eigenvector centrality measurement is PageRank, which is used by Google to rank web pages and is elaborated on below (Hanneman and Riddle, 2005).
2.4.3.2. PageRank

Pagerank was developed by Larry Page (Page et al., 1999) and tries to measure the relative importance of a linked “document” within a set. More literally, it is the probability distribution that represents the likelihood that someone clicking randomly at links will arrive at the “document”. PageRank acknowledges the fact that not everything in the network is equal to each other by transferring PageRank values of the connections. The overall structure of a network is thereby reflected in the PageRank value of a page. To obtain the PageRank of a document, an iterative process is performed with the end result being a numeric value representing the probability of landing on the page randomly.

Kwak (2010) and Zhang (2007) have used PageRank in a social context to find experts or opinion leaders. Zhang has used human raters to make a comparison with the PageRank results and found that they showed a high degree of correlation, which would lead to the conclusion that structural information can be used to determine social roles within a social network. Next to PageRank, Zhang (2007) also tested in-degree and the HITS algorithm for correlation with human raters and concluded that both scales can be used to predict social roles.

2.4.3.3. HITS

HITS (Hyperlink-Induced Topic Search (HITS)) is the predecessor of PageRank and was created in the context of web search engines with the purpose of finding the most relevant websites to a keyword search (Kleinberg, 1999). The HITS algorithm produces 2 components: hubs and authorities. Hubs are the webpages that are estimated to contain the most valuable links, whereas authorities contain the most valuable content. Translating this to a social network, this model is supposed to find authoritative persons having the most valuable information and hubs that are responsible for spreading the information. Although not many studies have performed the HITS algorithm in a social context, the method is commonly supported by a diversity of social network statistics tools (i.e. Gephi and SNAP).

2.4.3.4. Structural Holes

Burt (1992) relates the principle of structural holes (the absence of connections between social clusters) to opinion leaders in his book “Structural Holes: The Social Structure of Competition”. He explains that opinion leaders are brokers who carry information across boundaries between social groups. People with many structural holes are in the position to act as such a broker, because they are the ones that span the chasms between social clusters. Thus, people with structural holes are key to the spread of information in the network. Additionally, Burt says that the familiar two-step flow of communication is a compound of two network mechanisms: contagion by cohesion through opinion leaders gets information into a group and contagion by structural equivalence (does ego reside in the same structural position within a network as alter) triggers adoptions within the group.

2.4.3.5. Greedy Hill Climbing

In 2003, Kempe et al. approached the case of finding influencers with a greedy algorithm. Greedy algorithms work with one basic rule: at each stage make the optimal choice on a local perspective, without looking at the global consequences. Kempe et al. have used this to find persons in a network that are most likely to create a maximized spread within the network. Based on the greedy algorithm, Kempe et al. outperformed the above mentioned traditional degree centrality and closeness centrality in spread velocity.
2.5. Identification of influencers based on sociometrics

In this chapter we discuss which sociometrics are, based on scientific research, supposed to be able to identify opinion leaders and market mavens. By synthesis of theories from marketing and social network theory hypothesis are developed of the possible relationships between sociometrics and both influencer constructs.

2.5.1. Identifying Opinion Leaders from a network theory perspective

The basic characteristic of an opinion leader is based upon two aspects: opportunity to influence and opportunity to spread. Burt (1999) defines them in social network structures and states that opinion leaders need: (1) strong connections/ties and (2) being in a structural position in between social groups to be able to act as an information broker. The first characteristic is confirmed by Coleman (1988) in which Coleman states that the strong connections are needed to be able to influence others. This statement is confirmed by both Weimann (1982) and Brown and Reingen (1987) who state the following: “strong ties were also perceived as more influential than weak ties, and they were more likely to be utilized as sources of information for related goods” (pp. 350). Therefore, the first assumption is that opinion leaders can be identified by having strong connections, which can be found based on the frequency of interaction (Granovetter, 1978). Therefore we assume the following hypothesis:

HO1: Opinion leaders are positively associated with having strong ties in the social network

The second characteristic, being in a structural position to act as an information broker, determines the opportunity to spread information. As discussed earlier, Katz and Lazarsfeld (1955) noticed a two-step flow in information in which opinion leaders are the main pass-through of information from mass-media to the general public. Although this theory later on developed, the main premise of opinion leaders acting as brokers of information between different groups has always been the key concept (Robinson, 1976; Weimann, 1982; Valente, 1996; Nisbet, 2009).

In general, information is always likely to stay within a certain social group (Wu, 2003). Additionally, social networks tend to have people that are the connectors between social groups (Zaheer and Bell, 2005), that bridge structural holes (Burt, 1999). Therefore, the actors that bridge social groups have opportunity to act as a broker and control the information flow between the groups. Such a structural position within a network is associated with a high betweenness and closeness centrality (Kiss and Bichler, 2008; Lee, 2010). Therefore we assume the following hypothesis:

HO2: Opinion leaders are positively associated with betweenness centrality

HO3: Opinion leaders are positively associated with closeness centrality

HO4: Opinion leaders are positively associated with having structural holes in the social network

2.5.2. Identifying Market Mavens from a network theory perspective

Research on the structural network position of market mavens hasn’t been performed so far. Therefore, network theory has to be combined with descriptions about market mavens from a marketing perspective to make sound hypothesis about the identification of market mavens based on their structural position within a network.
Based on marketing theory, market mavens are (1) experts and (2) like to interact based on questions from others. Being an expert is associated with both PageRank and HITS (Page, 1999; Zhang, 2007), because PageRank and HITS are both created to find the “expert” page on the web that contains the answer to your questions/keywords. Therefore we assume following hypothesis:

**HM1:** Market Mavens are positively associated with authority (HITS algorithm)

**HM2:** Market Mavens are positively associated with PageRank

Secondly, market mavens are thought to be responsive in their communications. They are sought by others, because they’re experts and have information regarding the market (Feick & Price, 1987). Many people seek contact with them, making it likely for them to have many incoming connections. Therefore we assume the following hypothesis:

**HM3:** Market Mavens are positively associated with in-degree

### 2.1. Validity of Influencer Identification methods in Network Theory

The main reason to perform this research is to test the validity of the use of sociometrics to identify opinion leaders and market mavens. Therefore, main validation methods on sociometrics used till date and critique on the marketing surveys are discussed. The latter point is relevant, since it is implied that current marketing surveys are valid and hold the “final truth” on the identification.

#### 2.1.1. Validity of Sociometrics

The previously discussed sociometric scales are used in a diversity of scientific research, but have mainly been validated through the use of simulation (Nolker, 2005; Wuyts & Bulte, 2007; Iyengar et al., 2011). The current operational models for the simulation of social network diffusion all have the same basic core in which an actor can be either active (has adopted) or inactive (not adopted). It is somewhat assumed that if actors switch from inactive to active, that they have been influenced. Thus, with a successful spread, actors within a network will start to progress from inactive to active, based on the connections between the actors.

In research, 2 types of simulation models are widely studied and generally used for simulation:

**LT-Models** - LT-models have been proposed for the first time by Granovetter (1978) and Schelling (1971) and form the basis for many of the models that have followed (Watts et al., 2002; Young, 2002; Valente 1995 and Macy 1991). LT is an abbreviation that stands for “Linear Threshold” and in this case means the moment/value at which person X changes from inactive to active, and is based on the weighted number of neighbours that have to be active in order for person X to become active. The threshold differs from person to person and basically represents how easy or hard it is to influence that specific person.

**IC-Models** - The second type of model based on the diffusion within networks are the IC-models. IC means Independent Cascade, in which a cascade is represented by the steps of information flowing from persons recommending something to their friends, who in turn recommend something to their friends and so on (Kempe et al. 2003). In this case, a message starts at a person (or group of persons) who have the possibility “P” to activate its neighbours. An important premise to this model is that whether or a person adopts based on a connection to someone else, there are no second chances for that specific connection to create an adoption in a second attempt.
As can be concluded, these simulation models are very basic and use a lot of assumptions. Connections within social networks can be very complex and are very much simplified by these models. For example, in the LT-model diversities in the strength of the sender is not modelled, only the threshold of the receiver, which would mean influencers don’t exist. Another important example is that the values regarding the behaviour of actors are randomly distributed, whereas in real life the structural position implies certain behaviour (Granovetter, 1973; Coleman, 1988; Burt, 1992). A final example would be on the one-dimensional and non-historical states of users (adopted or not), whereas adoption typically is a process (Rogers, 1962) with accumulated information and/or experiences through multiple channels with multiple states before adopting.

2.1.2. Validity Marketing Surveys

To validate both opinion leaders and market mavens, instead of using simulations, self-report surveys from marketing theory can be used. Criticism in terms of validity in the use of self-report surveys is mainly based upon 2 arguments. Cook and Campbell (1979) have proven that subjects (a) tend to report what they believe the researcher expects to see, or (b) report what reflects positively on their own abilities, knowledge, beliefs, or opinions. Prove of the latter argument stems from Macdiarmid and Blundell (1998), who performed research based on data from several authors to find that women consistently under-report their dietary intake and weight. The most likely reason is that women are very concerned about their body weight, food and eating and like to report values that are socially acceptable.

To get a better perspective in the validity of the surveys, it is best to focus on the construct validity. Construct validity can be broken up into discriminant and convergent validity. Tests designed to measure the same construct should correlate highly amongst themselves to obtain convergent validity. In the case of opinion leadership, cross-validation of network theory with marketing theory would result in testing the construct validity of the construct of opinion leaders. Regarding discriminant validity, the construct being measured by a test should not correlate highly with different constructs. Discriminant validity of opinion leaders and market mavens has been tested by Feick and Price (1987), who found that opinion leaders, early purchasers/adopters and market mavens are distinct concepts through empirical research.

Finally, Corey (1971) has done a very interesting research about the validity of self-report on opinion leaders. In her research, Corey examines the characteristics of people that self-report through the use of a survey that they were opinion leaders. She finds that “the consumers it (the self-report survey) classifies as opinion leaders have, as a group, the same characteristics as those attributed to opinion leaders by the theoretical and empirical literature”. An important conclusion that states that self-report surveys are valid for the identification of opinion leaders.
2.2. Learnings outcomes
Influencers are capable of disproportionately influencing others and can enhance diffusion of information/adoption in a network. Because of these characteristics, they are in the interest of companies to enhance the spread and the strength of their messages. Two types of influencers have found to be of high importance regarding information diffusion: opinion leaders and market mavens. To be able to address them, research has been performed in both marketing and network theory to identify influencers.

Opinion leaders can reliably be identified by a survey from Flynn et al. (1996) and market mavens can be identified with the help of a survey from Feick and Price (1987). However, the research field on Network Theory has created “sociometrics” that define the structural position of actors in a network and are claimed to be useful in identifying specific types of persons. This method is more cost-effective than surveying for the purpose of identifying influencers and doesn’t have to deal with non-response.

Methods to validate these sociometrics are all based on simulation and have to deal with simplifications of complex systems. An important argument is that people are thought to go through a process before adoption with multiple different stages and states whereas the simulations only have 2. Moreover, one of the simulations uses random distribution to assign characteristics to actors, which is not in line with Network Theory, where a certain position implies certain characteristics. Finally, the other simulation doesn’t support the differences in people sending messages, only in receiving.

To get from face-validity of sociometrics to construct validity, additional research is necessary. This study proposes and proves that cross-validation with surveys developed by marketing are an option, since they produce reliable results and data on single persons can be combined from both identifying methods. This type of cross-validation has only partially been done in 2 studies, of which one only tested the basic metrics of in-degree and out-degree. The other study also included centrality measures, but used it in the context of children and toys. This leaves a gap for validation in an online context and also for an extension of the tested sociometrics. Secondly, it can add to the accumulated validity of the previous two research studies. By closing the gap, future research can focus on optimizing the ability to find opinion leaders/market mavens based on the structural position of a person within a network.

3. Method
To control for construct validity of sociometrics for the identification of opinion leaders and market mavens 2 data collection methods were used. The first method was used to collect social network data to calculate the sociometrics. The second method was surveying a random sample of the social network to collect information about opinion leadership and market mavenism. Afterwards statistical analysis was performed on the combination of datasets to perform validation between influencer constructs from network theory and marketing research.
3.1. Network data collection

As discussed, opinion leaders and market mavens are always associated to a particular topic. Therefore, one topic was selected based on a brainstorm of subjects that were relevant at the time of research. The Dutch House of Commons elections was chosen as a subject, because it would gather a sufficient amount of data and has a strong relationship to the opinion leadership construct, which was the area of interest during the first discovery of opinion leaders.

Network data was collected from Twitter based on all tweets that contained keywords related to the Dutch House of Commons elections (Appendix XV - Twitter Keywords). By setting up a custom coded infrastructure that connects to the Twitter API, collects, queues and processes all available tweets based on the specified keywords, a total dataset with more than 2 million tweets with a timeframe of 2 months up until the elections of the Dutch House of Representatives (3 July to 14 September 2012) was gathered from Twitter. Pre-processing of the collected data was needed to perform the correct statistical analysis on the data. Based on retweets, mentions and replies, separate databases were created to contain the connections between actors within the network, to be used for modelling the network and calculating sociometrics.

All data gathered from Twitter was filtered using language detection. Based on the LIGA algorithm (Tromp & Pechenizky, 2011) only Dutch tweets were preserved for the data set; all other tweets were removed. Additionally, manual scans of randomly chosen lists with 100 tweets were performed to detect large data pollutions from specific keywords or users. The final dataset contains 808,395 tweets.

Finally, an extract of the previously mentioned database with connections between actors in the network was used as an input for the program SNAP 1.10 (Stanford network analysis project) from the university of Stanford to calculate sociometrics based on the communications on Twitter that are resembled as relationships for the model of the social network on the elections of the Dutch House of Commons. The following sociometrics are calculated with the tool SNAP 1.10 from the Stanford University:

- PageRank (Page et al., 1999)
- Hubs (Kleinberg, 1999)
- Authorities (Kleinberg, 1999)
- Degree centrality
- Eigenvector centrality
- Betweenness centrality
- Closeness centrality
- Clustering Coefficient (Watts and Strogatz, 1998)

Additionally, general descriptive metrics on the intensity and type of communications (such as number of retweets) were collected and tie strength was computed manually based on the number of weeks there was interaction with the condition of existence of reciprocity in their communication.
### 3.2. Questionnaire data collection

The second dataset contains more than 500 responses based on a questionnaire including answers to questions about opinion leadership, market mavenism and demographics. In more detail, it includes both the questionnaires from Flynn et al. (1996) and Feick & Price (1987), as well as several control variables (Table 1: Elements of the Final Questionnaire).

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>How old are you?</td>
</tr>
<tr>
<td>Gender</td>
<td>What is your gender?</td>
</tr>
<tr>
<td>Income</td>
<td>What is your gross annual income?</td>
</tr>
<tr>
<td>Education</td>
<td>What is your level of education?</td>
</tr>
<tr>
<td>Occupation</td>
<td>What is your level of occupation?</td>
</tr>
<tr>
<td>Source credibility</td>
<td>Are you an active member of a political party?</td>
</tr>
<tr>
<td>Vote</td>
<td>On which political party did you vote during the past Dutch elections?</td>
</tr>
<tr>
<td>Strategical voting</td>
<td>Have you voted strategically in the past Dutch elections?</td>
</tr>
<tr>
<td>If yes:</td>
<td>Which party would have gotten your vote if you hadn’t vote strategically?</td>
</tr>
<tr>
<td>Opinion leadership</td>
<td>Questions from questionnaire by Flynn et al (1996)</td>
</tr>
<tr>
<td>Market mavenism</td>
<td>Questions from questionnaire by Feick &amp; Price (1987)</td>
</tr>
</tbody>
</table>

Table 1: Elements of the Final Questionnaire

The opinion leadership construct from Flynn et al. (1996) contains responses on a likert-scale ranging from 1 to 7, including 1=“Geheel mee oneens”, 2=“mee oneens”, 3=“Een beetje mee eens”, 4=“oneens/eens”, 5=“een beetje mee eens”, 6=“mee eens” and 7=“Geheel mee eens”. The 6-item survey from Feick and Price (1987) that was used to determine Market Mavenism contains responses on a likert-scale ranging from 1 to 7, including 1=“Geheel mee oneens”, 2=“mee oneens”, 3=“Een beetje mee eens”, 4=“oneens/eens”, 5=“een beetje mee eens”, 6=“mee eens” and 7=“Geheel mee eens”.

Because a political party is somewhat different from a product (e.g. one does not buy a political party; one votes on a political party), the questions from Flynn et al. (1996) and Feick & Price (1987) were interpreted and translated to Dutch and to a political context.

All data from the opinion leadership survey and the market mavenism survey were checked on the following conditions:

1. Contains survey results for at least opinion leadership and has one response per “ip-adres”/”user agent” combination for each survey token (unique anonymous ID to connect survey data to the Twitter data)
2. Completion time (at least more than 1 minute)
3. Acquiescent Bias (more than 2 points difference on average between positive and negative items)

Using a waterfall method, from the original set of 581 entries, 291 surveys remained after applying the first condition, 290 after the second and 242 after applying the last condition. From this set of 242 surveys with data for the opinion leadership construct, 204 contained survey data for the market mavenism construct since it was presented one page later in the survey.
4. Results
This section discusses the results of the statistical analysis performed to validate construct validity of sociometrics to identify both opinion leaders and market mavens. Both constructs are analysed separately, but using the same process.

4.1. Opinion Leadership
Sociometrics to identify opinion leadership are validated using results from the questionnaire by Flynn et al. (1996). First principal components analysis is used to reduce the number of variables from the survey. Afterwards the resulting variable is correlated to sociometrics based on a priori defined hypothesis. Additionally, more sociometrics and general descriptive metrics on the intensity and type of communications are tested.

4.1.1. Principal Components Analysis
To reduce the number of variables and to test for unidimensionality and construct validity a principal components analysis was performed on the data (Appendix VII – OL: Principal Component Analysis), with the first factor explaining 61.96% of the variance. Based on the scree plot and eigenvalues, unidimensionality can be confirmed. Secondly, the Kaiser-Meyer-Olkin measure of sampling adequacy was .848 (very good; Hutcheson & Sofroniou, 1999) which is above the recommended value of .6, Bartlett’s test of sphericity was significant ($\chi^2 (241) = 710,953, p < .001$) proving this is not an identity matrix and reliability of the construct is good ($\alpha = 0.876$; Appendix VIII – Opinion leadership: Reliability). No problems of multicollinearity were found and KMO statistics for individual variables are all sufficient (>0.5). The resulting unidimensional factor shows a non-normal distribution within the sample, which is confirmed by a K-S test (Appendix IX – Opinion Leadership: Normality). The distribution is skewed towards being an opinion leader.

4.1.2. Sociometrics and opinion leadership
To determine whether sociometrics are capable of identifying opinion leaders, the opinion leadership construct based on the 6-item survey from Flynn et al. (1996) is checked for covariance with the above discussed sociometrics. Both scatter plots and bivariate correlations are computed with the help of SPSS to determine a possible correlation between a sociometric and the opinion leadership survey. Because of the violation of normality (see above) and the relatively large sample size, covariate analysis is performed using spearman’s correlation coefficient.

4.1.3. Validation of sociometrics
Based on research from both network theory and marketing theory, four hypothesis were created in which sociometrics are related to opinion leadership. The following hypothesis are tested based on the dataset:

- **HO1**: Opinion leaders are positively associated with having strong ties in the social network
- **HO2**: Opinion leaders are positively associated with betweenness centrality
- **HO3**: Opinion leaders are positively associated with closeness centrality
- **HO4**: Opinion leaders are positively associated with having structural holes in the social network

Results of covariance analysis on three of four hypothesis are shown below. Betweenness centrality ($r_s = 0.247$, $p<0.01$) and closeness centrality ($r_s = 0.191$, $p<0.01$) show weak to low positive
correlations (Mason, Lind and Marchal, 1983), whereas having strong ties shows no significant relationship with opinion leadership ($r_s = 0.049$).

<table>
<thead>
<tr>
<th></th>
<th>OL factor</th>
<th>Correlation Coefficient</th>
<th>Sig. (2-tailed)</th>
<th>$R^2$ (% of ranking variance)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spearman’s rho</strong></td>
<td>OL factor</td>
<td>1.000</td>
<td></td>
<td>100.00%</td>
<td>241</td>
</tr>
<tr>
<td>Betweenness Centrality</td>
<td></td>
<td>.247**</td>
<td>.000</td>
<td>6.10%</td>
<td>241</td>
</tr>
<tr>
<td>Closeness Centrality</td>
<td></td>
<td>.191**</td>
<td>.003</td>
<td>3.65%</td>
<td>241</td>
</tr>
<tr>
<td>Having strong ties</td>
<td></td>
<td>.049</td>
<td>.357</td>
<td>0.24%</td>
<td>241</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 2: Results of hypothesis on covariance between sociometrics and Opinion Leadership survey

Scatter plots from betweenness and closeness centrality reveal no possibilities of other type of relationships (Appendix XI – Opinion Leadership: Scatterplots closeness and betweenness). Hypothesis H02 and H03 are therefore confirmed, but with weak to low correlations (Mason, Lind and Marchal, 1983). Hypothesis H01 is rejected; no significant relationship between opinion leadership and having strong ties could be found. Unfortunately, H04 couldn’t be calculated within the timeframe at the current moment because of the size of the dataset.
4.1.4. Explorative research

Additional to analysis of hypothesis, explorative research has been performed on correlations with other sociometrics and with general descriptive metrics on the intensity and type of communications available based on the dataset. Results are shown below.

4.1.4.1. Sociometrics – no hypothesis

Results of covariance analysis on the remaining sociometrics and opinion leadership are shown in Table 3, which shows there are significant positive relationships between the majority of sociometrics and the opinion leadership construct based on the survey. However, the correlation coefficient shows a weak to low correlation for the majority of the variables (Mason, Lind and Marchal, 1983). Most notable sociometrics that correlate with the opinion leadership construct are in-degree ($r_s = 0.304$, $p<0.01$), degree centrality ($r_s = 0.301$, $p<0.01$) and PageRank ($r_s = 0.288$, $p<0.01$). The results are not surprising, since there are significant strong correlations between the same majority of sociometrics (Appendix X – Sociometrics correlations; Mason, Lind and Marchal, 1983).

<table>
<thead>
<tr>
<th>Correlations</th>
<th>OL factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's rho</td>
<td></td>
</tr>
<tr>
<td>PageRank</td>
<td>.288**</td>
</tr>
<tr>
<td>Degree Centrality</td>
<td>.301**</td>
</tr>
<tr>
<td>Eigenvector Centrality</td>
<td>.259**</td>
</tr>
<tr>
<td>ClusteringCoefficient</td>
<td>.142</td>
</tr>
<tr>
<td>Hubscore</td>
<td>.260**</td>
</tr>
<tr>
<td>Authority</td>
<td>.260**</td>
</tr>
<tr>
<td>Indegree</td>
<td>.304**</td>
</tr>
<tr>
<td>Number of strong ties</td>
<td>.076</td>
</tr>
<tr>
<td>Outdegree</td>
<td>.269**</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 3: Results of explorative research on correlations between opinion leadership and sociometrics
Scatter plots with significant correlated sociometrics all show the same regarding the relation to the opinion leadership construct (Figure 4: Scatter plot of PageRank and Opinion Leadership). Most important conclusion based upon visualisation, is that at certain levels of most sociometrics, low scores on opinion leadership are relatively rare (green box), leaving people that are somewhat higher ranked as an opinion leader possible to target. Depending on the required opinion leadership strength, targeting rates increase from 14.9% to 25.7% (OL >= 1) or from 7.05% to 15.15% (OL>=1.2).

4.1.4.2. **General descriptive metrics**

Additional to the sociometrics, general metrics were available for analysis on the intensity and type of communications (Table 4: Covariance between traditional Twitter statistics and opinion leadership). The correlation coefficient shows a weak to low correlation for the majority of the variables (Mason, Lind and Marchal, 1983). Most notable metrics that correlate with the opinion leadership construct are retweeted ($r_s = 0.307, p<0.01$), and mentioned ($r_s = 0.306, p<0.01$).

<table>
<thead>
<tr>
<th>Correlations</th>
<th>OL factor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spearman’s rho</strong></td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td>OL factor</td>
<td>1.000</td>
</tr>
<tr>
<td>Followers</td>
<td>0.184</td>
</tr>
<tr>
<td>Friends</td>
<td>0.209</td>
</tr>
<tr>
<td>Statuses</td>
<td>0.100</td>
</tr>
<tr>
<td>Mention</td>
<td>0.262</td>
</tr>
<tr>
<td>Mentioned</td>
<td>0.306</td>
</tr>
<tr>
<td>Retweet</td>
<td>0.285</td>
</tr>
<tr>
<td>Retweeted</td>
<td>0.307</td>
</tr>
<tr>
<td>Reply</td>
<td>0.180</td>
</tr>
<tr>
<td>Replied</td>
<td>0.239</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4: Covariance between traditional Twitter statistics and opinion leadership
Since the variables above are absolute numbers on activity towards and from actors, it’s interesting to explore relative numbers. Table 5 shows that Retweeted per status (status meaning the number of sent tweets) and Degree per status are positively associated with opinion leadership (resp. \( r_s = 0.326, p<0.01 \) and \( r_s = 0.327, p<0.01 \)).

<table>
<thead>
<tr>
<th>Correlation</th>
<th>OL factor</th>
<th>Correlation Coefficient</th>
<th>Sig. (2-tailed)</th>
<th>R2 (% of variance)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s rho</td>
<td>OL factor</td>
<td>1.000</td>
<td>.</td>
<td></td>
<td>241</td>
</tr>
<tr>
<td>Retweeted per status</td>
<td></td>
<td>.326</td>
<td>.000</td>
<td>10.63%</td>
<td>239</td>
</tr>
<tr>
<td>Degree per status</td>
<td></td>
<td>.327</td>
<td>.000</td>
<td>10.69%</td>
<td>239</td>
</tr>
<tr>
<td>Retweeted per followers</td>
<td></td>
<td>.291</td>
<td>.000</td>
<td>8.47%</td>
<td>239</td>
</tr>
<tr>
<td>Mentioned per status</td>
<td></td>
<td>.275</td>
<td>.000</td>
<td>7.56%</td>
<td>239</td>
</tr>
<tr>
<td>Pagerank per status</td>
<td></td>
<td>.247</td>
<td>.000</td>
<td>6.10%</td>
<td>239</td>
</tr>
<tr>
<td>Retweeted per pagerank</td>
<td></td>
<td>.292</td>
<td>.000</td>
<td>8.53%</td>
<td>241</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5: Covariance between relative Twitter statistics and opinion leadership

### 4.1.4.3. Multiple Regression

Multiple regression was performed on the data to explore opportunities of predicting opinion leadership based on multiple variables. Collected control variables (age, income and level of education) were included as well as all sociometrics and general descriptive Twitter metrics. Because of high multicollinearity between variables (see above), factor analysis was used to reduce the number of variables. 2 extra factors were extracted from the data (Appendix XVI – Dimension Reduction):

- Twitter usage (reply, retweet and mention)
- Twitter popularity (replied, retweeted and mentioned)

A third factor involving centrality measures was found using factor analysis, but internal consistency was not sufficient (\( \alpha = 0.052 \)). Using stepwise addition and removal in combination with normalization of collinear variables the following model was derived (Table 6: Multiple Regression on Opinion Leadership):

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>-.200</td>
<td>.116</td>
<td>-1.727</td>
<td>.086</td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td>-.921</td>
<td>.377</td>
<td>-.156</td>
<td>-2.445</td>
</tr>
<tr>
<td>Active member in politics</td>
<td>.770</td>
<td>.128</td>
<td>.388</td>
<td>5.993</td>
</tr>
<tr>
<td>Retweeted per PageRank</td>
<td>3.796E-007</td>
<td>.000</td>
<td>.135</td>
<td>2.096</td>
</tr>
<tr>
<td>Friends per Degree</td>
<td>-.001</td>
<td>.001</td>
<td>-.149</td>
<td>-2.288</td>
</tr>
</tbody>
</table>

Table 6: Multiple Regression on Opinion Leadership
Table 7: Model summary of multiple regression on Opinion Leadership

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.496</td>
<td>.246</td>
<td>.230</td>
<td>.86232040</td>
</tr>
</tbody>
</table>

Active member in politics seems to behave as a mediating or confounding variable in relation to clustering coefficient and opinion leadership, since removing the variable renders clustering coefficient insignificant, while there is no indication of multicollinearity. Based on meaning of the variables mediating is most logic, for not all strong clusters within the network are necessarily associated with interest in politics.

Interestingly, clustering coefficient and “friends per degree” are both negatively associated to opinion leadership, meaning opinion leaders don’t reside in very strong clusters. This is somewhat in line with the hypothesis of structural holes.

For social sciences, the model has a relative good fit ($r^2 = .246$). However, this is mainly due to the variable “active member in politics”. Removing the mediating variable causes a drop in model fit of Δ$-0.131$ ($r^2=0.115$).

4.1.5. Conclusions on Opinion Leadership

Validation of the use of current sociometrics for the identification of opinion leaders was performed by covariance analysis of sociometrics with survey results on the opinion leadership construct by Flynn et al. (1996). Results confirm hypothesis HO2 which predicts a significant positive association between market mavenism and betweenness centrality ($r_s = 0.247, p<0.01$). Secondly, results confirm hypothesis HO3 which predicts a significant positive association between market mavenism and closeness centrality ($r_s = 0.191, p<0.01$). Interestingly, general descriptive metrics on the intensity and type of communications (such as the number of retweets) perform equal or better in covariance to opinion leadership. Metrics such as Retweeted per status (number of retweets relative to the number of tweets) or Degree per status (number of connections relative to the number of tweets) have a significant higher correlation to opinion leadership. Regarding identification and targeting of opinion leadership, there is an opportunity to target people scoring higher than average on metrics significantly positive related to opinion leadership. Depending on the required opinion leadership strength, chances can almost double. In regard to the multiple regression, it is interesting to see that being an active member in politics acts as a mediator towards the negatively associated clustering coefficient. In line with the theory, opinion leaders choose sides for a particular political party and don’t reside in very strong clusters, but are more likely to reside within a network of structural holes. Additionally, friends per degree is significantly positive associated with opinion leadership, which corresponds with theory about having strong connections.
4.2. Market Mavenism

Sociometrics to identify opinion leadership are validated using results from the questionnaire by Feick and Price (1987). First principal components analysis is used to reduce the number of variables from the survey. Afterwards the resulting variable is correlated to sociometrics based on a priori defined hypothesis. Additionally, more sociometrics and general descriptive metrics on the intensity and type of communications are tested.

4.2.1. Principal Components Analysis

To reduce the number of variables and to test for unidimensionality and construct validity a principal components analysis was performed on the data (Appendix XII – Market Mavenism: PCA), with the first factor explaining 65.1% of the variance. Based on the scree plot and eigenvalues, unidimensionality can be confirmed. Secondly, the Kaiser-Meyer-Olkin measure of sampling adequacy was .874 (very good; Hutcheson & Sofroniou, 1999) which is above the recommended value of .6. Bartlett’s test of sphericity was significant ($\chi^2 (204) = 656.701, p < .001$) proving this is not an identity matrix and reliability of the construct is good ($\alpha = 0.890$; Appendix VIII – Opinion leadership: Reliability). No problems of multicollinearity were found and KMO statistics for individual variables are all sufficient (> .5). The resulting unidimensional factor shows a non-normal distribution within the sample, which is confirmed by a K-S test (Appendix XIV - Market mavenism Normality). The distribution is skewed towards being a market maven.

4.2.2. Sociometrics and Market Mavenism

To determine whether sociometrics are capable of identifying Market Mavens, the Market Mavenism construct based on the 6-item survey from (Feick and Price, 1987) is checked for covariance with the above discussed sociometrics. Both scatter plots and bivariate correlations are computed with the help of SPSS to determine a possible correlation between a sociometric and the market maven survey. Because of the violation of normality (see above) and the relatively large sample size, covariate analysis is performed using spearman’s correlation coefficient.

4.2.3. Validation of sociometrics

Based on research from both network theory and marketing theory, four hypothesis were created in which sociometrics are related to market mavenism. The following hypothesis are tested based on the dataset:

- **HM1**: Market Mavens are positively associated with authority (HITS algorithm)
- **HM2**: Market Mavens are positively associated with PageRank
- **HM3**: Market Mavens are positively associated with in-degree

Results of covariance analysis are shown in Table 8 and show that there are significant positive relationships between all sociometrics and the market mavenism construct based on the survey. However, the correlation coefficient shows a weak to low correlation for the all of the variables (Mason, Lind and Marchal, 1983). Most notable sociometric that correlates with the market mavenism construct are degree centrality ($r = 0.224, p < 0.01$) and in-degree ($r = 0.222, p < 0.01$).
### Table 8: Covariance between sociometrics and Opinion Leadership survey

<table>
<thead>
<tr>
<th>Sociometric</th>
<th>Correlation Coefficient</th>
<th>Sig. (2-tailed)</th>
<th>R² (% of ranking variance)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's rho</td>
<td>1.000</td>
<td>.000</td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>Degree Centrality</td>
<td>.224**</td>
<td>.001</td>
<td>5.02%</td>
<td></td>
</tr>
<tr>
<td>Pagerank</td>
<td>.188**</td>
<td>.007</td>
<td>3.53%</td>
<td></td>
</tr>
<tr>
<td>In-degree</td>
<td>.222**</td>
<td>.001</td>
<td>4.93%</td>
<td></td>
</tr>
<tr>
<td>Authority</td>
<td>.193**</td>
<td>.006</td>
<td>3.72%</td>
<td></td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Scatter plots with the sociometrics all show the same regarding the relation to the market mavenism construct (Figure 5: Scatter plot of PageRank and Market Mavenism). Another conclusion based upon visualisation and analysis, is that at certain levels of all sociometrics (except clustering coefficient), low scores on market mavenism are relatively rare (green box), leaving people that are somewhat higher ranked as a market maven possible to target. Depending on the required market mavenism strength, targeting rates increase from 6.8% to 10.3% (MM >= 1) or from 5.4% to 7.6% (MM>=1.2)

![Figure 5: Scatter plot of PageRank and Market Mavenism](image-url)

---

**Figure 5: Scatter plot of PageRank and Market Mavenism**
4.2.4. Explorative research
Additional to analysis of hypothesis, explorative research has been performed on correlations with other sociometrics and with general descriptive metrics on the intensity and type of communications available based on the dataset. Results are shown below.

4.2.4.1. Sociometrics – no hypothesis
Results of covariance analysis on the remaining sociometrics and market mavenism are shown in Table 9, which shows there are significant positive relationships visible between sociometrics and the market mavenism construct based on the survey. However, the correlation coefficient that are significant show a weak to low correlation for the majority of the variables (Mason, Lind and Marchal, 1983). Most notable sociometrics that correlate with the market mavenism construct are out-degree ($r_s = 0.215$, $p<0.01$), eigenvector centrality ($r_s = 0.193$, $p<0.01$) and Hubscore ($r_s = 0.193$, $p<0.01$).

<table>
<thead>
<tr>
<th>Correlations</th>
<th>MM factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>Sig. (2-tailed)</td>
</tr>
<tr>
<td>Spearman’s rho</td>
<td>MM factor</td>
</tr>
<tr>
<td>Eigenvector</td>
<td>,193**</td>
</tr>
<tr>
<td>ClusteringCoefficient</td>
<td>,112</td>
</tr>
<tr>
<td>Hubscore</td>
<td>,193**</td>
</tr>
<tr>
<td>Betweenness</td>
<td>,112</td>
</tr>
<tr>
<td>Closeness</td>
<td>,166*</td>
</tr>
<tr>
<td>Out-degree</td>
<td>,215**</td>
</tr>
</tbody>
</table>

**: Correlation is significant at the 0.01 level (2-tailed).
*: Correlation is significant at the 0.05 level (2-tailed).

Table 9: Results of explorative research on correlations between market mavenism and sociometrics
4.2.4.2. **General descriptive metrics**

Additional to the sociometrics, some general metrics were available for analysis on the intensity and type of communications (Table 10: Covariance between traditional Twitter statistics and market mavenism). The correlation coefficient shows a weak to low correlation for the majority of the variables (Mason, Lind and Marchal, 1983). Most notable metrics that correlate with the market mavenism construct are retweeted ($r_s = 0.239$, $p<0.01$), and retweet ($r_s = 0.234$, $p<0.01$).

<table>
<thead>
<tr>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spearman's rho</strong></td>
</tr>
<tr>
<td>MM factor</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>MM factor</td>
</tr>
<tr>
<td>Followers</td>
</tr>
<tr>
<td>Friends</td>
</tr>
<tr>
<td>Statuses</td>
</tr>
<tr>
<td>Mention</td>
</tr>
<tr>
<td>Mentioned</td>
</tr>
<tr>
<td>Retweet</td>
</tr>
<tr>
<td>Retweeted</td>
</tr>
<tr>
<td>Reply</td>
</tr>
<tr>
<td>Replied</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 10: Covariance between traditional Twitter statistics and market mavenism

Since the variables above are absolute numbers on activity towards and from actors, it’s interesting to explore relative numbers. Table 11 shows that Degree per status (status meaning the number of sent tweets) and Retweeted per status are positively associated with market mavenism (resp. $r_s = 0.307$, $p<0.01$ and $r_s = 0.281$, $p<0.01$).

<table>
<thead>
<tr>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spearman's rho</strong></td>
</tr>
<tr>
<td>MM factor</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>MM factor</td>
</tr>
<tr>
<td>Retweeted per status</td>
</tr>
<tr>
<td>Degree per status</td>
</tr>
<tr>
<td>Retweeted per followers</td>
</tr>
<tr>
<td>Mentioned per status</td>
</tr>
<tr>
<td>Pagerank per status</td>
</tr>
<tr>
<td>Retweeted per pagerank</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

Table 11: Covariance between relative Twitter statistics and market mavenism
4.2.4.3. **Multiple Regression**

Multiple regression was performed on the data to explore opportunities of predicting market mavenism based on multiple variables. Collected control variables (age, income and level of education) were included as well as all sociometrics, general descriptive Twitter metrics and factors from factorial mentioned in opinion leadership research (4.1.4.3 Multiple Regression; Appendix XVI – Dimension Reduction).

Using stepwise addition and removal in combination with normalization of collinear variables the following model was derived (Table 12: Multiple Regression on Market Mavenism):

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td></td>
<td>-422</td>
<td>097</td>
<td></td>
<td>-4.326</td>
</tr>
<tr>
<td>retweetedperpagerank</td>
<td>6,033E-007</td>
<td>000</td>
<td>220</td>
<td>3,359</td>
<td>0.001</td>
</tr>
<tr>
<td>ActiefLidpp01</td>
<td>0678</td>
<td>127</td>
<td>348</td>
<td>5,330</td>
<td>0.000</td>
</tr>
<tr>
<td>followers_per_outdegree</td>
<td>000</td>
<td>000</td>
<td>-197</td>
<td>-3,025</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Table 12: Multiple Regression on Market Mavenism

Interestingly, “followers per outdegree” is negatively associated to market mavenism, which is somewhat in line with the role of market mavens that are responsive to questions from all people within the market of interest. The description of market mavens implies ties are weak ties, which would explain the low amount of followers in relation to the outdegree.

For social sciences, the model has a relative good fit ($r^2 = .216$). However, this is mainly due to the variable “active member in politics”. Removing the variable causes a drop in model fit of $\Delta 0.120$ ($r^2=0.113$).
4.2.5. Conclusion on Market Mavenism
Validation of the use of current sociometrics for the identification of market mavens was performed by covariance analysis of sociometrics with survey results on the market mavenism construct by Feick and Price (1987). Results confirm all hypothesis regarding market mavenism with weak to low relationships. Interestingly, general descriptive metrics on the intensity and type of communications (such as the number of retweets) perform equal or better in covariance to market mavenism. Metrics such as Retweeted per status (number of retweets relative to the number of tweets) or Degree per status (number of connections relative to the number of tweets) have a significant higher correlation to opinion leadership. Regarding identification and targeting of market mavenism, there is a very small opportunity to target people scoring higher than average on metrics significantly positively related to market mavenism. Finally, multiple regression shows that being an active member of a political party is the most important variable in determining Market Mavenism. This isn’t in line with typical market mavenism descriptions, because market mavens oversee the field of interest more than actually participate and choose sides.
5. Discussion
This section discusses the conclusion of results from the previous section and both the managerial and academic implications. Finally, limitations and recommendations for further research are discussed.

5.1. Conclusion
The study was conducted to validate the use of sociometrics for the identification of opinion leaders and market mavens. Data analysis based on network data from Twitter and an online survey was used to perform the validation.

The literature study showed the use of sociometrics in several scientific publications, in which they were used to identify certain influencers. Although they mention the concept of opinion leadership, the focus of these sociometrics is strong towards spread of information rather than influence. Based on marketing literature, hypothesis on relations between sociometrics and opinion leaders as well as on market mavens were produced to be validated.

Results largely confirm these hypothesis, but with low to weak correlations (Mason, Lind and Marchal, 1983) between the sociometrics and both constructs. Moreover, general descriptive metrics on the intensity and type of communications (such as number of retweets) perform equal or better in covariance to both opinion leadership and market mavenism.

Only two sociometrics have a notable difference in covariance between opinion leadership and market mavenism. PageRank and Betweenness have a stronger relationship to opinion leadership than towards market mavenism. All other sociometrics are somewhat alike in covariance towards both constructs, making PageRank and Betweenness the only variables that are somewhat capable of making a distinction between both constructs.

Another interesting finding is the importance of active membership of a political party in being opinion leader or market maven. Although data only stems from tweets about politics, being involved with the subject is an important variable for both constructs.

In general, one can state that the identification of either opinion leaders or market mavens is not valid through use of discussed sociometrics (alone). However, increasing the likelihood of targeting opinion leaders and/or market mavens is possible, because scatter plots show that a majority of sociometrics exclude a majority of non-opinion leaders and market mavens. Thereby potentially reducing costs of influencer marketing activities.
5.2. Managerial Implications

Based on this research, important managerial implications can be made. Most important implication is regarding the use of sociometrics to identify opinion leaders or market mavens. Current sociometrics do not suffice in identifying both influencers, and as such should not be used to identify either of them for use in influencer marketing.

However, there is an opportunity of cost reduction, since certain sociometrics can increase the likelihood of targeting one of both influencers. Based on the values from these sociometrics, one can exclude a large proportion of the population that is not very likely to contain opinion leaders or market mavens. As such marketing can reduce costs in examining the entire population to search for opinion leaders or market mavens with current tactics.

Finally, the use of general metrics available on Twitter (such as retweets) doesn’t differ much from more complex algorithms (such as PageRank) to identify both influencers. Because these metrics are easily obtained, whereas the more complex algorithms require more resources, use of the generally available metrics is to be advised.

5.3. Academic Implications

Important academic implications can be made based on this study. The main implication is the use of the construct of opinion leadership in network theory. The current use of the construct of opinion leadership in relation to sociometrics that indicate the importance of a node within a network in regard to the spread of information is not correct. Often, there is a reference towards this construct, whereas opinion leadership has an extra dimension that is not sufficiently addressed. The construct of opinion leaders also addresses an above average influence on others, which is not depicted by these models.

In return, marketing should be careful in adopting techniques from network theory to identify influencers for future research. Additional research should be necessary to better identify influencers and the different types of influencers based on data available in online social networks.
5.4. Limitations and recommendations for further research

An evident limitation to the current study includes the use of only one channel in the analysis. The network data is solely based on one online medium and doesn’t contain other online channels or offline connections and its characteristics of communication. Therefore, it might be possible that regarding the subject of politics, the constructed network misses important nodes that can potentially have an significant impact on the network and values of the used sociometrics.

Additionally, the sociometrics are not computed periodically to detect changes, but only once over the entire period of which data was collected (10 weeks). Because of timing issues and the possibility of people changing behaviour, certain actors might be wrongly classified based upon the sociometrics.

Thirdly, the conducted survey contained general items based upon scientific literature that don’t make a distinction between influence on the medium Twitter or influence on people in general. Control for these distinctions should be incorporated in next research, especially if it includes the use of multiple channels/media.

Finally, the context of politics in regard to market mavens is questionable. Whereas research on opinion leadership has its roots in research in the field of politics, market mavens have never been associated with politics. Moreover, the definition of market mavens given by Feick and Price (1987) is focused on consumer markets. Although a small part of the theory on which the existence of market mavens is derived stems from political research. Additional research is needed to verify the role of market mavens in the political area.

Based on this research, one might conclude that research on network theory has to redefine their constructs for important nodes in the spread of information. The focus within network theory is on the spread of information rather than on the actual influence a person has on others, which isn’t in line with marketing theory from which the construct of opinion leadership stems. Future research to identify opinion leaders and/or market mavens based on network information could therefore include dynamic (time based) qualitative data, such as the messages being send. For example, relatively new text analysis methods such as sentiment and emotional analysis could be used to detect opinion changes over time of actors or groups/cliques towards certain subjects, after which potential sources of influence could be determined.
6. References


doi:10.1086/266584


doi:10.1177/1075547008328797


Trepte, S., & Scherer, H. (2010). Opinion leaders – Do they know more than others about their area of interest? *Communications*, 35(2). doi:10.1515/comm.2010.007


7. Appendices

7.1. Appendix I: PageRank Calculations

\[ PR(A) = 0.15 + 0.85 \times \left( 0.15 + 0.85 \left( \frac{PR(A)}{2} + 0.15 + 0.85 \times \frac{PR(A)}{2} \right) \right) \]

\[ PR(A) = 0.15 + 0.85 \times \left( 0.15 + 1.5725 \times \frac{PR(A)}{2} + 0.1275 \right) \]

\[ PR(A) = 0.15 + 0.235875 + 0.668313 \times PR(A) \]

0.385875 = 0.331688 \times PR(A)

\[ PR(A) = 1.163369 \]
### Appendix II: Opinion Leadership Scale Items

Source: Flynn et al., 1996

**Opinion leadership scale items:**

<table>
<thead>
<tr>
<th>Flynn et al. (1996)</th>
<th>Translated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original</strong></td>
<td><strong>Translated</strong></td>
</tr>
<tr>
<td>My opinion on rock [fashion; environmentally friendly products] seems not to count with other people</td>
<td>Mijn mening over politieke zaken lijkt voor andere mensen niet van belang</td>
</tr>
<tr>
<td>When they choose a rock music recording [fashionable clothing; &quot;green&quot; products], other people do not turn to me for advice</td>
<td>Wanneer mensen hun keuze voor een politieke partij maken, vragen ze mij niet om advies</td>
</tr>
<tr>
<td>Other people come to me for advice about choosing cd’s and tapes [fashionable clothing; products that are good for the environment].</td>
<td>Andere mensen komen bij mij voor advies over politieke partijen</td>
</tr>
<tr>
<td>People that I know pick rock music [clothing; &quot;green&quot; products] based on what I have told them</td>
<td>Mensen die ik ken stemmen op een politieke partij gebaseerd op wat ik hen heb verteld</td>
</tr>
<tr>
<td>I often persuade other people to buy the rock music [fashions; &quot;green&quot; products] that I like</td>
<td>Ik overtuig vaak mensen om te stemmen op de partij die mij het meest bevalt</td>
</tr>
<tr>
<td>I often influence people’s opinions about popular rock [clothing; environmentally correct products].</td>
<td>Ik beïnvloed vaak de mening van anderen aangaande politieke partijen</td>
</tr>
</tbody>
</table>
### 7.1. Appendix III: Market Maven Scale Items

**Source:** Feick and Price (1987)

**Market Maven Scale Items:**

<table>
<thead>
<tr>
<th>Original</th>
<th>Translated</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like introducing new brands and products to my friends</td>
<td>Ik houd ervan om een nieuwe kijk op bestuurlijke en maatschappelijke vraagstukken te introduceren bij mijn vrienden</td>
</tr>
<tr>
<td>I like helping people by providing them with information about many kinds of products.</td>
<td>Ik houd ervan om mensen te helpen door ze informatie te verstrekken over diverse soorten bestuurlijke en maatschappelijke vraagstukken</td>
</tr>
<tr>
<td>People ask me for information about products, places to shop, or sales.</td>
<td>Andere mensen vragen mij informatie over bestuurlijke en maatschappelijke vraagstukken</td>
</tr>
<tr>
<td>If someone asked where to get the best buy on several types of products, I could tell him or her where to shop</td>
<td>Als iemand op zoek is naar informatie over bestuurlijke en maatschappelijke vraagstukken, zou ik hem/haar daaraan kunnen helpen</td>
</tr>
<tr>
<td>My friends think of me as a good source of information when it comes to new products or sales</td>
<td>Mijn vrienden zien mij als een betrouwbare bron van informatie waar het bestuurlijke en maatschappelijke discussies betreft</td>
</tr>
<tr>
<td>Think about a person who has information about a variety of products and likes to share this information with others. This person knows about new products, sales, stores, and so on, but does not necessarily feel he or she is an expert on one particular product. How well would you say that this description fits you?</td>
<td>Probeer u een persoon voor de geest te halen die informatie heeft over een grote diversiteit aan maatschappelijke en bestuurlijke onderwerpen en deze informatie graag deelt met anderen. Deze persoon weet alles over de betrokken personen, partijen en belangen, maar voelt zich niet noodzakelijkerwijs een expert op één specifiek onderwerp. In hoeverre voldoet u aan deze omschrijving?</td>
</tr>
</tbody>
</table>
### Appendix V: Literature Quality Overview

#### Literature Table

<table>
<thead>
<tr>
<th>Article</th>
<th>Journal/book title</th>
<th>Journal Rating</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borgatti, S. P. (1997)</td>
<td>None</td>
<td>Website</td>
<td></td>
</tr>
<tr>
<td>Burt, Ronald. (1992)</td>
<td></td>
<td>Book</td>
<td>11724</td>
</tr>
<tr>
<td>Burt, Rs. (1992)</td>
<td></td>
<td>Book</td>
<td>905</td>
</tr>
</tbody>
</table>

---

1. The Abdc 2010 rating; possible ratings: A*, A, B, C. A minimum of B for is used as quality measure.
2. Citations retrieved from Google Scholar (scholar.google.com)
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Journal</th>
<th>Year</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gladwell, M.</td>
<td>Book</td>
<td>5177</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Granovetter, M. S. (1973)</td>
<td>American Journal of Sociology</td>
<td>A*</td>
<td>22265</td>
<td></td>
</tr>
<tr>
<td>Lin, N. (1971)</td>
<td>Book</td>
<td>106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Title</td>
<td>Journal/Book Title</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Ryan, B., &amp; Gross, N. (1943)</td>
<td>Ryan, B., &amp; Gross, N. (1943)</td>
<td>Rural Sociology</td>
<td>1069</td>
<td></td>
</tr>
<tr>
<td>Trolldahl, V. C., &amp; Dam, R. V. (1965)</td>
<td>Trolldahl, V. C., &amp; Dam, R. V. (1965)</td>
<td>Public Opinion Quarterly</td>
<td>A 95</td>
<td></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Title</td>
<td>Year</td>
<td>Journal/Publication</td>
<td>Volume/Page</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>------</td>
<td>---------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Wasserman, S., &amp; Faust, K.</td>
<td>Social Network Analysis: Methods and Applications</td>
<td>1994</td>
<td>Book</td>
<td>14080</td>
</tr>
<tr>
<td>Weimann, Gabriel.</td>
<td>Public Opinion Quarterly</td>
<td>1991</td>
<td>A</td>
<td>133</td>
</tr>
<tr>
<td>Weimann, Gabriel.</td>
<td>The Influentials: People Who Influence People</td>
<td>1994</td>
<td>Book</td>
<td>235</td>
</tr>
<tr>
<td>Word of Mouth Marketing Association.</td>
<td>Influencers: Defined</td>
<td>2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wuyts, S., &amp; Bulte, C. van den.</td>
<td>Social Networks and Marketing (Marketing Science Institute)</td>
<td>2007</td>
<td></td>
<td>118</td>
</tr>
<tr>
<td>Young, P.</td>
<td>The Diffusion of Innovations in Social Networks</td>
<td>2002</td>
<td>Book</td>
<td>122</td>
</tr>
</tbody>
</table>
Twitter Survey

Beste Twitteraar,
Welkom bij een onderzoek van de Technische Universiteit Eindhoven naar de rol van twitters da tijdens de Nederlandse verkiezingen. U bent uitgenodigd om deel te nemen aan dit onderzoek dat ongeveer 3 minuten duurt om in te vullen.

U gegevens zullen met zorg worden behandeld en anoniem worden geanalyseerd. De Wet bescherming persoonsgegevens zal hierbij in acht worden genomen.

Heeft u vragen en wilt u contact met ons opnemen? Mail dan naar de onderzoeker. Ik dank u hartelijk voor uw bijdrage aan dit onderzoek.

Leeg onvoltooide enquête
Volgende
Stoppen, verwijder alle ingevulde antwoorden
If: Ja

De Nederlandse verkiezingen (1/3)

Wat is uw leeftijd?

Wat is uw geslacht?

Vrouwelijk  ☐  Mannelijk  ☐  Geen antwoord

Wat is uw opleidingsniveau (hoogst behaalde diploma)?

Maak uw keuze...

Wat is uw beroepsniveau (hoogst behaalde diploma)?

Maak uw keuze...

Wat is uw jaarlijks brutoirloon?

Maak uw keuze...

Bent u actief lid van een politieke partij?

Ja  ☐  Nee  ☐  Geen antwoord

Op welke partij heeft u gestemd?

Maak uw keuze...

Heeft u strategisch gekozen tijdens deze verkiezingen?

Ja  ☐  Nee  ☐  Geen antwoord

Een stem kan uitgebracht worden op basis van strategische overwegingen. Bijvoorbeeld om invloed uit te kunnen oefenen op de coalitieonderhandelingen. Ondanks de idealistische voorkeur ('voorkeur vanuit het hart') voor een andere partij.

Op welke partij zou u gestemd hebben indien u geen strategische overweging zou hebben gemaakt?

Maak uw keuze...
<table>
<thead>
<tr>
<th>Stelling</th>
<th>Geheel mee oneens</th>
<th>Mee oneens</th>
<th>Een beetje mee oneens</th>
<th>Oneens/eens</th>
<th>Een beetje mee eens</th>
<th>Mee eens</th>
<th>Geheel mee eens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mijn mening over politieke zaken lijkt voor andere mensen niet van belang</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wanneer mensen hun keuze voor een politieke partij maken, vragen ze mij niet om advies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andere mensen komen bij mij voor advies over politieke partijen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mensen die ik ken stemmen op een politieke partij gebaseerd op wat ik hen heb verteld</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ik overtuig vaak mensen om te stemmen op de partij die mij het meest bevalt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ik beïnvloed vaak de mening van anderen aangaande politieke partijen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Geheel mee eens</td>
<td>Mee eens</td>
<td>Een beetje mee eens</td>
<td>oneens/eens</td>
<td>Een beetje oneens</td>
<td>Mee oneens</td>
<td>Geheel oneens</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------</td>
<td>----------</td>
<td>---------------------</td>
<td>-------------</td>
<td>-------------------</td>
<td>-----------</td>
<td>--------------</td>
</tr>
<tr>
<td>Ik houd ervan om een</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>nieuwe lijst op bestuurlijke en maatschappelijke vraagstukken te introduceren bij mijn vrienden</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ik houd ervan om mensen te helpen door ze informatie te verstrekken over diverse soorten bestuurlijke en maatschappelijke vraagstukken</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Andere mensen vragen mij informatie over bestuurlijke en maatschappelijke vraagstukken</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Als iemand op zoek is naar informatie over bestuurlijke en maatschappelijke vraagstukken, zou ik hem/haar daarop kunnen helpen</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Mijn vrienden zien mij als een betrouwbare bron van informatie waar het bestuurlijke en maatschappelijke discussies betreft</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Probeer u een persoon voor de geest te halen die informatie heeft over een grote diversiteit aan maatschappelijke en bestuurlijke onderwerpen en deze informatie graag deelt met anderen. Deze persoon weet alles over de betrokken personen, partijen en belangen, maar voelt zich niet noodzakelijkerwijs een expert op één specifieke onderwerp. In hoeverre voldoet u aan deze omschrijving?</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Wat is uw opleidings/beroeps-niveau</td>
<td>Wat is uw jaarlijks Bruto-inkomen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagere School (Basisonderwijs)</td>
<td>Geen inkomen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laag (lbo, mavo, vmbo, mbo-1, avo-onderbouw [eerste drie jaren van havo en vwo])</td>
<td>€10.000 - €20.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midden (havo, vwo, mbo-2-4)</td>
<td>€20.000 - €30.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hoog (hbo/wo)</td>
<td>€30.000 - €40.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>€40.000 - €50.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>€50.000 en meer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Op welke partij heeft u gestemd?</th>
<th>Op welke partij zou u</th>
</tr>
</thead>
<tbody>
<tr>
<td>VVD</td>
<td>VVD</td>
</tr>
<tr>
<td>PVDA</td>
<td>PVDA</td>
</tr>
<tr>
<td>PVV</td>
<td>PVV</td>
</tr>
<tr>
<td>SP</td>
<td>SP</td>
</tr>
<tr>
<td>CDA</td>
<td>CDA</td>
</tr>
<tr>
<td>D66</td>
<td>D66</td>
</tr>
<tr>
<td>ChristenUnie</td>
<td>ChristenUnie</td>
</tr>
<tr>
<td>GroenLinks</td>
<td>GroenLinks</td>
</tr>
<tr>
<td>SGP</td>
<td>SGP</td>
</tr>
<tr>
<td>Partij voor de Dieren (PVDD)</td>
<td>Partij voor de Dieren (PVDD)</td>
</tr>
<tr>
<td>50PLUS</td>
<td>50PLUS</td>
</tr>
<tr>
<td>Piratenpartij</td>
<td>Piratenpartij</td>
</tr>
<tr>
<td>Partij voor de Mens</td>
<td>Partij voor de Mens</td>
</tr>
<tr>
<td>Soeverein Onafhankelijke Pioniers Nederland</td>
<td>Soeverein Onafhankelijke Pioniers Nederland</td>
</tr>
<tr>
<td>Partij van de Toekomst</td>
<td>Partij van de Toekomst</td>
</tr>
<tr>
<td>Democratisch Politiek Keerpunt</td>
<td>Democratisch Politiek Keerpunt</td>
</tr>
<tr>
<td>Libertarische Partij</td>
<td>Libertarische Partij</td>
</tr>
<tr>
<td>Anti Europa Partij</td>
<td>Anti Europa Partij</td>
</tr>
<tr>
<td>NXD</td>
<td>NXD</td>
</tr>
<tr>
<td>Blanco</td>
<td>Blanco</td>
</tr>
<tr>
<td>Ik heb niet gestemd</td>
<td>Ik heb niet hebben gestemd</td>
</tr>
<tr>
<td></td>
<td>Onbekend</td>
</tr>
</tbody>
</table>
### 7.4. Appendix VII – OL: Principal Component Analysis

PCA data on opinion leadership questionnaire

<table>
<thead>
<tr>
<th>Component Matrix&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Component 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mijn mening over politieke zaken lijkt voor andere mensen niet van belang</td>
<td>-0.689</td>
</tr>
<tr>
<td>Wanneer mensen hun keuze voor een politieke partij maken, vragen ze mij niet om advies</td>
<td>-0.800</td>
</tr>
<tr>
<td>Andere mensen komen bij mij voor advies over politieke partijen</td>
<td>0.817</td>
</tr>
<tr>
<td>Mensen die ik ken stemmen op een politieke partij gebaseerd op wat ik hen heb verteld</td>
<td>0.822</td>
</tr>
<tr>
<td>Ik overtuig vaak mensen om te stemmen op de partij die mij het meest bevalt</td>
<td>0.763</td>
</tr>
<tr>
<td>Ik beïnvloed vaak de mening van anderen aangaande politieke partijen</td>
<td>0.823</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

a. 1 components extracted.
### Total Variance Explained

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% of Variance</td>
</tr>
<tr>
<td>1</td>
<td>3,718</td>
<td>61,960</td>
</tr>
<tr>
<td>2</td>
<td>1,752</td>
<td>12,533</td>
</tr>
<tr>
<td>3</td>
<td>1,533</td>
<td>8,877</td>
</tr>
<tr>
<td>4</td>
<td>1,433</td>
<td>7,217</td>
</tr>
<tr>
<td>5</td>
<td>1,331</td>
<td>5,516</td>
</tr>
<tr>
<td>6</td>
<td>1,234</td>
<td>3,898</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

### KMO and Bartlett's Test

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</td>
<td>.848</td>
</tr>
<tr>
<td>Approx. Chi-Square</td>
<td>710,953</td>
</tr>
<tr>
<td>Bartlett's Test of Sphericity</td>
<td>df 15</td>
</tr>
<tr>
<td></td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td>.000</td>
</tr>
</tbody>
</table>
### 7.5. Appendix VIII – Opinion leadership: Reliability

#### Reliability Statistics

<table>
<thead>
<tr>
<th>Cronbach’s Alpha</th>
<th>Cronbach’s Alpha Based on Standardized Items</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>.876</td>
<td>.876</td>
<td>6</td>
</tr>
</tbody>
</table>

#### Item-Total Statistics

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Corrected Item-Total Correlation</th>
<th>Squared Multiple Correlation</th>
<th>Cronbach’s Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andere mensen komen bij mij voor advies over politieke partijen</td>
<td>19,4398</td>
<td>42,006</td>
<td>.568</td>
<td>.366</td>
<td>.872</td>
</tr>
<tr>
<td>Mensen die ik ken stemmen op een politieke partij gebaseerd op wat ik hen heb verteld</td>
<td>20,0871</td>
<td>37,680</td>
<td>.696</td>
<td>.605</td>
<td>.852</td>
</tr>
<tr>
<td>Ik overtuig vaak mensen om te stemmen op de partij die mij het meest bevalt</td>
<td>19,1992</td>
<td>37,352</td>
<td>.719</td>
<td>.627</td>
<td>.848</td>
</tr>
<tr>
<td>Ik beïnvloed vaak de mening van anderen aangaande politieke partijen</td>
<td>19,3859</td>
<td>39,146</td>
<td>.727</td>
<td>.535</td>
<td>.848</td>
</tr>
<tr>
<td>OL1_pos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OL2_pos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7.6. Appendix IX – Opinion Leadership: Normality

Tests of Normality

<table>
<thead>
<tr>
<th></th>
<th>Kolmogorov-Smirnov&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>df</td>
<td>Sig.</td>
</tr>
<tr>
<td>OL factor</td>
<td>.108</td>
<td>241</td>
</tr>
</tbody>
</table>

<sup>a</sup> Lilliefors Significance Correction

Mean = .00000
Std. Dev. = 1,06030
N = 241
## Appendix X – Sociometrics correlations

<table>
<thead>
<tr>
<th></th>
<th>MM factor</th>
<th>OL factor</th>
<th>Degree</th>
<th>Eigenvector</th>
<th>ClusteringCoefficient</th>
<th>Pagerank</th>
<th>Hubscore</th>
<th>Authority</th>
<th>Betweeness</th>
<th>Closeness</th>
<th>Number of strong ties</th>
<th>Strong degree</th>
<th>Outdegree</th>
<th>Indegree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's rho</td>
<td>1.000</td>
<td>.612**</td>
<td>.224**</td>
<td>.193**</td>
<td>.112</td>
<td>.188**</td>
<td>.193**</td>
<td>.193**</td>
<td>.112</td>
<td>.166</td>
<td>.098</td>
<td>.075</td>
<td>.215**</td>
<td>.222**</td>
</tr>
<tr>
<td>Correlation</td>
<td>.612**</td>
<td>1.000</td>
<td>.301**</td>
<td>.259**</td>
<td>.142</td>
<td>.288**</td>
<td>.260**</td>
<td>.260**</td>
<td>.247</td>
<td>.191</td>
<td>.076</td>
<td>.059</td>
<td>.269**</td>
<td>.304**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.001</td>
<td>.006</td>
<td>.110</td>
<td>.007</td>
<td>.006</td>
<td>.110</td>
<td>.018</td>
<td>.164</td>
<td>.164</td>
<td>.164</td>
<td>.164</td>
<td>.164</td>
</tr>
<tr>
<td>N</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
</tr>
<tr>
<td>Correlation</td>
<td>.224**</td>
<td>.301**</td>
<td>1.000</td>
<td>.922**</td>
<td>.063</td>
<td>.978</td>
<td>.923</td>
<td>.923</td>
<td>.847</td>
<td>.742</td>
<td>.618</td>
<td>.498</td>
<td>.890</td>
<td>.908</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.001</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.028</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.003</td>
<td>.242</td>
<td>.242</td>
<td>.242</td>
<td>.242</td>
<td>.242</td>
</tr>
<tr>
<td>Correlation</td>
<td>.193**</td>
<td>.259**</td>
<td>.922**</td>
<td>1.000</td>
<td>.187**</td>
<td>.868**</td>
<td>.100**</td>
<td>.100</td>
<td>.765</td>
<td>.818</td>
<td>.610</td>
<td>.505</td>
<td>.889</td>
<td>.793</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.006</td>
<td>.000</td>
<td>.000</td>
<td>.004</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Correlation</td>
<td>.112</td>
<td>.142**</td>
<td>.063</td>
<td>.187**</td>
<td>1.000</td>
<td>.037</td>
<td>.191**</td>
<td>.191**</td>
<td>-.101</td>
<td>.024</td>
<td>-.070</td>
<td>.007</td>
<td>.178**</td>
<td>-.024</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.006</td>
<td>.000</td>
<td>.000</td>
<td>.004</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.110</td>
<td>.028</td>
<td>.331</td>
<td>.004</td>
<td></td>
<td>.568</td>
<td>.003</td>
<td>.003</td>
<td>.118</td>
<td>.714</td>
<td>.279</td>
<td>.911</td>
<td>.006</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>.188**</td>
<td>.288**</td>
<td>.978**</td>
<td>.868**</td>
<td>.037</td>
<td>1.000</td>
<td>.869**</td>
<td>.869**</td>
<td>.897**</td>
<td>.741**</td>
<td>.609**</td>
<td>.481**</td>
<td>.858**</td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>.193**</td>
<td>.260**</td>
<td>.923**</td>
<td>1.000**</td>
<td>.191**</td>
<td>.869**</td>
<td>1.000</td>
<td>1.000</td>
<td>.767**</td>
<td>.811**</td>
<td>.610**</td>
<td>.505**</td>
<td>.889**</td>
</tr>
<tr>
<td>Pagerank</td>
<td>Sig. (2-tailed)</td>
<td>.007</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.568</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>.193**</td>
<td>.260**</td>
<td>.923**</td>
<td>1.000**</td>
<td>.191**</td>
<td>.869**</td>
<td>1.000</td>
<td>1.000</td>
<td>.767**</td>
<td>.811**</td>
<td>.610**</td>
<td>.505**</td>
<td>.889**</td>
</tr>
<tr>
<td>Hubscore</td>
<td>Sig. (2-tailed)</td>
<td>.006</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.003</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>.193**</td>
<td>.260**</td>
<td>.923**</td>
<td>1.000**</td>
<td>.191**</td>
<td>.869**</td>
<td>1.000</td>
<td>1.000</td>
<td>.767**</td>
<td>.811**</td>
<td>.610**</td>
<td>.505**</td>
<td>.889**</td>
</tr>
<tr>
<td>Authority</td>
<td>Sig. (2-tailed)</td>
<td>.006</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.003</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>.112**</td>
<td>.247**</td>
<td>.847**</td>
<td>.765**</td>
<td>.101**</td>
<td>.897**</td>
<td>.767**</td>
<td>.767**</td>
<td>1.000</td>
<td>.637**</td>
<td>.580**</td>
<td>.424**</td>
<td>.762**</td>
</tr>
<tr>
<td>Betweenness</td>
<td>Sig. (2-tailed)</td>
<td>.110</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.118</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>.166**</td>
<td>.191**</td>
<td>.742**</td>
<td>.818**</td>
<td>.024</td>
<td>.741**</td>
<td>.811**</td>
<td>.811**</td>
<td>.637**</td>
<td>1.000</td>
<td>.559**</td>
<td>.463**</td>
<td>.732**</td>
</tr>
<tr>
<td>Closeness</td>
<td>Sig. (2-tailed)</td>
<td>.018</td>
<td>.003</td>
<td>.000</td>
<td>.000</td>
<td>.714</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td>.098</td>
<td>.076</td>
<td>.618**</td>
<td>.610**</td>
<td>-.070</td>
<td>.609**</td>
<td>.610**</td>
<td>.610**</td>
<td>.580**</td>
<td>.559**</td>
<td>1.000</td>
<td>.635**</td>
<td>.654**</td>
</tr>
<tr>
<td>Number of strong ties</td>
<td>Sig. (2-tailed)</td>
<td>.164</td>
<td>.242</td>
<td>.000</td>
<td>.000</td>
<td>.279</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>289</td>
<td>358</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
7.8. Appendix XI – Opinion Leadership: Scatterplots closeness and betweenness centrality
7.9. Appendix XII – Market Mavenism: PCA

Component Matrix

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ik houd ervan om een nieuwe kijk op bestuurlijke en maatschappelijke vraagstukken te introduceren bij mijn vrienden</td>
<td>.799</td>
</tr>
<tr>
<td>Ik houd ervan om mensen te helpen door ze informatie te verstrekken over diverse soorten bestuurlijke en maatschappelijke vraagstukken</td>
<td>.856</td>
</tr>
<tr>
<td>Andere mensen vragen mij informatie over bestuurlijke en maatschappelijke vraagstukken</td>
<td>.812</td>
</tr>
<tr>
<td>Als iemand op zoek is naar informatie over bestuurlijke en maatschappelijke vraagstukken, zou ik hem/haar daaraan kunnen helpen</td>
<td>.856</td>
</tr>
<tr>
<td>Mijn vrienden zien mij als een betrouwbare bron van informatie waar het bestuurlijke en maatschappelijke discussies betreft</td>
<td>.813</td>
</tr>
<tr>
<td>Probeer u een persoon voor de geest te halen die informatie heeft over een grote diversiteit aan maatschappelijke en bestuurlijke onderwerpen en deze informatie graag deelt met anderen. Deze persoon weet alles over de betrokken personen, partijen en belan</td>
<td>.694</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

a. 1 components extracted.
## Total Variance Explained

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% of Variance</td>
</tr>
<tr>
<td>1</td>
<td>3,905</td>
<td>65,081</td>
</tr>
<tr>
<td>2</td>
<td>1,607</td>
<td>10,119</td>
</tr>
<tr>
<td>3</td>
<td>1,543</td>
<td>9,052</td>
</tr>
<tr>
<td>4</td>
<td>1,392</td>
<td>6,528</td>
</tr>
<tr>
<td>5</td>
<td>1,321</td>
<td>5,349</td>
</tr>
<tr>
<td>6</td>
<td>1,232</td>
<td>3,872</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

## KMO and Bartlett's Test

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</td>
<td>.874</td>
</tr>
<tr>
<td>Approx. Chi-Square</td>
<td>656,701</td>
</tr>
<tr>
<td>Bartlett's Test of Sphericity</td>
<td>df 15</td>
</tr>
<tr>
<td>Sig.</td>
<td>.000</td>
</tr>
</tbody>
</table>
### Appendix XIII - Market Mavensim: Reliability

#### Reliability Statistics

<table>
<thead>
<tr>
<th>Cronbach's Alpha</th>
<th>Cronbach's Alpha Based on Standardized Items</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>.890</td>
<td>.891</td>
<td>6</td>
</tr>
</tbody>
</table>

#### Item-Total Statistics

<table>
<thead>
<tr>
<th>Item</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Corrected Item-Total Correlation</th>
<th>Squared Multiple Correlation</th>
<th>Cronbach's Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ik houd ervan om een nieuwe kijk op bestuurlijke en maatschappelijke vraagstukken te introduceren bij mijn vrienden</td>
<td>24.9853</td>
<td>32.901</td>
<td>.700</td>
<td>.547</td>
<td>.872</td>
</tr>
<tr>
<td>Ik houd ervan om mensen te helpen door ze informatie te verstrekken over diverse soorten bestuurlijke en maatschappelijke vraagstukken</td>
<td>25.1127</td>
<td>31.500</td>
<td>.776</td>
<td>.645</td>
<td>.859</td>
</tr>
<tr>
<td>Andere mensen vragen mij informatie over bestuurlijke en maatschappelijke vraagstukken</td>
<td>25.5833</td>
<td>32.146</td>
<td>.713</td>
<td>.564</td>
<td>.870</td>
</tr>
<tr>
<td>Als iemand op zoek is naar informatie over bestuurlijke en maatschappelijke vraagstukken, zou ik hem/haar daaraan kunnen helpen</td>
<td>25.0441</td>
<td>32.693</td>
<td>.772</td>
<td>.614</td>
<td>.861</td>
</tr>
<tr>
<td>Mijn vrienden zien mij als een betrouwbare bron van informatie waar het bestuurlijke en maatschappelijke discussies betreft</td>
<td>25.1961</td>
<td>33.548</td>
<td>.717</td>
<td>.561</td>
<td>.869</td>
</tr>
<tr>
<td>Probeer u een persoon voor de geest te halen die informatie heeft over een grote diversiteit aan maatschappelijke en bestuurlijke onderwerpen en deze informatie graag deelt met anderen. Deze persoon weet alles over de betrokken personen, partijen en belan</td>
<td>25.8676</td>
<td>33.672</td>
<td>.583</td>
<td>.344</td>
<td>.891</td>
</tr>
</tbody>
</table>
7.11. Appendix XIV - Market mavenism Normality

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Kolmogorov-Smirnova</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic df Sig.</td>
<td>Statistic df Sig.</td>
</tr>
<tr>
<td>MM factor</td>
<td>.130 204 .000</td>
<td>.918 204 .000</td>
</tr>
</tbody>
</table>

a. Lilliefors Significance Correction
### Appendix XV - Twitter Keywords

<table>
<thead>
<tr>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>VVD</td>
</tr>
<tr>
<td>PvdA</td>
</tr>
<tr>
<td>CDA</td>
</tr>
<tr>
<td>PVV</td>
</tr>
<tr>
<td>#SP</td>
</tr>
<tr>
<td>D66</td>
</tr>
<tr>
<td>GroenLinks</td>
</tr>
<tr>
<td>ChristenUnie</td>
</tr>
<tr>
<td>SGP</td>
</tr>
<tr>
<td>PvdD</td>
</tr>
<tr>
<td>#TON</td>
</tr>
<tr>
<td>@VVD</td>
</tr>
<tr>
<td>#MenS</td>
</tr>
<tr>
<td>PPNL</td>
</tr>
<tr>
<td>Piratenpartij</td>
</tr>
<tr>
<td>#OBP</td>
</tr>
<tr>
<td>PvdT</td>
</tr>
<tr>
<td>SOPN</td>
</tr>
<tr>
<td>@PvdA</td>
</tr>
<tr>
<td>@CDA</td>
</tr>
<tr>
<td>@PVV</td>
</tr>
<tr>
<td>@SPnI</td>
</tr>
<tr>
<td>@D66</td>
</tr>
<tr>
<td>@groenlinks</td>
</tr>
<tr>
<td>@christenunie</td>
</tr>
</tbody>
</table>
7.13. Appendix XVI – Dimension Reduction

Twitter usage metric

### KMO and Bartlett's Test

<table>
<thead>
<tr>
<th>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</th>
<th>.465</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approx. Chi-Square</td>
<td>817,801</td>
</tr>
<tr>
<td>Bartlett's Test of Sphericity</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>3</td>
</tr>
<tr>
<td>Sig.</td>
<td>.000</td>
</tr>
</tbody>
</table>

### Total Variance Explained

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% of Variance</td>
</tr>
<tr>
<td>1</td>
<td>2,425</td>
<td>80,819</td>
</tr>
<tr>
<td>2</td>
<td>.551</td>
<td>18,376</td>
</tr>
<tr>
<td>3</td>
<td>.024</td>
<td>.805</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

### Component Matrix

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mention</td>
<td>.981</td>
</tr>
<tr>
<td>Retweet</td>
<td>.926</td>
</tr>
<tr>
<td>Reply</td>
<td>.777</td>
</tr>
</tbody>
</table>

Extraction Method:
Principal Component Analysis.
a. 1 components extracted.

### Reliability Statistics

<table>
<thead>
<tr>
<th>Cronbach's Alpha</th>
<th>Cronbach's Alpha Based on Standardized Items</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>.745</td>
<td>.877</td>
<td>3</td>
</tr>
</tbody>
</table>
Twitter popularity metric

### KMO and Bartlett's Test

<table>
<thead>
<tr>
<th>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</th>
<th>.587</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approx. Chi-Square</td>
<td>784,656</td>
</tr>
<tr>
<td>Bartlett's Test of Sphericity</td>
<td>df</td>
</tr>
<tr>
<td>Sig.</td>
<td>.000</td>
</tr>
</tbody>
</table>

### Total Variance Explained

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% of Variance</td>
</tr>
<tr>
<td>1</td>
<td>2,636</td>
<td>87,858</td>
</tr>
<tr>
<td>2</td>
<td>.320</td>
<td>10,678</td>
</tr>
<tr>
<td>3</td>
<td>.044</td>
<td>1,464</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

### Component Matrix

| Component | Component | 1      | Mentioned | .983 | Retweeted | .933 | Replied | .894 |

Extraction Method: Principal Component Analysis.
a. 1 components extracted.

### Reliability Statistics

<table>
<thead>
<tr>
<th>Cronbach's Alpha</th>
<th>Cronbach's Alpha Based on Standardized Items</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>.796</td>
<td>.930</td>
<td>3</td>
</tr>
</tbody>
</table>
Centrality metric

### KMO and Bartlett's Test

| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | .720 |
| Approx. Chi-Square | 3380.624 |
| Bartlett's Test of Sphericity df | 3 |
| Sig. | .000 |

### Total Variance Explained

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvectors</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% of Variance</td>
</tr>
<tr>
<td>1</td>
<td>2,862</td>
<td>95,400</td>
</tr>
<tr>
<td>2</td>
<td>.138</td>
<td>4,600</td>
</tr>
<tr>
<td>3</td>
<td>1,734E-006</td>
<td>5,779E-005</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

### Component Matrix

<table>
<thead>
<tr>
<th>Component</th>
<th>Eigenvector</th>
<th>Pagerank</th>
<th>Authority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.989</td>
<td>.952</td>
<td>.989</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.  
a. 1 components extracted.

### Reliability Statistics

<table>
<thead>
<tr>
<th>Cronbach's Alpha</th>
<th>Cronbach's Alpha Based on Standardized Items</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>.052</td>
<td>.976</td>
<td>3</td>
</tr>
</tbody>
</table>
7.14. Appendix XVII – Multiple regression OL

### Variables Entered/Removeda

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables Entered</th>
<th>Variables Removed</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>friendsperdegree, ClusteringCoefficient, retweetedperpagerank, ActiefLidpp01b</td>
<td></td>
<td>Enter</td>
</tr>
</tbody>
</table>

a. Dependent Variable: OL factor  
b. All requested variables entered.

### Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.496a</td>
<td>.246</td>
<td>.230</td>
<td>.86232040</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), friendsperdegree, ClusteringCoefficient, retweetedperpagerank, ActiefLidpp01

### ANOVAa

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression</td>
<td>46,488</td>
<td>4</td>
<td>11,622</td>
<td>15,629</td>
</tr>
<tr>
<td>1</td>
<td>Residual</td>
<td>142,771</td>
<td>192</td>
<td>.744</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>189,258</td>
<td>196</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Dependent Variable: OL factor  
b. Predictors: (Constant), friendsperdegree, ClusteringCoefficient, retweetedperpagerank, ActiefLidpp01
### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
<td>Zero-order</td>
<td>Partial</td>
</tr>
<tr>
<td>(Constant)</td>
<td>-.200</td>
<td>.116</td>
<td></td>
<td>-1.727</td>
<td>.086</td>
<td></td>
</tr>
<tr>
<td>ClusteringCoefficient</td>
<td>-.921</td>
<td>.377</td>
<td>-.156</td>
<td>-2.445</td>
<td>.015</td>
<td>-.121</td>
</tr>
<tr>
<td>ActiefLidpp01</td>
<td>.770</td>
<td>.128</td>
<td>.388</td>
<td>5.993</td>
<td>.000</td>
<td>.409</td>
</tr>
<tr>
<td>retweetedperpagerank</td>
<td>3.796E-007</td>
<td>.000</td>
<td>.135</td>
<td>2.096</td>
<td>.037</td>
<td>.223</td>
</tr>
<tr>
<td>friendsperdegree</td>
<td>-.001</td>
<td>.001</td>
<td>-.149</td>
<td>-2.288</td>
<td>.023</td>
<td>-.255</td>
</tr>
</tbody>
</table>

*a. Dependent Variable: OL factor*

### Collinearity Diagnostics

<table>
<thead>
<tr>
<th>Model</th>
<th>Dimension</th>
<th>Eigenvalue</th>
<th>Condition Index</th>
<th>Variance Proportions</th>
<th>(Constant)</th>
<th>ClusteringCoefficient</th>
<th>ActiefLidpp01</th>
<th>retweetedperpagerank</th>
<th>friendsperdegree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Constant)</td>
<td>ClusteringCoefficient</td>
<td>ActiefLidpp01</td>
<td>retweetedperpagerank</td>
<td>friendsperdegree</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2.932</td>
<td>1.000</td>
<td>.03</td>
<td>.04</td>
<td>.04</td>
<td>.03</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>.877</td>
<td>1.829</td>
<td>.00</td>
<td>.01</td>
<td>.08</td>
<td>.09</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>.548</td>
<td>2.313</td>
<td>.01</td>
<td>.50</td>
<td>.07</td>
<td>.33</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>.441</td>
<td>2.577</td>
<td>.00</td>
<td>.36</td>
<td>.65</td>
<td>.16</td>
<td>.08</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>.202</td>
<td>3.811</td>
<td>.96</td>
<td>.09</td>
<td>.16</td>
<td>.38</td>
<td>.23</td>
<td></td>
</tr>
</tbody>
</table>

*a. Dependent Variable: OL factor*
Model without mediating “active member in politics”

**Model Summary**

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.340(^a)</td>
<td>.115</td>
<td>.104</td>
<td>.94077870</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), friendsperdegree, ClusteringCoefficient, retweetedperpagerank
b. Dependent Variable: OL factor

**Coefficients**

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>.006</td>
<td>.105</td>
<td>.057</td>
<td>.955</td>
</tr>
<tr>
<td>1</td>
<td>ClusteringCoefficient</td>
<td>-.386</td>
<td>-.065</td>
<td>-1.052</td>
</tr>
<tr>
<td></td>
<td>retweetedperpagerank</td>
<td>5.724E-007</td>
<td>.000</td>
<td>3.236</td>
</tr>
<tr>
<td></td>
<td>friendsperdegree</td>
<td>-.002</td>
<td>-.222</td>
<td>-3.548</td>
</tr>
</tbody>
</table>

a. Dependent Variable: OL factor
### 7.15. Appendix XIII – Multiple regression MM

#### Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R$ Square</th>
<th>Adjusted $R$ Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.483&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.233</td>
<td>.221</td>
<td>.85299553</td>
</tr>
</tbody>
</table>

<sup>a</sup> Predictors: (Constant), followers_per_outdegree, ActiefLidpp01, retweetedperpagerank

<sup>b</sup> Dependent Variable: MM factor

#### ANOVA

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>$F$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>40,993</td>
<td>3</td>
<td>13,664</td>
<td>18,780</td>
<td>.000&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>1 Residual</td>
<td>134,606</td>
<td>185</td>
<td>.728</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>175,599</td>
<td>188</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Dependent Variable: MM factor

<sup>b</sup> Predictors: (Constant), followers_per_outdegree, ActiefLidpp01, retweetedperpagerank

#### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>$t$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>-.422</td>
<td>.097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 retweetedperpagerank</td>
<td>6.033E-007</td>
<td>.000</td>
<td>.220</td>
<td>3.359</td>
</tr>
<tr>
<td></td>
<td>ActiefLidpp01</td>
<td>.678</td>
<td>.127</td>
<td>.348</td>
</tr>
<tr>
<td></td>
<td>followers_per_outdegree</td>
<td>.000</td>
<td></td>
<td>-.197</td>
</tr>
</tbody>
</table>

<sup>a</sup> Dependent Variable: MM factor
### Coefficients\(^a\)

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>-.422</td>
<td>.097</td>
<td>-4.326</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>retweetedperpagerank</td>
<td>6.033E-007</td>
<td>.000</td>
<td>3.359</td>
<td>.001</td>
<td>.248</td>
</tr>
<tr>
<td></td>
<td>ActiefLidpp01</td>
<td>.678</td>
<td>.127</td>
<td>5.330</td>
<td>.000</td>
<td>.396</td>
</tr>
<tr>
<td></td>
<td>followers_per_outdegree</td>
<td>.000</td>
<td>.000</td>
<td>-1.97</td>
<td>.03</td>
<td>-.208</td>
</tr>
</tbody>
</table>

\(^a\) Dependent Variable: MM factor

### Collinearity Diagnostics\(^a\)

<table>
<thead>
<tr>
<th>Model</th>
<th>Dimension</th>
<th>Eigenvalue</th>
<th>Condition Index</th>
<th>Variance Proportions</th>
<th>(Constant)</th>
<th>retweetedperpagerank</th>
<th>ActiefLidpp01</th>
<th>followers_per_outdegree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Constant)</td>
<td>retweetedperpagerank</td>
<td>ActiefLidpp01</td>
<td>followers_per_outdegree</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2,288</td>
<td>1,000</td>
<td>.06</td>
<td>.07</td>
<td>.07</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.978</td>
<td>1,529</td>
<td>.00</td>
<td>.00</td>
<td>.04</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.457</td>
<td>2,237</td>
<td>.00</td>
<td>.52</td>
<td>.60</td>
<td>.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.276</td>
<td>2,878</td>
<td>.93</td>
<td>.41</td>
<td>.29</td>
<td>.01</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Dependent Variable: MM factor
### Residuals Statistics

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Value</td>
<td>-2.3274374</td>
<td>1.1060599</td>
<td>0.0351379</td>
<td>0.46695683</td>
<td>189</td>
</tr>
<tr>
<td>Residual</td>
<td>-3.18438625</td>
<td>1.96572435</td>
<td>0</td>
<td>0.84616235</td>
<td>189</td>
</tr>
<tr>
<td>Std. Predicted Value</td>
<td>-5.060</td>
<td>2.293</td>
<td>0</td>
<td>1.000</td>
<td>189</td>
</tr>
<tr>
<td>Std. Residual</td>
<td>-3.733</td>
<td>2.304</td>
<td>0</td>
<td>0.992</td>
<td>189</td>
</tr>
</tbody>
</table>

a. Dependent Variable: MM factor

**Histogram**

Dependent Variable: MM factor

**Scatterplot**

Dependent Variable: MM factor