Political Competition and Social Media:

Can Facebook Change the Status Quo of Finnish Politics?

Aleksi Suuronen

UNIVERSITY OF TAMPERE

Faculty of Social Sciences

Master's thesis in Public Choice

March 2018

DESCRIPTION

Title of the master's thesis	**Political competition and social media:** **Can Facebook change the status quo of Finnish politics?**	
Author	Aleksi Suuronen	
Supervisor	Professor Katri K. Sieberg	
Time of Completion	March 2018	
Place of Completion	University of Tampere, Faculty of Social Sciences	
	Pages: 102	Pages of appendices: 3

ABSTRACT

The rise of multiple social media platforms such as Facebook and Twitter have sparked a debate within political sciences on the degree to which these new platforms influence the overall political system. For some, the expected impact is assumed to diversify the number of people who wield and influence political power, while for others the impact concentrates and solidifies power to those who could be considered as political elites.

The purpose of this thesis is thus to investigate the underlying assumptions surrounding this debate and to construct a theoretical understanding of how social media can influence political power relations. The research conducted in this area has thus far mainly focused on studying Twitter and not many studies exist that have examined Facebook. Moreover, considerable gaps exist within the Finnish context, and no studies exist there that have explicitly examined what direct political gains social media platforms are able to give to its users.

This thesis aims to fill these gaps by adopting a case study strategy on the 2015 parliamentary elections held in Finland by using quantitative techniques to estimate the impact that Facebook has had on the electoral performance of those candidates who have chosen to use it in their campaign.

The results of this study show that Facebook is indeed linked with electoral gains, and especially the ability of candidates to acquire Facebook 'likes' is a significant factor that can boost their electoral performance. Furthermore, not all candidates are equal in this, and it is non-incumbent politicians who have gained the most from Facebook during these elections.

These results highlight the fact that social media in general and Facebook in particular are important subjects of political study and that they do capture a resource that the candidates can use to impact political outcomes. Moreover, this has implications for the future of the status quo of political competition within Finland and might change how politics and deliberation are conducted in the future.

Keywords: Facebook, Social Media, Elections, Finland

CONTENTS

1. INTRODUCTION

In 2008 Barrack Obama was running for the highest office in the United States, and the internet and social media was an integral part of his strategy to getting there. By using various methods from grassroots campaigning, fundraising to supporter mobilization, the way Obama utilized social media has become a prime example of how the platform can be used for electoral gains and later for more effective governance (Katz & Jain, 2013). Later, in 2016, a similar pattern would emerge, although the candidates could not have been more different from one other. While Donald Trump can be criticized on multiple fronts as a politician, one aspect in which he is universally given credit for is his masterful use of Twitter during his campaign. In contrast to Hilary Clinton Trump's campaign was significantly less funded, the candidate was much less politically experienced and multiple scandals during the campaign trail lead most observers to believe that Trumps days were numbered and his presidency a near impossibility (Allen & Parnes, 2017). And yet, it did happen. It would be an overstatement to say that Twitter was the main element contributing to Trump's victory, but it allowed him to directly communicate with his supporters and, furthermore, to create controversial Tweets that perpetually gave him free coverage within the media. Actions such as these cannot be discounted as meaningless either.

In other parts of the world, the events of 2011 and the Arab Spring gave fuel to the idea of cyberoptimism and lead some scholars to claim that "Twitter revolution" was a reality in many Arab states. Twitter, so the argument goes, gives citizens new tools to organize protests and crystalized public resentment towards governing regimes in authoritarian states. This, in turn, leads to stronger oppositional forces and, eventually, to regime changes and to democratization (Diamond & Plattner, 2012). Furthermore, social media has not only been claimed to mobilize oppressed citizens in authoritarian regimes; it accomplishes this in western democratic states as well. A study conducted in collaboration between political scientists and Facebook staff claims that a single message on Facebook's platform can mobilize hundreds of thousands extra voters to the polls on election day. On 2nd of November 2010, during the US congressional elections, approximately 60 million users were shown a message of their friends who had clicked the "I voted" button, which was a new feature introduced by the experiment. The message was claimed to directly increase turnout by approximately 60,000 votes and a further 280,000 people were indirectly nudged to the polls (Bond et al., 2012).

Although all of the above are mainly anecdotal evidence about the influence of social media to the political sphere, it is by no means a controversial stance to claim that social media matters in politics. Social media platforms are worthy to take under serious investigation within the political context by the sheer number of global users alone. According to Statista[1], the number of global social media users has reached 3.02 billion by January 2017, which at the time accounted for over 40% of the world population[2]. However, social media matters to politics not only because of the sheer number of people using these platforms but because it presents a shift in how people communicate with one other. As I will argue during this thesis, social media platforms and the functionality brought by the Web 2.0 era have changed the nature of information dissemination, how that information is organized and processed by individuals, and how the same information is further distributed and modified outside from its original sender. These changes in the flow of information have a direct impact on politics because it redefines who controls, consumes and interprets information and thus meaningfully changes political calculus and power distribution.

Indeed, the importance of social media has been noted by the academic field, and a lot of ink has been spilled on, for example, who uses social media, to what purposes it is being used for and how it affects wider populations. The topic can be problematic since it is characterized by rapid change and novelty. Social media is not a very old phenomenon, and it is very much a moving target. Facebook launched just over 13 years ago in 2004 and Twitter a few years later in 2006. Although the relative age of social media depends to some degree on how social media is defined, it is still young and somewhat limited by the amount and specifically by the focus of scholarly attention allocated to it. Furthermore, social media platforms change and develop constantly, which makes it hard for academics to keep up with it. For example, the algorithms that govern what is seen in the personal Facebook Feeds have changed drastically multiple times, and, furthermore, new features are continuously under development in various social media platforms in order to attract new users and to stay competitive against competition. Further problems and challenges for academic study arise from the fact that completely new platforms emerge and old ones die on a regular basis which creates caution to devote too much time to a single platform.

[1] https://www.statista.com/
[2] Data taken from www.worldometers.info

Many more challenges exist, but despite this social media is a phenomenon that I believe must be studied and efforts can be made to overcome the challenges related to it. The potential influence social media –and the internet in general– presents for politics means that we cannot have a full understanding of our contemporary political systems without incorporating it into the political theory. Unfortunately, while numerous empirical studies have been made about social media and politics, the theoretical side is not nearly as developed. Many previous studies have concentrated their efforts to describing social media use by various political actors, explaining social media adaption rates or investigating the content created within the platforms, but only a few have systematically tried to create a theoretical understanding of social media use that could explain how social media use changes political behavior, what goals various actors have and why certain social media strategies are more efficient than others. Additionally, the empirical side also has many gaps, which is mostly due to limitations regarding data gathering methods. For example, studies focusing on Twitter tend to be overrepresented because the platform is much friendlier to a scholar who wishes to use its data for research purposes.

This thesis aims to fill these gaps on two fronts: by formulating an underlying theory on why social media can have an impact on political competition and by empirical investigation on how accurate the theory is in the context of Finnish parliamentary elections. Theoretical work on politics and social media is regrettably undersized undertaking, which is understandable given the age and fast-changing nature of the topic. However, it is necessary to start building a theoretical framework for political social media use because without it we are limited in our understanding how influential social media truly is to politics and how it functions. There is no obvious path where such a framework should start, but my approach is to lean upon social choice theory and epidemiology in order to explicitly lay out the assumptions I have about the behavior of individuals and to construct a starting model that aims to explain why and how social media activity is influential. Without going too much into detail at this point (further details will follow throughout this thesis), the two basic assumptions behind my work are, firstly, that politicians are mainly interested in gaining votes in order to influence policy and, secondly, that ordinary citizens are rational actors in the sense that they wish to maximize their utility, *ceteris paribus*, and will avoid outcomes that are harmful to them.

A couple of things needs to be said about these assumptions before we can continue. For one, the rationality assumption can justifiably be criticized right off the bat by pointing out that it can be unrealistic and a poor description of human behavior. Recent advancements in behavioral economics have shown multiple ways in which the traditional understanding of rationality by economists does not hold up (such as when people make decisions under uncertainty[3]), which has led some to believe that such public behavior like voting is inherently irrational[4]. There is certainly merit to this train of thought, but it can lead to unjustified dismissal of rational action. While the traditional understanding of rationality in economics can be and should be criticized, it should not lead to the complete dismissal of the concept of rationality itself. Indeed, individual behavior can be seemingly irrational, but rational behavior can exist in the sense that behavior is consistent and aims to accomplish a certain goal within a context and with limitations. This is apparent within the concept of utility, which does not translate to strictly monetary gains –although that is how it is mostly presented– but rather it describes an outcome that is desirable to the individual. If we abandon the concept of rationality we lose or at least severely diminish the ability to understand and predict behavior because without rationality it is difficult to formulate goals and consistent strategies to achieve those goals[5]. Rather than abandoning the rationality assumption, my goal is to update and understand it within the context of political use of social media.

Furthermore, the assumption that politicians are selfish vote-seekers who mainly aim to acquire influence can likewise be criticized. After all, politicians can have multiple other motivations for their actions besides strictly aiming to gain more power. However, if one aims to provide a predictive theory of potentially effective uses of social media, the main objectives of action must nevertheless be specified. If vote-seeking politicians is not the right assumption, what would be the alternative? It seems to me that it is hard to argue that more altruistic motivations, such as producing policies that are beneficial for the largest portion of the population or the simple desire to have a positive impact on the society, are either too vague or too optimistic to serve as an explanatory variable for behavior. For one, such motivations are almost impossible to

[3] See, for example, the work of Kahneman and Tversky, *Prospect Theory: An Analysis of Decision Under Risk* (1979).
[4] Such claims are put forth by political scientists like Bryan Caplan (2011), who argued in his book *The Myth of the Rational Voter* that voters are irrational in the political sphere and have systematically biased opinions.
[5] Since it is not the objective of this thesis to combat the many criticisms pushed against public choice theory, I will not spend more time in justifying the approach. Such an undertaking would simply take too much space and would ultimately be counterproductive. For more interested and critical readers, I strongly suggest the work of Dennis Mueller: *Public Choice III* (2003) and specifically the last chapter for the more critical ones.

operationalize in a meaningful way since they vary greatly from one individual to another, are extremely hard to measure, and it is unclear how these motivations could explain larger trends at the aggregate level. Furthermore, it is somewhat naïve to think that vote-seeking politician is a poor assumption to explain the behavior of politicians. This is due to the simple fact that politicians who are not mainly interested in gaining votes tend not to stick around for long. If a politician ignores his prospects at winning the next election, he will very soon find himself outside of office.

Theory is only as good as how well it explains or predicts behavior, and as such no theory can stand on a solid ground without empirical evidence supporting it. To this end, this thesis also presents empirical investigation that is aimed at testing the assumptions behind the theoretical framework constructed during this thesis. Data from the 2015 Finnish parliamentary elections functions as the testing ground in this case, and a variety of statistical models are used to explain the vote distribution among the candidates who have used social media during the elections. The starting point is that social media activity by the candidates can be used to acquire votes in a way that is independent of other factors. In other words, social media activity in itself can influence voters to give their votes for a candidate and it is not merely a proxy measure of some other factor. The focus of this thesis is thus placed on individual political candidates during elections, and the theoretical framework is ultimately built with this context in mind. This will surely place some limitations how widely the theory can be utilized outside from its context, but it will serve as a starting point from which it can be expanded and developed further.

In summary, the main aim of this thesis is then to build a theoretical framework from which it is possible to understand why political actors use social media during elections and why it can influence voters to their side. This framework is then tested by investigating how politicians use and benefit from various online platforms to gain electoral support in the context of Finnish elections with the focus on candidates' use of social networking sites. It is especially important to formally estimate how much social media activity directly translates to votes because, as Gibson and McAllister also claim (2015, p. 530), it is not enough to observe who are actively or creatively using social media for political gain, but we also need to see that it directly impacts political outcomes. If there is no evidence supporting this, then it is highly questionable how great of an impact social media has for political life and whether it is truly an important topic worthy of being

a subject of study by itself or whether it is merely a new communication tool with little to no novel impact on politics.

2. THEORETICAL BACKGROUND AND RESEARCH QUESTIONS

The emergence of interactive social media platforms such as Facebook and Twitter have undoubtedly had an impact on how social interaction is conducted within societies. It could be argued, for example, that online interaction is increasing in contrast to face-to-face interaction since it is online where we now spend ever larger amounts of time presenting what we like, who we support, and who we are. Thus, the way individuals interact with one other is increasingly transferred from direct face-to-face communication into the online sphere where the rules and boundaries of interaction can be vastly different from their offline counterparts.

The political and the public sphere have not escaped this phenomenon, and it is also in these areas that the new online platforms are incorporated in the way both politicians and citizens interact with each other and within the political system. However, while many recognize the increasing importance of online interaction in political life, it is little understood, and we still do not know whether this presents a profound change in how politics is conducted or whether it is merely enforcing existing rules and structures through yet another new medium. After all, almost all new communication technologies have been accompanied by optimistic voices who have claimed that this new technology will transform how we communicate, will create a new era of increased democratization, and will give more power to individual citizens (Wu, 2010). It is then worthwhile to first explain where social media came from and to summarize the existing theories on how social media is suspected to impact the political sphere.

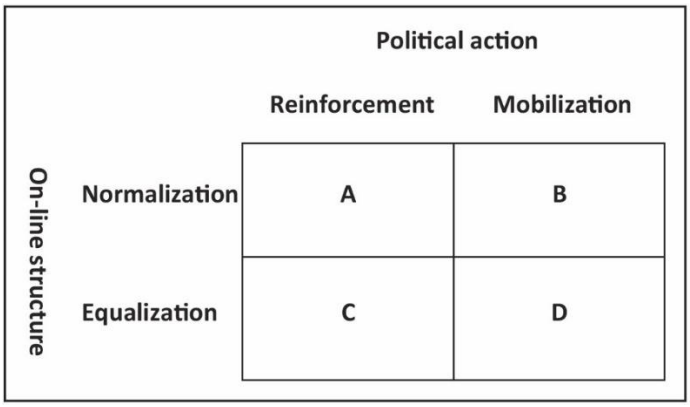
The theoretical background of social media has come largely from the discussion surrounding the differences and implications of the so-called Web 2.0 when compared to its earlier 1.0 version. The term “Web 2.0” was popularized by Tim O´Reilly and Dale Dougherty (O’Reilly, 2005) with the purpose of capturing how the logic and interaction between users of the World Wide Web and the Web itself have fundamentally changed from the early days of its implementation. During the Web 1.0 era, users could mainly engage with static websites and passively receive information from them. In contrast to this, Web 2.0 not only enabled users to access vast amounts of

information, it also introduced the ability for users to create and share content themselves in a manner that does not require expert-level information or specified skills. In essence, this gave the ability to virtually anybody with internet access to communicate and interact globally and instantly with each other with little to no expertise or financial resources. In countries where the internet penetration rates are between 80 and 100 percent, and where smartphones and other ICT devices are widespread, this ability has had profound effects on how people and citizens live and act in their everyday lives.

In the realm of political science and public behavior, the communicative abilities brought by Web 2.0 have sparked a debate on what the impact of such abilities is on political life. Indeed, research on campaigning and political competition has produced four distinct theses on the effects of social media, which are usually used in opposition to each other. To illustrate these theses, Kim Strandberg (2013) has constructed an analytical framework (see Figure 1) which can be used as a theoretical tool to assess the relative impact of social media on the political status quo.

The fourfold framework represents two distinct domains where social media can have an impact and consequently change the nature of political competition and action. The first domain, the on-line structure, is highly related to elections and electoral performance research, and it asks whether social media and online campaigning tools normalize or equalize relations between different political actors. According to the equalization thesis, online campaigning creates a more even playing field for the actors involved because it enables actors with limited resources to reach wider audiences and to interact with more voters than they could reach solely through offline efforts.

Figure 1. Strandberg's Analytical framework



Source: Kim Strandberg, 2013, p. 1335

Much of the traditional campaigning strategies are expensive to use effectively and can require large amounts of expertise and manpower, which are only available to political elites. In contrast, campaigning online and on social media gives cheap and easy-to-use campaign tools to candidates whose resources and campaigning expertise are relatively limited, thus giving them more of a fighting chance against the elites. However, the alternative view is that online campaigning does not necessarily differ significantly enough from its offline counterparts to break the status quo and will rather reinforce existing structures (Norris, 2001, pp. 218–221). This normalization thesis justifiably casts doubt on the ease-of-use of social media and its claimed low requirement of resources. Gaining large groups of followers and communicating effectively to large audiences becomes more complex as the scope increases and will most likely demand further resources as more and more individuals try to reach the same groups of people on the same platforms. As such, using social media effectively might not be as cheap and easy as it may initially appear, and the candidates with limited funding might still be left behind due to the more extensive resources of bigger parties and political elites.

The debate between these two theses are still going strong, and empirical findings have found support both for equalization and normalization. For example, Gibson and McAllister (2015) have examined the issue using Australian surveys of election candidates conducted between 2001 and 2010. The study found that in contrast to the web 1.0 era, when most of the empirical evidence pointed towards normalization (Gibson & McAllister, 2015, pp. 530–531), in the Web 2.0 era smaller parties, such as the Green Party, have been able to achieve significant electoral gain through the use of social media platforms in contrast to larger parties. For the authors, this in part supports the equalization thesis and suggests that where Web 1.0 was dominated by large parties and political elites Web 2.0 gives underdogs more opportunities.

On the other hand, in Finland the findings are less encouraging. By examining the 2011 Finnish parliamentary elections, Strandberg (2013) found that while the use of various social media sites by a multitude of different candidates was extensive, the analysis leaned towards favoring the normalization thesis. According to the logistic regression analysis, it appeared that the highest odds for a candidate to exhibit extensive use of social media is associated with being an incumbent, having high levels of education, and being a member of a major party. Furthermore, in another

study focusing on multiple parliamentary elections from 2003 to 2015, Strandberg (2016) found that the most active social media users come from the bigger parties rather than from the smaller or medium-sized ones, and it was also noticeable that the experienced incumbent politicians, who were acting as members of parliament during the elections, were considerably more active users than non-incumbents. From these findings, Strandberg concludes that the longer trends, at least in Finland, seem to point towards the normalization thesis and that it is precisely the candidates from the bigger parties that are dominating campaigning over the internet and social media (Strandberg, 2016, pp. 106–113).

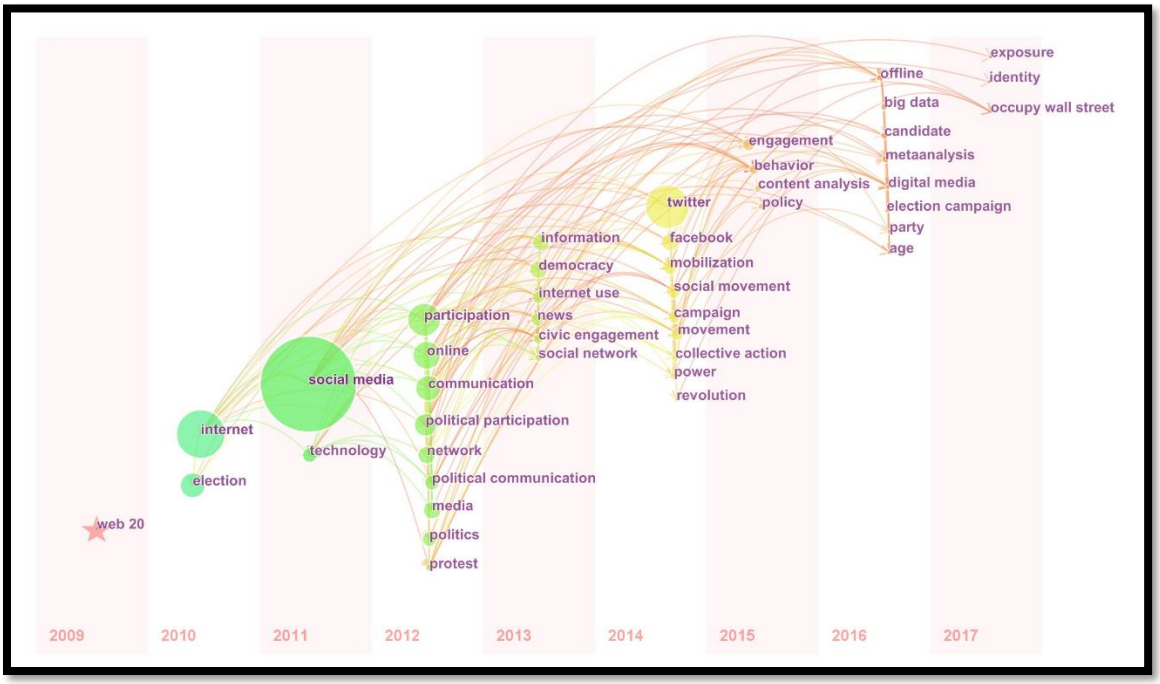
The second pair of the framework looks at social media from the viewpoint of voters. According to the so-called mobilization thesis, online interaction has the potential to provide information, organizational tools and more options for political engagement to those citizens who are marginalized or out of reach of political arenas and thus unable to participate in public life in a meaningful way (Norris, 2000, p. 218). As a result, social media can nurture new forms of political activity in which previously inactive groups will eventually be drawn into the public life and civic engagement. In opposition to this, the reinforcement thesis states that the claimed new forms of online political activity will, at most, be modest due to the fact that access and the required skills to use online outlets are divided within societies. Furthermore, even if skills and access were evenly distributed, online political engagement will still most likely be practiced by those who are already politically active and motivated. Hence, the new online possibilities for political engagement will do nothing more than reinforce the established patterns of activity and participation.

If Strandberg's analytical framework is considered in the sense of contrasting and mutually exclusive ideal types, then the possible outcomes A through D (see again Figure 1) can give a sense of how much social media has changed the status quo of everyday politics. In the event of outcome A, the significance of social media seems rather small, and it will most likely merely maintain the already existing gap between political underdogs and political elites. In the middle-ground of outcomes B and C, the significance can be considered to be modest, depending on the relative scale of the change, while in outcome D, social media can have a rather large significance.

In most cases, political research on social media tries to evaluate in which of these four compartments we are at the moment or whether we are moving from one compartment to another,

although research has by no means exclusively concentrated on providing answers to the normalization-equalization debate. In fact, the sheer range of different topics is quite vast and touches upon many aspects of political life. For example, in 2017 the Web of Science core collection comprised a total of 354 studies on social media in the political science category, from which 122 are from the United States, 55 from the United Kingdom, and 24 from Germany. The range of different topics can be seen in figure 2, which shows the most often used keywords as nodes where the size of the node is proportional to the number of times it has appeared within a research article. Links between the nodes indicate when the keywords have been used together within the same article, referred to as co-occurrence, and the years below the figure show the first time the keyword was used. As the figure shows, the articles included within the Web of Science core collection have studied social media in connection with elections, participation, campaigning, protests, and news, to name a few, with elections being a popular topic ever since 2010, participation and communication from 2012, and Twitter and Facebook more specifically focused on from 2014 onwards.

Figure 2. Co-occurrence network of keywords in political research of social media



Note: Based on 354 articles from the Web of Science core collection. The research articles are from the political science category acquired by the search operator "social media" OR "social networking sites".

It is then clear that social media is seen as an important research subject, and many researchers are contributing to advance our knowledge on social media within the political sphere. However, the existing research, by and large, is fragmented and has not received its fair amount of theoretical work that could connect many of the previous findings to an overarching framework and explain why social media use can benefit political actors. For example, studies have found that young politicians and women are highly active on social media and consistently aim to incorporate it into their political careers (e.g. Jungherr, 2016; Larsson & Kalsnes, 2014), but it is not clear why the young and women are more active or whether or not it is plausible to assume that their active social media presence will translate to political gains and, if it does, why this is the case. Furthermore, while Strandberg's (2013) analytical framework gives us a useful tool to conceptualize many relevant questions regarding social media and politics, it gives us little guidance on how to answer these questions. In short, without a theory to guide us, the political study of social media is reduced to mere observational studies that cannot explain the phenomenon we are interested in.

Given the scope of the topic, it is not obvious where this theoretical work should take place, but a promising starting point does present itself in democratic elections where political power and influence is most clearly seen. After all, it is during elections when the status quo can be challenged, and political actors are in the most immediate need to capitalize on their political resources and transform them into votes. If social media is shown to have an impact here, it clearly benefits political actors and can thus shape and influence political behavior and institutions. In this regard, a few studies have investigated directly how social media activity has impacted electoral outcomes (e.g. Baxter & Marcella, 2013; Kruikemeier, 2014; Spierings & Jacobs, 2014; Williams & Gulati, 2008), but within Finland, unfortunately, no such studies exist. This might come as a surprise, since, as Strandberg has observed (2013, pp. 1335–1336), several different factors make Finland an excellent country to study. Firstly, the relative role of individual candidates during campaigns is rather important in Finland. In the Finnish electoral system proportional representation is used in multimember districts where the voters can cast preferential votes for individual candidates who represent parties or electoral coalitions. In effect, this makes the elections not only a competition between parties but also a competition between individual candidates (Sundberg, 1997, p. 72). Secondly, Finland has both a high level of societal internet and social media penetration. Thus, a potentially large segment of the Finnish population is skilled in online interaction and is highly likely to be reached through online efforts. Furthermore, Finland

also has well-established data on both the use of social media for political purposes by the voters and the political candidates. These factors then make Finland an excellent case study, where the potential direct gains from social media use by the political candidates are likely to be observed if they exist at all.

Thus, the research questions that this thesis aims to answer are closely related to investigating electoral success and to Finland as a testing ground. More precisely, the objective of this thesis is to answer one main research question:

(1) How has social media affected the status quo political competition within Finland from the viewpoint of political candidates?

Although this question could be gauged from the viewpoint of citizens and voters as well, in the interest of narrowing down the scope of this thesis it is reasonable to concentrate solely on political candidates and leave the reinforcement and mobilization theses from Strandberg's Analytical framework out from the analysis. Moreover, even though political uses of social media surely are not solely constrained to politicians, they are the individuals that have the most immediate impact on the political system. To study them thus gives us the fastest path at gaining an intuition as to how social media impacts the political sphere.

In addition to this main question, this thesis sharpens its focus through multiple sub-research questions that further narrow the scope of the thesis and connect it to the broader issues surrounding the normalization v. equalization debate:

(1.1) Has social media activity impacted the electoral success of those political candidates who have integrated it into their campaign?

(1.2) Are there differences between these candidates such that the benefit gained from social media clearly benefits some candidates more than others?

(1.3) If the answers to questions 1.1 or 1.2 are positive, are there mechanisms internal or external to social media that can explain them?

These questions are important to answer since they allow us to make inferences that help us to answer the main research question of this thesis. For social media to have an impact on the status quo of political competition, we need to first establish that social media can indeed result in some clear benefits. This is the point of the first sub-question. This line of thought once again echoes

Gibson and McAllister (2015, p. 530) who believe that it is not enough to observe who are using social media, but we also need to see that it directly impacts political outcomes if we are to argue that social media has a role to play within the political sphere. The second sub-question is more directly aimed at the equalization v. normalization debate and attempts to investigate who the political actors that benefit the most from social media are. Are they incumbent political elites or marginal challengers? After all, social media can only equalize political competition if it helps marginal actors more than incumbents. The final sub-question is geared towards setting a theoretical framework from which these results can become understandable and comparable outside of the Finnish context. Although all political systems are different from one other and the likelihood of social media impacting the political sphere varies from one country to another, we need a formal mechanism that is common from one context to another. Without it, it becomes almost impossible to set social media into the larger study of political science and incorporate it into the overall theory. What I argue here is that social media should not be solely an independent area of study, but its most valuable contribution to our understanding materializes only after it is connected to a larger field.

To answer these questions, five concrete objectives will be set that will serve as a structure for the remainder of this thesis:

1. To define the main topic of interest during this thesis, and to answer what social media is and how it differs from other communication technologies
2. To establish a theoretical framework on how social media functions as a campaigning tool to be used by political candidates
3. Estimating the impact that social media had on the vote share of the individual candidates who incorporated it to their campaign during the Finnish 2015 parliamentary elections.
4. Investigating whether differences and patterns exist on how social media impacted different candidates
5. Analyzing the results in terms of the normalization vs. equalization debate

To elaborate further on these objectives, in this thesis it will first be necessary to define the topic and explore just what social media is and how it fits in the historical context of the development of political communication. Taken at face value, social media seems to represent a new way of political communication and could be used as a medium to reach voters both at the individual level

by directly communicating with voters and at the national level by broadcasting information to a selected population who have chosen to receive messages from a specific candidate. As it turns out, defining social media is not a straightforward task, and in consequence it becomes difficult to map out just how it can be used for political purposes. The task of the following chapter is thus to define our subject matter and to explore what tools, both familiar and new, it providers for political actors.

Next, once the topic is properly defined, the thesis will move to establish a formal theory of how social media activity can translate into political capital. As we will see, social media can be used in numerous ways to impact the electoral success of various candidates, but there is at least one aspect common to all of them. As Jacobs and Spierings have argued recently (2016), the casual path to electoral gains through social media can be complex, a mixture of the different opportunities, characteristics and contextual aspects of any given political system, but establishing a connection between citizens is usually the first step in this chain. For social media to be useful to any political actor, they must first reach their potential audience. Only after this connection is made can social media be used, for example, as a crowd-funding tool as Obama did in 2008 (Katz & Jain, 2013). Without going in further details, it is argued during this thesis that the way social media establishes and spreads these connections is the key mechanism that can explain electoral success within social media and this mechanism is analogous to how diseases spread among populations.

As objectives 3 and 4 are more empirical in nature, once the theoretical side of the thesis is completed I will proceed to analyze the Finnish 2015 parliamentary elections, their outcomes, and how the candidates used social media as a campaigning tool. This is the task of chapters six and seven where I will use candidate level data gathered from multiple sources and put the formulated theory under statistical tests, which will either validate the theory or suggest alternative paths for further study. Finally, the last objective brings the findings and the overall work of the thesis together and provides answers and conclusions to the research questions.

3. DEFINITION AND THEORETICAL USES FOR POLITICAL COMMUNICATION

Although social media has not been around as a concept and a phenomenon for long, quite a few different platforms have emerged that are considered to be examples of it. From the late 1990s onward, social media platforms began to take shape and services such as Blogger (1999), Wikipedia (2001), Myspace (2003), Facebook (2004) Flickr (2004), YouTube (2005), and Twitter (2006) began to offer web-based tools that created new ways in which individuals could communicate online (Van Dijck, 2013). Although all of the above-mentioned platforms, and many others, have received scholarly study, it is surprising that, despite the growing interest to the topic, many of these studies tend to leave out the specifications of social media undefined and tend to rely on direct examples to illustrate what they actually mean by social media. The issue, however, might be more difficult than it initially appears. Identifying just what constitutes as social media is not a simple matter much in the same way as it is hard to define an abstract concept such as pornography. In most cases, it seems to be enough to state that "I know it when I see it" (Jacobellis, 1964).

The issue has not yet created a significant problem mostly due to the fact that a vast majority of social media studies have studied a single platform, which in most cases means studying either Facebook or Twitter. However, since the main goal of this thesis is to provide a theoretical understanding of political social media use, defining social media is a step that cannot be left unattended. While platforms such as Facebook and Twitter can seem as too dominant to disappear, many other online platforms have come and gone without many remembering them once they are left out of use (Myspace comes readily to mind). The study of social media thus cannot be a study of one or two specific platforms but must be the study of the overarching phenomenon above all else.

3.1 What is social about social media?

Gayo-Avello (2016) provides us a useful starting point which we can use to define the subject at hand. By contrasting two different attempts to define the phenomenon, Gayo-Avello shows how social media has both evolved as a concept, what remains constant within it, and what we have

perhaps largely missed in it. The first definition comes from Wikipedia written in 2006, where social media was defined as:

> *Social Media is the term used to describe media which are formed mainly by the public as a group, in a social way, rather than media produced by journalists, editors and media conglomerates.* (as cited in Gayo-Avello, 2016, What do I talk about when I talk about social media?, para. 2)

The second definition came almost ten years after from Hogan and Melville (2015):

> *Social media represent a set of communication practices that can typically be described as 'many-to-many.' In contrast to broadcast media, consumers are typically also producers. In contrast to in-person communication, audiences are often ambiguous or underspecified.* (p. 421)

Both definitions rely heavily on the notion of "user-generated content" where the consumers or the public is itself creating and distributing content to other individuals without involving a third intermediary party. This certainly is an important aspect of social media, but it does not seem that it is a sufficient one. After all, many tools developed during and after the so-called "Web 2.0 revolution" fit the definitions: bulletin board systems, synchronous online chat, multi-user real-time virtual worlds, and e-mail lists (Gayo-Avello, 2016). Thus, it is likely that both definitions have important omissions that they fail to include and, indeed, Gayo-Avello (2016) has identified two of them. Firstly, they do not place any emphasis on the importance of communication networks. Social media platforms allow their users to choose their communication networks to a varying extent by, for example, friending people on Facebook, following them on Twitter or subscribing to them on YouTube. As Gayo-Avello (2016, What do I talk about when I talk about social media?, para. 7) argues, we cannot talk properly about social networks without this characteristic. The second aspect is related to the fact that neither of the above definitions addresses the social part of social media. While the first Wikipedia entry hints at the social aspect by stating that media is formed "in a social way", it does not continue to elaborate what this implies or why sociality is inherent within social media. Indeed, one might argue, as Nancy Baym (2015) does,

that there actually is nothing especially social about these online platforms since "any medium that allows people to make meaning together is social" (p. 1). Social media is merely yet another way of creating meaning such as postcards, television shows, newspapers, books or telephones have been for a long time.

For the purposes of this thesis, the final definition that Gayo-Avello (2015) arrives at works well at highlighting what social media is and how it can be used. In short, social media is "*any communication tool that allows users to consume, share and create multimedia contents which are addressed to unspecified audiences, in a potentially many-to-many fashion*" (What do I talk about when I talk about social media?, para. 14). This definition, while quite broad in nature, makes some important modifications to the earlier ones. It still keeps the concept of user-generated content and places emphasis on the many-to-many form of communication, but it does not make assumptions about the inherent "social" nature of the medium. Perhaps ironically this implies that the most important and novel part of social media is not social but is rather related to the way information and content are created and distributed among users.

Furthermore, the definition also recognizes that while social media can be used to create and share user-generated content, it does not necessarily need to revolve around it. Social media can also be used to share and consume multimedia that is not originally created by the user but comes from professional sources such as newspapers or other multimedia companies. Indeed, professionally created news stories are quite commonly shared on Twitter and Facebook. Lastly, while the form of communication can have the potential to be characterized as many-to-many, but it also facilitates one-to-one and one-to-many forms of communication. Applications such as chatting in Facebook messenger or WhatsApp are good examples of one-to-one communication, and while posting on Facebook allows for multiple people to engage in conversation, the original poster can use it in a more traditional way and only send content to be consumed but will not further engage with the consumers.

Within the political sphere, the most influential and most often studied examples of such social media platforms are Twitter and Facebook. There are multiple reasons why these two platforms are so dominant, but within the political sciences perhaps the biggest reason is related to the number of users who are on those two platforms and to the representativeness of their userbases.

According to Statista (2017), during 2017 Facebook was by far the most used social media platform with approximately 2,047 million users globally, followed by YouTube with 1,500 million and WhatsApp with 1,200 million. Twitter is only 10th on the list, but, as previous studies have found, it is widely used by politicians and candidates (Jungherr, 2016), which makes it logical for political scientists to study it. According to Jose Van Dijck (2013), both of these platforms further represent a special type of social media that have unique uses and aims. Dijck calls them Social Networking Sites (SNS), which primarily promote interpersonal contacts between individuals or groups. The aim of SNS is to create personal, professional or geographical connections and it encourages weak ties with people who have little direct face-to-face interaction. Other types of social media platforms exist, but since platforms other than SNS are only rarely used by political actors (e.g. Strandberg, 2016) the focus here will be on SNS, and the terms "social media" and "social networking sites" will be used interchangeably from now on. With these caveats in mind, this is what I generally mean by social media during this thesis.

3.2 Social media as a political tool

How, then, does social media fit within the wider context of political communication? Although it is easy to imagine why social media could be useful as a political tool, the various functions it provides are not alone anything new. Political actors already have at their disposal many of the tools that social media provides, such as personal websites for one-to-many communication, email for direct engagement, and chat rooms or blogs for more interactive and many-to-many forms of communication. However, one could argue that social media presents something fundamentally new in political communication not because it offers anyone novel function but rather because it combines many of the previously available functions into an integrated platform, which is easy and free to use and thus is readily available to everyone from politicians to voters. It is this holistic property of social media that can potentially have an impact on political communication.

Jacobs and Spierings (2016) argue that the main characteristics of social media can be summarized by five major aspects. First, the messages send on social media are unmediated. That is to say, that the way content and political messages spread within social media is not restricted, altered or prevented by gatekeepers or journalists who select what messages are newsworthy. By and large, even though most social media platforms impose some guidelines to what content can be published, individual users are autonomous in deciding what they post on the platform. Second,

social media platforms are also personal, even more so than many of the Web 1.0 applications. Almost all social media profiles are linked to an individual and allow multiple different ways for that individual to visualize and narrate his private and professional life. Moreover, since most profiles are in the personal control of an individual, the style of messaging and communication can be given a personal touch which reflects who that individual perceives himself to be. Third, one of the clearer ways in which social media differs from Web 1.0 applications is in its degree of interactivity. As the Jacobs and Spierings (2016) note, the option to involve others in your communication is typically a core feature of social media platforms. Various platforms accomplish this in different ways, such as directly commenting on the posts or tweets of other users, pulling users to the conversation via "@-mentions" in Twitter, "Tagging" in Facebook, or allowing quick reactions that indicate some form of engagement, such as liking or favoriting posts. In effect, this transforms users from passive receivers to active creators and influencers. Fourth, in respect to the political sphere, perhaps the most cited characteristic of social media is its cheapness and user-friendly interface. Indeed, the authors believe that the cheap and easy-to-use factors of social media open up the potential gain deriving from social media to virtually everyone, since much of the potential of traditional online tools is diminished by the high costs and expertise required to operate them. Lastly, the speed and volume of information diffusion within social media are miles beyond that of traditional media. Since social media connects its users to large networks by allowing users to "friend" or "follow" others, it creates a large networked environment where users are usually able to see what their friends have created, liked, shared or commented upon. This opens up the potential for information to spread not only through first-degree contacts but also first-, second-, third-degree contacts and beyond. This feature can be extremely powerful and, as the authors note, if the "six degrees of separation" theory (Travers & Milgram, 1967) is indeed correct, and everyone is reachable within six steps, then information on social media can diffuse to a large number of people across all social strata, given that they are in the platform. (Jacobs and Spierings, 2016, p. 22.)

Seen through this lens, social media appears to serve a multi-function in political communication and competition and can be placed somewhere between the local and the national. To provide some historical context, Pippa Norris (2000, pp. 137–161) explains that the evolution of campaigning has gone through three distinct phases where changes in communication technologies have played

a major part driving the change from one phase to another. In a short summary, political campaigning has progressed from a premodern phase, which was prevalent during the mid-19th century until the 1950s, to a modern phase that started to take shape during the 1960s and lasted until the 1980s. The final postmodern phase began in the 1990s and continues to evolve, although some scholars have argued that new technologies such as social media are driving our societies to another still unidentified phase (Vergeer, Hermans, & Sams, 2011). During the premodern phase, campaigns were largely local and decentralized, where voters held stable social and partisan alignments that made it difficult for any politician to change pre-existing voting preferences. Thus, the main function of the party organization was to energize and mobilize their voter base to go to the polls and encourage their friends and family to do the same. The main methods used to do this were based on personal networks of volunteers and face-to-face candidate-voter communication. In addition, partisan newspapers favorable to the party ideology were used for information dissemination, which was reinforced with local posters, pamphlets, and radio broadcasts.

The eventual progression moving the campaign methods away from the premodern phase was largely driven by the rise of television and regular publications of opinion polls. During the 1950s to the 1980s, the rising availability of TV-sets gradually sifted the main locus of communication from the partisan newspapers towards the national television broadcasts of the biggest TV-channels and focused on the main national evening news. Consequently, the main task of the party organizations was now to achieve favorable coverage in the mainstream evening news, current affairs programs and leadership debates on television. The effect was reinforced by the fact that in most nations only few television stations were available to the public to watch and the politically relevant programs were tightly scheduled to specific time slots. This limited coverage helped party leaders to control and to lead increasingly national campaigns where the electorate participated mainly via passive spectatorship mediated through television. Increasingly, voters were less likely to encounter laboursome forms of political communications, such as face-to-face discussions or local meetings, and both their partisan alignments and social cleavages became less stable.

The progression towards postmodern campaigning was again led by new developments in communication technologies, but also by the fragmentation of the mediascape (Chadwick, 2017). During the 1990s, a more diverse selection of different news sources appeared in the advent of satellite and cable television, talk radio, and 24/7 news bulletins. Furthermore, the internet

provided new opportunities for party-voter interaction. Due to the new communication capabilities, campaigns remained nationally organized but allowed decentralized operations. Regular opinion polls and focus groups helped the party organization to identify the most likely voters and to target them through direct mail, targeted ads, and narrowcasting within specific TV-channels and timeslots. Norris (2000) argues that this symbolized a return to the localized and interactive forms of political communications, which was somewhat lost during the modern era. As she explains, developments such as political discussion groups on the net, party intranets, the availability of interactive political websites, and political uses of e-mail or list-servers all represent a midway point between local and national campaigning, and between passive and active participation. (pp. 147–149.)

However, while the advent of the postmodern era greatly increased the available tools for political actors to communicate with the electorate, the fragmentation of information sources and the need to accurately identify and target potential voters made campaigning extremely expensive. Whereas before party organizations could rely largely on their existing resources and on the labor of their own staff and volunteers, the postmodern era requires professional staff to organize national strategies, coordinate decentralized operations and to conduct focus group research. Understandably, the rising costs concentrated political power to larger parties and well-funded individuals who could bear the costs of campaigning. (Norris, 2000, pp. 137–161.) It is still debatable whether or not social media is another development within communication technology that, like the television before, will drive the change towards a completely new stage of political communication, but it is certainly arguable that it offers many benefits for the candidates who have the ability to use them effectively. Perhaps the main benefit that social media provides is that it helps to manage the fragmented media environment of the postmodern era and reduces the costs associated with it. As an example, Facebook could be considered as a digital manifestation of the old reliance on volunteers to spread information. Whenever a candidate creates a post on Facebook it initially spreads towards the portion of the population that has voluntarily decided to receive information from that candidate. If this population then only consumed the message contained in the post and did not further interact with it with the functions provided by the platform, the spread of the message might end there. However, since one of the main function of social media platforms is to allow its users to further engage with the messages they receive, the users can help the

message to spread further to their own personal networks. This is done explicitly through functions such as sharing but can happen also implicitly without a clear intention for the user to do so. Liking a post on Facebook, for example, makes it more likely for that post to appear in the News Feeds of other users who are connected within the same network (DeVito, 2017). In both instances, the users are spreading information about the candidate without monetary compensation with or without their knowledge.

Indeed, for some this mode of user-based information diffusion is one of the main aspects of social media that sets it apart from older technologies like the radio or television. One of the proponents of this idea are Jason Gainous and Kevin M. Wagner (2014) who argue in their book: *Tweeting to power* that all previous advancements in political communication were mainly about increasing efficiency and reaching a greater number of people through one-directional channels. In the older technologies, the audiences are passive recipients of political messages who have little direct influence on how and what information continues to spread to other viewers. The logic of information diffusion was mainly dictated by a small group of individuals who controlled the main broadcasting channels and who studied readily identifiable groups and decided how to best reach and influence them (Wu, 2010). Social media, however, changes this dynamic and opens up a whole new form of two-way communication, which was not previously realistically available for mass-audiences. Instead of passively receiving information and indirectly influencing its spread as a member of a homogeneous group, in social media both the receivers and producers of information can communicate with each other, which allows for the receivers to directly and instantly indicate his or her preferences and to gain direct control on what content is deemed as worthy of redistributing further and even have an impact on how information is processed in the older media channels (Chadwck, 2017, pp. 70–102). Furthermore, this dynamic also allows individuals to increasingly manipulate what kind of information and communication they are exposed to, allowing for selective exposure.

As Gainous and Wagner (2014) argue, this leads to a new kind of communication logic, because whereas the producers of traditional media content had large control over the message and framing, social media is more user-driven and shifts more of the control from the media producer to the users and consumers of media content (Chadwck, 2017, pp. 70–102). As a consequence, communication becomes more complex and versatile than it was before. Furthermore, traditional

tools, such as newspapers or television broadcasts, are limited on how much content can be distributed to them and how many issues they can handle at a time. Social media, on the other hand, places no theoretical limits on the amount of content that can be distributed, nor does it place geographical constraints. Social media also transforms and adds new features in a much more rapid phase, which allows for new tools to emerge as the platforms develop. Indeed, since new tools can be continuously added to social media platforms, they can perform all the same functions that the traditional media outlets have previously done and to go beyond them:

> *At its most basic level, social media captures all of the elements of the previous mediums. In operation, social media can distribute everything that television, magazines, radio, and newspapers do, and in a more timely manner with an easily accessible and interactive interface on demand (…). Beyond being simply a compilation of the previous mass media, the social media presents a mass, multidirectional conversation. One can respond to a video with a video. One can discover an issue, research that issue and respond to it, respond to the responses or even chat about it, and then distribute it to networks of other people from any of a multitude of mobile or fixed computing devices.* (Gainous and Wagner 2014, p. 7).

Considering how vast this range of different functions is within social media platforms, it becomes apparent just how versatile and endlessly complex tool it can be. It is no surprise then, as Gayo-Avello (2015, Social media in electoral campaigns) has noted, that the tool has been used in numerous ways by political actors, such as to organize campaigns, raise funds, boost grassroots support, and to persuade undecided voters (e.g. Cogburn and Espinoza-Vasquez, 2011; Harfoush, 2009; Jaeger et al., 2010); get feedback from the electorate (e.g., Meeks, 2003) and to engage in deliberation with the electorate (e.g. Enli and Skogerbø, 2013; Ross and Bürger, 2014). However, there is considerable variation within the use of social media for political purposes across countries, and politicians often do not use various platforms to their full potential by neglecting their followers or by avoiding interactions with them (Gibson et al., 2013; Roper, 1999; Stromer-Galley, 2003). In broad terms, Foot et al. (2007) argue that there are four main practices which are commonly used by political actors during elections across countries both specifically in social media and more generally in the internet.

First, there is *Informing* where political actors use online tools to provide information about the central figures of specific campaigns, policy, and ideology relevant information that they use during public discourse and about the election process itself. Examples of this can be seen in candidates' biographies on the net, announcements on their positions on different issues and when candidates provide information about when, where and how to vote. Second, *Involving* refers to actions that are meant to get users to engage with the campaigners. Features such as calendars about future campaign events, the option to subscribe to additional sources of information, outlets for volunteering or contributing money are examples of these actions. Third, *Connecting* refers to any attempt by the campaigners to bring users towards other political actors. These attempts can be purely cognitive, like contrasting candidates to their challengers, or by providing hyperlinks that lead to further connection channels, such as email. Lastly, *Mobilizing* aims to transform users into volunteers and advocates. Like in the premodern era of campaigning, this can include giving materials to be distributed either online or offline and involve tools for supporters to publicly express their support for the candidate. (Foot et al., 2007.)

In addition to these, Gayo-Avello (2015, Agenda setting) has also noted that social media is well suited for agenda setting. In this regard, social media can be appealing to many marginal political actors, since it allows for them to bypass the traditional media, which they cannot access either due to lack of resources or inherent lack of interest by the media to cover them. Twitter, in particular, is useful in this regard, since political tweets can inject issues and views into the public domain that journalist may feel the need to cover (Graham, 2012).

Since the possibilities offered by social media are indeed vast, it is reasonable to assume that there is a significant learning curve associated with political social media use. Thus, political actors need time and learning to use social media effectively as a political tool. Moreover, it can be the case that the usefulness of social media will vary among different political parties, actors, and elections. The tools offered might not be useful to a party that relies on a portion of the electorate that has only a marginal presence on social media, for a candidate who does not know how to cope with the multiple functions that social media offers, or in high-profile elections where competition is conducted mostly in other mass-media platforms. Thus, every political actor needs to conduct an internal cost-benefit analysis and decide how heavily to invest in social media. Since the results of this analysis are likely to depend on the actor in question and on the time period, no single answer

can be given how actively social media will be eventually used nor how much it benefits those who use it.

Furthermore, the mechanisms that drive the decision for political actors to invest in social media are dynamic in nature and change alongside with the overall population that uses social media for political and other purposes. Jacobs and Spierings (2016) suggest that this dynamic is captured by what they call the Motivation-Resource-Based Diffusion Model. The model presumes that at first, when social media is still a novelty, it mainly benefits political actors from smaller parties[6] because they lag behind electorally and can be expected to have more incentives to try new communications and campaigning tools. This motivation drives them to experiment with social media, which will eventually lead to greater efficiency and incrementally greater benefits from its use. In contrast, larger parties are in more secured positions, and since their current efforts seem to be sufficient to keep them in their large party status, they see little reason to adopt new tools in their early stages. During this early-adoption phase it can then be expected that some form of equalization takes place and several smaller parties benefit from social media while the bigger parties remain at their old methods. Roughly, this situation is prevalent when approximately 10–15% of the entire population uses social media in some form. Once some time has passed and early-adopters are seen to benefit from social media, other bigger and better-financed parties take notice and encouraged by the positive results are motivated to invest in it. This process is further magnified when a larger portion of the population starts to have various uses of social media and the public who are on social media becomes more diverse.

The situation has now crucially changed because the dominant position of the large parties is threatened, and the environment transitions to a widespread-diffusion phase. During this phase, the large parties are likely to observe that an increasing number of their voter-base is on social media and that, even worse, their minor party rivals have taken positions in there while their own presence is insignificant in comparison. To balance the situation, bigger parties are then expected to take a larger interest in the platform. The problem faced by these larger parties, however, is that

[6] This reference to "small parties" is a bit of an oversimplification of the arguments advanced by Jacobs and Spierings (2016). For one, the theory they have advanced does not simplistically imply that all smaller parties will benefit from social media or have more motivation to start using it. Rather, the fact that multiple parties who are more prone to start using social media due to their ideological foundations or other charasteristics are also quite often marginal parties. This is especially the case with the so-called "postmaterialist" parties.

success on social media is largely a practice of trial and error, and it is plausible that minor parties are at an advantage to gain greater benefits due to their gathered experience during the early adaption phase. Fortunately, this experience can be bought and transferred to parties that have the resources to acquire them. For example, parties who have the necessary resources can hire professional consultants, social media managers, and to push their ads and posts via paid advertising. As Jacobs and Spierings (2016, p. 33) put it: "Resources enable [big parties] to catch up", and hence the term "resource" is in the name of the model. In contrast to early-adoption, it can be expected that normalization rather than equalization will take place during this time because larger, mainstream political actors adopt social media at a high speed. Finally, what Jacobs and Spierings call the laggard diffusion phase, almost everyone is to some extent on social media and even the most traditional and change-resistant parties will start to adopt it. However, these laggards often lack the resources and expertise to benefit from social media and will thus remain behind, normalizing their position as a smaller party. This model will then eventually predict that in the beginning social media leads to equalization for some, but at the end normalization for most actors.

It is important to point out that this model is implied to work both at the aggregated party level and down below at the individual candidate level. In theory, social media can provide an arena to which marginalized political groups can gain ground because the gatekeepers that kept them out of the game are largely absent (Blumler & Kavanagh, 1999; Enyedi, 2008).Thus, especially lesser-known politicians and traditionally underrepresented groups have an incentive to invest in social media in order to gain more name recognition, build alternative networks, and try to enter into the traditional media (Lasorsa, Lewis, & Holton, 2012). These individual-level dynamics are equally important as are the aggregate party level dynamics. Indeed, I largely agree with Jacobs and Spierings (2016, p. 11), who state that the power balances between individual politicians and intraparty competition are crucial in shaping the degree and practices of personalization and representation and thus fundamental to the functioning of our democratic systems. If we leave the individual level outside of the equation, we risk underestimating the impact that social media may have.

Given these possibilities, it seems plausible that social media represents a fundamental change in political communication – at least once enough time and learning have passed – and perhaps, as a result, has changed how political competition is conducted among western democracies. The mere possibilities and existence of these tools, however, does not directly imply that political

actors will start to use them or will eventually gain from their usage. Furthermore, as illustrated above, while social media is a versatile tool, and in most cases free to use, it is also extremely complex and requires significant effort from the political actors who need to learn how to make the best use of it. The issue is further complicated by the fact that different actors can have different situations, access to resources, and degrees of pre-existing political power, which ultimately change how social media is used. It is then beneficial to further investigate the current empirical evidence we have on how the existing platforms are currently used, who uses them, and what direct impact this usage has thus far had on electoral gains. The next chapter reviews the current literature on what we know in this respect, and it is later used to work our way towards an overarching theoretical model that can be used to explain some of the results the literature has previously arrived at. To elaborate further, the model adopts the view from Jacobs and Spierings that individual-level interaction is equally important as are party level dynamics and aims to explain how individual politicians acquire political capital through social media that they can transform to votes on election day. This work is needed since while the Motivation-Resource-Based Diffusion Model created by Jacobs and Spierings allows us to explain why certain political actors start to invest in social media at various stages it says little on why and when such investments can be beneficial. Before going too far ahead, however, we will move on to the literature review.

4. LITERATURE REVIEW

As previously mentioned, social media has received a fair amount of scholarly attention during the past ten years. Since political campaigns are one of the more direct channels through which political actors can influence the political sphere, it is no surprise that elections are one of the main research areas of political social media research. In general terms, research related to elections can be broken down to three main topics: explaining social media adaptation rates, analyzing content and patterns of use, and estimating the impact of social media on electoral outcomes. It seems logical to first explore who are the political actors that first start to use social media, or, in the terms of Jacobs and Spierings (2016), have the motivation to start investing in it. Thus, we shall start there and later on move towards content and patterns of use and end up with the impact this has had on election outcomes.

4.1 Social media adaptation rates

First, many researchers, especially in the early years of social media research, have been interested in how, what, and why politicians start to use social media for political purposes. This question has become ever more salient as political actors around the western world have increasingly started to have social media accounts. Many studies have documented an increasing trend of social media adaption in the UK (Jackson & Lilleker, 2009), Germany (Marcinkowski & Metag, 2014), France (Koc-Michalska, Gibson, & Vedel, 2014), Denmark (Hansen & Kosiara-Pedersen, 2014), and in Norway (Kalnes, 2009), for example. Theoretically speaking and based on earlier research on online campaigning and personal website use of political candidates, it could be argued that certain individual characteristics of political candidates (such as sex, race, and age), party voter base and regional power relations between parties can explain whether or not political actors start to adopt social media as a political tool. For example, a study by Esterling et al. (2005) observed that in the United States the age of the candidate and district characteristics explain a sizable portion of the variance of the use and quality of personal websites. In general, younger individuals and candidates in more competitive districts have had the most visible online presence. Furthermore, the study also noted that the candidates in rural areas are more prone to have quality websites, which might be an attempt to overcome the difficulty related to communicating with a geographically dispersed constituency.

In the age of social media, the importance of the competitiveness of elections has likewise been used to explain social media adaption (Lassen & Brown, 2010). In highly competitive elections candidates need to use resources extensively to remain relevant and visible to their constituency, and many effective tools to accomplish this are either highly expensive or only available to incumbent and experienced candidates. For example, paid advertising on television or on newspapers can have a positive effect on electoral gains, but only a minority of the candidates can afford them (Ansolabehere, Iyengar, Simon, & Valentino, 1994; Huber & Arceneaux, 2007; Ridout, Shah, Goldstein, & Franz, 2004). Thus, any tool that gives candidates with relative minor resources a cheap way to communicate with voters and to advertise themselves should be appealing. Furthermore, the more uncertain the election outcome is for an individual candidate, the more likely it is that all available tools will be utilized in the fear of losing elections. Empirical

findings support this line of thought; in the US 2006 midterm congressional elections candidates running in competitive races were the most likely to update their Facebook profile and it had a significant effect on whether or not a Senate candidate campaigned on Facebook (Williams & Gulati, 2007).

External factors specific to a certain time period can also play an important role in explaining why and when certain candidates choose to use social media. One well-documented example is the minority or opposition party status. Members of such parties can be drawn to alternative media outlets when traditional media has either lost interest in covering them or are unable to gain media coverage because ruling parties dominate media landscape (Lassen & Brown, 2010; Vergeer & Hermans, 2013). In countries where voters can cast preferential votes to specific candidates, the above argument can be extended to individual candidates and their personal characteristics. Just like a minority party can have difficulties in gaining media attention from traditional outlets, individual candidates that belong to minority groups can face similar difficulties.

The empirical findings, however, do not seem to support the suggested effect from a minority status. In the Netherlands, for example, the candidates running in the 2009 European parliament were more likely to adopt Twitter if they came from parties that were successful in past elections, suggesting that minority status is not a predictor of social media adaptation. Furthermore, the same study found that candidates who used Twitter belonged to the top quarter of the party candidate list, which further casts doubt on whether a minority status actually leads to social media use. (Vergeer, Hermans, & Sams, 2011.) It could be argued that this is a special case of intranational European elections, but similar findings have been reported on national elections as well. Vergeer & Hermans (2013) investigated whether candidates from less established and smaller parties were more likely to use Twitter as a campaigning tool during the 2010 Dutch general elections, but found no empirical evidence to support this hypothesis. The same study did, however, find that candidates who came from parties that had lost seats in the last general elections were more likely to subscribe to Twitter. This could indicate that other external factors other than minority status could be more important in explaining social media adaptation. Parties that are experiencing heavy losses could be more willing (or more desperate) to use any kind of new tools in an attempt to find new ways to reach voters since the old tools seem to be losing their effectiveness.

The ideology of the party has also been suggested as an explanatory variable. One of such arguments come from Vergeer & Hermans (2013), who argue that since the ideology of the party reflects their ideas about how society should be organized, what societal goals should be pursued, and how these ideas should be realized, it will also factor in how these parties utilize new media. According to Vergeer & Hermans (2013), this can be seen within liberal parties who appear to be early adopters of online campaigning tools. This can be partially explained by their liberal attitude that allows their candidates to design their campaign more freely and in a more individualized way, which in turn makes social media platforms such as Twitter a natural tool to use. There seem to be some merit to this hypothesis, since there are empirical findings showing that liberal parties have been the most extensive users of Twitter (Vergeer, Hermans, & Sams, 2011) and that conservative parties use social media less frequently (Baxter & Marcella, 2013), although there have not been many more studies that try directly to use ideology to explain social media adaptation. Furthermore, in their own analysis, Vegeer & Hermans (2013) found that ideology had no effect on whether or not Twitter was used as a campaign tool, suggesting that this argument does not stand on a solid foundation.

The most robust factors that have been found to meaningfully impact adaptation rates seem to be more closely related to a handful of personal traits related to individual candidates. Firstly, it seems that the uppermost elite candidates are more prone to start using social media as a political tool. This finding is supported by the fact that past electoral success of the party, larger finances, and upper list position give higher probabilities for an individual party member to adopt and use social media platforms (Lassen & Brown, 2010; Williams & Gulati, 2007; Verger, Hermans, & Sams, 2011; Vegeer & Hermans, 2013). Other important characteristic seems to be, perhaps unsurprisingly, the age of the candidate. In general, younger candidates are more likely to have a social media account (Lassen & Brown, 2010). At least two explanations can be made on the relevance of age the simplest of which is that younger people are overall more likely to use social media, and because of this social media is a more familiar tool to use for younger candidates. The other explanation is related to the expected voter base of young candidates. Young politicians are likely to target young voters who are more reachable through social media, which could mean that the potential gains of social media use are greater for younger candidates.

However, as it has been pointed out by Hoffman, Suphan, and Meckel (2016), examining the characteristics of individual candidates, parties, and campaign conditions does not tell us why certain politicians choose to utilize social media or explain what their motivations are to do so. Hoffman et al. (2016) aimed at filling this informational gap in their study of the 2011 Swiss parliamentary elections. The study found three distinct sources of motivation: promotion, information seeking, and entertainment. The most important motive out of the three is the promotion motive, which includes the candidates desire to create attention for relevant topics and distributing political messages, thereby increasing the personal profile of the candidate. The second most important motive is "information seeking", which includes actions related to collecting information on the political climate, seeing what others are talking about and understanding the electorates' interests. The final, and least important, motive encompasses entertainment uses, such as having fun and passing the time.

Taken together, these findings suggest that social media is a normalizing force within the political sphere and will more likely reinforce existing power relations rather than break it down. From the viewpoint of rational choice theory, this phenomenon should not come as overly surprising, given if we assume that social media use can result in electoral gains. Those political actors with most resources, experience, and status within the party are better situated at taking the full advantage from new campaigning tools, since, as already explored above, they have the ability to hire personnel to first learn how to use them and later manage them. It is then natural that vote-seeking politicians will use any tool that gets them more votes and those who remain at the top are the ones who are able to identify and sufficiently use promising new tools. This fits fairly well to the motivation-resource-based diffusion model from Jacobs and Spierings (2016), in that while some marginal actor might benefit from social media in its early days, later these differences are evened out and incumbents will use social media platforms in increasing numbers.

Lastly, apart from the personal candidate characteristics, the observed differences in the aggregated levels of social media adaptation within different parties seem to be more closely related to the demographic layout of the voters, rather than to the ideology of the party as such. It is likely that the reason why progressive and liberal parties have been both more likely to have and

to use social media accounts in contrast to conservatives is a result of what voters these parties target. Liberal and progressive parties tend to target younger, urban and more tech-savvy voters, a group who plausibly are also more likely to be on social media and to use it to search political information. Conservatives, on the other hand, target vastly different groups who are less likely to use social media and thus the expected gains from such platforms are likely to be minimal.

4.2 Analyzing the intensity and patterns of use

Once political actors start to use social media, the next logical step is to investigate how often and in what ways such tools are used. In this regard, we are lucky to have a multitude of findings from different countries and elections. Starting from the United States, by far the most popular social media platform used by political actors is Twitter, although it took some time to gain popularity among politicians. In 2009 well below half of the members of Congress used the platform. At the time there were a total of 159 members using Twitter: 39 Senators (39%) and 120 Representatives (28%). Furthermore, Twitter was mainly used to distribute one-directional information, which often can be characterized as self-promotion where the candidate sends, for example, tweets with links directing the user to blog posts written by the candidate. A second popular activity related to either announcing a type of activity the politician was currently engaging in or updating real-time their current location. This type of activity usually included tweets such as where the Congressperson had for lunch and where, updates when he is having meetings with constituents, corporations and other groups, and information from and about events they attended. Together these two types make up 82% of the 200 most recent tweets from the Congresspersons listed on TweetCongress[7] on 6th of February, 2009 (69 people in total). While one of the early hopes placed on Twitter argued that it would enable a more personal connection with citizens and politicians, in these early stages communication and interaction was a rather limited activity. Direct communication with constituents accounted only for 7.4% of the tweets (338 in total), suggesting that the platform was mainly used for self-promotion. (Golbeck, Grimes, & Rogers 2010.)

Meanwhile, Australian politicians exhibit much of the same patterns of behavior, and between 2009 and 2010 political actors in Twitter were much more likely to use it for broadcasting rather

[7] A grass-roots organization aimed at increasing interaction between the congress and US citizens through twitter: http://www.tweetcongress.org/about

than for conversation. A study by Will J. Grant, Brenda Moon and Janie B. Grant (2010) observed that, within the 10-month observation period, most of the tweets sent by political actors were isolated statement without reference to any other tweet or Twitter user. More conversational tweets (replies and retweets) were clearly a less popular activity overall. Interestingly, politicians from the Green party represented a clear exception to this case. In general, Green politicians send more conversational tweets, although mostly to their own party members, and additionally followed more people on twitter than politicians from other parties. There is no clear reason why this difference was observed, although Grant suggests that this might either result from the strategic attempt by the greens to maximize the reach their tweets were able to get, which suggests better intuition on how Twitter is used effectively or due to the less professionalized cohort of politicians within the party. The pattern of Twitter behavior by the Greens is much closer to the average Australian citizen, which can lead us to believe that many Green politicians are somewhat closer to them than to professional politicians.

In the Nordic countries, the patterns of use do not largely deviate from the above findings. The activity among Swedish and Norwegian politicians on Facebook and Twitter seem to be rather low outside of elections. A sample of 570 politicians shows that during non-election periods the median amount of posts or tweets are close to one per day. However, those who are active seem to contribute a significant amount of the overall content created by politicians, and the distribution of the number of tweets and Facebook posts is highly skewed. (Larsson & Kalsnes, 2014.) The same was also true in the Australian context, where a minority of politicians sent over 1,000 tweets within a 10-month period but most contributed with only one tweet or none at all (Grant et al., 2010).

Interestingly, those who were active social media users in the Nordic context seemed to be, as Larsson and Kalsnes put it, in ‘challenger’ positions. The most active politicians tended to be younger, in opposition and out of the political limelight (Larsson & Kalsnes, 2014). This contrasts with the early research on social media adoption, where the well-positioned politicians were more likely to open up a social media account. Perhaps continuous and active use of these accounts is something that challenger politicians have a greater interest in and perceive more benefits from it than the established politicians do, at least in the Nordic context. Alternatively, the result can arise from differing needs. Opening a social media account requires only minimal effort, whereas

maintaining it is something that requires a constant use of resources, which is something that the established politicians might want to spend elsewhere. As mentioned previously, challenger politicians frequently face more restrictions on how they can engage in political competition and remain relevant to the public. As such, social media is one of the few tools they can use, whereas established politicians have many alternatives and are thus more likely to spend their efforts elsewhere.

During elections the patterns of use remain partly the same, but some changes appear as well. For one, the intensity of use increases significantly, which is perhaps not a surprising result since all politicians have more incentives to communicate with voters. For example, Vergeer, Hermans, and Sams (2011) show that the candidates running for the Dutch 2009 European Parliament were more active on Twitter during the campaigning season than outside of it, which suggest that the platform is mainly used as a campaigning tool. Furthermore, it appears that a more social approach to campaigning becomes increasingly difficult as the size of the network grows: the more people follow the candidate the less likely it is that the candidate will also follow their supporters. This again provides confirmation that social media is used rather unsocially by politicians and that the platforms mainly exists to inform citizens. Interestingly, Vergeer, Hermans, and Sams (2011) point out that those parties that could potentially benefit the most from the platform had only little interest in it. Candidates that came from small and newer parties (such as Libertas, Newropeans, Solidara and Liberal Democrats) were among the less active parties on Twitter.

During the 2010 Swedish parliamentary elections, Larsson & Moe (2012) provide further evidence that patterns of use differ during elections. Whereas challenger politicians were more active users outside of campaigns, during them it appears that Twitter is an outlet mainly for the elite political actors or for individuals affiliated with mainstream media or political life in general. Furthermore, the overall communication method during these elections was in favor of one-directional communication: merely 7% of the messages in the sample of tweets were replies. While this holds true in general, some minority parties presented a more conversational style of tweeting. For example, the Pirate party and Feminist party were more active in both directly mentioning another Twitter user and within the retweet networks, which are used for interaction rather than for one-way broadcasting.

In the United Kingdom, similar findings largely correspond with the results from the above-mentioned studies. For example, in Scotland, during 2010, the candidates running for the UK Parliament rarely responded to the comments posted by voters in their Facebook pages, and there was a general reluctance to respond to "difficult" policy questions or critical comments. Furthermore, the content published by the candidates themselves frequently lacked any direct references to policy issues. (Baxter & Marcella, 2012.) Likewise, results from the country overall showed that the main strategy employed by the candidates was to use Twitter for one-way communication and to inform the public about the campaign and party affairs, such as about events, strategies, promotion, and polling. Policy issues remained largely at the background and only three percent of tweets were considered as position taking where a candidate posted his/her opinion, argument or the party position on a political issue (Graham, Broersma, Hazelhoff, & van't Haar, 2013).

Similar findings were uncovered by Parmelee and Bichard (2012), who studied a sample of 4,174 tweets collected from candidates in 12 competitive 2010 US races from 1st of October to 2nd of November. The findings showed that tweets regarding candidates' ideology or describing policy/issue positions was the least popular tweet within the sample (contributing only 4.7%). During these elections, by far the most popular style of communication was giving general information about the campaign. As Parmelee and Bichard (2012) suggest, this emphasis is likely a way for the candidate to keep voters aware and involved with campaign activities and progress. Furthermore, tweets were also likely to contain mostly negative information regarding opponents. Apparently, Twitter can provide a platform where many candidates feel comfortable spreading negative messages about their competitors in the hopes of disqualifying them in the eyes of the electorate. This could be, however, a unique characteristic of the political environment within the United States, and it remains to be seen how popular negative social media campaigning is outside of it.

As in other countries, some notable differences between parties do exist within the United Kingdom and on the patterns of their social media use. Mainly, there were a group of candidates who were significantly more active in interacting with voters and aimed at mobilizing, helping and consulting them. As Graham et al. observe (2013), these candidates largely came from the Labor party and from the Liberal Democrats, whose percentage of interactive Tweets were 47 and 42

percent respectively. Graham et al. (2013) believe that these findings illustrate that Twitter can be used to involve people in the political process by two ways: either broadcasting information on the campaign or providing a platform for interaction and mobilization.

Together, the studies presented above highlight at least three different aspects how social media has been used by political actors in recent years. First, it is abundantly clear that political actors most of the time do not use social media platforms to interact or to provide two-way communication with citizens and voters. The vast majority of tweets and Facebook posts are used for one-way communication and politicians only rarely answer to other users when they try to interact with them. This holds true almost regardless of what country election or time period we are talking about. However, it is important to focus on the fact that while this holds true in the aggregate, some clear differences are found between different groups. Candidates coming from smaller, liberal and younger parties have shown more interactive approach towards their political social media behavior and this might work in their favor. Lastly, what politicians talk about on social media is largely disconnected from policy discussions. The content created by politicians is in many cases characterized as self-promotion, announcements on the campaign trail or updates on what they are currently doing, where they are at and who are they with. While policy is not entirely absent, it does fade into the background and is perhaps even actively avoided by some politicians (Baxter & Marcella, 2012). In this light, social media is far from the digital public sphere some have hoped it would become (Shirky 2011). Here, it seems, is not the place where politicians and citizens can come together to discuss to one other nor policy to be debated among differing viewpoints.

4.3 Investigating impact

As previous studies have shown us, social media is being used as a political tool, more actively by some than by others, but these observations alone do not shed much light on how this activity has impacted the political sphere. If anything, the findings presented seem to suggest that not much has changed after all; a clear majority of politicians are not encouraged to interact more with voters, and social media is used mostly for one-directional broadcasting and self-promotion much in the same way as television and online campaigning has been used before (Bimber, 2003). However, there are some exceptions to this and, although the mode of communication is not all that different

from what has been seen before, the fact that many candidates from minority parties and from challenger positions are the most active social media users and use interactive methods could lead to changes in the balance of power between candidates. However, for this to be true, social media activity needs to have a significant impact and bring political gains for its users. As the following studies illustrate, in many countries this is indeed what we see.

As early as 2006, a study from the US midterm elections has shown that social media is associated with electoral performance, although the relation seems to be indirect. A study conducted on the 2006 US Congressional candidates investigated whether or not Facebook activity had a significant impact on the candidates' vote shares (Williams & Gulati, 2007). Using multivariate regression, where Facebook Activity was coded as a "1" if a candidate accessed and updated his or her profile in any way and coded a "0" if s/he did not, the study found that activity in itself did not seem to have a significant impact. When controlling for a number of electoral variables, it was found that incumbents or challengers who updated their Facebook profiles did not perform any differently than those who did not. However, while activity in itself did not seem to be significant, the number Facebook members who registered as a supporter of the candidate[8] had a significant effect on a candidate's final vote shares, particularly for open-seat candidates. The pattern of this effect was characterized by diminishing returns; increasing the number of supporters from 100 to 200, for example, adds approximately 1.1% to an incumbent's vote share, but to add another 1.1% requires 200 more supporters and yet another 1.1% requires 400 additional supporters. For Williams and Gulati (2007), these results suggest that rather than having a direct electoral impact, Facebook is capturing the underlying enthusiasm and intensity of support for a candidate in the form of Facebook supporters. In this case, the members of a community who are more enthusiastic about their choice are more likely to publicize that support and then take the time to do it, which manifest itself within Facebook.

Continuing with the investigation, Williams and Gulati looked further to presidential elections in the US and studied the impact Facebook had on the 2008 presidential nominating contests. As before, multivariate regression models were used to explain each candidate's popular vote share for all state nominating contests. The findings here are similar with respect to congressional

[8]At the time Facebook offered a "Endorsements" feature that allows its users to indicate support for a U.S. political candidate, which can be shared with the user's friends or publicly with everyone.

elections, and Facebook support was again found to be an important indicator of electoral success that is independent of traditional measures such as expenditures, media coverage, and organizing activities (such as campaign events). However, while direct activity yet again was not an important variable in itself, the absence of it could diminish the impact that the number of Facebook supporters had. For those candidates who made little effort to use Facebook and integrate it into their campaign strategy, the number of Facebook supporters had a near zero impact on the vote share. (William & Gulati, 2008.) This implies that, at least for high profile elections, activity within Facebook does matter and supporters alone do not guarantee electoral gains.

While it is interesting to see that social media variables are associated with electoral gains, the above studies do not tell us what type of social media activity is most likely to result in favorable outcomes. In this regard, Parmelee and Bichard (2012) have shown that the type of content candidates spread in Twitter also appears to have some influence on election outcomes. The analysis conducted on 4,174 tweets indicated that the types of tweets send by campaign losers featured topics relating to campaign strategy, ideology, or personal issues. On the other hand, campaign winners were more likely to feature replies, '@' mentions and to have URL links. Parmelee and Bichard (2012) suggest that these results can be explained by arguing that voters are not all that interested in seeking information about the personality or ideology of candidates in Twitter but are rather seeking ways and information on how to get personally involved. Thus, the reason why the winning candidates are more likely to send out replies, '@' mentions and URL links is that these kinds of tweets offer at least some level of increased interactivity and engagement for the voter who as a consequence feels more involved in the campaign.

In Europe too, the relationship between social media activity and electoral gains has been explored. In the Netherlands, Kruikemeier (2014) has explored the relationship between the style of usage of Twitter by political candidates and whether or not this usage has transformed to electoral support for these specific candidates. Looking at a sample of 40,957 political tweets from 177 candidates during the 2010 Dutch national elections (the sample was collected during the first 3 months before the elections and approximately 2 months after the elections), Kruikemeier (2014) found that Twitter use during the elections had a significant effect on the number of votes they received, although, as similarly observed in the US, followers seem to matter more than active use. On average, the dummy variable indicating whether or not a candidate used Twitter during the

campaign was associated with a 10% increase in preferential votes and the dummy was able to give a 9% increase in the variance explained after control variables such as incumbency, sex of the candidate, candidates' prominence in the media, and position in the party list were controlled for. It is telling, however, that once this simple dummy is replaced with the total number of tweets a candidate created during the campaign, the impact of this variable is not only insignificant but also negative, which implies that excessive tweeting can actually hurt the candidate. What seems to be more important, is the number of followers the candidate has, which is associated with a 17.1% increase in votes for a point change in the average amount of followers a candidate had during the campaign.

Furthermore, Kruikemeier (2014) also studied different patterns of Twitter use and their effects on preferential votes and found that more interactive users (candidates who used more '@' mentions) are found to gain more votes, although both the impact and significance is lower than the variable measuring the average number of followers. Kruikemeier (2014) argues that this result indicates that interactivity is one of the reasons why social media activity can bring electoral gains. Following an argument from Lee and Shin (2012), Kruikemeier (2014) proposes that this kind of interactivity induces social presence, which is defined as the degree to which individual feels that another communicator is present and there is an opportunity to engage in an actual conversation. Social presence consequently leads to higher intention to vote for a candidate who is able to create it. In short, interactivity leads individuals to support a candidate because it leads to the feeling that the candidate is closer to you. Unfortunately, while '@' mentions did indeed seem to have an impact, it is not known to whom these mentions were made. It is quite possible that the candidates are not interacting with potential voters, but are rather communicating with other party members, their supporters or with rival candidates. Evidence for this can be found in the US where '@' mention feature was regularly used by candidates to mention the Twitter accounts of their endorsers (Parmelee & Bichard, 2012). If this represents an overarching trend in other countries as well, this pattern of behavior would be indicative of something quite different than increased social presence among voters.

An alternative viewpoint comes from Spierings and Jacobs (2014) who argue that social media can have two main effects for preferential votes: a direct and an interaction effect. Firstly, the direct effect implies that social media can serve as a showcase or a campaign poster. Platforms such as

Facebook allow users to add personal details about themselves and as such allow for personally-centered campaigns that give voters information on the life and interests of the politician. In addition, merely having a social media account can signal to voters that the candidate is up-to-date with current trends and thus is shown to be a "modern" candidate. As Spierings and Jacobs (2014) note, this direct effect implies that the more followers a candidate has the larger the effect on votes can be since following acts as an implicit endorsement of the candidate. Secondly, the interaction effect assumes that a mere social media presence is not enough. In addition to having a large number of followers, the candidates need to maintain communication with them as well. In short, this line of argument proposes that followers only materialize to electoral gains when a candidate actively mobilizes them.

These effects were explored within the same 2010 Dutch national elections from which Kruikemeier also drew his sample, but this time another social media platform specific to the Dutch context was used alongside with Twitter for the analysis. The dataset used by Spierings and Jacobs (2014) include all 493 candidates of the 10 parties that had won at least one seat in the parliament and data from their Hyves[9] and Twitter social media profiles. The results from the data showed that the direct effect of having a Twitter account is significant but limited in comparison with Hyves. Every 1,000 followers on Twitter yield some 190 extra votes while the same number of followers in Hyves gives 1,343 bonus votes, but this is not statistically significant. Spierings and Jacobs (2014) suspect that the lack of significance in Hyves is due to the highly skewed data: 69 % of all Hyves followers came from just two candidates. While followers alone do not seem to do much, in contrast with the findings that Kruikemeier (2014) arrived at, social media activity can increase its impact. According to the ordinary least squares results (OLS), the interaction term between the number of followers and the number of tweets sent was statistically significant and each tweet published during the campaign gives 11 extra votes per 1,000 followers. Similar results were found in Hyves users and with a somewhat larger effect. From this, the authors conclude that social media can have a modest effect on the number of preference votes a candidate receives, but this impact is the largest when the social media profiles are actively used.

[9] Hyves was the largest social media platform in the Netherlands at the time and is equivalent of Facebook with broadly similar functions (Spierings & Jacobs, 2014). In the Netherlands, Hyves had roughly 8 million users, Facebook 4.5 million users and Twitter around 2.5 million users (Oosterveer, 2011).

While the studies above demonstrate that social media is indeed associated with electoral gains, they leave us a puzzle to ponder upon. First of all, since all of the above studies have been cross-sectional and quite plausibly subject to omitted variable bias, it is unclear whether social media variables capture an independent effect that can have a direct impact on electoral gains or are they, as Williams and Gulati (2007) have suggested, merely a proxy for the underlying enthusiasm and intensity of support for a candidate. However, the fact that in a number of cases social media activity is a significant variable in explaining vote shares and the finding that the interaction effect between the number of tweets and number of followers substantially increases the impact on preferential votes are promising indicators. These results show that while social media can capture underlying support for the candidate, that support can be maximized through active use of social media platforms and thus political use of social media is an independently important element in political competition. Indeed, looking back at the Obama campaign, one of the reasons why the campaign was successful was due to mobilizing uses of social media platforms that were used to inspire further support among followers, organize supporters towards grassroots campaigning and canvassing or acquire further campaign donations (Katz & Jain, 2013).

Whatever the case might be, by the very least these results show that social media is an important topic of study within the political sphere and it clearly does capture some kind of resource that political actors can acquire and transform to political gains. Based on these previous studies, it seems to be the case that during elections social media is mainly used as a one-way communication channel to the electorate, where candidates give information about their personalities, campaign activities, and self-endorsements. However, in many instances the level of interactivity varies among different politicians and some, such as many green party candidates, utilize the interactive features much more than others. Moreover, the interactive features have also been found to be associated with candidates who are the winners of elections and especially '@' mentions on Twitter can have a direct impact on the number of votes a candidate receives. Thus, it is highly plausible that interactivity is one of the main variables that determine whether or not a candidate can successfully use social media for electoral gains.

However, the reason why interactivity should matter is open to debate. Both Kruikemeier (2014) and Parmelee & Bichard (2012) hypothesize that the interactive features bring voters closer to the campaigns, either by making the individual candidate feel closer to voters or by giving voters more

tools to participate directly with the campaign, but there are other plausible reasons. One alternative explanation, which is not often brought up, is highly related to the algorithms that govern how each social media platform works and what content is shown to what types of users and how many will be able to see it. As explained earlier, one of the main novel changes that social media has brought to political communication is the increased importance of individual users, who can now impose greater control both on what kind of information they receive and what content is deemed worthy to be redistributed to other users. Thus, it is clearly important to consider the impact of social media not only through the actions of candidates but also through how other users interact with the content that is produced by the candidates. Interactivity, then, is not only an indication of social presence but is also an important variable that controls how content spreads within a social media network. Producing a @ mention, for example, subjects all individual users who follow the Twitter account to which the mention was made and thus increasing the reach the tweet has (Wang et al., 2013).

Indeed, it is precisely here where I believe we encounter one of the fundamental building blocks that allow us to explain why social media activity can transform into electoral gains. Whatever the eventual usage of the platform may result in, the success of it is fundamentally dependent on who is able to observe the efforts of the candidate. If a candidate invests in social media, but cannot attract attention to it, it is likely that these efforts will be in vain. Thus, the main resource that a candidate must have to be successful on social media is the level of visibility that his messages are able to achieve within the electorate. Following this logic, it is important to have a better understanding how the algorithms of social media platforms decide what content is seen to other users and how this content can be spread from one user to another. For practical reasons it is not desirable to go through in great detail every algorithm that has been created to numerous different social media platforms, and, furthermore, for theoretical purposes, it is enough to concentrate on a single platform. The main arguments created here, in the context of a single platform, can be transferred and altered, if necessary, to other contexts as well. Since Facebook is the most often used platform for both citizens and political actors within Finland (Strandberg, 2016) we will then turn our attention to it and explore in greater detail how it distributes content to users. Using this information, we can then finally arrive at a formal model that illustrates how the platform provides electoral gains.

5. REACHING VOTERS THROUGH INFECTION

To develop further the idea of visibility as a necessary condition for successful political use of social media, a study from Taina Bucher (2012) illustrates the algorithmic power that Facebook has on its users and explains why visibility and attention are the main ways in which Facebook allocates influence. After this, DeVito (2017) gives us further understanding how the News Feed algorithm works within Facebook, which enhances our intuition on what is shown in the Feed and how other users spread various content from user to user, either explicitly or implicitly. Indeed, the way information spreads in Facebook can be rather similar as for how diseases spread among populations. Like diseases, content on Facebook spreads through effective contacts with individuals who have encountered social media content and interacted with it. Liking a post on Facebook, for example, allows for the post to spread out to the personal network of the person who interacted with it (DeVito, 2017). Actions such as these are thus similar to an infection where one individual first spreads the disease to people who he is in contact with and then expands even further as these new infections are able to spread the disease through additional contacts.

This mode of information diffusion is surprisingly similar to the premodern phase of campaigning described by Norris (2000, pp. 137–161) where campaigners heavily relied on volunteers who would go canvassing from door to door, and who also aimed to influence their friends and relatives. What differs, however, is that many of the users who interact with social media content on Facebook are not, in most cases, entirely aware that the act of liking a post will spread the content further or did not intentionally meant to spread it. After all, liking something on Facebook can have multiple other motivations behind it, and the desire to give the post more visibility is just one possible motivation. As such, campaigners on Facebook need to consider how “infectious” their messages are, and, furthermore, they need not rely on volunteers who knowingly and actively will contribute to the campaign. It is enough for them to get other users to engage, to interact with the messages they send and to get their political diseases out to the world. To see how this helps us to understand social media within the political sphere, we have to first discuss in more detail how Facebook and its algorithms work.

5.1 Algorithms and prisons, how visibility is allocated within Facebook

By building on the theory of visibility and power created by Michel Foucault, Bucher (2012) argues that social media represents a space of "constructed visibility" where space is purposefully designed such that individuals are made seeable and, furthermore, seeable in a specific way. In the famous example of Jeremy Bentham's prison Panopticon, the prisoners residing in the prison are made visible by the observatory tower at the center of the circular building, and while one person in the tower can observe any individual prisoner, none of the prisoners can see who is in the tower or, more crucially, whether or not anyone is in the tower at all. A space constructed in this way projects power over the prisoners who are forcefully made visible and must therefore always act as if they are being observed. (Foucault, 1975.)

For Bucher (2012), this kind of material or technical structuring of visibility is especially interesting and relevant in terms of social media. Like the Panopticon, social media platforms are designed to make objects visible or invisible based on a criterion that the owners of the platform ultimately decide. In the Panopticon the threat of visibility impacts the prisoner's behavior to self-regulate themselves and internalize the rules of the prison. Furthermore, the threat is felt and distributed equally to all prisoners, since all prisoners are equally likely to be seen, and thus the mechanism of constructed visibility works in the same way for each individual. Bucher (2012) contrasts this to the non-physical space in Facebook, where the visibility of individuals is likewise constructed with a purpose, but where individuals are not all in equal position. The algorithm that dictates who or what is seen in the individual news feeds of other Facebook users is a complex mechanism that accounts, among other things, how strong relationships individual users have to one other, how users interact with content and what Facebook itself decides should acquire greater attention.

For individual users this is important since whereas the prisoners in Panopticon are living under the threat of visibility in Facebook the opposite is true: users are under a threat of invisibility. This is especially true for politicians who need visibility like a resource which they can later capitalize as electoral support or to agenda setting within the public discourse. Indeed, a whole industry has emerged around "News Feed Optimization" where marketers, media strategists, and PR firms advice how to remain relevant and visible on Facebook and on other platforms. However, in order to acquire visibility, politicians need to follow the logic embedded in the architecture of Facebook

and to learn how to act in the space of constructed visibility. It is here where we can start to see one of the ways in which social media platforms can shape and influence the political sphere. If social media tools become crucial for political actors, then the platform can decide and guide the type of behavior that is needed from those actors by rewarding them with increased visibility and, as a consequence, with more political capital.

The way Facebook's algorithm works is rather complex, and there is no public knowledge to which we could consult to learn accurately how the Facebook's News Feed functions or what determines what content is seen by individual users. Facebook keeps the algorithm governing the process out from the public view and even those within Facebook may not have a clear picture of how it works at the individual level (Napoli, 2014). However, while the exact procedure remains a mystery to most of us, some public information on the most general functions are available from sources such as Facebook's Newsroom blog, Facebook's patent filings, and from Facebook's Securities and Exchange Commission (SEC). Perhaps the most efficient way to summarize what we know about the Facebook News Feed algorithm is to start with one of its earlier version, which was in use until sometime around 2013.

Before it was declared "dead" in 2013 (McGee, 2013), the general understanding was that the News Feed was governed by an algorithm called "EdgeRank". EdgeRank was governed by two main elements: *Objects* and *Edges*. Objects refer to any status updates, uploaded pictures or other items that can be inserted into the feed. These objects, in other words, are the content that is created and shared within Facebook. Edges, on the other hand, are any relational interactions that other users make with Objects, for instance through "liking" or "commenting". The system works with these two components and assigns ranks to the edges, hence the name, and determines the overall likelihood that an object appears on an individual user's feed.

As Bucher (2012) explains, EdgeRank is like the algorithmic editorial voice of Facebook that determines what is shown to users by drawing on different factors relating to the Edges. While we do not know all variables that determine the rank of an Edge, at least three different components were understood to be in the key roles. The first variable is referred as *Affinity*. This value aims to capture the nature of the relationship between any two individual users. One important factor that determines this value is the degree to which two users are in interaction with one other. For example, checking a specific user's Facebook profile on a regular basis or sending private

messages with him will increase the user's Affinity score with respect to that person. The second value is *Weight*. Not all edges are equal, and some types of interaction are weighted more by Facebook than others. The Weight is determined by multiple factors, such as on how popular the edge is overall and how important Facebook itself considers it to be, but in most instances, or at least according to Bucher (2012), the weight Facebook places on different edges such as likes, comments, and shares seem to be based on the level of effort. The logic behind this is fairly simple: chatting requires more effort than commenting, which in turn requires more effort than liking, which requires more than just viewing. Thus, the more effort one spends the more important that edge should be. The final factor is *Time Decay* which simply refers to the recency of the edge. Here older Edges are considered less important than new ones.

The final EdgeRank is calculated based on the multiplication of these three measures: Affinity, Weight and Time Decay. Each object receiving higher ranked edges will have better chances at appearing in the news feed. This system is based on the assumption that not all users are equal, and it matters who you are friends with, what type of content you typically interact with, and what types of edges you personally create. For example, a user who typically likes content with videos in them will receive more videos in the future and if a number of your friends are interacting with a particular object, it is very likely that you will see it too. What is more, the weight given to certain Edges is directly shaped by Facebook based on their internal goals, which can change over time and thus change how the news feed works in an any given time period. In this way, the algorithm is not only a model that works through fixed assumptions on what the users want, but it is also a future-orientated in that it can be used to direct action to specific ways that are seen as valuable by Facebook. This can be seen in the way Facebook promotes new products, such as the "Questions" feature or "Places" by ranking interactions with them higher than others. (Bucher, 2012.)

However, as noted above, the EdgeRank is apparently dead (McGee, 2013). Facebook has continually moved away from the simple weight measures used in the EdgeRank to a more machine learning approach where the content one sees on Facebook is continuously influenced on how each individual behaves and interacts in Facebook. In a telephone interview Lars Backstrom, an Engineering Manager for News Feed Ranking at Facebook, estimated that there are as many as "100,000 individual weights in the model that produces News Feed." (McGee, 2013). Affinity,

Weight and Time Decay are not irrelevant in this system and are still important factors, but many more have equal importance.

Since Facebook prefers to keep most of the factors included in the current form of the algorithm secret, it is difficult to evaluate how far away Facebook has moved from EdgeRank, and thus it is almost impossible to know for sure what users can do to maximize their visibility on the platform. However, while the algorithm itself is out of reach, Facebook still publishes press releases, blog posts, presentations, patents and other documents that are valuable in giving us further intuition on what are the main attributes governing the feed. In this regard, a study conducted by DeVito (2017) is of great value since it is one of the most recent attempts by scholars to study how Facebook's News Feed works. In his study, DeVito conducted a content analysis from the main sources of public documents that Facebook itself has created that deal with the News Feed. The aim of the analysis was to investigate the core aims of the algorithm. The result of this analysis produced a list of algorithmic values that most likely influence the current News Feed. These values, in descending order of importance, are: (1) friend relationships, (2) explicitly expressed user interests, (3) prior user engagement, (4) implicitly expressed user preferences, (5) post age, (6) platform priorities, (7) page relationships, (8) negatively expressed preferences and (9) content quality.

As DeVito notes, all nine of these algorithmic values can be reduced down to personal significance where Objects are deemed more important if the algorithm thinks, for example, that it has a high impact to the user, the user has explicitly expressed interest in it (such as liking political pages) or if the user's friends follow similar patterns. Thus, this modern News feed is a constantly updated, personalized model that updates its outputs based on prior behavior, the behavior of friends and based on the sub-groups of users the algorithm decides you personally belong to. (DeVito, 2017).

Out of all the values the algorithm considers, friend relationships seem to be the most important factor. One of the more interesting documents pointing to this direction comes from the technical patents. For example, the main features of the News Feed are described in detail in the US patent 8,583,690 (Sittig & Zuckerberg 2013), where it is explained that the timeline module uses relationship data to generate personalized stories within the News Feed. As DeVito (2017) argues, this, and many other similar patents, confirm the fact that friend relationships are the core algorithmic value that decides what content is seen, and that it also moderates and informs all the other values. In short, those with whom the user is connected with, have higher affinity ratings

with, and exhibit the same kind of behavior are important influencers of what appears in the users personal News Feed; the objects that these individuals interact with will have a high probability of appearing in the individual user's Feed.

Taken together, although the Facebook algorithm can be complex and nuanced, it seems that the most important factor that candidates need to take into account when they aim to acquire visibility on Facebook is the degree to which other users interact with content that they create. Time and again, multiple different sources indicate that relationships between different users are the main catalysts that dictate how far messages from a political candidate are able to reach. Although Facebook can be used in a more traditional manner, such as creating a large degree of followers who voluntarily receive messages and then broadcasting to them in a one-directional manner, creating communication strategies that encourage users to interact and engage with the content makes it more likely that political messages spread further from this first-degree follower base and to reach far wider audiences. In this way, information spreads like a disease and continue to spread from user to user who continue to interact with it. Although we cannot accurately know just how this process works and how an individual message spreads in a particular network, we can, however, attempt at creating generalizations from the information available and use it to gain intuition on how social media has been used before, what potential it offers and what strategies can be used to acquire visibility on Facebook.

From this point onwards, I will continue with the disease analogy and use it for the purpose of a more explicit and accurate theory building. It turns out that epidemiological modeling is not an unfamiliar approach in social media studies, and it is there where I draw tools to explain how activity on Facebook can transform into visibility and consequently into political currency.

5.2 Political candidate as a disease – SEI model of candidate spread

In epidemiological modeling, individuals are usually divided into various compartments depending their epidemiological status (e.g. susceptible, exposed, infected, etc.). Individuals in each compartment transit between them via mutual contact at given average rates, such as from exposed to infected. The main purpose of these kinds of models is to describe and predict the average progression of the disease by tracking how many people are in each compartment. The same method is also used outside the realm of epidemiology and has been applied to model more

abstract phenomena, such as the spread of ideas (Bettencourt et al., 2006) and spread of news and rumors on Twitter (Jin et al., 2013), for example.

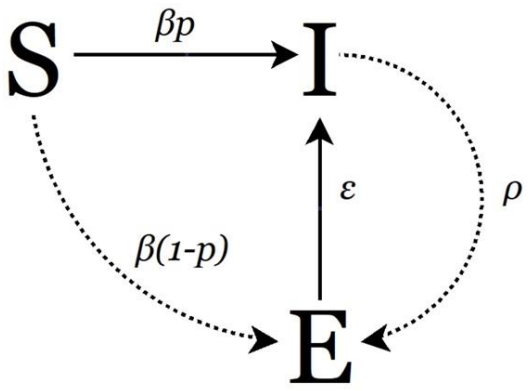
Similarly, political candidates campaign on social media can be viewed in this way. Once a candidate creates content on a social media platform and posts it online, it first reaches the voluntary followers who have decided to receive content from the candidate. This initial number may be large enough to have an electoral impact, but as exemplified by Facebook, this need not be a necessary condition. If the content a candidate posts is contagious enough, other users will start interacting with it, which leads them to become infected and for the content to spread further to second-degree networks and can create new infections along the way. Formalizing this relationship into a mathematical form gives us a useful tool that allows us to investigate how political candidates are able to reach potential voters and how and why individual candidates can differ in this endeavor.

To create such a formal model, I will build upon the foundation of epidemiological modeling. The model starts with a candidate C_i (the metaphorical disease) who spreads through a population N which is divided in to three compartments: the susceptible (S) who have not yet heard of the candidate, the exposed (E) who have heard of the candidate from social media, but do not interact with the content created by this candidate, and the infected (I) who interact and spread the content further. Using Facebook as an example, an individual is considered to be susceptible if he or she has not liked the official Facebook page of the candidate and has not interacted with any of the candidate's posts. In other words, the susceptibles have not encountered the candidate on Facebook in any way. Exposed are defined as individuals who have read at least one post or other content created by the candidate or have liked his official Facebook page or made a Friend connection with him and have thus indicated explicitly that they want to receive messages from the candidate. Lastly, infected are individuals who are reacting to the content created by the candidate through actions such as liking, sharing or commenting. Although in practice sharing, liking, and commenting will have a varying impact on how widely the interacted content spreads, in the interest of simplicity the scope of their spread is assumed to be the same on average in this simplified model[10].

[10] Furthermore, we do not have accurate information available on how these different actions impact the spread of information, which is likely to further depend heavily on other variables such as friend relationships, age of the

The process through which candidate C_i spreads to the population N is similar to Jin et al's. (2013) process of how news and rumors spread on Twitter, but with significant differences. The process is described in figure 3, where E and I recruit from the susceptible compartment S with different probabilities:

Figure 3. SEI model of candidate spread



- At the start of the process, the infected compartment has at least one infected, the candidate C_i, who creates and spread content denoted as c. As the number of published content grows, the interaction between the infected and susceptibles can lead to two different outcomes: the susceptibles to become infected or exposed.
- The first arrow leading from S to I in figure 3 describes the attempt by the infected to recruit from the susceptible compartment at the rate β which is defined as the effective contact rate. Since neither contact with diseases nor with posts on Facebook deterministically always leads to infections (a post can appear on a News Feed, but completely disregarded or unnoticed by the user) the parameter β is further compiled from the total contact rate (the total number of potential voters reached per one unit of content regardless of whether or not they actually noticed the content), denoted γ, multiplied by the probability that the content is consumed, denoted l. Thus, β is the effective contact rate: $\beta = \gamma l$.

content and how many users have overall interacted with the content. Thus, it is not possible to put precise assumptions into this model. However, it could be assumed that on average the scope of the impact will be roughly greater for some reactions (or edges in the EdgeRank terminology). For example, it could be assumed that sharing will result in a wider spread than commenting and commenting to wider spread than liking. This assumption will be explained further below and tested empirically during this thesis.

- However, in most cases consuming content does not directly lead to individuals immediately spreading it further. Because of this, susceptibles move directly to become infected only at probability p and at $1 - p$ move to the exposed compartment.
- If exposed, the parameter ε describes further effective contact with the infected, through which the exposed can transit to the infected compartment. In the context of Facebook, it is reasonable to assume that $\beta < \varepsilon$ since the already exposed individuals are likely to be more willing to engage with content that comes from a familiar source. Furthermore, being exposed to a candidate in Facebook usually implies that the exposed individual has friends who actively engage with the candidate or that the individual has liked the candidates official Facebook page. Both of these increase the priority, and thus the effective contact rate, that the Facebook algorithm assigns to the candidate's content in the individuals News Feed.
- In most cases, infections do not last indefinitely, and at some point individuals leave the infected compartment. In social media too, individuals usually do not continue to spread every piece of content published by a candidate. Thus, the parameter ρ denotes the rate at which individuals stop spreading content and move to the exposed compartment instead.
- In most epidemiological models the system has *vital dynamics* which determine the rate of which the population exists and enters the compartments via deaths and births. However, in this application the covered time period is assumed to be rather short: the average time of an election cycle. Thus, the total population size N is assumed to be constant.

Table 1. Parameter definitions of the SEI model

Parameter	Definition
β	S-I contact rate
ε	E-I contact rate
ρ	Departure rate from I to E
p	S->I Probability given contact with the Infected
$1-p$	S->E Probability given contact with the Infected

The process is described mathematically by the following set of ordinary differential equations where the sizes of each compartment are defined as derivatives with respect to the number of content published.

$$\frac{\partial S}{\partial c} = -\beta S \frac{I+z}{N} \quad (1)$$

$$\frac{\partial E}{\partial c} = (1-p)\beta S \frac{I+z}{N} - \varepsilon E \frac{I+z}{N} + \rho I \quad (2)$$

$$\frac{\partial I}{\partial c} = p\beta S \frac{I+z}{N} + \varepsilon E \frac{I+z}{N} - \rho I \quad (3)$$

$$N = S + E + I + z \quad (4)$$

In standard epidemiological models, the proportion of infectious contacts is $\frac{I}{N}$ and thus the rate of increase for infectious individuals is $\beta S \frac{I}{N}$. However, since the process in this instance is described in the terms of number of content published and not in terms of time, and since the candidate creating the content is himself considered to be infected, the term includes a constant z and becomes $\beta S \frac{I+z}{N}$. The constant z not only describes the candidate himself, but also any additional personnel or other devoted followers who will spread content created by the candidate indefinitely. While parameter β controls the rate at which the candidate becomes known to the susceptibles, the term $\varepsilon E \frac{I+z}{N}$ determines the rate of increase of infected from the exposed compartment and the term ρI the rate at which the infected stop spreading content.

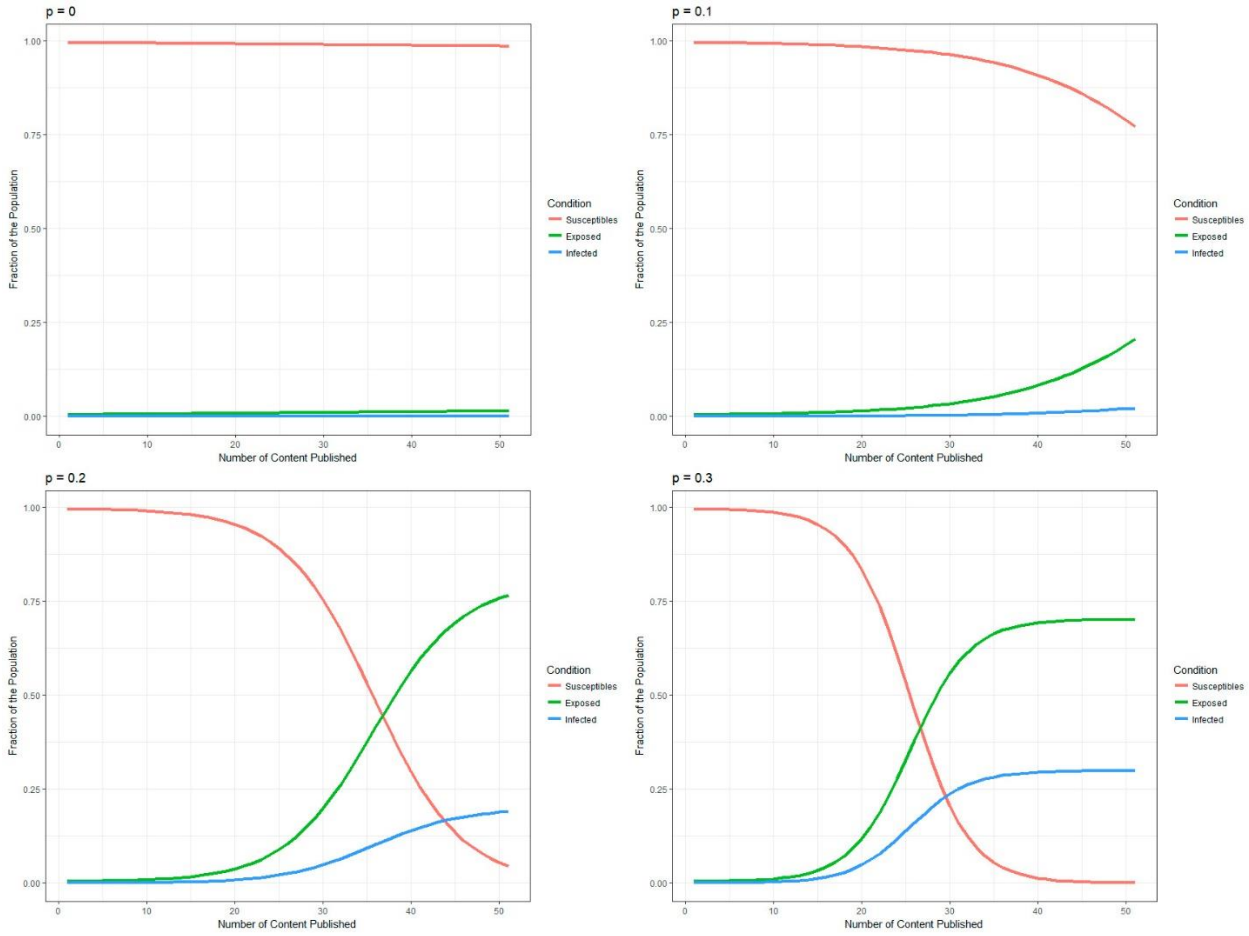
From equations 1-3 it becomes clear how effective different social media campaign strategies are when the goal is to reach a maximum number of potential voters. Consider a population of 10,000 individuals where the candidate C_i is initially the only one belonging to the infected compartment ($z = 1$) and 50 individuals are initially exposed to him. When the following parameters are set at $p = \varepsilon = \rho = 0$ and the candidate is forced to reach voters solely through his own initial contacts and he cannot create infectious individuals who would help the candidate to spread his messages, the rate at which susceptible are requited to the exposed compartment is rather linear and the

infected population remains constant at 1. Thus, values change only within equation 1 and 2, and equation 2 is reduced to $\frac{\partial E}{\partial c} = \beta S \frac{1}{N}$.

In contrast, when parameter p is allowed to vary at $0 < p \leq 1$ the increase of the infected population can have drastic consequences for the diffusion of knowledge about the candidate. Figure 4 displays the population changes of each of the compartments with the above set initial values, with different values for p and where $\varepsilon = \rho = 0$ and $\beta = 1$. As it can be seen from the figure, when $p = 0$ the candidate needs to have a significantly high effective contact rate because he alone is spreading messages to the overall population. With higher values of p the candidate can gain more visibility within social media even with a low value for β, and with probability approximately larger than .30, the candidate is able to reach the total population once he has send approximately 50 messages to the population. Naturally, this is only an extreme case where the infected will continue to spread content indefinitely, and in practice parameter ρ will control the effect the number of additional infected can have.

By using these equations, we can define the differences between different candidates and how effective their social media campaigns can be in terms of the number of voters reached, and thus how likely it is that the platform can change the status quo political competition. Firstly, it seems natural to assume that various candidates have varying degrees of how "infectious" they are and how skillful they are at creating engaging content. In other words, their value for β will differ and is likely to be context dependent as well, as the previous research has shown us. For example, one might assume that at certain elections politicians who are critical of immigration will be more infectious when the political climate at the time places high prominence to issues revolving around foreign and immigration policy. Furthermore, others will differ due to more constant conditions, such as being born at a time when social media usage was common and thus have higher understanding and ability to use such platforms effectively.

Figure 4. SEI model simulation



Note: Initial population values start at z = 1, S = 10 000, E = 10 and I = 1. The rest of the parameters are set at: $\varepsilon = \rho = 0, \beta = 1$.

Likewise, the probability that susceptibles will become infected is largely dependent on the type of content that a candidate decides to produce. In various social media platforms, the type of content a user creates will have a significant impact on how likely it is that it will be shared. For example, when a published content includes a URL link to further content, it is generally more likely to be retweeted on Twitter (Suh et al., 2010). Beyond of the characteristics that a candidate has and how he operates on social media, the spread of a candidate also depends on the characteristics of the audience he targets. For example, not everyone is on social media, and those candidates whose support base either does not use or are not active on social media will be less likely to become infected.

Finally, although social media is partly seen as an attractive campaign tool due to the perceived low costs associated with it, campaign funds can make a difference. Funds can be spent in almost

every area discussed thus far, such as paying a consulting company to advice or to directly create engaging content. Funds can also be used to hire personnel to manage candidate's social media accounts, establish a constant number of infected who will always spread the content to their networks, and to oversee what topics and type of content are likely to be shared and interacted with at any given moment during the campaign. Considering that these options are available, it can be the case that only candidates with large enough funds are the ones who will benefit the most from social media platforms. If this is the case, then social media will most likely reinforce the status quo rather than break it.

The following section will consider the arguments put forth thus far about the status quo changing nature of social media and formalizes them through the constructed SEI model. These serve as initial hypotheses on what effects social media campaign has had on Finnish elections and the subsequent section will test them with regression analyses based on multiple datasets gathered on the 2015 parliamentary elections.

5.3 The normalization v. equalization debate revisited, again

From the perspective of the SEI model, the normalization v. equalization debate and the previous findings related to it can be interpreted in a new light. For one, the model suggests that there are two main ways in which social media can be used as a political tool. The first way is to rely on the passive exposed part of the population and attempt to achieve as large effective contact rate (value for β) as possible. The situation is familiar to the days of modern campaigning described by Norris (2000, pp. 137–161), where mass communication was used to spread messages in the evening news during prime-time television. Social media can be similarly used, if the candidate has a large pool of exposed individuals who have, for example, liked the official Facebook page of the candidate. In effect social media then functions as the press agency of the candidate and, provided that the exposed have favorable opinions of the candidate, they can be mobilized to vote on election day, spend financial aid to help the candidate or go door to door to campaign for the candidate. Indeed, as we have seen, the number of social media followers have been significant in explaining the vote share that individual candidates have received (Kruikemeier, 2014; Spierings & Jacobs, 2014; Williams & Gulati, 2007; Williams & Gulati, 2008). Furthermore, this effect is amplified if the candidate sends more messages through social media (Williams & Gulati, 2008; Spierings &

Jacobs, 2014). This, according to the SEI model, can be explained by arguing that sending out more messages leads to increased visibility of the candidate and, in general, the more visible the candidate is during elections the more likely it is that he will find some voters who are willing to support him. In addition, increased number of messages also imply more effort from the candidate and can be indicative of active ways to mobilize the support base of the candidate or attempts to influence the perceptions of voters. While this effort does not automatically lead to electoral support, it is an important precondition for the candidate to gain full advantage of his followers and is thus directly linked to electoral gains.

These aspects lead us to expect that, in general, political candidates who have more individuals in the exposed population and send more political messages on social media will perform better during elections. In the context of Facebook, the exposed can be operationalized as the number of people who have liked (or friended) an official Facebook page of a political candidate, which allows us to formulate the first two hypothesizes on how social media activity transfers to electoral gains:

1) The number of Facebook Friends will have an impact on the number of votes a political candidate receives on election day
2) This impact increases as the candidate sends more messages on Facebook

The second way in which social media can be used as a political tool relies more heavily on the infected population and on the probability that the messages send out by the candidate will be shared and interacted with (as in when the parameter p varies from $0 < p \leq 1$). As we have seen, one of the clear ways in which social media differs from traditional media is in the way messages are distributed among populations. Here social media differs particularly from traditional media in that the effective contact rate in social media is much more dependent on the candidate himself. To elaborate further on this, one of the reasons why television commercials, for example, are expensive for candidates to create is because they have to be send out to independent broadcasters who control what content is shown on their communication channels and who make their revenue by selling the attention of their established viewership. In the context of the SEI model, the broadcaster can credibly demonstrate that they can achieve a high effective contact rates with a multitude of different voters on which the candidate can capitalize on, but only if the price is right (Wu, 2016). On social media, on the other hand, that attention and visibility must be acquired by

the candidate himself. Here, the audience does not exist as a given, and the candidate needs to spend effort and resources to create the audience necessary to achieve effective contact rates that are worthwhile for campaining purposes.

However, social media has its advantages. In particular, unlike traditional broadcasters, social media has the ability to create infectious users which can compensate for a relatively small contact rate when compared to professional broadcasters. As the SEI model demonstrates, if a candidate is able to create content that is engaging enough there is a possibility that it will quickly spread far beyond the initial exposed population and reach voters that might not be reachable through other ways. As the number of the infected grows and they further recruit from both the susceptible population and from the already exposed, the reach that a message created on Facebook can accomplish grows exponentially as the effective contact rate is multiplied by each new additional infected. This demonstrated in figure 4 where the rate at which the susceptible population diminishes becomes greater as the probability of creating infected individuals becomes larger. Intuitively, this mechanism is fairly simple. If we think of the infected population as copies of the candidate himself, then each new infected will multiply the original parameter β and leads to exponential growth[11].

It is easy to see why this is useful for the candidates. The infected do not (usually) require payments from spreading the candidate's messages and if the number of infected is sufficiently high, even relatively low values for β can spread political messages widely with fraction of the costs associated with traditional media. Furthermore, this allows political actors to use social media in more efficient ways and should be especially optimal for self-promotion, which is, as we have seen in the literature review, one of the most often used method to use social media. Since self-promotion aims to influence the perceptions that voters have on the candidate, increasing the infected population should be one of the most promising ways to maximize its impact. Focusing solely on the exposed population is limited in that those who have explicitly decided to receive political Facebook messages from the candidate are very likely to already view the candidate favorably. Thus, spending more resources on self-promotion on them is counterproductive and unnecessary. Concentrating efforts to increace the infected population, on the other hand, can reach

[11] It is true that this assumption will most likely not hold in real life and that is likely that political candidates will have, on average, a bigger reach than ordinary citizens have. However, this does not distract from the main point of the model that the infected act as catalysts for the ability of political messages to spread in an exponential manner.

voters who would not normally receive any information about the candidate and who are on the fence who to vote for.

This leads us to expect that not only will increase in the size of the infected population have an impact on electoral outcomes, its effect should as well be greater than increasing the exposed population. In order to test this, the infected population is operationalized as the number of people who have liked, commented or shared any post on Facebook which has been created by a political candidate. With this, we can arrive at three more hypothesizes to be tested:

3) Those candidates who have a greater ability to acquire likes, comments, and shares on Facebook will perform better electorally than those candidates who do worse in this respect
4) This impact will be, on average, greater than the impact associated with the exposed population outlined in hypothesis 1
5) However, as for the exposed population, this impact will be greater as more posts are sent in Facebook

The reasons for the fifth hypothesis are largely the same as they are in the second hypothesis, in that the number of messages indicate effort and is an important precondition for electoral impact, but the infected are also completely dependent on the messages that the candidate sends in Facebook: by the very definition the infected can only exist if there is content on which individuals can interact with.

The infected, however, are likely to differ from one other, and some will have a larger network to which political messages can spread than others have. Thus, the type of infected that the candidate is able to acquire will impact his average effective contact rate. Moreover, the public knowledge about the Facebooks News Feed algorithm leads us to suspect that liking, commenting and sharing have differing impacts on how they help Facebook content to spread within networks. Because of this, we should expect that different types of infected will be more optimal in spreading political content to the overall population than others, and candidates could create more efficient social media strategies by intentionally aiming to create certain types of infections. Unfortunately, a complete study on how this mechanism works is beyond the scope of this thesis and, furthermore, the data necessary to investigate this in detail is not publicly available. However, the research described earlier by Bucher (2012) suggest that the importance Facebook places on different

interactions is based on the level of effort: chatting requires more effort than commenting, which in turn requires more effort than liking, which requires more than just viewing. Thus, the more effort a user has to spend on interaction, the greater the spread of that content should be. In addition, sharing is a function explicitly meant to spread external content to the individual network of the users, which suggest that shares should be the most effective type of infection. From these, I expect that:

6) The impact that the infected population will have on the vote shares of individual candidates will depend on the type of interaction from which the infection is created from
7) This impact will be largest for infections resulted from Facebook shares, followed by comments and will be the smallest for likes

Hypotheses one through seven establish the main mechanisms how social media activity can transform into visibility and, as a result, to opportunities for electoral gains. These can then be used to gain answers to the question does social media have an impact to the Finnish political competition and has it affected the electoral success of those political candidates who have integrated it to their campaign. However, they do not allow us to directly tackle the normalization v. equalization debate and remain silent on who are the political candidates that benefit the most from social media. In this regard, previous empirical findings and theoretical work from other studies and are of great use and give us feasible assumptions to be tested, which can be interpreted through the SEI model.

Firstly, it seems to be the case that one of the most common characteristic of an electoral candidate who frequently benefit from social media and who usually do not belong to the political elite are younger cohorts. As we have seen in the literature review, younger politicians are actively using various social media platforms and are quite plausibly the ones who can gain the most from such platforms. The reasons for this were previously hinted at earlier in the thesis, but in terms of the SEI model this pattern is understandable through two main mechanisms. First, the younger a politician is the more likely it is that he or she has been using social media in their pastime and become familiar with its functions, norms of use and intuition on how to create engaging content. Thus, it is reasonable to assume that younger cohorts will have an increased ability to achieve higher levels of effective contacts, where the messages they send are consumed and not simply

ignored by their network. Similarly, greater use-experience of social media can plausibly extend to a greater ability to create content that other users want to like, share and comment on, which creates infected individuals who spread their messages further. If hypotheses 1 and 3 hold true, then this should lead to young politicians to gain more from social media and provide an advantage to the older incumbent politicians.

This argument can be generalized a bit further. Since the underlying reason why younger politicians are able to gain more from social media is derived from their greater experience, it can be established as an assumption on its own: those actors who have used social media longer than others will have greater knowledge what messages are consumed and interacted with on social media. This, in turn, gives them a better opportunity to receive electoral gains from their social media use. In the Finnish context, it has been established that those candidates who use social media for political purposes during campaigns are not, as the normalization thesis would imply, marginal political actors, but rather are highly educated, incumbents, and come from the major Finnish parties (Strandberg 2013; 2016). These findings favor the normalization thesis and suggest that the incumbents are more experienced in social media campaigning and should thus be better able to gain electoral benefits from them.

These considerations lead to the following hypothesizes:

8) Younger political candidates will gain more electoral benefits from Facebook than older candidates
9) Incumbents will gain more electoral benefits from Facebook than non-incumbent candidates

The final hypothesis set in this thesis is closely related to the most often cited benefit of social media: its nature as a cheap and easy-to-use tool for political communication. Throughout this thesis, we have seen both empirical and theoretical results that cast considerable doubt on how easy it really is to use social media for political gains and how cheap it will be in the end. Social media, truly, requires large amounts of effort, time and expertise to be used effectively and this effort might not be achievable by marginal political actors. As Jacobs and Spierings (2016) have argued, incumbent politicians and large parties with extensive resources can buy expertise and personnel to manage their social media accounts, which can both lead to impact almost every parameter in the SEI model. Money can be used to increase the effective contact rate and the

probability of creating infected individuals by consulting companies that specialize in social media platforms who can advise how to create as engaging content as possible. It can also be used to maintain a steady population of infected who continue to spread political messages in exchange for a fee.

Taking these options into consideration, it is hypothesized that:

10) Financial resources are able to boost the impact that campaigning on Facebook has on electoral outcomes

These hypotheses 1 through 10 are investigated with a dataset that includes information on the Facebook activity of those candidates who actively used Facebook for their 2015 campaign. In combination with information on election outcomes from Statistics Finland, multiple OLS regression analyses are used to estimate and test the validity of the outlined hypotheses. The following section presents the dataset in detail, after which the results of the data analysis is presented.

6. RESEARCH METHODS AND DATA

In order to put the SEI model to the test, this thesis adopts a case study research strategy and uses the 2015 Finnish elections as a testing ground. As outlined by Strandberg (2013, pp. 1335–1336) multiple reasons make Finland an excellent country to study and test social media theories. For one, the Finnish population is using social media tools in high numbers and Finnish political actors have started using them variedly from 2011 onwards (Strandberg, 2016). Hence, in Finland social media is an attractive political tool and quite plausibly has a great influence on the Finnish electorate. Furthermore, since the questions of this thesis are ultimately concerned with causal relationships and aim to make broader generalizations to larger population, the chosen data analysis method is an empirical, statistical analysis conducted on the Finnish population who have run in parliamentary elections. Other elections could have been considered, but since the parliamentary elections are the most impactful in terms of their political outcomes, it is preferable to first concentrate on them. If Facebook has an impact on these elections, it is highly plausible that the platform can influence the Finnish political system and the overall status quo of Finnish political competition.

The main method used to test the hypotheses outlined above relies on multivariate OLS regressions, which can be used to establish whether or not a statistical relationship exists between the candidates social media activity and their final vote counts. More precisely, the OLS regressions are designed to test the assumptions suggested by the SEI model by operationalizing the model's parameters as observable candidate level variables. In this case, the main depended variable is the natural logarithm of the vote count that the candidates received after the elections. Since the raw vote count is a highly skewed and the majority of the candidates received extremely low counts and very few are able to collect staggering numbers in comparison, this creates significant problems for the OLS regressions which are very responsive to the influence of outliers. The log-transformation eases this problem by creating a distribution that resembles the normal distribution more closely. In addition to this technical point, the log-transformation also makes the results more comparable to other elections and to other countries. Since the log-transformed depended variable can be interpreted in terms of percent changes (how many percent the candidates vote count would increase given a certain set of independent variables), it is easier to compare these relative percent changes to other contexts where the number of voters, election districts and the number of total votes are different.

The main dataset used in this study comes from the 2015 Finnish parliamentary elections and includes data on all of the 2,146 candidates who were running at the time. However, since the main interest in this thesis involves only those candidates who have active social media accounts and since Facebook is solely focused on, the main analysis will be conducted on the sub-population of those candidates who had an active Facebook account. Including all candidates would in effect inflate the numbers of passive Facebook users, since any candidate who does not have a Facebook account would automatically receive zero values on any Facebook variables. This would lead to biases in the statistical analyses due to artificially inflated zero values, and thus the non-active candidates are excluded.

Statistics Finland provides information on Finnish election results, which shows that four parties collected the vast majority of the overall votes during the 2015 elections. The four most successful parties accumulated approximately 73.5% of the total vote out of which the Centre Party, National Coalition and the Finns Party eventually formed a coalition government after the elections. This data is supplemented with data from the election compass created by the national broadcasting

company Yle. The election compass is an online tool that helps voters to gain information on the candidates and to make their choice between them. The election compass includes an optional questioner that the candidates can fill out and give voters information on their background and political views. This data is particularly useful because it allows us to have multiple control variables on the candidates that can be used in the upcoming analysis. More information on these variables will be provided further below.

6.1 Main variables

The main independent variables for this study come from Facebook's application programming interface (API), which can be used to access the social media activity and performance of every candidate that had a public Facebook profile on the platform. This data was originally collected by a private data analytics company *99Analytics* and was used for a multidisciplinary research project made in collaboration with the University of Helsinki, Helsinki Institute for Information Technology (HIIT), and the University of Jyväskylä (Nelimarkka, 2015). The dataset includes social media data from the candidates Facebook activity (how many posts they created); how many friends (followers) the candidates had at any given period during the data collection; and on their social media performance in the form of how many Facebook likes, comments and shares each of their posts received.

The parameters of the SEI model are operationalized based on the above mentioned social media variables. Mainly, the overall activity of a candidate is defined as the average number of Facebook posts that the candidate sent per day, and the candidates exposed population is defined as the median number of Facebook friends that the candidate had during the campaign. Since the number of friends that a candidate has varies considerably during the observation period and is particularly influenced by outliers who have extremely large friend counts, the median was chosen over the mean to avoid problems related to highly skewed data. Furthermore, since the SEI model assumes that each message sent by the candidate has a possibility to create infectious individuals who spread the message further, and since it is assumed that infections resulting from likes, comments and shares will have a different impact on the dependent variable, the ability of a candidate to create infections is operationalized as three different variables: the mean number of likes, shares, and comments that a candidate received for every Facebook post sent. It is true that this

operationalization will not completely capture the described mechanism suggested by the SEI model, and it is plausible to assume that many candidates have a particular set of Facebook users who regularly like, comment or share their posts. Thus, each of these three variables can have significant overlap between them and having high average like or share rates do not necessarily indicate the ability to create a high number of individual infections. However, with the available data, this is the best possible proxy to use as the ability of the candidate to create infections, and it is reasonable to assume that these three variables correlate highly with the infectious population each candidate was able to create during the campaign. Moreover, other centrality measures besides the mean could be considered, such as the median, but this can create a significant problem for OLS regressions. This point will be elaborated further below.

As it has been already noted by Spierings and Jacobs (2014), the problem with cross-sectional studies such as this one is that it is difficult to have any definite answers on the issue of causality. Even if the social media variables presented above will prove to have a significant impact on the number of votes that each candidate receives, this relationship is extremely likely to be riddled with omitted variable bias which leads us to observe mere correlations instead of true causations. Candidates who create successful social media campaigns and then perform well electorally should be wary of drawing straightforward causal connections between the two because it is very likely that it is not the whole story. Instead of any causal mechanism, the relationship can be established because, for example, these candidates are well-known due to their pre-existing popularity and this factor contributes positively both to their electoral performance and to their popularity on social media. In short, a simple bivariate analysis omits important alternative explanations which act as lurking variables to which both electoral success and social media success is correlated with. It is not possible to completely rule out this omitted variable problem with the cross-sectional dataset at our current disposal, but data from the Yle election compass provides us with multiple control variables that at least mitigate the problem. Following Spierings and Jacobs (2014), this study aims to control for the most often cited variables that contribute to the electoral success of a candidate and to see whether social media still has an impact even after alternative influences are controlled for. In this regard, this study controls for the overall experience a candidate has, his or her position on the candidate list, the size of the party from which the candidate comes from, and the available campaign resources that the candidate can use for campaigning.

The experience of candidates is important to include into the OLS models at least for the following reasons. First, those candidates who have experience in running on prior elections have an advantage in their recognizability. Previous campaigns allow candidates to build up their presence within the electorate and, even if they are not voted into office, they will most likely be more recognizable to those voters who pay attention during elections than those candidates who are running for the very first time. Secondly, this experience also provides important knowledge on how campaigns are run effectively and what activities are most important to prioritize. Moreover, this experience can accumulate over time and some candidates can even use a familiar campaign team that knows the candidate and has experience from multiple campaigns. This knowledge and continuity can prove to be vital for the eventual success of the candidate and will likely influence how effectively the candidate is able to use social media for campaigning purposes.

As with experience, the position of the candidate on the candidate list can provide significant gains for electoral performance. The logic behind this is fairly simple and is based on the assumption that a significant portion of the voting population will make use of shortcuts to make their voting decision easier to make. Since most voters are faced with a decision problem between multiple different candidates, most voters will not take the time to familiarize themselves with every available option. Instead, voters may first filter out most candidates based on their preferred parties and then observe these candidates present in the list starting with the first one and working their way down until they have found one or more most preferred candidates. If this process is followed by a substantial number of voters, it gives a significant advantage to those who first appear on the list. Additionally, studies have found that for some this process is indeed followed, but the initial filtering is further conducted by only voting for the female candidates (e.g. Wauters, Weekers, & Maddens, 2010). In this case, the individual candidate per se is not as important as it is to vote for a female candidate. These voters often look for the first woman on the party list and simply vote for that candidate.

It is clear that campaign funding greatly impacts electoral performance and that this is the main resource that any candidate has. Those with more money to spend will have more options to conduct campaign activities, advertise themselves in multiple media channels, and to hire professional staff to help them to win votes. This benefit easily extends to social media, where, as has been discussed previously, funds can be spent to create more engaging content, increase the

spread of the messages artificially through paid advertising, and to hire personnel who maintain the candidate's social media profile continuously. As such, it is important to control for the available funding that each candidate has available and to see whether social media can have an impact that is independent of these funds.

Furthermore, the size of the party to which the candidate belongs can give meaningful resources to the candidate in the form of knowledge and campaign infrastructure. Those who come from larger parties benefit from the guidance and help of the party officials and can tap into the collective resources that the party organization has to offer for its individual candidates. Indeed, interviews with the staff of the major Finnish parties suggest that the party organizations can try to systematically help their candidates in social media campaigning and to transfer the cumulative knowledge they have about the platform into their members and candidates. If then the main party is heavily invested in social media campaigning, it is likely to encourage all of its candidates to use it and, if the resources exist, to provide help on its use. (Railo & Ruohonen, 2015.)

These main control variables are gained from two sources: from Statistics Finland and from the election compass constructed by Yle. The overall experience was captured by two dummy variables that tap into the candidates incumbent status and their prior experience in campaigning. The first dummy variable indicates whether or not the candidate was a member of the parliament during the election (1 = yes) and the second dummy indicated if Statistics Finland showed that the candidate was running during the previous 2011 election (1 = yes). The list position effects were likewise captured through two variables. The first is an integer number that indicates what position each candidate had on their party lists, 1 indicating first place. The second variable is a dummy that shows if the candidate is a first woman on the list (1 = yes). Candidates' resources are gained from the election compass which provides the self-reported campaign funds of the candidate. The compass allowed candidates to indicate their funding level on an ordinal scale with the following options: "under 1,000€", "1,000–5,000€", "5,000–10,000€", "10,000–20,000€", "20,000–50,000€", and "over 50,000€". However, since only few observations exist in the extreme values and only few candidates reported spending over 50,000€ (as it can be seen in the tables below), the funding variable was recoded as a variable with three levels where the above values are combined into larger categories. Thus, the funding variable includes the following values: "under

5,000€", "5,000–20,000€", and "over 20,000€". This ensures that each level has enough observations for the statistical analyzes used in the later stages of this study.

Furthermore, the size of the candidate's party is included as an ordinal variable with values "large", "medium", and "small". This categorization is based on the number of seats that the parties held after the previous 2011 elections. Traditionally, Finland has three big parties that have largely remained dominant within the parliament, but the 2011 elections represented a significant change in this long-running trend. After the elections, the Finnish party system could be divided into three major clusters based on the number of seats they held: the Centre Party, National Coalition, Social Democratic Party, and the Finns Party. Together, these four parties held 160 seats out of the 200 available in the parliament (80%). After this major cluster, four parties were able to acquire seats from the remaining 20 places: the Left Alliance, Green League, Swedish People's Party, and the Christian Democrats. Thus, these two clusters constitute as "large" parties and "medium" parties respectively. In addition, the Finnish party system has a number of smaller parties that have remained outside of the parliament. During the 2015 elections, a total of nine parties placed candidates that did not have any seats in the parliament at the time. These parties are assigned to the "small" parties. In addition to these main control variables, the following variables are also included to account for the most common socioeconomic factors that the candidates possess: their gender, education, and age.

6.2 Data and descriptive statistics

Before we continue to the main analysis of this thesis, it worthwhile to first explore the main variables in detail, describe the dataset, and to show the descriptive statistics of the overall population. The main reason for us to do so is associated with the self-reported nature of the Yle election compass data and with technical difficulties regarding the social media data collected by 99Analytics. In both cases, the unique characteristics of these two data sources has led to missing observations and resulted in a dataset that falls short of being a full sample. In order to estimate the external validity of the results arriving from this dataset, then, we need to examine how representative the dataset is of the overall population we desire to study. Unfortunately, we do not have many variables at our current disposal from which we have full information on all of the candidates who were running during the 2015 elections. However, we do know the age, party,

gender, mother tongue, and vote count of each individual candidate from Statistics Finland. In addition, the Yle election compass provides us further background information from the control variables described in the previous chapter.

Table 2 first shows the descriptive statistics on the entire population of the 2,146 candidates who were running in the 2015 election. The table shows that the candidates were mostly men, possessed a degree from high education, self-reportedly spend less than 5,000€ in the campaign, and came mostly from the four largest parties. In particular, this table draws attention to two issues regarding the data that is used in this thesis. The first issue arises from the fairly large portion of NA's in the education and funding variables: approximately 20% out of all candidates did not wish to reveal information either on their education or on their funding. Both of these variables are gained from the Yle election compass, and thus the NA's are caused by candidates who either did not participate in the election compass at all or did not wish to answer these particular questions.

The second issue is apparent in the number of Facebook addresses that the candidates reported in the Yle election compass. The Yle data shows that only approximately 69% of the candidates reported having a Facebook address. This proportion does not match the official estimate of candidates who had a Facebook profile during the elections, which is 73% (Strandberg, 2016). The disparity between these two numbers might be due to the self-reported nature of Yle data. Some candidates might not have revealed their profile addresses in the election compass questionnaire, forgot to put it in or only created a profile after the questionnaire was filled out and did not update the questionnaire later on. At any rate, the difference is only four percent and does not cause any major concerns for the validity of the Yle data. However, the presence of missing cases in the education and funding variables cause a greater concern and could lead to biases in the regression analyses if it is not caused by a random process and is related to value of the variable from which the data is missing (for example, when a candidate is less likely to report his funding if he has large funds to report). However, since the current project is mainly concerned with those candidates who use social media, the missing observations are only a problem if they extend to the sub-section of the population who are active social media users.

***Table 2*.**
Descriptive statistics on the background variables of the 2015 parliamentary election candidates

Variable		n	Share of all candidates
Gender	F	847	39.47%
	M	1299	60.53%
Education	Primary	48	2.24%
	Vocational	350	16.31%
	Upper secondary	187	8.71%
	Higher education	1018	47.44%
	Other	81	3.77%
	NA's	462	21.53%
Funding	Under 1 000 €	450	20.97%
	1 000-5 000 €	496	23.11%
	5 000-10 000 €	251	11.70%
	10 000-20 000 €	245	11.42%
	20 000-50 000 €	245	11.42%
	Over 50 000 €	16	0.75%
	NA's	443	20.64%
Party Size	Large	861	40.12%
	Medium	721	33.60%
	Small	533	24.84%
	No party	31	1.44%
List Position	1	180	8.39%
	2	156	7.27%
	3	148	6.90%
	4	137	6.38%
	5	129	6.01%
	Other	615	65.05%
Dummy Variables	Personal website	1491	69.48%
	Facebook profile	1475	68.73%
	Twitter profile	1031	48.04%
	Campaign experience	743	34.62%
	First woman on the list	104	4.85%
	Incumbent	168	7.83%

Indeed, the problem seems to be somewhat alleviated when only the candidates from which we have social media data are considered. From this subpopulation, only 9.71% did not report their planned funding expenditures, and 10.7% did not wish to disclose their educational backgrounds. In order to explore possible connections with other variables from which we have full data, the education and funding variables were recoded as dummy variables showing who answered these questions, and then Chi-Square test was conducted on each of the other control variables outlined

in the previous chapter. The associated p-values were all non-significant, which suggests that the remaining missinges is not related to any of the other variables we have information on. It is quite possible that some other factors influence candidates not to reveal this information, but due to lack of data, it is not possible to investigate further connections. We are left to use these variables with due caution and to take solace from the fact that only a relatively few values are missing.

Further issues that need to be addressed come from the social media data gathered by 99Analytics. Since this thesis is mainly interested in the effects of campaigning, the time periods falling outside the campaigning season are left out of the analysis. The final observation period, then, covers approximately four months starting from 1st of January and ending at 18th of April, the final day before the elections. From this period, 99Analytics gathered as many Facebook addresses from those political candidates who used it in their campaign as possible. However, some candidates Facebook addresses were missed or could not be retrieved during the data collection due to technical difficulties. Because of this, the dataset is not complete and includes a total of 1,128 candidates, which amounts to 52.6% of the entire candidate population. Given that the official estimate of candidates who had a Facebook profile during the elections is 73% (Strandberg, 2016) this suggests that 20.4% of the desired population is missing. However, since these missing observations are mostly due technical difficulties, it is likely that the missing observations are missing at random and will thus not skew the regression results. Nevertheless, these missing cases should be examined more closely and to see how well the used dataset represents the population of those candidates who had Facebook profiles during their campaign.

The confidence on the available dataset is increased by the fact that the population from which we have full data on seems to resemble those candidates who reported having a Facebook account quite well. Table 3 shows the distributions of the main background variables from those candidates who reported having a Facebook account in the Yle election compass (the main population) and those from which we have full data, i.e. those candidates from which we have social media data and who reported their funding and education levels in the Yle election compass, (the sub-population). None of the observed differences between these proportions and averages goes over by a factor of two and the ratio of these values keep fairly close to 1.0 in most cases.

Table 3.
Comparison between the main population and the sub-population

Variable		Main population	Sub-population	Sub/Main
Gender	F	42.1%	47.2%	1.12
	M	57.9%	52.8%	0.91
Age	mean	43.0	40.6	0.94
	median	42.0	39.0	0.93
Vote Count	mean	1383.3	1861.4	1.35
	median	478.5	860.5	1.80
Language	Finnish	92.3%	92.4%	1.00
	Swedish	6.3%	6.9%	1.09
	Other	1.4%	0.7%	0.52
Education	Primary	2.0%	1.1%	0.53
	Vocational	18.2%	16.2%	0.89
	Upper secondary	11.5%	12.2%	1.06
	Higher education	64.0%	66.9%	1.05
	Other	4.3%	3.6%	0.83
Incumbent	No	90.6%	91.8%	1.01
	Yes	9.4%	8.2%	0.87
Party Size	Large	54.7%	47.0%	0.86
	Medium	28.6%	44.0%	1.54
	Small	16.7%	9.0%	0.54
Funding	Under 1 000 €	21.2%	15.4%	0.73
	1 000-5 000 €	30.0%	31.8%	1.06
	5 000-10 000 €	16.0%	17.4%	1.09
	10 000-20 000 €	15.5%	15.6%	1.01
	20 000-50 000 €	16.2%	18.4%	1.14
	Over 50 000 €	1.2%	1.4%	1.21

However, the clearest difference between these populations comes from the size of the party; candidates from the smaller parties are clearly underrepresented in the sub-population. There are almost less than half fewer candidates from smaller parties in the sub-population than there is in the main population and, correspondingly, candidates coming from the medium-sized parties are over-represented in the sub-population. This causes the clearest threat to the external validity of this study and it should be kept in mind that the lack of small party candidates can bias the upcoming results. However, since further data collection is not an option, a simple listwise deletion was performed to handle the missing value problems. This results in a final dataset that includes

844 observations, which remains as a sizable sample size and can be used to estimate OLS coefficients.

A set of selected background and social media variables of these 844 observations are summarized in table 4, which shows descriptive statistics on the age, final vote counts and Facebook variables of the candidates included in the final dataset. In comparison to table 2, which shows the total population of the running candidates, those with active Facebook profiles seem to be, on average, two years younger with a mean age of approximately 41. In addition of being younger, candidates with active Facebook accounts seem to perform better electorally as well since their average number of votes received is almost 400 higher than in the overall population. In terms of their Facebook activity and performance, most candidates sent around 100 messages during the campaign from 1st of January until the final day before elections in 18th of April. The intensity of these messages was fairly low, however, and most candidates sent only one message per day. In contrast, most Swedish and Norwegian politicians exhibit the same level of activity during non-campaigning periods and tend to send one tweet or post per day (Larsson & Kalsnes, 2014). Since in most countries social media use is the most passive during the non-campaigning period, it is likely that one tweet or post per day is a rather inactive messaging pattern within the Nordic context. Interestingly, the descriptive statistics in table 4 seem to suggest that most candidates did not deviate much from this overall messaging pattern. Both the range and standard deviation in the total number of posts and average number of posts send per day remain fairly low, and it is common that most candidates did not deviate from the average number of total posts more than by 93 posts. In addition, only one candidate sent on average more than seven posts per day whereas most sent only one.

Table 4. Descriptive statistics of selected social media and background variables

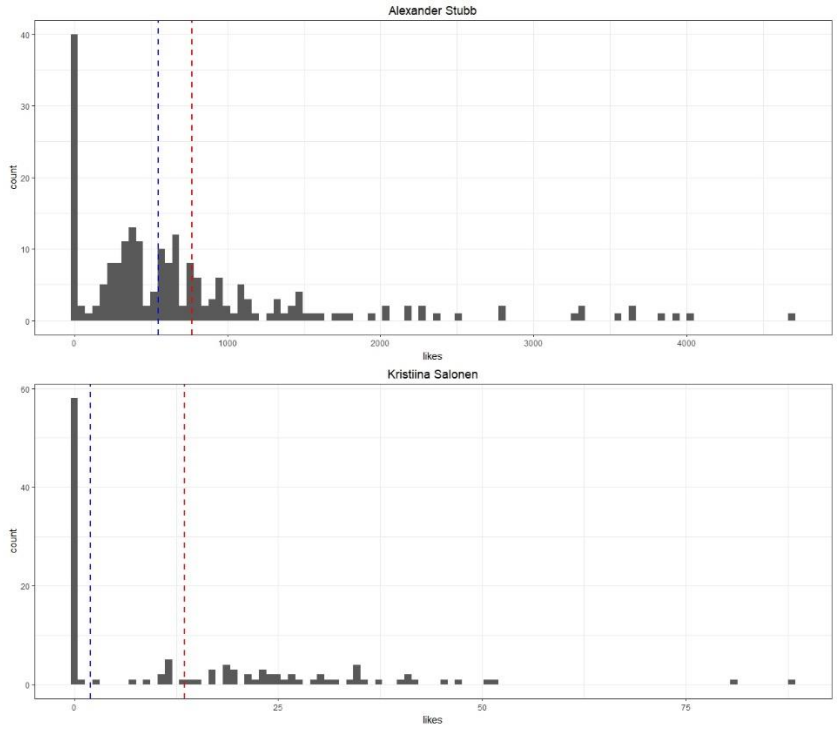
	n	mean	median	range	sd	skew
Age	844	40.59	39.00	54.00	11.41	0.32
Votes	844	1861.43	860.50	30743.00	2651.69	3.69
Total posts	844	113.90	90.00	822.00	92.69	1.99
Average post per day	844	1.04	0.83	7.54	0.85	1.99
Total likes	844	1970.61	735.00	70695.00	4685.79	7.14
Total comments	844	142.07	52.00	4179.00	328.10	6.43
Total shares	844	108.76	38.00	4279.00	282.36	8.02
Friends	844	509.74	270.00	10324.00	916.12	6.30

However, while the level of activity was largely consistent among all candidates, the same cannot be said about their performance. The number of likes, comments, and shares each candidate received during the observation period deviated largely from one other, and, for example, where one candidate could gather more than 70,000 likes during the campaigning period the majority could gather only about 735 likes. Furthermore, the figures in table 4 suggest that likes are the most common response to the posts send out by the candidates. The median of the total number of comment and shares are fairly close to one other, 52 and 38 respectively, whereas the median number of likes is over 14 times greater than either of the two. It seems that commenting and sharing takes more effort and consideration from voters who in response distribute them more carefully to the candidates. In the context of the SEI model, this suggests that the infectious population is divided into at least two sub-populations in Facebook, and it is likely that those who are turned into the infectious population by creating likes are different in terms of their social media behavior. Furthermore, if hypothesis 7 holds true and comments and shares can spread content further than likes can, then these smaller numbers might be misleading. It remains to be seen whether a greater impact can compensate for the smaller numbers of comments and shares. Similarly, the number of friends that the candidates had deviated greatly from one other. Most candidates had just 270 friends, or exposed in terms of the SEI model, but some had significantly more than that, which is apparent in the skewness of the friend distribution.

As mentioned previously, in order to operationalize the ability for individual candidates to increase their infectious population through their Facebook activity, the average number of likes, comments and shares are used to represent the average performance for each individual candidate. As it can be seen from table 4, however, each of the social media variables are highly skewed and especially the number of shares that the candidates were able to acquire for their messages has a long range and an extremely long right tail, suggesting very few large numbers of shares for few candidates and a multitude of candidates that only received a couple of shares. This skewed distribution can be problematic for further analysis because it can distort averages due to a couple of extremely high values. This is an especially acute problem when calculating the mean, and this gives us reason to consider alternative centrality measures that are not as sensitive to outliers. One obvious alternative would be to use the median over the mean since it does not place as much weight to few high values as the mean does. However, in this case, the median is not without its problems. The main issue here is that the median contains a lot less information and undervalues candidates

who are able to create few successful posts in terms of likes, comments or shares. To see why this is, consider two plots displayed in figure 5, which show the like distributions of two candidates, and their mean and median values. The first plot shows the like distribution from Alexander Stubb who was by far the most popular candidate in Facebook during the 2015 elections and accumulated a staggering 164,106 likes during the observation period, over two times more than the next best performing candidate who collected a total of 70,697 likes. The figure shows the mean value of Stubb's likes in red and the median value in blue.

Figure 5. Like distribution of Alexander Stubb and Kristiina Salonen



As the plot shows, Stubb indeed has a few outlier values which gathered considerably more likes than most of Stubb's posts, but it is worth noting that at the same time by far the most common value is a round zero. Stubb sent a total of 214 posts on Facebook during the observation period, but 40 of them (18.7%) did not perform very well and did not receive a single like from other Facebook users. As such, the high values at the right end of the tail pull the mean towards them, but at the same time the zero values pull it back towards the center. The median is considerably smaller than the mean (547 and 766.85 respectively), but this is mostly due to the zero values

which have a high influence on the median. The influence of zero values might not represent a considerable problem if Stubb was a representative example of most candidates Facebook performance. However, most candidates could not muster the same kind of support and follower base as Stubb could.

Kristiina Salonen represents a more typical case, whose Facebook metrics are fairly close to the overall average. She posted a total of 118 messages on Facebook, approximately one per day, and received 1,597 likes for her efforts. As can be seen from her graph, the mean and median number of her likes give us quite different information. While her median of two likes would suggest that she did not perform very well on Facebook, the distribution on the graph shows that this low number is mainly due to the high number of zero values.

As Salonen demonstrates, social media performance, at least in Facebook, is a hit and miss endeavor, where most of the posts do not resonate with voters, but a few of them are able to make a considerable impact after many trials and errors. The median obscures this fact in many cases and can even completely hide the few successes that candidates have had on Facebook and result in zero values with zero information on their performance. This point is even more crucial for the other two metrics. Comments and shares are extremely rare, and most posts do not generate them at all. The median, then, would hide their impact in a vast majority of cases. Indeed, the magnitude of this problem is quite severe for comments and shares since the median would turn positive mean values into zero median values in almost 90% of all cases! Based on this reasoning, the mean value is preferred over the median and will be used as an operationalization for the candidates ability to create infectious users.

7. EMPIRICAL INVESTIGATION: FINDINGS AND DISCUSSION

Now that all main independent variables are justified and described, it is time to turn to the regression analyses and start explicitly testing the hypotheses outlined above. This testing is done through multiple OLS regressions, which test whether there is an observable connection between the social media variables used in this study and the log-transformed vote counts of the individual candidates included in the final dataset of 844 observations. The structure of this chapter follows closely hypothesizes 1 through 10 and aims to provide answers to each of them in turn. The first

set of regressions are designed to test the overall reliability of the used variables and to see whether or not our main social media variables are able to explain the log-transformed vote counts even after the control variables are included as explanatory variables. Then, the stability of the model is evaluated, and its reliability confirmed in that it follows the basic assumptions behind OLS regressions. Once we have established confidence in the model, the hypothesizes are tested explicitly via a multiple sets of statistical tests on the regression coefficients regarding our main social media variables.

To start this investigation, table 5 shows the unstandardized coefficient estimates, standard errors[12], and the associated t-Statistics for the regressors of the first two models which include the control variables and the main social media variables. As can be seen from the first Control Model, even before any of the social media variables are included to the regression, the control variables already explain a sizable portion of the variance associated with the log-transformed vote counts. The Control Model has both normal and adjusted R-squared values at around 66%, indicating that the included nine control variables are already able to explain well over half of the variance associated with the vote distribution.

Out of all of the control variables included in the Control Model, by far the most impactful variable is the level of campaign funding the candidate reported spending. Since the main dependent variable is in logs, the estimated coefficients indicate a percent change in votes for a unit increase in the specified independent variable. In the case of our funding variable, in contrast to those candidates who spent less than 5,000€, those who spend around 5,000 to 20,000€ could expect to see almost 90% boost in the number of votes they received and those who spent more than 20,000€ could increase their vote count by 157%. Similarly, both the incumbent status and campaigning experience provided a helpful edge over those who were non-incumbents and did not have experience from the previous 2011 campaign. Incumbent status provided a significant 104% boost and experience around 30% boost.

[12] Almost of the models reported in this thesis displayed a degree of heteroscedasticity, which can influence the standard errors of the coefficients and thus influence their associated p-values. However, each regression was run twice with the normal standard errors reported by the stats v3.4.3 package in the R programming language and with heteroscedasticity-consistent White standard errors reported by the lmtest package (Zeileis, 2017). The standard errors did not deviate to a large degree in any of the regression models and the associated p-values always remained the same regardless of which of the two standard errors were used. Thus, the reported standard errors are calculated with the usual formula and do not adjust for heteroscedasticity.

Table 5. First set of regression models

	Dependent variable: log(*votes*)	Control Model			SM Model		
		Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic
SM variables	Average Post Per Day				0.128***	0.030	4.258
	Friends				0.000	0.000	1.539
	Average Likes				0.011***	0.003	3.983
	Average Comments				-0.020	0.033	-0.600
	Average Shares				0.029	0.017	1.640
Control Variables	Age	-0.010***	0.002	-4.016	-0.006*	0.002	-2.530
	Male	-0.142**	0.055	-2.582	-0.123*	0.051	-2.407
	List Position	-0.007	0.004	-1.910	-0.009*	0.004	-2.549
	First Female	0.100	0.133	0.751	0.118	0.122	0.964
	Incumbent	1.040***	0.110	9.420	0.766***	0.107	7.170
	Campaign Experience	0.296***	0.061	4.891	0.268***	0.056	4.778
	Education: Primary	0.241	0.288	0.839	0.206	0.265	0.778
	Education: Vocational	-0.060	0.151	-0.400	-0.079	0.139	-0.565
	Education: Upper Secondary	-0.073	0.161	-0.456	-0.163	0.148	-1.104
	Education: Higher Education	0.194	0.141	1.368	0.110	0.130	0.845
	Funding: 5,000-20,000 €	0.893***	0.066	13.497	0.717***	0.063	11.437
	Funding: Over 20,000 €	1.567***	0.086	18.143	1.273***	0.084	15.201
	Party Size: Medium	1.036***	0.096	10.739	0.941***	0.090	10.429
	Party Size: Large	1.439***	0.105	13.714	1.419***	0.098	14.477
	(Intercept)	5.262***	0.194	27.150	7.239***	0.217	33.295
	n	844			844		
	R-squared	0.667			0.7195		
	Adjusted R-squared	0.661			0.7133		
	F-statistic	123.4***			116***		

*p < .05; **p < .01; ***p < .001.

Note: The models include three dummy variables: First Female (First Female in candidate list = 1), Incumbent (Candidate served as an MP during the election = 1), and Campaign Experience (candidate was running in the previous 2011 elections = 1). Education, Funding and Party Size are categorical variables, where reference categories are:
Education = *Other*; Funding = *less than 5,000 €*; Party Size = *Small*.

Interestingly, male candidates seemed to be at a somewhat disadvantage during the 2015 election and in general received 14% fewer votes than women did. Here, it is important to recall that the population on which these regressions are based on include only those candidates who have active Facebook accounts. Indeed, there were considerably more women in this population than there was overall among all candidates during the 2015 elections. it could be the case that this difference is reflected in the regression estimates and perhaps filtering out all of the non-Facebook users

corresponds to other background variables which select highly performing women. Furthermore, there seems to be no connection between electoral performance and education; those with lower educational degrees performed no worse than those with degrees from high education. The list position effects similarly seem to have no impact.

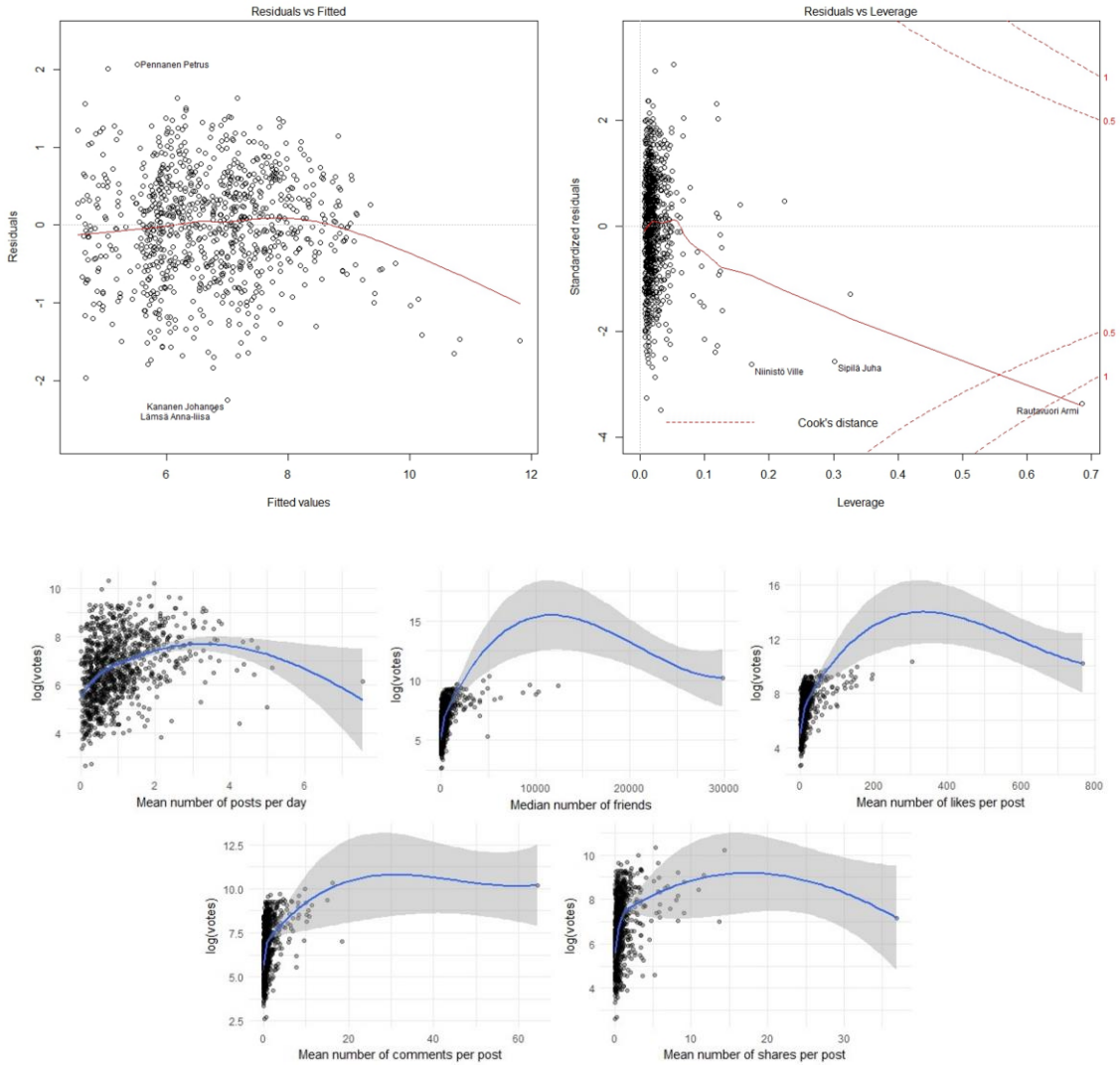
When we turn our attention the second model which includes our main social media variables, we can see some notable patterns. First, the five social media variables are able to increase the performance of the model. The adjusted R-squared increases by additional 5.21% which is a statistically significant increase at $p < .001$ level according to F-test which compared these two models. Furthermore, all of the control variables retain their significance levels except for age and gender variables. Both are reduced to be significant only at the very lowest $p < .05$ level and the magnitude of their coefficients are reduced as well. Some further changes appear within the list position effects since list position turns into a significant variable, although at a very low level. It remains to be seen how robust this variable is during alternative regression models.

However, the more important findings are not seen in the estimate changes within the control variables but are present within our main variables of interest. Even though our five social media variables seem to be able to improve the model, only two of these variables seem to have any impact on the vote distributions: the average number of posts that a candidate sent per day and the mean of likes their posts received. However, a closer inspection of the model residuals reveals that the SM model is riddled with problems that cast doubt upon the reliability of this simplified model.

The main problem is that multiple observations with unusually high vote counts are performing much worse than linear OLS model above would suggest. Figure 6 shows the Residual vs Fitted plot in the upper right corner, which shows the residuals from the social media model and contrasts them with the fitted values from the same model. What is clearly apparent from this plot is that a number of observations at the right end are distributed in a pattern that is not completely random, which casts doubt upon how appropriate the basic assumptions behind OLS regression are. Furthermore, it seems that the model is very prone to the influence of outliers and three candidates (Niinistö, Sipilä, and Rautavuori) have a significant leverage over the model and greatly influence the coefficient estimates, which can be seen in the Residuals vs Leverage plot at the upper left corner of figure 6. The plot shows the associated Cook's distance values for each observation and

indicates cases that are beyond the suggested threshold (1 in this case) and highlights that especially candidate Rautavuori has a large impact on the model.

Figure 6. Diagnostic and scatterplots of the SM Model



All of this leads to the conclusion that this model is poorly specified and that the coefficient estimates cannot be fully trusted. Moreover, the fact that the model is highly influenced by few extreme values and given that a number of observations are performing worse than a linear relationship would suggest, gives us reason to suspect that the relationship between the included social media variables and log-transformed votes is non-linear and that their effect will eventually reduce as their values increase. Indeed, looking at the plots presented at the bottom corner of figure

6 shows the relationship between logged votes and the five social media variables is highly non-linear and closely resembles a concave function. Thus, the reason behind the non-significant variables in the previous social media model is likely to be due to functional misspecification and needs to account this non-linear relationship. Multiple different options exist to accomplish this, but this study opts to continue with logarithmic transformations[13]. In effect, transforming all of the social media variables shifts their distribution to the right and creates a distribution that follows a normal distribution more closely. This reduces the impact that the few outlier values have to the model and transforms the model specification into a log-linear model where the linear assumption of the OLS should hold more reasonably.

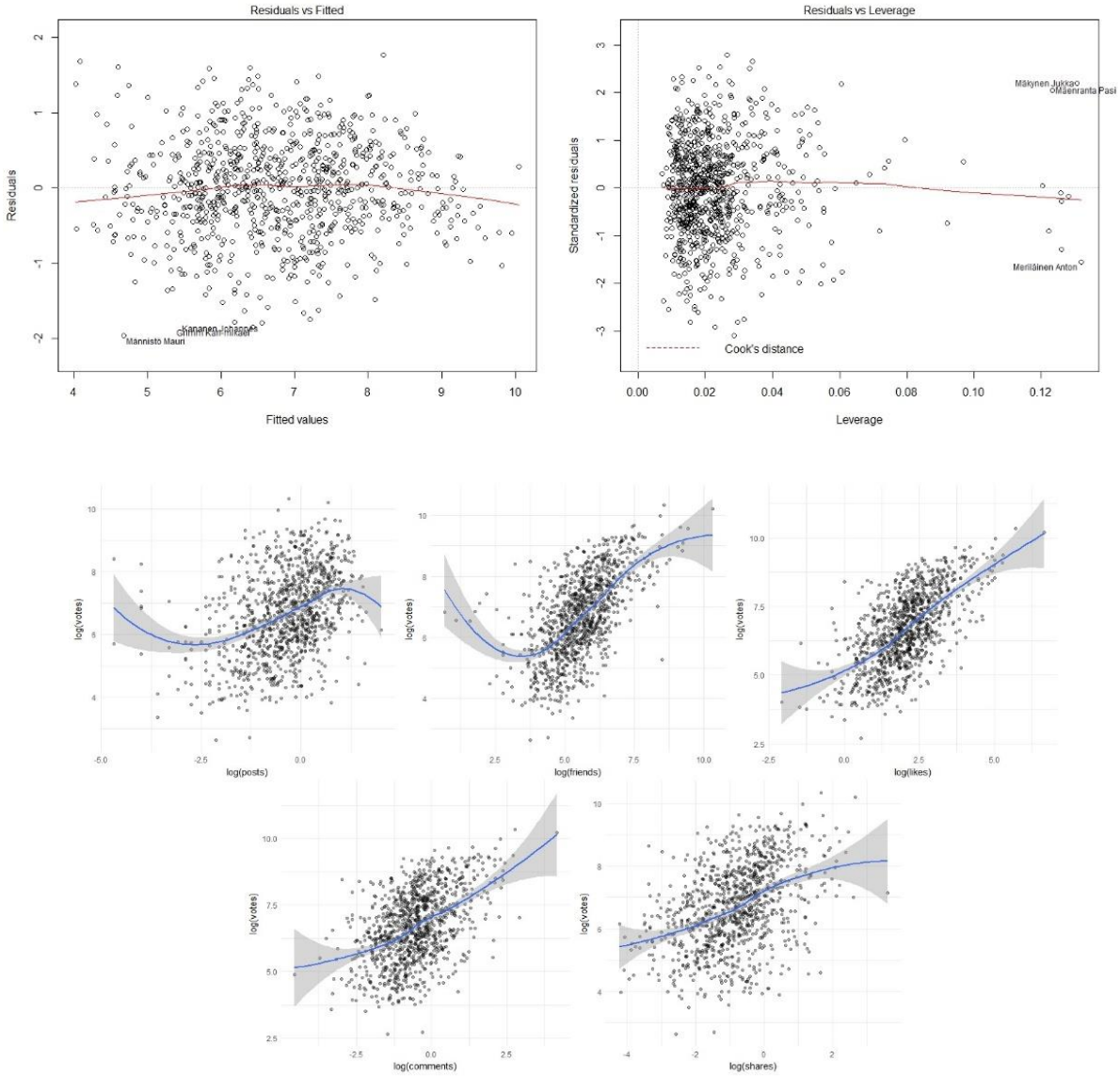
However, the log-transformation of the main independent variables brings some interpretational difficulties. Because the independent variables no longer represent their original values, the interpretation of the regression coefficients changes from the first SM model presented above. In short, regressing a log-transformed independent variable to a log-transformed dependent variable can be interpreted as how many percentages the dependent variable changes for one percent change in the independent variable. Or, using econometric terminology, this represents the degree of elasticity between these two variables. This added complexity to the model is explored in more detail below, but, in the interest of the hypotheses to be tested with these models, for now it is sufficient to concentrate on the direction of the relationship between our variables and on their compared magnitudes.

Indeed, as the figure 7 shows, the log-transformation is able to improve the stability of the model and helps both with the linear assumption and with reducing the influence of outliers. The residuals of this log-linear model are more randomly distributed than the simple linear model, outliers do not have as much influence on the coefficients, and the relationship looks, more or less, linear for likes, comments, and shares. However, posting intensity and the number of friends the candidates had on average are not as clear-cut as the rest of the social media variables, as is shown by the scatter plots in the bottom corner of figure 7. The few observations at the left end of these two

[13] Other solutions do exist. For one, adding a number of polynomial variables in addition to the main social media variables, such as $Friends^2$, could account for the existing non-linearity. However, this would result in multiple additional regressors which can cause multicollinearity problems and decrease the accuracy of the model or even prevent us from making inferences which coefficients are significantly larger from zero. Using log-transformation on the independent variables, however, allows us to keep using the same number of regressors and to avoid the aforementioned problems.

plots suggest that some candidates were able to acquire fairly large number of votes, even though they posted relatively infrequently on Facebook or did not possess a large following of Facebook friends.

Figure 7. Diagnostic and scatterplots of log-linear SM Model



Furthermore, the plots that include the post and friend variables suggest that a dual relationship exists within these two variables: at first, the relationship is negative and turns upward only after a certain threshold of Facebook activity or friend following is achieved. Thus, while only few observations deviate from the overall positive pattern in the post and friend variables, it is worthwhile to explicitly test the hypothesis that even the log-linear model does not completely

meet the linearity assumption and has a negative relationship first and turns into a positive one only later on. To test this, table 6 shows the regression estimates of two log-linear social media models where the first assumes that all of the five social media variables impact votes linearly and the second model tests the assumption that the log(posts) and log(friends) variables still have a non-linear impact.

In comparison with the first social media model from table 5, the first noticeable differences in the log-linear models presented at table 6 is that where the original model improves the explained variance by 5.21% the log-transformed model is able to go further and improve upon the control-only model by 8.95% (statistically significant at $p < .001$). Furthermore, three out of the five social media variables are now turned into statistically significant at a very high .001 level. As before, it seems that being active on Facebook does pay off. Increasing the number of posts a candidate sends per day, say, by 10% has an estimated increase of 1.5% for received votes. Not a huge impact, but still beneficial. In addition, this relationship seems does not seem to have additional complexity, since the added quadratic term in the second logged SM model is not significant.

However, considering previous findings from the Netherlands and from the United States, it is interesting that the amount of Facebook friends does not seem to have any independent impact on the vote shares of the candidates in any of the models presented thus far. This variable is able to reach a significant impact only after adding a quadratic term in the second logged SM model, but it does so in the very lowest .05 level and still the impact is only significant for the quadratic term. Thus, the impact that the amount of Facebook following the candidates have seems to be spurious and it is questionable how far-reaching claims can be made on the significance of this variable based on this data. Nevertheless, the results strongly indicate that the importance of Friends on Facebook is much lower for electoral performance in Finland than the importance of Followers is on Twitter in countries such as the Netherlands and the United States. Whether this is due to country-specific differences or is more related to the underlying differences between the user base and interface of Facebook and Twitter remains to be seen and cannot be answered with the data at hand. Indeed, this finding reveals a potentially worthwhile avenue for further research and gives us further reason to suspect that there are systematic differences between Facebook and Twitter in the electoral context that should be investigated further with dual data from both platforms.

Table 6. Second set of regression models

	Dependent variable: log(*votes*)	Control Model			1st Logged SM Model			2nd Logged SM Model		
		Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic
SM Variables	log(*Posts*)				0.151***	0.031	4.821	0.177***	0.038	4.683
	log^2(*Posts*)							0.029	0.021	1.405
	log(*Friends*)				0.064	0.038	1.689	-0.242	0.128	-1.889
	log^2(*Friends*)							0.029*	0.012	2.468
	log(*Likes*)				0.307***	0.053	5.842	0.297***	0.053	5.612
	log(*Comments*)				0.000	0.038	-0.011	-0.012	0.039	-0.316
	log(*Shares*)				0.099***	0.029	3.411	0.099***	0.029	3.414
Control Variables	Age	-0.010***	0.002	-4.016	0.001	0.002	0.384	0.001	0.002	0.421
	Male	-0.142**	0.055	-2.582	-0.118*	0.049	-2.424	-0.113*	0.048	-2.341
	List Position	-0.007	0.004	-1.910	-0.010**	0.003	-3.089	-0.010**	0.003	-3.098
	First Female	0.100	0.133	0.751	0.000	0.116	0.857	0.111	0.116	0.956
	Incumbent	1.040***	0.110	9.420	0.731***	0.098	7.442	0.704***	0.100	7.071
	Campaign Experience	0.296***	0.061	4.891	0.297***	0.055	5.440	0.283***	0.055	5.181
	Education: Primary	0.241	0.288	0.839	0.291	0.247	1.179	0.281	0.246	1.142
	Education: Vocational	-0.060	0.151	-0.400	-0.009	0.132	-0.068	-0.008	0.131	-0.061
	Education: Upper Secondary	-0.073	0.161	-0.456	-0.124	0.140	-0.882	-0.130	0.134	-0.928
	Education: Higher Education	0.194	0.141	1.368	0.153	0.123	1.241	0.145	0.123	1.182
	Funding: 5,000-20,000 €	0.893***	0.066	13.497	0.518***	0.061	8.490	0.513***	0.061	8.442
	Funding: Over 20,000 €	1.567***	0.086	18.143	0.921***	0.083	11.110	0.908***	0.083	10.969
	Party Size: Medium	1.036***	0.096	10.739	0.717***	0.095	7.587	0.747***	0.095	7.893
	Party Size: Large	1.439***	0.105	13.714	1.212***	0.100	12.115	1.247***	0.100	12.429
	(Intercept)	7.740***	0.231	33.452	4.531***	0.258	17.537	5.287***	0.404	13.084
	n	844			844			844		
	R-squared	0.667			0.7562			0.7588		
	Adjusted R-squared	0.661			0.7505			0.7526		
	F-statistic	123.4***			134.5***			123.1***		

*p < .05; **p < .01; ***p < .001.

Note: The models include three dummy variables: First Female (First Female in candidate list = 1), Incumbent (Candidate served as an MP during the election = 1), and Campaign Experience (candidate was running in the previous 2011 elections = 1). Education, Funding and Party Size are categorical variables, where reference categories are: Education = *Other*; Funding = *less than 5,000 €;* Party Size = *Small.*

Taken as a whole, the results related to the Friend variable gives us a reason to doubt the first hypothesis formulated in this thesis and it seems that Facebook followers have a minimal impact at best and a completely zero effect at worst.

However, the Friend variable could still have a meaningful impact on votes through interaction. As it was the case in the study conducted by Spierings and Jacobs (2014), adding an interaction term to the regression model can test the hypothesis that the relationship between the number followers the candidate has on Facebook on the log-transformed vote counts is different for those candidates who are more active in creating posts or tweets. In short, those who are more active on social media can mobilize their followers and gain electoral gains from this interaction. In this case, this amounts to arguing that: $\log(\text{votes}) = \beta_0 + \beta_1 \log(\text{posts}) + \beta_2 \log(\text{friends}) + \beta_3(\log(\text{posts}) \times \log(\text{friends}))$. Here it is assumed that the impact that the Friend and Post variables have are dependent on each other and can thus be used to test hypothesis 2. However, while the regression output is omitted here (the specific results can be found in the appendix), the interaction term between Facebook posting activity and number of Facebook friends is not found be statistically different from zero. This result suggests that, unlike in many of the previous studies, even increasing the effort a candidate spends on Facebook does not seem to be able to mobilize followers to vote on election day. Hypothesis 1 and 2 are thus not supported.

The same, however, cannot be said on hypothesizes 3 and 4. As it can be seen from table 6, the ability to create infectious users through Facebook likes and shares have a clear impact on the vote distribution. This is most notable in the ability to generate likes for created posts: a 10% increase in how many likes a candidate is able to acquire for each of the posts created is associated with 3.07% increase in votes, over two times more than what could be expected from solely creating more posts per day. The first logged SM model also reveals a significant impact on the average number of times each post is shared. However, the impact is the lowest of the three significant variables, and the same 10% increase in shares results only to approximately 1% increase in votes. These results support hypotheses 3 and 4, and it indeed seems to be the case that those candidates who are able to acquire more users to the infectious population have a better chance to gain votes on Facebook. Moreover, this impact clearly outperforms the impact that the number of friends (or the number of exposed) the candidate had during the campaign. As for hypothesis 5, the results are largely the same as with the Friend variable. Adding an interaction term between how many

times a candidate sends posts per day and how many likes or shares a candidate received on average is not statistically significant (estimates in the appendix). Increased activity is then not able to boost the influence that likes and shares have, which leads to not give support to hypothesis 5.

Perhaps the most intriguing result thus far comes when the relative impact of likes, comments and shares are compared. As hypothesis 6 assumes, these three interactions do indeed behave quite differently from one other and have a vastly different impact on the vote distribution of the candidates. However, the arrived results are almost completely opposite to the ones predicted by hypothesis 7. The effect of creating likes does not have the smallest impact. It is by far the most influential social media variable within our models and it is certainly not dwarfed by the influence of creating comments. Indeed, comments do not seem to have any impact whatsoever. The beta coefficient for the comment variable stays non-significant for both models and cannot be established to differ from zero. Thus, receiving a consistently high number of comments for each post created does not seem to be related to electoral performance. This finding casts some doubt upon arguments which tend to emphasize the importance of interaction and communication during campaigning on social media. Rather than creating conversation, a more optimal path to gain votes on Facebook is to acquire as much attention and reactions in the form of Facebook likes and shares as possible. From this, hypothesis 6 is supported while hypothesis 7 is partially supported.

In the context of the SEI model, this makes intuitive sense in that both likes and shares are the clearest factors that determine the spread that a post on Facebook is able to achieve. While comments can be assumed to require more effort form users, the evidence that they impact the reach of Facebook posts is far from conclusive and is based solely on the work from Bucher (2012), which is already several years old – a lifetime in the lifecycle of software algorithms. The results found here, then, can be due to the relatively weak weight that the Facebook's News Feed algorithm has placed on them during these elections or from a number of other changes and tweaks Facebook has done with the News Feed.

Other plausible explanations, of course, also exist. For one, whereas likes contain a relatively low amount of information on the intentions that the user had when generating the like reaction, comments can be extremely noisy variables and indicate any number of different factors on the campaign performance that a candidate was able to maintain on Facebook. In short, it is difficult

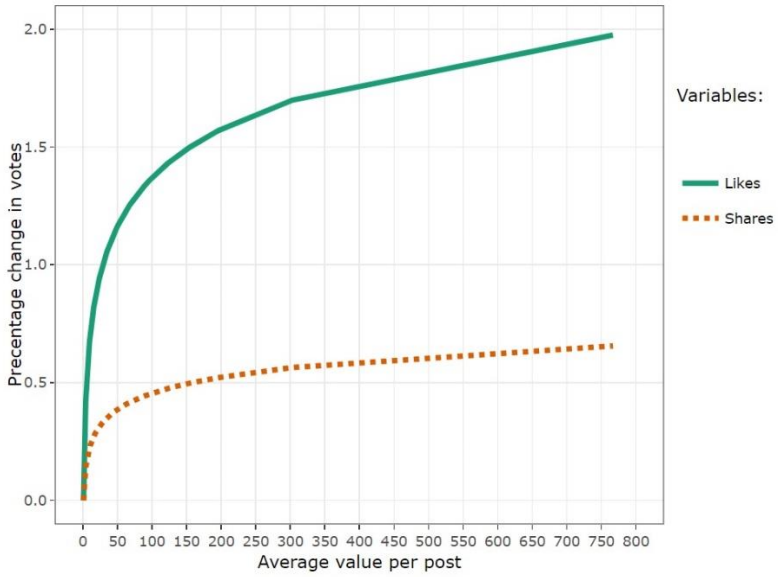
to tell what exactly a high number of comments per post actually reveals about the candidate or about the campaign. It could mean that the candidate has been able to tap into very relevant issues that resonate with multiple different voters and that the campaign successfully facilitated a space to discuss these issues on Facebook. On the more negative end on the realm of possibilities, a high level of commenting activity can as well be indicative of increased outrage felt towards the views of the candidate or of negative feedback that the candidate has gotten from scandals during the campaign trail. We cannot tell which is which from just raw comment counts. This can result in multiple different effects that average each other out and leave us the impression that comments do not matter electorally. As is often the case, more data is needed to explore this matter in more detail and the content of the comments themselves need to be investigated in order to distinguish positively contributing comments from those that most likely indicate harm for the electoral prospects of the candidate.

Since likes and shares are the most clearly influential social media factors contributing to the vote distribution, it is worthwhile to investigate their impact more closely. As previously mentioned, the log-transformation of both dependent and independent variables change the interpretation of the coefficient estimates, and they cannot be directly interpreted in their original units. One of the biggest reasons for this is that the log-linear model does not describe a constant change in y for a unit change in the original values of x. This is because the impact that the log-transformed units have on y depends on the original value of x, which, in this case, decrease as x gets larger. In other words, increasing the average number of likes a candidate received per post from 10 to 20 will have a greater impact than increasing it form, say, 20 to 40. The same remains true for the impact that shares have. To illustrate the impact of the like and share variables, rather than describing the impact that any conceivable change in x values have on y, it is more intuitive to show a plot that describes how a unit change on the average number of likes and shares a candidate receives per post impacts their vote distribution. This is seen in figure 8, where the x-axis shows the mean value of likes and shares in their original units and the y-axis shows how large percent change in votes can be expected from these values. The figure is derived from the second logged SM model and uses actual values of the mean number of likes per post that the candidates have within the dataset.

The figure illustrates how quickly the impact of likes diminish as their values grow larger, and at the start attempts to gain more likes per post can result in substantial gains. Achieving more than

one percent boost to votes can be accomplished by as little as 50 likes per post. However, this impact quickly diminishes, and getting a 1.5 percent boost already takes approximately 150 likes per post. Even if a candidate is able to match the staggeringly popular Alexander Stubb in terms of how many likes each of the individual posts is able to collect (Stubb's average is 766.85) this does not even pass a 2 percent threshold. In this light, the ability for likes to generate votes is an increasingly difficult task to achieve. Since most candidates could manage averages of 7.78 likes per post, it looks like the help Facebook provided during the election was less than 1 percent for most candidates.

Figure 8. The impact of Likes and shares on votes



In addition, if gaining votes through likes is a cumbersome task, it is even more difficult with shares. Even if a candidate could manage to gain 50 shares for each of the posts that he sends, which is unlikely given that the even the best performing candidate in this regard could only establish a mean of 36.9 shares per post, this would result in a less than 0.5 percent increase in votes. Moreover, the result found here are somewhat doubtful as the range given in figure 8 well exceeds the maximum value of the average number of shares present in the dataset. The model, thus, does not give reliable estimates on what happens after the highest mean of 36.9 is surpassed, but it illustrates just how likely – or rather how unlikely – a candidate is to gain a significant electoral advantage if his main social media strategy is to generate a lot of shares. Of course, it could be the case that shares can have a greater impact if candidates become better at generating

them. Thus, the potential of shares to impact election outcomes can be a dormant force waiting to be used by a skillful candidate, but based on these findings this seems like an unlikely option.

The findings thus far give us good intuition on the main mechanisms by which Facebook activity is connected with electoral performance, but they do not tell us much about differences between groups. Although the control variables show that being an incumbent and having larger campaigning funds are connected with larger vote shares, alone they do not imply that they can boost the impact that Facebook has. Rather, the previous models merely show that Facebook can have an impact which is independent of these variables. Hence, in order to test the assumptions made in the last three hypotheses (8, 9 and 10), one last set of regression estimates are presented in table 7, which shows three separate interaction models based on the first logged SM model presented in table 6. The three models are designed to test whether candidates with different age, incumbent status or campaign funding have different slopes associated with the most impactful like and share variables[14].

If, for example, the coefficient in the interaction term is significantly different from zero between incumbent status and the logged like variable, this implies that those who are incumbents gain differently from Facebook likes in contrast to non-incumbents. To avoid problems arriving from multicollinearity, each of the models only uses one of the three mentioned control variables (age, incumbent status, and funding) and first adds an interaction term between the control variable and the like variable and then again the same interactions for the share variable. To make this procedure clear, to test whether an interaction effect exists between different levels of funding, for example, and the amount likes and shares a candidate received on average, two separate regressions were run: first with an interaction term between the funding variable and the like variable, and then another between the funding variable and the share variable. This procedure was then repeated for age and incumbent status[15].

[14] The precense of an interaction effect within the post variable was investigated aswell, but this always resulted in a non-significant regressor estimates.

[15] This one-by-one approach was selected because including multiple interactions within the same regression resulted in a hihgly correlated regressors, which prevented from making inferences on which interaction terms were significant as the model could not separate the independent impact that each of the regressors had on the model.

Table 7. Third set of regression models

	Dependent variable: log(votes)	Age Model			Funding Model			Incumbent Model		
		Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic
SM Variables	log(*Posts*)	0.149***	0.031	4.768	0.152***	0.031	4.859	0.152***	0.031	4.887
	log(*Friends*)	0.064	0.038	1.687	0.066	0.038	1.729	0.066	0.038	1.740
	log(*Likes*)	0.397***	0.094	4.223	0.338***	0.058	5.795	0.324***	0.053	6.174
	log(*Comments*)	0.002	0.039	0.041	0.002	0.039	0.044	0.012	0.038	0.300
	log(*Shares*)	0.101***	0.029	3.476	0.100***	0.029	3.445	0.099***	0.029	3.423
	log(*Likes*) x Age	-0.002	0.002	-1.160						
	log(*Likes*) x 5,000-20,000 €				-0.074	0.058	-1.277			
	log(*Likes*) x Over 20,000 €				-0.038	0.074	-0.516			
	log(*Likes*) x Incumbent							-0.240**	0.078	-3.091
Control Variables	Age	0.006	0.005	1.205	0.001	0.002	0.408	0.001	0.002	0.378
	Male	-0.116*	0.049	-2.384	-0.117*	0.049	-2.405	-0.125*	0.048	-2.579
	List Position	-0.010**	0.003	-3.032	-0.010**	0.003	-3.126	-0.010**	0.003	-3.106
	First Female	0.100	0.116	0.861	0.106	0.116	0.909	0.096	0.116	0.829
	Incumbent	0.746***	0.099	7.531	0.741***	0.100	7.379	1.475***	0.260	5.679
	Campaign Experience	0.295***	0.055	5.402	0.299***	0.055	5.480	0.294***	0.054	5.416
	Education: Primary	0.295	0.247	1.195	0.299	0.247	1.213	0.317	0.246	1.291
	Education: Vocational	-0.008	0.132	-0.064	-0.006	0.132	-0.046	0.007	0.131	0.053
	Education: Upper Secondary	-0.131	0.140	-0.936	-0.135	0.141	-0.957	-0.106	0.140	-0.760
	Education: Higher Education	0.152	0.123	1.236	0.151	0.123	1.223	0.166	0.123	1.352
	Funding: 5,000-20,000 €	0.519***	0.061	8.504	0.670***	0.134	5.006	0.500***	0.061	8.212
	Funding: Over 20,000 €	0.926***	0.083	11.158	0.990***	0.205	4.838	0.912***	0.083	11.042
	Party Size: Medium	0.703***	0.095	7.374	0.694***	0.096	7.210	0.695***	0.094	7.374
	Party Size: Large	1.199***	0.101	11.910	1.183***	0.103	11.484	1.189***	0.010	11.905
	(Intercept)	4.354***	0.300	14.516	4.495***	0.261	17.226	4.513***	0.257	17.554
	n	844			844			844		
	R-squared	0.757			0.7566			0.759		
	Adjusted R-squared	0.751			0.7504			0.753		
	F-statistic	127.9***			121.7***			129.6***		

*p < .05; **p < .01; ***p < .001.

In almost all cases, the coefficient in the interaction term could not be established to differ significantly enough from zero. Adding interaction terms in the logged post variable is non-significant in all of the control variables, and the logged share variable is only significant with the Incumbent status variable, although only at a low $p < .05$ level. In the interest of compactness, table 7 only shows the interaction term results from the logged like variable, since only here any clear differences are found and including all of the other regression models would take considerable amount of space without providing much additional information.

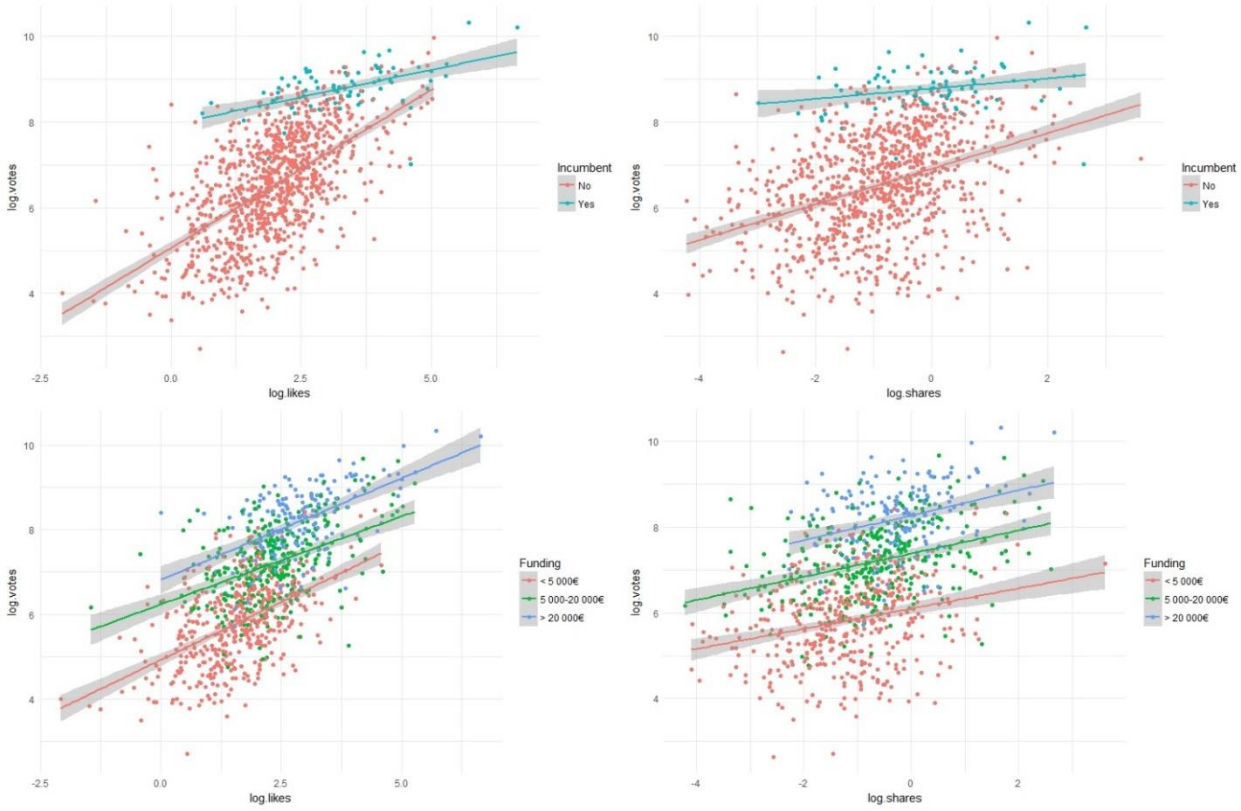
As it can be seen from the table, the coefficient associated with the interaction terms are non-significant in all cases with one notable exception. The interaction between incumbent status and log-transformed likes is significant (at .01 level) and, moreover, it shows a negative relationship! According to the regression estimates, the impact that likes have on the vote shares for incumbents is $0.324 - 0.240 \times Incumbent$. Since the Incumbent variable is 0 for non-incumbents, this implies that non-incumbents are able to get a 0.324% boost on their votes for a percent change in the average number of likes they receive, while incumbents receive only a 0.084% boost. In other words, the non-incumbents are gaining almost four times more from likes than incumbents are. A similar pattern is apparent for shares, though the difference is not as great as it is for likes.

Indeed, a graphical examination confirms this intuition. Figure 9 shows the linear relationship of logged likes and shares to logged vote counts for both incumbents and non-incumbents on the top of the figure and below the impact of the funding variable is similarly placed for contrast. These four plots show that while funding impacts the intercepts of the two social media variables, it does not affect the slopes themselves. Those candidates with bigger funding are automatically guaranteed to receive more votes, but the increase they can have from Facebook likes is the same for all funding levels. Incumbents, on the other hand, start from a higher intercept, but non-incumbents are able to compensate with a steeper slope.

Considering the non-linear impact that likes have on electoral performance, this result is perhaps not all that surprising. On average, incumbents perform better electorally and are more popular on Facebook, but since the impact that Facebook has on votes declines as higher values are reached incumbents belong precisely to the group where the impact of Facebook has reduced the most. If the assumptions behind the SEI model are correct, this is perhaps explainable after factors external to social media are considered. It is plausible that incumbents have multiple other channels through

which they can interact with voters and they often already possess a high degree of recognizability. Thus, the added benefit that Facebook can bring to them is minimal and the comparatively unknown non-incumbents are able to gain more from visibility on Facebook because they have few other sources to gain attention. This seems like a plausible assumption, but it should be tested with further data. Social media data alone cannot give us further insight on the credibility of this, and further studies should expand its scope outside of social media and include variables form alternative media channels.

Figure 9. The interaction effects of the Incumbent and Funding variables



Taken together, the interaction terms revealed that the major difference between groups resides within the incumbent status. Neither the age of the candidate nor the funding level was able to boost the impact that Facebook had. This suggests that younger voters are not more apt at using Facebook for campaigning purposes and that funding alone does not help candidates to gain more from the platform. The funding variable, however, is somewhat problematic due to its self-reported nature and inaccurate measurement. Because of this, the variable is not precise and might not even

be accurate in that it does not measure what we wish it should measure. Moreover, the variable does not tell us whether or not the candidates actually allocated their funds to social media campaigning or did they mostly spend them in other activities. Hence, we should be wary of making too broad interpretations based on this particular result. The incumbent effect, however, remains intriguing and is one of the few empirical findings that favor the equalization thesis. It remains to be seen whether this is a robust result and holds in further elections, but for now it gives some hope for the equalization thesis after a long run of studies which mainly favor normalization. With these considerations, hypothesis 8 and 10 are not supported, but hypothesis 9 is partially supported.

Thus, the final result of these regression estimates is that Facebook can indeed be associated with electoral gains and it seems that, at least in these particular elections, that non-incumbents are enjoying most from these gains. This is a rather surprising result given all of the other studies which have come before and showed opposite results. As the literature review above has shown, most candidates who first start adopting social media can be considered as elite politicians who have the greatest resources and are placed in the political limelight. On the other hand, many studies have shown that there are some clear differences in the behavioral patterns of social media use by different candidates, which are more favorable towards marginal political actors. Indeed, as William & Gulati (2008) and Spierings and Jacobs (2014) have shown, more active social media users can gain more from the platform than passive users which suggests that normalization is possible if marginal actors are the once who exhibit more intensive use patterns. However, none of these previous studies have considered how different functions of diverse set of platforms impact the effectiveness of social media strategies for electoral gains and have, at most, looked at the number of followers or friends and the level of activity a candidate has spent during the campaigning season. Especially Facebook is a rare subject of study in this regard, and no studies have attempted to include its internal performance metrics – likes, comments, and shares – to the analysis.

What these results show is that if these metrics are left out from the analysis, the impact that social media has to the electoral context can be vastly underestimated. Furthermore, the results give some support for the SEI model outlined above. Although not all of the assumptions that were made prior to the empirical investigation hold under statistical testing, they do confirm the most

fundamental building blocks of the model. If we assume that visibility is the first necessary condition for a candidate to gain from social media, then we should expect that those who can neither attain it through first-degree friend networks or from infectious interactions will not be able to gain full benefits from the medium regardless of the amount of effort they put into it. Indeed, the results show that increasing effort on Facebook, sending more posts per day, is not nearly as impactful as it is to get more likes for the same amount of posts sent.

Rather than discrediting the SEI model, the results, then, suggest alterations and improvements which can be made to it. For one, it seems that the "level of effort" hypothesis is not solid and it is dubious whether comments or shares create a wider reach in contrast to likes because users need to spend more effort on them. However, at the very least, the results highlight that likes should be considered as a special variable which should not be mixed with comments or shares. Moreover, since likes take the least amount of effort by other users, it might be the case that they work more closely to what the SEI model suggest than the other two variables do. Likes can be given without careful consideration when to do so and thus Facebook users are like to push the like button merely when they "feel" to do so. In contrast, shares and comments require more clicks and types from the user who is consequently more likely to contribute them only in special circumstances. What these special circumstances are, is an important avenue for further research, and it could be the case that they represent untapped potential for effective social media campaigning for those candidates who are able to figure out what posts get shared and commented upon. At this point, the low overall levels of these interactions either tells us that candidates just do not know how to get them, or that voters are just overall less likely to be interested in commenting or sharing political content.

In terms of the normalization and equalization debate, the results presented here give us reason to hope that Facebook can be an equalizing force. While incumbents can be expected to have a higher vote share overall, they do not get as significant boosts from the most important Facebook variables than non-incumbents do. Furthermore, in spite of a number of different ways in which financial resources can be used to influence the reach that Facebook can achieve, the funding level of a candidate did not seem to influence how effective Facebook was. Although the used funding variable cannot distinguish how much a given candidate used resources specifically for social media, it nevertheless gives us preliminary intuition on the matter. This, of course, does not mean

that financial resources do not matter in campaigning on Facebook. One alternative explanation is that those political candidates with large resources had multiple of other areas where they could use them and opted not to prioritize social media. This can easily change in future elections. As Jacobs and Spierings (2016) argue, if bigger parties and elite candidates feel that social media threatens their status, they can use their extensive resources catch up with more experienced users. Perhaps incumbents were not yet threatened by social media during 2015 but will be in 2019 when the next parliamentary elections are held. If so, funding can play a greater role then than it has in 2015.

Thus far, we have advanced on almost all of the research objectives set forth at the beginning of this thesis. We have sought answers on how we can define social media, established a theoretical framework on how it can function as a campaigning tool, estimated the impact that social media had on the vote share of individual candidates, investigated whether differences and patterns existed on how social media impacted different candidates, and analyzed these results from the normalization v. equalization debate. It is thus time to return to the original research questions formulated at the beginning of this thesis, which is the task of the final concluding chapter.

8. CONCLUSIONS

The main research question that drove the formulation of all subsequent questions and research goals of this thesis asked whether social media can truly have an impact on the status quo of Finnish politics. All new advancements in communication and information technologies tend to be accompanied by enthusiastic and positive expectations from all of the new possibilities that can be now imagined. The same holds true with social media, and while it is truly characterized by innovative and novel ways to communicate with vast audiences, it is far from certain that the hopes placed in it will materialize as imagined. Indeed, Tim Wu (2010) perfectly describes why communication technologies can inspire such optimism in their early days:

> *When in the course of human affairs things go wrong, the root cause is often described as some failure to communicate, whether it be between husband and wife, a general and a front-line commander, a pilot and a radio controller, or among several nations. Better communications, it is believed, lead to better mutual understanding, perhaps a recognition of a shared humanity,*

> *and the avoidance of needless disaster. Perhaps it is for this reason that the advent of every new technology of communication always brings with it a hope for ameliorating all the ills of society.* (p. 36)

As Wu (2010) continues to describe, the advent of the radio in the United States at the start of 19th century made people hope it would bring the remote federal government closer to the citizens, improve public discourse and even create a more cultured society. The reality, as it always is, proved to be much murkier, and all of the hopes and expectations placed on this new technology proved not to work out quite as imagined. Now at the age of social media, it is no surprise that this new medium has yet again given fuel to optimistic voices and is surrounded with hopeful theories, such as the equalization thesis or the mobilization thesis. However, we should treat these voices with caution and thoroughly investigate their claims before jumping to conclusions. While social media may indeed be a force for a more equal political competition and give voices to marginalized groups, multiple other forces are at play simultaneously, which make the reality the murky place that it is. As newspapers, radio and television exemplify, we have been at this juncture of a new technological age multiple times before, and we would be wise recall all of the hopes placed in them and continue to remember how those hopes were eventually realized as the new technologies matured and became old technologies.

During this thesis, one of the most underlying expectation placed on social media is that it truly can have an impact on the lives of individuals and further influence our institutions and political systems. Similarly, both the normalization and equalization theses have this expectation as well, although the direction of the impact is vastly different. Hence, it is here where we must first shed light and see how murky the reality really is. As has been done in the United States and the Netherlands before, there seem to be strong grounds to argue that social media has had an impact within the Finnish electoral context. Those who use social media and perform well there are indeed associated with direct electoral gains in terms of the number of votes they receive. Moreover, not every candidate is equal in this endeavor. Rather, it seems to be the case that non-incumbent politicians can benefit more from social media platforms. Although the reason behind this cannot be fully investigated within this study, a plausible hypothesis is that incumbents can already gain visibility from multiple other sources, such as from news stories that cover their daily political activities. In most cases, non-incumbents do not have the resources or necessary pre-established

recognizability to be confident that a sizable portion of the electorate is already familiar with them. Social media, then, can be the sole source for visibility that they can acquire and, if used properly, it can help them to gain more votes.

These answer the first two sub-questions formulated at the start of the study, but the final sub-question "*If the answers to questions 1.1 or 1.2 are positive, are there mechanisms internal or external to social media that can explain them?"* is not as easily answered. The main reason for this has to do with the fact that the data used during this thesis is collected almost solely from social media, and only from Facebook at that, and it has next to no information on forces external to it. As the hybrid media system theory states (Chadwick 2017), our current media system cannot be fully understood through any one medium. Rather, the current system is a network of multiple fractured media systems that interact, compete, and intersect with one other constantly. This is especially topical when studying social media since it is one of the most versatile and interconnected medium out there. Social media can be used to read news from traditional broadcasters, distribute campaigning videos by political candidates, organize political events at a moments notice, and much, much more. Thus, how individuals use social media cannot be understood without knowing how it is used in relation to other external media channels and consequently we cannot have a full understanding on how it affects politics without taking external factors into account. Unfortunately, more data is needed for us to go further in this direction, which is currently out of reach for this thesis.

We can, however, start making preliminary observations and hypothesizes on the main internal mechanisms through which social media functions and gives its users political capital. While the SEI model is by no means meant to be the only mechanism imaginable, it illustrates the first and absolutely necessary condition that political actors must meet in order to use the platform effectively for political purposes. Before political candidates can induce social presence with voters, before they can promote themselves and their actions, and before they can try to influence the public agenda through social media, they need to be made visible within the platform. In social media, no candidate is guaranteed to have this resource *a priori*. How this visibility is allocated to users is one of the most important mechanisms we must understand before we can have a full picture on how social media benefits political actors and who can benefit the most from it.

This thesis has made few steps towards this goal and furthered our understanding on the importance of social media to the Finnish political system. While social media activity in itself can have a meaningful impact to the electoral prospects of candidates, by far the most influential factor is the internal functions and interactions that Facebook allows its users to use, such as liking or sharing posts. As Bucher (2012) and DeVito (2017) have illustrated, these interactions do influence the visibility that a post in Facebook can achieve. While the magnitude of the impact was revealed to be rather small for individual candidates, in the aggregate it can sum up to significant changes. Since the Finnish electoral system uses proportional representation, which allocates seats proportional to the total number of votes that a party receives, even a 2% increase to the vote share of individual candidates can have consequences for the formation of the parliament if multiple candidates within the same party can achieve this boost. Party officials could then increase the influence their party has by systematically teaching and guiding their candidates how to use social media for campaign purposes in the hopes that this increased ability will bring them a boost to their overall vote share and reward them with extra seats within the legislature.

The importance of these internal functions is something that has not been studied before in the context of Facebook and political campaigning. Many previous studies have focused at examining how raw posting frequency and the number of Facebook Friends or supporters contribute to candidates electoral success, but, due to data collection or other constraints, have ignored other relevant Facebook metrics. This study has shown that both the raw posting frequency and the number of Friends can have a rather marginal effect and are dwarfed by the internal Facebook functions, such as creating likes to Facebook posts. Moreover, some of the results revealed in this thesis cast doubt upon the few arguments which state that social media can influence voters by creating tighter interactions between the candidates and voters (e.g. Kruikemeier, 2014; Parmelee & Bichard, 2012). According to this line of thought, since social media can facilitate intimate and one-to-one communications with voters, the platform can be used to draw voters closer to candidates and their campaigns, which in turn makes voters feel more involved and more inclined to give their only vote to the candidate who has been the most successful in interaction with them. However, the results of this thesis have shown that the ability to generate conversation and comments within Facebook is not associated with electoral gains.

This gives us reason to reconsider the plausibility of the interaction argument. Indeed, recalling the studies outlined in the literature review chapter, multiple examples from the United States and Europe have shown that most politicians are not particularly interested in the interactive and conversational form of campaign. This suggests that even if more interactive social media approach could be helpful electorally, most candidates cannot practice it because it would exhaust too many resources, are afraid of negative unintended consequences from an uncontrolled communication environment or find other campaigning tactics more efficient to spend resources on, for example. Indeed, it could be the case that the relative unimportance of comments within this thesis is due to the fact that they are extremely scarce and are just not created in the degree in which they can have a meaningful impact. Whichever the case might be, at the very least the results of this thesis show that interactive or conversational campaign is not necessary in order to have political gains from social media use. Pushing the like button is certainly not as interactive or communicative as it is to write comments and is a rather a one-directional and simplistic way to communicate between candidates and voters. The candidate cannot, for example, further interact with a voter who has liked his post or know why the like button was pushed in the first place. If the visibility argument advanced in this thesis holds true, it might not even matter why the like button was pushed. That is to say, there is no such thing as bad publicity within Facebook; all that matters is that your posts are infectious enough to get reacted to and in this way spread to as many people as possible.

However, before broad generalizations are made based on this study, its faults and shortcomings must be recognized. The clearest issue with the results of this study comes from the fact that it is impossible to have a definite conclusion on the issue of causality. What this thesis proposes, after all, is that there exists a causal connection between Facebook and electoral performance, but we cannot be sure of this with the cross-sectional data at our current disposal. Even after controlling for various alternative explanations, we cannot tell in which way the arrow of causality goes. Does good performance on Facebook result in better electoral prospects or is it merely the case that electorally strong candidates often have active social media followers who continuously like and share the posts of their favorite candidate? If the latter is closer to the truth, then the impact of likes and shares that we have discovered is not due to direct causal connection but rather they are an indication of the underlying enthusiasm felt towards the candidate.

In order for us to be more confident in the causal interpretation, we need to collect better data which allows us to bypass the ever-present omitted variable problem through Instrumental Variable methods, for example, by conducting controlled experiments where we have more control over the research environment, or by focusing on cases where it is possible to collect time series data. In this regard, focusing only on vote shares is problematic since it does not allow us to adopt a time series approach because elections are usually held after a long time-interval between elections during which the candidates also change. This creates non-comparable datasets which are not easily used for analysis. The alternative is to shift our focus to other areas where we can have multiple observations from different time periods from the same individuals or entities. One plausible option is to focus on the parties themselves rather than on individual candidates. Regular opinion polls published by various sources allow us to track the changes that occur on the level of support that a party enjoys and to test whether social media can affect how popular the party is. This can prove to be a promising avenue for further research.

The second major issue within this thesis comes from one of the most common problems related with regression models: not including all relevant control variables. In this case, an obvious issue with the used variables is that they do not control for the influence of pre-existing fame or recognizability of the candidates. The more famous or recognizable the candidate is, the more likely it is that the candidate will perform well during the elections because the sheer number of people who can plausibly consider voting for him is greater than it is for less-known individuals. Moreover, the same can be said about social media performance and the more well-known the candidate is the more likely it is that he has a large social media following and more users who can interact and react to the messages he sends on any given platform. Thus, this factor clearly influences both the dependent variable and our main independent variable of interest, which likely leads to overestimating the influence of our social media variables since part of the impact they have can be due to the fame of the candidate and does not come from social media itself.

It could be argued that the influence of fame is somewhat controlled by including a variable indicating incumbent status because incumbents are usually the most well-known political figures within any political system. However, there are multiple of other ways in which a candidate can be recognizable that has nothing to do with politics. This is true for actors, athletes, and other media personalities, for example. This is a problem that the political study of social media needs

to solve, and we need to consider what solutions arrive at the best results. We can, as Jacobs and Spierings (2014) have done, try to control for fame by collecting secondary data from various news sources and count the number of time a candidate has appeared within the news prior to the election period. Further paths come from a more qualitative approach and device coding strategies where candidates are coded either well-known or not well-known based on the judgment of ordinary citizens who must recognize the candidate only based on the name and face of the candidate. One more alternative path is to use other sources where only accomplished or well-known individuals appear. Wikipedia is one of such sources that could be considered since having a Wikipedia article that remains on the site for a long period of time can be considered as an indication of fame. All of these solutions can help with the omitted variable problem, but for now the uncertainty it brings to this study must be kept in mind alongside with the results of this study.

Finally, the nature of the research subject brings some considerable complications for the external validity of the study. Since most algorithms that function beneath various social media platforms are vastly different with different internal functions and logic, we cannot draw straightforward conclusions on the overall of influence of social media to the Finnish political system based solely on data from Facebook. Moreover, even within a single platform the algorithm can change almost from day to day which makes the study of Facebook, for example, in different time periods almost a study of two completely different topics. This is the main reason why the SEI model is as general as possible and does not make any prior assumptions on the degree to which different internal functions impact the spread of messages. However, it is precisely because of the huge variety and volatility of social media platforms why we need a generalized model such as the SEI model. While the specific mechanism through which visibility is gained within social media tends to change over time, the fact that all social media can be thought of as a network of individual users who influence how information spreads within the overall network gives us the necessary solid ground to which we can root overarching theories. The importance of individual users for the dissemination of information is the greatest in social media than it has ever been in any of the prior communication technologies that have come before, and the SEI model aims to capture this main mechanism.

While the disease analogy is nothing new when discussing social media, it has not been used explicitly through a systematic model that describes the relationship in detail, and it is here that

this thesis gives its major contribution within the research field. Going forward, the model can be improved, and our intuition strengthened on the specific differences between different social media platforms by conducting further research with dual data on multiple platforms at the same time. Indeed, the differences between Twitter and Facebook is something that has not been examined closely thus far, and most studies use data only from one of the two platforms.

Even after taking account the above-cited shortcomings of this thesis, the results we have arrived at are nevertheless intriguing and further gives us confidence that social media is an important area of study for political scientists. Together with the numerous studies which have come before, this thesis highlights the fact that social media is connected with direct political gains for those political actors who have chosen to use it as a political tool. Moreover, it is plausible that the importance of social media will continue to grow in the future and its use is certainly not declining among candidates nor ordinary citizens. The implications that this brings are numerous, and it is difficult to tell how social media continues to shape the political sphere. While the future is always difficult to predict accurately and there is no way of telling how social media matures, with these kinds of studies we can start to picture different set of outcomes and trends that social media can bring.

For one, if the main arguments in this thesis are standing on a solid ground, the way internal social media functions such as the like button shapes our political discourse and outcomes can be worrisome. If there truly is no such thing as bad publicity on Facebook, then the ultimate measure in which political actors measure their success on the platform is based on the amount of reactions they are able to generate. It matters not what one posts, how accurate the information within it is or how carefully it is formulated. What matters is that it gets noticed, interacted with and shared with others. In this kind of environment it is easy to imagine a form of deliberation that does not focus on difficult policy issues, thoughtfully considerations of current issues, or solid arguments for or against particular topics. Rather, the environment favors easily digestible, thought-provoking and perhaps even poorly confirmed information. Indeed, as a recent study from MIT shows, those news which contain confirmed and solid information spread six times slower on Twitter than those news which contain false or unconfirmed information (Vosoughi, Roy and Aral 2018).

Moreover, the fact that the most popular social media platforms are corporate entities that have not been constructed with deliberation or political communication more generally in mind but are

rather primarily concerned with the underlying logic of the market system and with profit maximization should give us some pause. This aspect is highly relevant to the functioning and health of a democratic system because if such social networks are becoming ever more important for the eventual success of political candidates and if at the same time political actors are becoming incrementally reliant on them to gain a voice, the internal logic of social media companies can become embedded to the political process. Like the Panopticon described by Michel Foucault, social media platforms are designed to make objects and users visible or invisible based on a criterion that the owners of the platform ultimately decide. If this visibility is increasingly acquired through social media and translates to better chances to perform well electorally, it not only influences who wields power but also impacts the very heart of any democratic system: how democratic deliberation is conducted and who gets heard. This is more than enough reason for us to continue to study social media.

9. REFERENCES

Allen, J., & Parnes, A. (2017). *Shattered: Inside Hillary Clinton's doomed campaign*. New York, NY: Crown.

Anders Olof Larsson, & Hallvard Moe. (2012). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society, 14*(5), 729-747. doi:10.1177/1461444811422894

Ansolabehere, S., Iyengar, S., Simon, A., & Valentino, N. (1994). Does attack advertising demobilize the electorate? *American Political Science Review, 88*(4), 829-838.

Baxter, G., & Marcella, R. (2012). Does Scotland ‘like’ this? Social media use by political parties and candidates in Scotland during the 2010 UK general election campaign. *Libri, 62*(2), 109-124.

Baxter, G., & Marcella, R. (2013). Do online election campaigns sway the vote? A study of the 2011 Scottish parliamentary election. *Libri, 63*(3), 190-205.

Baym, N. K. (2015). Social media and the struggle for society. *Social Media + Society, 1*(1), 205630511558047. doi:10.1177/2056305115580477

Bettencourt, L. M. A., Cintrón-Arias, A., Kaiser, D. I., & Castillo-Chávez, C. (2005). The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications, 364,* 513-536. doi:10.1016/j.physa.2005.08.083

Bimber, B. (2003). *Information and American democracy: Technology in the evolution of political power*. Cambridge: Cambridge University Press.

Blumler, J. G., & Kavanagh, D. (1999). The third age of political communication: Influences and features. *Political Communication, 16*(3), 209-230.

Broersma, M., & Graham, T. (2012). Social media as beat: Tweets as a news source during the 2010 British and Dutch elections. *Journalism Practice, 6*(3), 403-419.

Bucher, T. (2012). Want to be on the top? algorithmic power and the threat of invisibility on Facebook. *New Media & Society, 14*(7), 1164-1180.

Caplan, B. (2011). *The myth of the rational voter: Why democracies choose bad policies*. Oxfordshire: Princeton University Press.

Chadwick, A. (2017). *The hybrid media system* (2nd ed.). New York, NY: Oxford University Press. doi:10.1093/oso/9780190696726.001.0001

Cogburn, D. L., & Espinoza-Vasquez, F. K. (2011). From networked nominee to networked nation: Examining the impact of web 2.0 and social media on political participation and civic engagement in the 2008 Obama campaign. *Journal of Political Marketing, 10*(1-2), 189-213.

DeVito, M. A. (2016). From editors to algorithms: A values-based approach to understanding story selection in the Facebook news feed. *Digital Journalism, 5*(6), 753-21. doi:10.1080/21670811.2016.1178592

Diamond, L., & Plattner, M. F. (2012). *Liberation technology: Social media and the struggle for democracy.* Baltimore, MD: John Hopkins University Press.

Dijck, J. V. (2013). *The culture of connectivity: A critical history of social media*. Oxford: Oxford University Press.

Enli, G. S., & Skogerbø, E. (2013). Personalized campaigns in party-centred politics: Twitter and Facebook as arenas for political communication. *Information, Communication & Society, 16*(5), 757-774.

Enyedi, Z. (2008). The social and attitudinal basis of political parties: Cleavage politics revisited. *European Review, 16*(3), 287-304.

Esterling, K., Lazer, D. M., & Neblo, M. A. (2005). Home (page) style: Determinates of the quality of the house members' web sites. *International Journal of Electronic Government Research (IJEGR), 1*(2), 50-63.

Foot, K. A., Schneider, S. M., Kluver, R., Xenos, M., & Jankowski, N. W. (2007). Comparing web production practices across electoral web spheres. In R. J. Kluver Nicholas, K. Foot & S. Schneider (Eds.), *The internet and national elections; A comparative study of web campaigning* (pp. 243-260). London and New York: Routledge.

Foucault, M. (1975). *Discipline and punish: The birth of the prison*. New York, NY: Vintage Books.

Gainous, J., & Wagner, K. M. (2014). *Tweeting to power: The social media revolution in American politics*. New York, NY: Oxford University Press.

Gayo-Avello, D. (2016). *Politics and social media.* Unpublished manuscript.

Gibson, R. K., & McAllister, I. (2015). Normalising or equalising party competition? assessing the impact of the web on election campaigning. *Political Studies, 63*(3), 529-547.

Gibson, R., & Cantijoch, M. (2013). Conceptualizing and measuring participation in the age of the internet: Is online political engagement really different to offline? *The Journal of Politics, 75*(3), 701-716.

Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the U.S. Congress. *Journal of the American Society for Information Science and Technology, 61*(8), 1612-1621. doi:10.1002/asi.21344

Graham, T., Broersma, M., Hazelhoff, K., & van't Haar, G. (2013). Between broadcasting political messages and interacting with voters: The use of Twitter during the 2010 UK general election campaign. *Information, Communication & Society, 16*(5), 692-716.

Grant, W. J., Moon, B., & Busby Grant, J. (2010). Digital dialogue? Australian politicians' use of the social network tool Twitter. *Australian Journal of Political Science, 45*(4), 579-604.

Hansen, K. M., & Kosiara-Pedersen, K. (2014). Cyber-campaigning in Denmark: Application and effects of candidate campaigning. *Journal of Information Technology & Politics, 11*(2), 206-219.

Harfoush, R. (2009). *Yes we did! An inside look at how social media built the Obama brand*. Berkeley, CA: New Riders.

Hoffmann, C. P., Suphan, A., & Meckel, M. (2016). The impact of use motives on politicians' social media adoption. *Journal of Information Technology & Politics, 13*(3), 239-256.

Hogan, B., & Melville, J. (2015). *Social media* (2nd ed.). Oxford: Elsevier.

Huber, G. A., & Arceneaux, K. (2007). Identifying the persuasive effects of presidential advertising. *American Journal of Political Science, 51*(4), 957-977.

Jackson, N. A., & Lilleker, D. G. (2009). Building an architecture of participation? political parties and web 2.0 in Britain. *Journal of Information Technology & Politics, 6*(3-4), 232-250.

Jacobellis v. Ohio, 378 U.S. 184 (1964). Retrieved from https://supreme.justia.com/cases/federal/us/378/184/case.html

Jacobs, K., & Spierings, N. (2016). *Social media, parties, and political inequalities*. New York, NY: Palgrave Macmillan.

Jaeger, P. T., Paquette, S., & Simmons, S. N. (2010). Information policy in national political campaigns: A comparison of the 2008 campaigns for president of the United States and prime minister of Canada. *Journal of Information Technology & Politics, 7*(1), 67-82.

Jin, F., Dougherty, E., Saraf, P., Cao, Y., & Ramakrishnan, N. (2013). Epidemiological modeling of news and rumors on Twitter. Paper presented at the *Proceedings of the 7th Workshop on Social Network Mining and Analysis,* 8.

Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics, 13*(1), 72-91. doi:10.1080/19331681.2015.1132401

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*(2), 263-91

Kalnes, Ø. (2009). Norwegian parties and web 2.0. *Journal of Information Technology & Politics, 6*(3-4), 251-266.

Katz, J., Barris, M., & Jain, A. (2013). *The social media president: Barack Obama and the politics of digital engagement*. New York, NY: Palgrave Macmillan.

Koc-Michalska, K., Gibson, R., & Vedel, T. (2014). Online campaigning in France, 2007–2012: Political actors and citizens in the aftermath of the web. 2.0 evolution. *Journal of Information Technology & Politics, 11*(2), 220-244.

Kruikemeier, S. (2014). How political candidates use Twitter and the impact on votes. *Computers in Human Behavior, 34*(Supplement C), 131-139. doi:10.1016/j.chb.2014.01.025

Larsson, A. O., & Kalsnes, B. (2014). 'Of course we are on Facebook': Use and non-use of social media among Swedish and Norwegian politicians. *European Journal of Communication, 29*(6), 653-667.

Lasorsa, D. L., Lewis, S. C., & Holton, A. E. (2012). Normalizing twitter: Journalism practice in an emerging communication space. *Journalism Studies, 13*(1), 19-36.

Lassen, D. S., & Brown, A. R. (2010). *Twitter: The electoral connection?*. Los Angeles, CA: SAGE Publications. doi:10.1177/0894439310382749

Lee, E., & Shin, S. Y. (2012). Are they talking to me? cognitive and affective effects of interactivity in politicians' twitter communication. *Cyberpsychology, Behavior, and Social Networking, 15*(10), 515-520.

Marcinkowski, F., & Metag, J. (2014). Why do candidates use online media in constituency campaigning? an application of the theory of planned behavior. *Journal of Information Technology & Politics, 11*(2), 151-168.

Maurice Vergeer, Liesbeth Hermans, & Steven Sams. (2013). Online social networks and micro-blogging in political campaigning: The exploration of a new campaign tool and a new campaign style. *Party Politics, 19*(3), 477-501. doi:10.1177/1354068811407580

McGee, M. (2013). EdgeRank is dead: Facebook's news feed algorithm now has close to 100K weight factors. Retrieved from https://marketingland.com/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors-55908

Meeks, L. (2003). Blogs and politicians. Retrieved from http://politicking03.blogspot.com/

Mueller, D. (2003). *Public choice III*. New York, NY: Cambridge University Press.

Napoli, P. M. (2014). Automated media: An institutional theory perspective on algorithmic media production and consumption. *Communication Theory, 24*(3), 340-360.

Nelimarkka, M. (2015). Digivaalit 2015. Retrieved from https://www.hiit.fi/digivaalit-2015

Norris, P. (2000). *A virtuous circle: Political communications in postindustrial societies*. New York, NY: Cambridge University Press.

Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the internet worldwide*. New York, NY: Cambridge University Press.

Nowak, M., & Eckles, D. (2014). “Determining user personality characteristics from social networking system communications and characteristics.” US20140074920 A1.

O'Reilly, T. (2005). What is web 2.0. Retrieved from http://oreilly.com{file}

Parmelee, J. H., & Bichard, S. L. (2012). *Politics and the Twitter revolution: How tweets influence the relationship between political leaders and the public*. Lanham, MD: Lexington Books.

Railo, E., & Ruohonen, R. (2016). Vaalikampanjat ja julkisuuden agenda. In K. Grönlund, & H. Wass (Eds.), *Poliittisen osallistumisen eriytyminen – eduskuntavaalitutkimus 2015* (pp. 76-94). Helsinki: Oikeusministeriö.

Ridout, T. N., Shah, D. V., Goldstein, K. M., & Franz, M. M. (2004). Evaluating measures of campaign advertising exposure on political learning. *Political Behavior, 26*(3), 201-225.

Robert M Bond, Christopher J Fariss, Jason J Jones, Adam D I Kramer, Cameron Marlow, Jaime E Settle, & James H Fowler. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature, 489*(7415), 295. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22972300

Roper, J. (1999) 'New Zealand political parties online: The world wide web as a tool for democratization or political marketing', in C. Toulouse and T. Luke (Eds.), *The Politics of Cyberspace*. London: Routledge. pp. 69–83.

Ross, K., & Bürger T. (2014). Face to face (book) social media, political campaigning and the unbearable lightness of being there. *Political Science, 66*(1), 46-62.

Shirky, C. (2011). The political power of social media: Technology, the public sphere, and political change. Foreign Affairs, *90*(1), 28-41.

Sittig, A., & Zuckerberg, M. (2013). "Managing information about relationships in a social network via a social timeline." US8583690 B2.

Spierings, N., & Jacobs, K. (2014). Getting personal? the impact of social media on preferential voting. *Political Behavior, 36*(1), 215-234. doi:10.1007/s11109-013-9228-2

Strandberg, K. (2013). A social media revolution or just a case of history repeating itself? the use of social media in the 2011 Finnish parliamentary elections. *New Media & Society, 15*(8), 1329-1347.

Strandberg, K. (2016). Ehdokkaiden ja kansalaisten internetin ja sosiaalisen median poliittinen käyttö vuosien 2003–2015 eduskuntavaaleissa. In K. Grönlund, & H. Wass (Eds.), *Poliittisen osallistumisen eriytyminen – eduskuntavaalitutkimus 2015* (pp. 95-116). Helsinki: Oikeusministeriö.

Stromer-Galley, J. (2003). Diversity of political conversation on the internet: Users' perspectives. *Journal of Computer-Mediated Communication, 8*(3), 0.

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. Paper presented at the 2010 IEEE Second International Conference on Social Computing. Date of Conference: 20-22 Aug. 2010. doi:10.1109/SocialCom.2010.33

Sundberg, J. (1997). Finland: Candidate choice and party proportionality. In A. Reynolds, & B. Reilly (Eds.), *The international IDEA handbook of electoral system design* (pp. 72-73). Stockholm: International Institute for Democracy and Electoral Assistance.

Travers, J., & Milgram, S. (1967). The small world problem. *Psychology Today, 1*(1), 61-67.

Vergeer, M., & Hermans, L. (2013). Campaigning on Twitter: Microblogging and online social networking as campaign tools in the 2010 general elections in the Netherlands. *Journal of Computer-Mediated Communication, 18*(4), 399-419. doi:10.1111/jcc4.12023

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science,* 359(6380), 1146-1151.

Wang, B., Wang, C., Bu, J., Chen, C., Zhang, W. V., Cai, D., & He, X. (2013). Whom to mention: Expand the diffusion of tweets by @ recommendation on micro-blogging systems. *Proceedings of the 22nd international conference on World Wide Web* (pp. 1331-1340). ACM.

Wauters, B., Weekers, K., & Maddens, B. (2010). Explaining the number of preferential votes for women in an open-list PR system: an investigation of the 2003 federal elections in Flanders (Belgium). *Acta Politica*, 45(4), 468–490.

Williams, C. B., & Gulati, G. J. (2007). Social networks in political campaigns: Facebook and the 2006 midterm elections. Paper presented at the *Annual Meeting of the American Political Science Association, 1*(11), 49-62.

Williams, C., & Gulati, G. (2008). What is a social network worth? Facebook and vote share in the 2008 presidential primaries. Paper presented at the 104th Annual Meeting of the American Political Science Association, Boston, MA

Wu, T. (2010). *The master switch: The rise and fall of information empires.* London: Atlantic Books.

Wu, T. (2017). *The attention merchants: The epic struggle to get inside our heads*. London: Atlantic Books.

Zeileis, A. (2017). lmtest: *Testing Linear Regression Models*. R package. version 0.9-35.

10. APPENDIX

Appendix 1. Interaction models between post intensity and friend, like, comment, and share variables

	Dependent variable: log(votes)	Friend Model			Like Model			Share Model		
		Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic
SM Variables	log(*Posts*)	0.048	0.136	0.356	0.118*	0.058	2.024	0.146***	0.037	3.963
	log(*Friends*)	0.071	0.039	1.822	0.063	0.038	1.670	0.064	0.038	1.678
	log(*Likes*)	0.304***	0.053	5.783	0.313***	0.053	5.868	0.308***	0.053	5.821
	log(*Comments*)	0.000	0.039	-0.009	-0.002	0.039	-0.041	0.000	0.039	-0.008
	log(*Shares*)	0.099***	0.029	3.402	0.098***	0.029	3.365	0.097**	0.030	3.186
	log(*Likes*) x log(Friends)	0.019	0.024	0.772						
	log(Post) x log(Likes)				0.017	0.026	0.659			
	log(Post) x log(Shares)							-0.005	0.022	-0.238
Control Variables	Age	0.001	0.002	0.392	0.001	0.002	0.413	0.001	0.002	0.382
	Male	-0.116*	0.049	-2.379	-0.115*	0.049	-2.364	-0.118*	0.049	-2.422
	List Position	-0.010**	0.003	-3.039	-0.010**	0.003	-3.070	-0.010**	0.003	-3.093
	First Female	0.103	0.116	0.889	0.103	0.116	0.886	0.100	0.116	0.856
	Incumbent	0.733***	0.098	7.457	0.733***	0.098	7.456	0.731***	0.098	7.429
	Campaign Experience	0.293***	0.055	5.366	0.294***	0.055	5.378	0.297***	0.055	5.437
	Education: Primary	0.295	0.247	1.197	0.291	0.247	1.178	0.290	0.247	1.174
	Education: Vocational	-0.007	0.132	-0.055	-0.009	0.132	-0.067	-0.009	0.132	-0.070
	Education: Upper Secondary	-0.123	0.140	-0.878	-0.122	0.140	-0.867	-0.125	0.141	-0.890
	Education: Higher Education	0.153	0.123	1.240	0.152	0.123	1.230	0.152	0.123	1.237
	Funding: 5,000-20,000 €	0.515***	0.061	8.430	0.515***	0.061	8.412	0.518***	0.061	8.489
	Funding: Over 20,000 €	0.917***	0.083	11.020	0.915***	0.084	10.959	0.922***	0.083	11.106
	Party Size: Medium	0.726***	0.095	7.623	0.727***	0.096	7.593	0.716***	0.095	7.571
	Party Size: Large	1.225***	0.102	12.072	1.227***	0.102	11.977	1.210***	0.101	12.017
	(Intercept)	4.480***	0.267	16.785	4.507***	0.261	17.257	4.531***	0.259	17.525
	n	844			844			844		
	R-squared	0.756			0.7563			0.756		
	Adjusted R-squared	0.750			0.7504			0.750		

F-statistic	127.7***	127.7***	127.6***

*p < .05; **p < .01; ***p < .001.

Note: The models include three dummy variables: First Female (First Female in candidate list = 1), Incumbent (Candidate served as an MP during the election = 1), and Campaign Experience (candidate was running in the previous 2011 elections = 1). Education, Funding and Party Size are categorical variables, where reference categories are: Education = Other; Funding = less than 5,000 €; Party Size = Small.

Appendix 2. Interaction terms with log-transformed share variable

Dependent variable: log(votes)		Age Model			Funding Model			Incumbent Model		
		Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic	Coefficient	SE	t-Statistic
SM Variables	log(*Posts*)	0.150***	0.031	4.810	0.149***	0.031	4.769	0.150***	0.031	4.843
	log(*Friends*)	0.064	0.038	1.687	0.063	0.038	1.666	0.064	0.038	1.690
	log(*Likes*)	0.307***	0.052	5.844	0.313***	0.053	5.941	0.313***	0.052	5.995
	log(*Comments*)	-0.001	0.038	-0.019	0.004	0.037	0.105	0.010	0.038	0.260
	log(*Shares*)	0.223**	0.079	2.811	0.132***	0.034	3.843	0.110***	0.029	3.794
	log(Shares) x Age	-0.003	0.002	-1.681						
	log(Shares) x 5,000-20,000 €				-0.070	0.045	-1.539			
	log(Shares) x Over 20,000 €				-0.096	0.066	-1.473			
	log(Shares) x Incumbent							-0.224**	0.072	-3.114
Control Variables	Age	-0.002	0.003	-0.712	0.001	0.002	0.323	0.001	0.002	0.387
	Male	-0.114*	0.049	-2.348	-0.115*	0.049	-2.372	-0.122*	0.048	-2.517
	List Position	-0.010**	0.003	-3.063	-0.011**	0.003	-3.198	-0.010**	0.003	-3.172
	First Female	0.089	0.116	0.766	0.104	0.116	0.891	0.099	0.116	0.856
	Incumbent	0.742***	0.098	7.545	0.745***	0.098	7.564	0.693***	0.099	7.032
	Campaign Experience	0.295***	0.054	5.422	0.297***	0.055	5.458	0.292***	0.054	5.378
	Education: Primary	0.295	0.246	1.197	0.312	0.247	1.263	0.301	0.245	1.225
	Education: Vocational	-0.010	0.132	-0.077	0.003	0.132	0.026	-0.005	0.131	-0.041
	Education: Upper Secondary	-0.126	0.140	-0.898	-0.123	0.140	-0.876	-0.113	0.140	-0.807
	Education: Higher Education	0.154	0.123	1.253	0.166	0.123	1.347	0.153	0.122	1.248
	Funding: 5,000-20,000 €	0.515***	0.061	8.447	0.451***	0.074	6.110	0.509***	0.061	8.376

Funding: Over 20,000 €	0.920***	0.083	11.105	0.857***	0.090	9.524	0.907***	0.083	10.973
Party Size: Medium	0.710***	0.094	7.511	0.729***	0.095	7.702	0.708***	0.094	7.532
Party Size: Large	1.206***	0.100	12.059	1.211***	0.100	12.104	1.199***	0.100	12.034
(Intercept)	4.648***	0.267	17.389	4.557***	0.259	17.626	4.554***	0.257	17.713
n	844			844			844		
R-squared	0.757			0.7572			0.759		
Adjusted R-squared	0.751			0.751			0.753		
F-statistic	128.2***			122***			129.6***		

*p < .05; **p < .01; ***p < .001.

Note: The models include three dummy variables: First Female (First Female in candidate list = 1), Incumbent (Candidate served as an MP during the election = 1), and Campaign Experience (candidate was running in the previous 2011 elections = 1). Education, Funding and Party Size are categorical variables, where reference categories are: Education = Other; Funding = less than 5,000 €; Party Size = Small.