E-voting in Estonia:
Technological Diffusion and Other Developments Over Ten Years (2005-2015)

Mihkel Solvak
&
Kristjan Vassil

E-voting in Estonia:Technological Diffusion and Other Developments Over Ten Years (2005 - 2015)

Mihkel Solvak
Kristjan Vassil

Johan Skytte Institute of Political Studies
University of Tartu

In cooperation with

Estonian National Electoral Committee

2 0 1 6

Published by the Johan Skytte Institute of Political Studies,
University of Tartu,
Ülikooli 18, 51003 Tartu, Estonia
E-mail: skytte@ut.ee

UNIVERSITY OF TARTU
Johan Skytte Institute of
Political Studies

In cooperation with Estonian National Electoral Committee,
Lossi plats 1a, 15165 Tallinn, Estonia
E-mail: info@vvk.ee

Cover & internal design
by: Pärt Ojamaa

Contents

List of Figures

List of Tables

Acknowledgements

Publication of this book was initiated and supported by the Estonian National Electoral Committee. Academic research that has provided empirical material for the book has been supported by various research grants from the following sources: Marie Curie post-doctoral research grant (nr: ERMOS53), Estonian Research Council personal research grant (nr: PUT523), EU structural grant "Verifiable Internet Voting – Event Analysis and Social Impact" (nr: 3.2.1201.13-0018), University of Tartu Development Fund, Estonian Science Foundation grant (nr: ETF8676).

We acknowledge the support Estonian National Electoral Committee, Ministry of Economic Affairs and Communications, Estonian Information System Authority, Tartu University, Johan Skytte Institute of Political Studies, Archimedes Foundation, Estonian Research Council, Cybernetica AS and European University Institute in studying Estonian e-voting at various periods during 2005-2015.

Numerous colleagues deserve special thanks for their support and helpful discussions, especially Priit Vinkel, Alo Heinsalu, Tarvi Martens, Epp Maaten, Heiki Sibul and Robert Krimmer. We also highlight the helpful role of Council of Europe, Michael Remmert and Prof. Alex Trechsel. Prof. Trechsel's part in providing the groundwork for scientific study of Estonian e-voting deserves special attention. He initiated the first surveys on e-voting in Estonia and was among the first to provide analytical insights into e-voting behaviour. Studying internet voting in Estonia over ten years by now would not have been possible without his contribution.

Foreword

Toomas Hendrik Ilves

President of the Republic of Estonia

We live in an era in which people are moving rapidly ahead in using IT, where digital technology dominates our everyday lives and with each day ever more so. When we established our independence we were a country in transition with limited resources. On the other hand, it was also at the start of the Internet revolution and this seemed the way we ought to go. We were hoping to increase our functional size by computerizing the country as much as possible. No legacy is an advantage in some cases.

In Estonia we can see a version of the interconnected and computerized future that is inextricably a part of the fundamental operations of society: 30% of the of participating voters cast their ballot online, nearly 100% of prescriptions and tax returns are done online, as are almost all banking transactions. Estonians have given more than 270 million digital signatures. Common e-services such a universal electronic ID for both public and private sectors are widely used and the whole of ICT infrastructure in a country should be regarded as an "ecosystem" in which everything is interconnected.

The technology itself is not sufficient to solve economical development problems, but it will become effective if combined with other reforms. Today we realise that our openness to technological changes has certainly helped our country to succeed. We have simply done things that would also be easy for others. The digital prescription and digital signature, as well as Internet elections could be available all across Europe and elsewhere. Nowadays many countries realize that strong remote voter authentication is an immense practical problem that has to be dealt with before they can consider deploying any online voting system.

Voters' trust is crucial, in general. If people do not trust their government, they will not trust voting systems either. If they do not trust computers and e-services, they will dislike Internet voting as well. Estonia's experience has been positive. We have built enough security into our e-services, be they private or public, and people do not have major negative experience with them. Thus, they still trust them.

It is hard to bring strong evidence beforehand that an IT system is secure. That makes it difficult to convince politicians or security experts to promote Internet voting in different corners of the world. The more digitized we are, the more vulnerable we are. It is therefore crucial to understand that security should not be seen as excuse or an additional cost but as an enabler, guarding our entire digital way of life.

The most effective means to be genuinely secure and to be safe from attacks is to go back to the pen and paper. That is one kind of solution, but it will not happen in Estonia. New features, for example, like the opportunity to verify whether his or her vote has been cast and counted increases voters' trust in this technology and it matters more than using the opportunity itself. Only a small amount of people verify their votes, but more important is the knowledge that one can check whether your vote was taken into account as it was supposed to.

Inclusiveness is vital and Internet is a great tool for it. Estonia's experience makes me optimistic. Since 2005, we have allowed online voting in 8 national, municipal and European elections. One might think that this would benefit the young urban elite, yet research has shown that there is no demographic or urban-rural divide. The pensioner living in a small village is just as likely to vote online, Skype with her family and stay in touch with her doctor remotely as are her grandchildren living in the city.

During the long decline of voter turnout in modern democracies, the question of how to motivate citizens to participate in elections has remained on the agenda. In recent years academia have become less optimistic about the Internet's ability to promote political participation and voter turnout. They argue that although in theory the Internet may lower the costs of electoral participation, strengthen democratic practices and include the disengaged into civic life, there seems to be little empirical support for these claims.

I am glad that by this book Estonia will add a valuable paragraph to the global history of Internet voting. The book provides an empirically-based account of the behavioral aspects of Internet voting. Moreover, if other countries are planning to modernize their elections, they will find valuable evidence how much the effects of modernization depend on a country's political, technological and social context.

Chapter 1

Introduction

KRISTJAN VASSIL

Estonia's use of modern information and communication technologies in public sector and for governance has placed the country at the forefront of states that are aiming to modernize their public sector and provide transparent governance. Numerous online public services are available to Estonian citizens and residents including digital identification, digital signatures, electronic tax filing, online medical prescriptions and internet voting. Driven by convenience, most of the services offer efficiency in terms of money and time saved for both the users and the public institutions. For example, selling a car in Estonia can be done online within less than 15 minutes, filing an online tax declaration takes the average person no more than five minutes, and participating in elections via internet voting takes on average 90 seconds.

The number of online public services that governmental offices offer to their "customers" are widely accepted and used by Estonian citizens and residents. Digital identification, the foundation stone of modern digital democracy, is compulsory for all citizens. In 2014 digital IDs were used more than 80 million times for authentication and 35 million times for digital transactions, significant numbers in a country with a population of only 1.3 million. Ninety-five percent of all income tax declarations are filed online, and every third citizen voted online in the last two elections in 2014 and 2015.

Regarding user attitudes and behaviour, survey evidence suggests that online governmental services are regarded as trustworthy and reliable. Citizens expect their public services and governmental offices to see their online presence not as a choice, but as a strategic and inevitable part of their day-to-day operations.

Yet, surprisingly little is known about how Estonian e-government in general and internet voting in particular has had an impact on an individual's behavioural? In this book we address precisely this question. Moreover, throughout the book we make use of extensive empirical material to substantiate and enrich the discussion on Estonian internet voting with evidence. However, before proceeding to specific findings, we will first set the scene regarding particular context of Estonia as well as the related technological environment.

BASIC FACTS ABOUT ESTONIA

After regaining independence in 1991, Estonia has become a full member of the European Union and NATO (both in 2004), the OECD (since 2010), and the Eurozone (since 2011). Estonia is a parliamentary democracy, with a Prime Minister as head of the government.

- Population: 1 294 236
- Area: 45 227 km^2
- Currency: Euro (since 2011)
- Capital: Tallinn

1.1 Internet voting in Estonia

In 2005 Estonia became the first country in the world to have nation-wide local elections where people could cast binding votes over the internet. This world premiere was followed by successful implementation of e-voting at all levels of elections: local, national and European. As of 2016, Estonia has held eight elections over ten years, where people could cast legally binding votes over the internet.

ELECTIONS IN ESTONIA

Since the restoration of independence in August 1991, Estonia has held 17 elections at either local, national or European level. With a population of about 1.3 million (according to the 2011 census), the size of the Estonian electorate is below 1 million[a] with the exact size depending on the election. The size of the electorate varies due to the fact that non-citizens are eligible to vote in local, but not in national elections. Voter turnout levels are comparable to other European countries, which together with institutional development, economic freedom and low levels of corruption, EU and NATO membership, as well as that of the Euro-zone, have made Estonia a consolidated and developed democracy.

[a] Average size of the electorate over all 17 elections is 921 594

The use of internet voting in Estonia has grown in a similar fashion to the general diffusion of technology, where only a few technology enthusiasts adapt to emerging technologies, but with the passage of time more and more users opt for the new technology with a subsequent spread across the population.

With regard to e-voting, the share of e-voters in the first e-enabled elections was very low, i.e only less than 2% of all votes were cast online. That is, every 50^{th} vote was cast over the internet. This number increased however by on average approximately 4.3 percentage points with each subsequent election and reached an all-time high in 2014 when every third vote was cast online. Figure 1.1 the growth of e-voting in relative and absolute terms since 2005.

There are two important policy implications to be highlighted on the basis of Figure 1.1. First, unlike the practice and theory of technological innovation diffusion prescribes, the growth of internet voting did not follow,

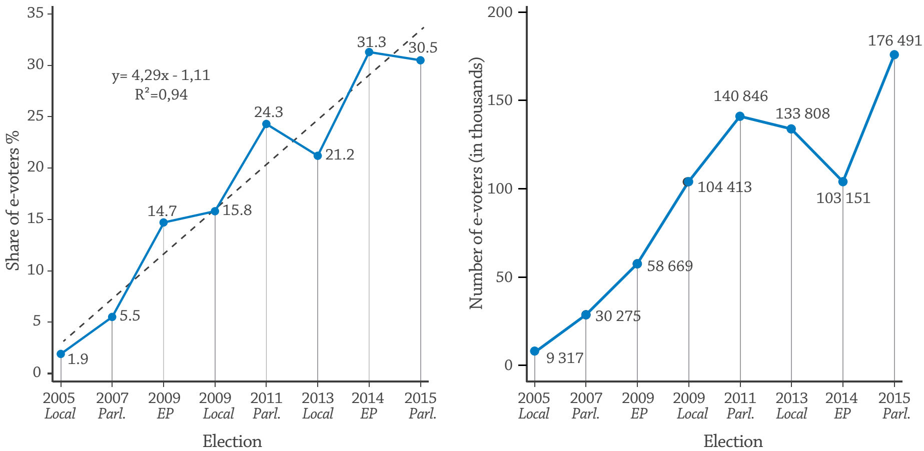

Figure 1.1: The relative share and absolute number of e-voters

at least in Estonia's case, an exponential pattern, but rather a linear one. This means that the conversion from paper-ballot voters to e-voters was almost constant over time, i.e. there were no rapid growth periods at certain thresholds. Second, the growth of e-voting similar to other technological innovations required sufficient time before it started to spread and appeal to the masses. In other words, due to its slow take-off pace at the beginning, governments adopting e-voting practices should not decide immediately after the first few internet voting trials on whether or not to continue to offer internet voting and whether it appeals only to a homogeneous subpopulation of technology enthusiasts or does it also attract voters that are less savvy with computers? Research has shown, and we dedicate a whole subsection to this topic below, that at least three elections that include internet voting are required before this new voting technology starts to diffuse among the electorate and engage voters from heterogeneous backgrounds.

This, and several other behavioural questions, bear relevant policy implications for countries investigating how to implement internet voting in their respective context. Using evidence from official election data, e-voting system log data and individual level survey data, this book addresses the following behavioural issues regarding internet voting:

1. Who are the internet voters, how do they differ from the general electorate, and how has their profile changed over time?

2. What is the impact of internet voting upon voter turnout?

3. What is the impact of vote verification on voter trust towards the system of internet voting?

4. Is internet voting politically neutral, or does it induce political bias?

In addition to behavioural consequences, we were interested in the contextual, institutional and historical background of Estonian internet voting in order to understand the key preconditions that led to the successful implementation of internet voting. Among others, we ask:

1. What were the technical and institutional preconditions for Estonian internet voting?

2. How did internet voting evolve in Estonia?

Setup of Estonian e-voting

The feasibility of e-voting in Estonia is based on the widespread internet penetration and use of digital ID cards. These credit card size personal identification documents allow citizens and residents to digitally sign documents and use private and governmental online services that require secure authentication. They also allow citizens to cast legally binding digital votes highly securely. Participation in the electronic ballot requires a computer with an internet connection and a ”smart-card reader”. Card readers are available for less than 10 euros at computer shops and supermarkets. Citizens may also access e-voting in public libraries or community centres, in fact any place with a secure internet connection. As of 2011, citizens can also electronically identify themselves with a so called ”Mobile-ID”, which requires a special mobile phone SIM card with security certificates and two pin codes. With Mobile-ID setup citizens can officially identify themselves using only their cell phone. The ID card is however still the most widespread method of digital identification.

E-voting is available during the advance voting period via a website hosted by the Estonian National Electoral Committee (2005-2011). In order to vote online, people are required to insert their digital ID card into a smart reader connected to an internet equipped computer. Next, they need to download a voting app which is a standalone program for Estonian e-voting. Using their ID-card and a four-digit pin (PIN1), the user has to first identify themself to the system, after which the system checks whether the voter is eligible according to age and citizenship to vote in the election. If affirmative, the e-voting system displays the list of candidates in the voter's district (Figure 1.2).

Voters can then browse the list of candidates and decide for whom to vote for. In order to cast an e-vote, the voter has to choose a candidate and provide a separate five-digit pin (PIN2) to vote. When certified correctly, the electronic vote is cast and sent to the server where it will be counted at an appropriate time, i.e. as prescribed by the procedures for online voting (Figure 1.3). [1]

Figure 1.2: Screenshot of the list of candidates displayed to the user after their eligibility has been checked using the authentication part of the digital ID (PIN 1).

The technical setup of the internet voting system is derived from the traditional way a person votes from outside of the polling district of their

[1] Refer to `http://vvk.ee/public/dok/E-voting_concept_security_analysis_and_measures_2010.pdf`

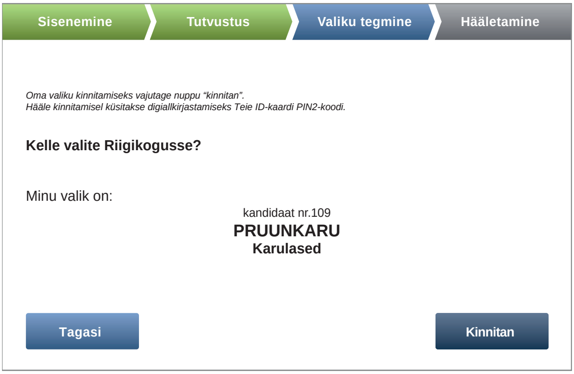

Figure 1.3: Screenshots of confirming the vote choice by using the transactional part of the digital ID (PIN 2).

residence, i.e. the postal voting. With postal voting, a two-envelope system is used to cast a vote. The inner envelope contains a ballot with the voter's vote choice, but has no identification markings. The outer envelope contains the voter's identification information. When sent to the ballot station, the information on the outer envelope is used to verify the voter's eligibility to vote and if confirmed, the inner envelope will be separated from the outer envelope and put into the ballot box for counting.

The system of internet voting in Estonia works in a similar fashion (see Figure 1.4 for a graphical representation). The downloaded e-voting app encrypts the vote (PIN1). The encrypted vote can be regarded as the vote contained in the inner, anonymous envelope. After this the voter gives a digital signature to confirm their choice (PIN2). By digitally signing the vote, the voter's personal data or outer envelope is added to the encrypted vote. Before the ascertaining of voting results during the evening of the Election Day, the encrypted votes and the digital signatures (i.e. the data identifying the voter) are separated. Then the anonymous e-votes are ”opened” and counted. The system opens the votes only after the personal data is removed.[2]

The first five Estonian elections were reasonably similar for the user-end.

[2]Source and more information available at: http://vvk.ee/voting-methods-in-estonia/

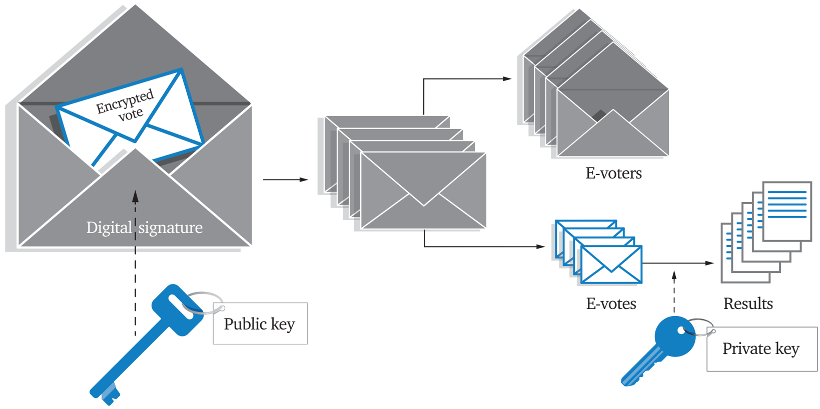

Figure 1.4: The two-envelope concept behind the Estonian internet voting system

The only marked difference was the length of time during which e-voting was available: three days in 2005 and 2007, and seven days from 2009 onwards. As of 2009 e-voters needed to download a voting program instead of voting via a web-embedded application. In 2013, a vote verification feature was introduced to the e-voting system that allowed voters to verify whether their electronic vote was received when cast using a smartphone or tablet. Other than these differences, in most other ways the eight e-enabled elections have been reasonably similar, providing a valid baseline for comparison of the dynamics of user behaviour between elections. On the technical side, e-voting only requires internet access and a minimum level of computer literacy, both of which are not universal in Estonia. However, the act of voting itself is no more difficult than other online activities, such as banking or shopping.

Secrecy of the vote and vote verification

An often-debated issue in terms of internet voting is the question of how to ensure vote secrecy in unsupervised environments. Because internet voting does not ensure that voters cast their votes alone, the validity of internet voting must be demonstrated on other grounds.[3] To ensure that the voter is expressing their true will, they are allowed to change their electronic vote by voting repeatedly (electronically) during advance polls or by voting at the polling station during advance polls.[4] This mechanism ensures that the vote buyer or coercer will not know for sure which ballot will be eventually counted rendering vote buying or coercing meaningless.

Following concerns regarding secrecy and security, the Estonian Electoral Committee established the following principles to which the internet voting systme must adhere:[5]

- **Time framework of e-voting:** e-votes may be cast during seven days, from the 10th until the 4th day before the Election Day.
- **Possibility to recast an e-vote:** during the e-voting period a voter can e-vote as often as they wish, but only the last e-vote is counted.
- **Primacy of ballot paper voting:** if a voter who has already e-voted goes to the polling station during the advance polls and casts their vote using a paper ballot, then the e-vote is cancelled. After this, the voter cannot recast their vote electronically or using a paper ballot.
- **Similarity of e-voting to regular voting:** e-voting adheres to the election acts and general election principles and customs. Thus, it is uniform and secret, only eligible voters may vote, every person may cast only one vote and it should be impossible for voters to know which way someone voted. The collecting of votes must be secure, reliable and verifiable.
- **An e-voter shall vote themself:** Using another person's ID card (or mobile-ID) for voting and transfer of the card's PIN codes to another person is prohibited. In order to avoid security risks, only a trusted

[3] https://www.ndi.org/e-voting-guide/internet-voting

[4] Source and more information available at: http://vvk.ee/voting-methods-in-estonia/

[5] Retrieved from: http://vvk.ee/voting-methods-in-estonia/

computer should be used, e.g. either owned by the voter or by a person the voter trusts.

The most recent and technologically advanced response to security concerns is vote verification. Piloted during the local elections of 2013 and fully implemented from the 2014 European elections onward, vote verification enables Estonian e-voters to verify whether their vote was cast as intended. Effectively, vote verification makes it possible to detect whether the computer is infected with malware that changes the e-vote or has blocked an e-vote. The process of vote verification involves the usage of a smart device (a smartphone or a tablet) equipped with a camera and internet connection. After the voting process a QR-code is displayed in the voting application and using a smartphone with a QR-code reader a vote verification app allows the voter to verify their vote. About 4% of all e-voters used vote verification for the last European Parliamentary elections in 2014.

Impact upon turnout

When the first pilots of internet voting were conducted during the early 2000s, the implicit hope was that the modernization of voting technologies would counter the declining levels of voter turnout in Estonia. However, this argument was never at the forefront of the political agenda for introducing e-voting. Rather, it was seen as an additional means to increase the convenience by which citizens can participate in political life and therefore constituted an extension of an already started motion to develop modern e-governance. Indeed, in many ways participation in an election over the internet manifests itself as the ultimate form of digital governance.

Still, an investigation of the turnout patterns in Estonian elections before and after the introduction of internet voting in 2005 is merited in order to assess whether turnout has changed or not. Because elections at different levels vary in terms of their salience and thus, turnout, we separated the turnout rates for national, local and European elections and display their trends over time in Figure 1.5.

With respect to national elections the general trend until 2005 was a de-

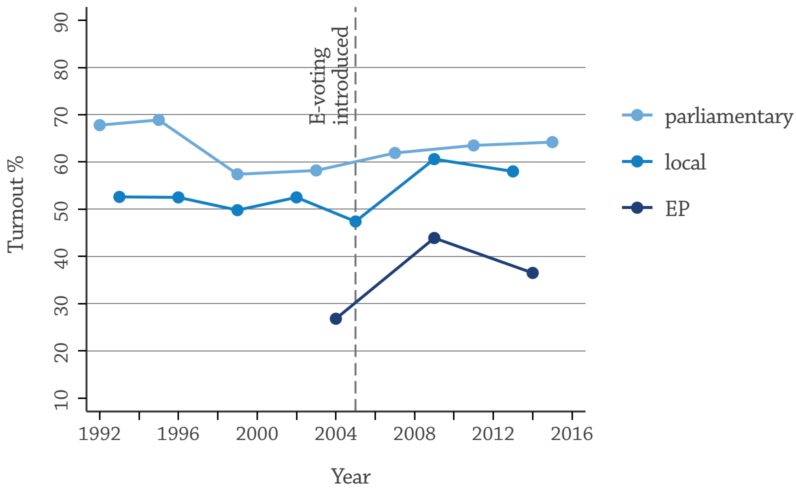

Figure 1.5: Turnout levels in Estonia, before and after the introduction of internet voting

clining one. It has often been argued that the first formative elections since the restoration of independence were unusually high due to the high sense of political efficacy among the citizens and therefore the decline afterwards was to be expected. Yet, from the 2007 national elections onwards, i.e., the first elections with the option of online voting, the turnout has incrementally, but steadily increased.

The same holds true for local elections, though with an even more pronounced pattern. Unlike the national elections, the turnout in local elections was almost constant up until 2002 (hovering at around 52%). Since 2005 however the turnout level has risen about 10 percentage points.

Regarding the European Parliamentary (EP) elections, we have refrained from making any inferences. This is primarily because we only have data for three elections and EP elections are clearly second-order in terms of saliency to the electorate throughout in Europe.

In light of this we can clearly say that turnout levels improved after the introduction of e-voting in Estonia. However, we advise caution in interpreting these findings in causal terms. It is not clear whether the modernization of voting technologies or some other relevant event or societal process has led to improved levels of turnout.

1.2 Structure of the book

The first two chapters of the book start off by setting a scene about the context and setup of Estonian e-government ecosystem, its foundational components, applications and outcomes. In the first chapter we outline primary prerequisites that make the e-government possible and explore some of the most relevant outcomes in terms of users, connected institutions, data repositories, etc. The second chapter looks in detail into the history of internet voting, its emergence and early political debates. It also outlines important legal aspects of internet voting that govern its normative environment.

From the third chapter onwards we delve into behavioural aspects of internet voting. We being by looking at the diffusion patterns of Estonian internet voting and ask "What is the profile of a typical e-voter and how has it changed over time?". The fifth chapter is dedicated to anonymized internet voting log files and further explores the notion of a typical internet voter and related characteristics. Sixth chapter investigates the impact of internet voting on individual level mobilization and posits an empirically validated mechanism for why internet voting fails to substantially increase voter turnout. In seventh chapter we show that internet voting has the potential to lower the cost of electoral participation. Eight chapters provides evidence to the 'stickiness' concept of internet voting. It means that internet voting as compared to regular voting is more habit-forming and that people who begin to e-vote hardly every switch back to paper-based voting or abstention. In ninth chapter we explore how the vote verification has influenced trust toward the system of internet voting. In the final empirical chapter we investigate whether internet voting is politically neutral.

Ultimately, the eleventh chapter summarises our main findings, provides a range of policy recommendations and outlines the future developments of Estonian internet voting. The book also features an extensive list of technical appendices for those interested in our inferential strategy. As the main purpose of this book was to provide an evidence based analytical account on behavioural mechanisms of Estonian internet voting, it is mandatory that our analytical strategy and empirical choices are transparent and available for scrutiny.

Chapter 2

The Estonian e-government ecosystem

Kristjan Vassil

Estonian e-governance is an intertwined ecosystem of institutional, legal and technological frameworks that jointly facilitate independent and decentralized application development by public and private institutions to offer public services digitally.

The most crucial components of Estonian e-governance are the digital identification of citizens, a digital data exchange layer and ultimately, a layer of applications developed by different public and private institutions. Figure 2.1 summarizes the role and function of each of the components of Estonia's ecosystem of e-government.

The critical component of any functional e-governance system is the digital identification of citizens and residents. However, digital authentication in itself requires several institutional, legal as well as societal preconditions. First, an institutional setup needs to address the issues of national identifi-

Academic reference: "Estonian E-Government Ecosystem: Foundation, Applications, Outcomes" by Kristjan Vassil. Background paper for the ***World Development Report 2016***, World Bank, Washington, DC.

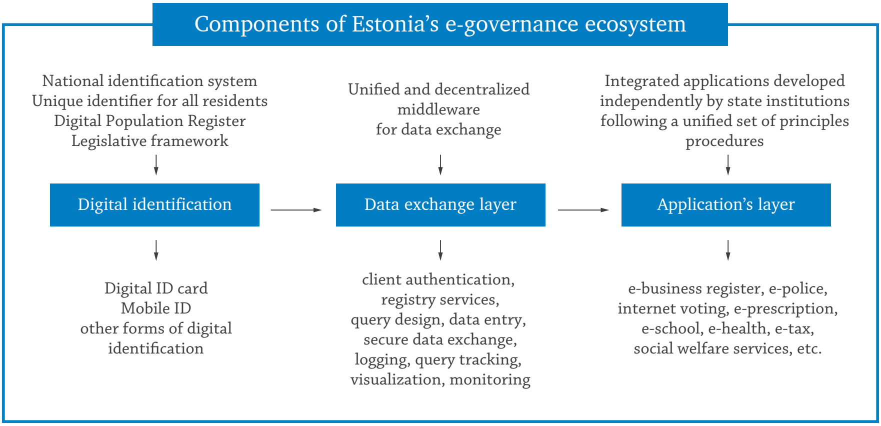

Figure 2.1: Components of Estonia's e-governance ecosystem

cation system where a single numeric identifier is used for the entire population (instead of many based on different services). Additionally, personal data should be stored in an electronic data repository, with the Estonian equivalent being the Estonian Population Register. And finally, a normative environment should explicitly regulate the process by which institutions, individuals and companies can request and receive access to any information stored in government databases.

Another functional component is the unified de-centralized data exchange layer that ensures standardized mechanisms for data gathering, structuring and storage, and any subsequent application development. Moreover, effective e-government would profit from a middle-ware system that would minimize - similarly to the Estonian X-Road - repetitive data collection, improve interconnectedness of the state's databases and avoids the time-consuming dealing with paper data entry and verification.

Finally, when digital identification and a data exchange layer are provided by the state, different institutions should be given the opportunity to develop an extension of their own services into the digital realm. Furthermore, if developed, the ecosystem of eventual services would facilitate a growth in usar numbers. The latter will require a substantial amount of time, about 5–7 years in case of Estonian. By implication, countries should not decide immediately after making their first steps into e-governance about its efficiency and impact. As the Estonian evidence shows, it is only after a first few years of intensive work that growth in usage, efficiency and impact will surface.

2.1 Foundation of the ecosystem: digital identification

Estonia's success making their public services available online is first and foremost based on the widespread use of electronic identification cards (see Figure 2.2). Since 2002 about 1.2 million of these credit-card sized personal identification documents have been issued, allowing citizens to digitally identify themselves and sign documents or perform actions.

ID-cards are compulsory for all citizens and they are equally valid for digital and physical identification. Due to their convenient size (unlike a passport they fit into a regular wallet) they are often used as the only identification document that people carry around. Physically, they are valid for identification in Estonia, but more importantly, they are also valid for travel in most European countries. Thus, in addition to their primary functionality – digital identification – ID-cards can be effectively used as replacements for traditional identification documents.

Figure 2.2: Estonian electronic ID-card

The digital functionality of the ID-card is based on an electronic chip and the two pin codes supplied with the card. By using a smart card reader and a computer connected to the internet, citizens can use two core functionalities provided by the ID-card, both of which are essential to the development of e-government, i.e. personal authentication (related to the four-digit PIN1) and digital signatory (related to the five-digit PIN2).

The first pin-code allows citizens to authenticate their identity to the corresponding e-service. This is a first step that provides a basic enabling infrastructure to furnish personalized services and information via online means. Many services run entirely on an authentication only basis, i.e. reviewing individual health records, checking the validity of one's car insurance or reviewing the list of political candidates in a voter's district. The second pin-code is used to sign documents or approve transactions online. For example, acquiring an insurance policy, confirming the submission of a tax declaration, or casting a vote in an election.

Functionally it is important to distinguish between authentication and signing as they enable different kind of services. Internet voting is perhaps the best example to illustrate the difference. When voting online, citizens download a voting app to their computer and upon a request from the system have to first identify themselves using the ID-card and the first pin-code. Next, the system checks whether the voter is eligible to vote in these elections and if the answer is affirmative, displays a list of candidates in their district. No digital signature has thus far been required. However, in order to cast an e-vote, the second pin-code – the signing function - is used to confirm the voter's choice. The latter is the transactional part of the citizen-state communication. When performed correctly, the electronic vote is sent to the server and will be counted at the appropriate time as prescribed by the procedures for online voting.

The difference between authentication and transaction is pivotal, as in virtually all aspects of public e-services users are required to use the digital ID either for authentication, a digital signature or both. A very similar example is applicable to online banking, where customers first use their PIN1 to enter their account, check the balance, browse their assets and so on.

However, should they intend to make an actual money transfer (a transactional part of the service), PIN2 is used. The major institutional precondition for digital identification is the national identification system, which helps to uniquely identify Estonian residents; and the Population Register that is the largest central data repository for personal data and family events.

2.2 Legal framework and data protection

In addition to technology and architecture, the Estonian e-government ecosystem is strongly regulated by legal instruments that provide a framework for security and protection of the personal data stored within the Population Register and other relevant government data repositories. Jointly these norms regulate the process by which institutions, individuals and companies can request and receive access to information stored in government databases and thereby build new public e-services by using the information already stored in the state's databases.

The legal framework is designed to work seamlessly with the technological solutions of e-government. For example, when new public e-services are developed it is legally not permitted to design systems that store the same data in different repositories. In practical terms this means that if a citizen's age is stored in the Population Register, it will be retrieved automatically for checking their eligibility for e.g. voting or driving, but not collected additionally by the system of internet voting. A simplified example though it may be, it shows how information stored in one repository can be reused by another. Moreover, Estonia's Public Information Act [6] prohibits establishing separate databases for the collection of the same data. In practice it means that state institutions cannot repetitively ask from a citizen the same personal information if it is already stored in any of the data repositories connected to the X-Road. This is an example of interconnectedness between enabling technologies and regulatory acts designed to work for a common goal, i.e. better citizen-sate interaction.

[6] https://www.riigiteataja.ee/en/eli/522122014002/consolide

These norms are most relevant with regard to Estonian e-governance and jointly provide the foundation for the entire range of application development, data protection and security issues in the realm of e-governance are listed in Table 2.1.

Table 2.1: Legal norms that provide the institutional background

Act or Decree	Brief summary
Personal Data Protection Act (1996)	The aim of this Act is to protect the fundamental rights and freedoms of people regarding the processing of their personal data, above all the inviolable right to a private life. This Act provides: 1) the conditions and procedures for the processing of personal data; 2) the procedure for the exercise of state supervision upon the processing of personal data; 3) liability for the violation of the requirements for the processing of personal data.[7]
Public Information Act (2000)	The purpose of this Act is to ensure that every person has the opportunity to access information intended for public use, based on the principles of a democratic and social rule of law and an open society, and to create opportunities for the public to monitor the performance of public services.[8]

Continued on next page

[7]https://www.riigiteataja.ee/en/eli/512112013011/consolide

[8]https://www.riigiteataja.ee/en/eli/514112013001/consolide

Table 2.1 continued from previous page

Act or Decree	Brief summary
Population Register Act (2000)	This Act provides for the composition of data in the population register and the procedure for the introduction and maintenance of the population register, the processing of data and access to data in the population register, the entry of data on residence in the population register and exercise of supervision over the maintenance of the population register.[9]
Digital Signatures Act (2000)	This Act provides the conditions necessary for using digital signatures and digital seals, and the procedure for exercising supervision over the provision of certification services and time-stamping services.[10]
Electronic Communications Act (2004)	The purpose of this Act is to create the necessary conditions for the development and promotion of electronic communications networks and electronic communications services without giving preference to specific technologies and to ensure the protection of the interests of users of electronic communications services by promoting free competition and the purposeful and just planning, allocation and use of radio frequencies and numbering.[11]

[9] https://www.riigiteataja.ee/en/eli/516012014003/consolide

[10] https://www.riigiteataja.ee/en/eli/530102013080/consolide

[11] https://www.riigiteataja.ee/en/eli/ee/Riigikogu/act/501042015003/consolide

2.3 Usage of digital ID

Given the fact of how fundamentally digital IDs are linked to online public services, our next question was to investigate how frequently they are used and what is its pattern of diffusion in society? In order answer this question, we acquired aggregate log-file data on the usage of digital IDs over time. [12]

The digital ID project started as early as 1998, when Estonia sought solutions as to how to digitally identify their citizens. By 1999 a viable project in the form of the current ID-card was proposed and the legal framework to enable digital identification was set up over the following years. In 2000 the Identity Documents Act and the Digital Signatures Act, the two most important bills that regulate the use of digital IDs, were passed in parliament. The first states the conditions to which an ID-card must adhere, but most importantly states that the ID-card is compulsory for all Estonian citizens. The latter Act states the conditions for a state-governed certification registry, which is fundamentally linked to the functioning of the digital ID-card system.

Following these events, the first ID-cards where issued in January 2002. Since then about 1.24 million digital ID-cards have been issued. By June 2015 digital ID-cards had been used about 353 million times for personal identification and 222 million times for digital signatures. The average annual growth rate over the 11 years (from 2003 to 2015) amounts to 7.4 million authentications, and about 3.5 million signatures per year (notice however that the growth is non-linear so the average growth rate should be interpreted with caution), see Figure 2.3. In other words, the growth rate for digital identification is about two times more rapid than for signatures. But the growth in ID-card users has been all but swift and instantaneous.

When the project started in 2002 the electronic use of digital ID's remained low for almost five consecutive years. As can be seen from Figure 2.3, the situation changed in late 2007, after which the usage of ID-cards started to grow rapidly and reached an all-time high in 2015. Why did it

[12] We thank Liisa Lukin and Tanel Kuusk of the Certification Centre for their assistance and help in acquiring the usage data on digital IDs.

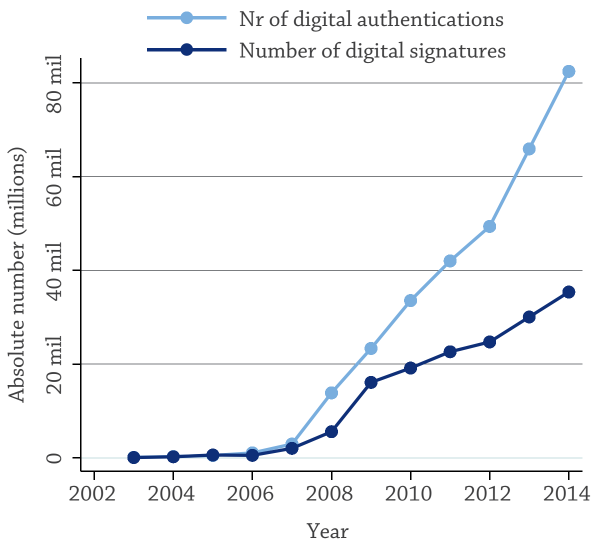

Figure 2.3: Growth of digital authentications and signatures over time (from 2002 until 2014). Source: Certification Centre

take so long and what triggered the sudden growth in 2007?

In order to respond to this question, we need to understand the socio-economic context in which Estonia was in during the early 2000s. At the beginning of the new millennium, less than one third of Estonians had used the Internet. A study about digital divide carried out in 2002, characterized Estonia as a country with relatively few internet users, limited access to computers and a growing, but still insufficient number of public internet access points (Kalkun and Kalvet, 2002). Most people still could not afford computers nor internet connection in their homes. However, the major obstacle to making use of these new technologies was not the lack of infrastructure *per se*, but the lack of skills and motivation of the general populace (ibid). As noted by the authors of the study:

> *The main barrier in Estonian society is the fact that the possibilities offered by the Internet are not associated with personal needs. It is believed that "computers are not for me"* (Kalkun and Kalvet, 2002, p. 9)

Indeed, the early 2000s characterized Estonia as a society with a gaping digital divide, i.e. an effectively disconnected society. The low usage of ID-cards throughout the 2000s was reflected by this situation.

Proactive policies

Although the problem of a digital divide is primarily a societal one, the solution was provided by a rather unexpected source: the private sector. Namely, the low number of internet users and paper-based means of identifying turned out to be an even larger headache for the banking and the telecom industries. Indeed, it was the banking sector that wanted to replace the physical code cards[13] with a more reliable and secure way of identifying their customers; and it was also in their interest to move the bulk of financial transactions out of their physical offices and into the digital realm. Because the banking, as well as the telecom sector, had prospered in comparison to other industries at the time, they also had the financial means to support a larger vision of how to help raise awareness of modern information and communication technologies amongst society as a whole.

Several coinciding events, comprised of the state's involvement in the development of ID-cards, the telecom and banking industry's concerns regarding reaching out to the society and previous attempts to digitize the country's educational infrastructure (the Tiger Leap project[14]) provided fertile ground for one of the largest public-private partnership projects in Estonia to this day: the Look@World project. The goal of this project was ambitious, yet simple: to promote the spread of internet use among the Estonia population. With four founders from the private sector (the two largest banks and two

[13] A card that lists dozens numbered codes, usually four to six digits long, one of which the user has to enter in the identification process. Which code exactly needs to be entered is alternated between different sessions.

[14] Tiger's Leap was a governmental project that started in 1997 with the goal to substantially increase the investments into the development and expansion of personal computers and network infrastructure in Estonia, with a particular emphasis on education. The primary outcome of the project was the provision of Internet access and computer labs to all Estonian schools.

telecom companies) and the government, the consortium initiated a 40 million Estonian kroon project (equivalent to 2.5 million euros) over the course of two years (2001–2002). As an outcome of this massive partnership, more than 100 000 individuals (i.e. about 10% of Estonian adult population), were taught to use internet and communication technologies (ICTs). The project raised the number of public internet access points from 200 in 2001 to about 700 in 2004. Through the years, the consortium worked hand-in-hand with the public sector and in 2001 the Look@World consortium members agreed to facilitate the widespread use of ID-cards, with the private banks given the right to issue digital ID-cards in their offices from the following year.

The involvement of private banks was pivotal with regard to the success of the ID-card, both regarding societal awareness and the actual distribution of cards. First and foremost, their contribution lies in the unconditional support for the ID-card infrastructure as the primary mechanism to identify their customers. If the banks had not made such a strong case for the support of the ID-card and continued to advertise the new system, the transition would have taken considerably more time. More importantly, it shows that the spread of a new identification mechanism is strongly related to the number of services that it makes available to people.

In particular, when people realized that their banks preferred digital-IDs for personal identification, that it was more secure and convenient, they had a real motivation to replace their old identification methods with the digital-ID. When income tax declaration was moved online, the Estonian Tax and Customs Board promised a swifter review of declarations and (if applicable) a quick tax return. In fact, tax returns for those using e-declarations was less than a month, rather than the typical 3–6 months period for those who submitted their tax declarations offline. Such incentives, alongside the growing number of new e-services, provided a sustainable environment for the increased use of digital identification.

In other words, by using the new digitized public services, people gradually learned that using their digital identity cards provides them with a quicker and more efficient channel to interact with private and governmental services. Growth in digital-ID usage does not therefore operate in isolation, but is clearly dependent on the amount of services rendered to the public for which digital-IDs can be used. Consequently, this explains the low take-off pace during the early years of the ID-card (see Figure 2.3).

In order to demonstrate that usage of ID-cards is indeed closely related to the number of services provided, we also acquired data from the Estonian Information Authority for a number of data repositories used to render public e-services. Figure 2.4 shows a strong positive association between the number of services (as measured by the number of data repositories) and the number of personal digital authentications. The time span covered in the data covers the period between 2003 and 2015.

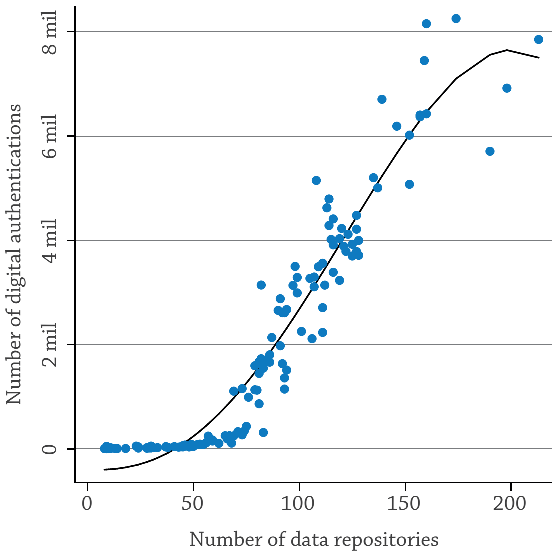

Figure 2.4: Relationship between the number of data repositories and digital authentication.

As one can see from Figure 2.4 until there were 75 data repositories online the growth of digital authentications is very slow and almost linear. Further data examination showed that the adding of the 75th data repository to the Estonian e-governance ecosystem in late 2007 coincided with a significant increase in the number of digital authentications (see also Figure

2.3. From Figure 2.4, we can additionally infer that it takes considerably more than just one, two or even dozen data repositories to be connected to the ecosystem of e-government before the spread of digital identification significantly picks up.

We believe that the lag effect in ID-card usage is applicable to many other domains of e-governance, and thus constitutes a more universal law on the diffusion of technology in the public sector. Technological innovations do not jump from no-usage to full-usage overnight. Instead, their spread is slow at the beginning, but when managed appropriately, after a certain period their usage takes off and turns into rapid, often exponential, growth.

However, digital identification is only one necessary precondition for a functioning e-government. The second, as relevant as a digital-ID, is related to enabling technological infrastructure, i.e. to provide a secure and reliable technological environment where different state institutions are incentivized to work together, share their data and develop online services. Estonia's solution to this fundamental question was to develop the X-Road, a data exchange layer that makes it easy for any institution - public or private - to make their services available online. In the next section, we look at the functional part of X-Road and using the data collected by it, estimate its economic impact.

2.4 The X-Road

As technology is the primary enabler of e-governance, the critical question is how to ensure secure communication between scattered governmental databases and institutions that use different procedures and technologies to deliver their services. Estonia's solution to this problem was to develop the X-Road, a secure internet-based data exchange layer that enables the state's different information systems to communicate and exchange data with each other.

X-Road serves as a platform for application development whereby any state institution can relatively easily extend their physical services into the electronic environment. For example, if an institution, or a private company for that matter, wishes to develop an online application it can apply to join the X-Road and thereby automatically obtain access to any of the following services (amongst others): client authentication (either by ID-card, mobile-ID or the banks' internet authentication systems), authorization, registry and query design services to various state managed data depositories and registries, data entry, secure data exchange, logging, query tracking, visualization environment, and central and local monitoring. These services provide vital components for the subsequent application design. Therefore, X-Road offers a seamless point of interaction between those extending their services online and different state-managed datasets and services. The conceptual logic of the X-Road is depicted in Figure 2.5.

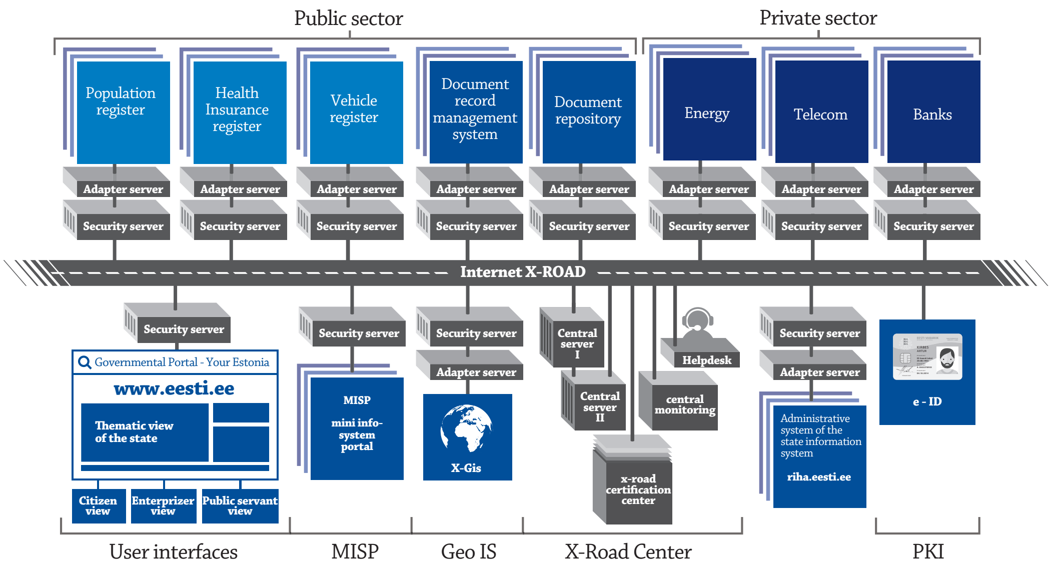

Figure 2.5: Schematic structure of X-Road as a data exchange layer between governmental and private services

Another important feature of X-Road is its decentralized nature. X-Road is a platform, an environment, for efficient data exchange, but at the same time it has no monopoly over individual data repositories that belong to the institutions that join the X-Road. Moreover, by its very design, X-Road requires every joining institution to share their data with the other members if required and necessary. As such, every joining institution and developed application can use the data stored in other institutions' repositories and is even legally encourage to do so in order to avoid repetitive collection of client information. Because data sharing enables the development of more convenient services than those institutions would be able to pull off single-handedly, the system implicitly incentivizes the reuse of data. The incentive works because such a collective process allows for a seamless and more efficient user experience and thus increases the interest from both state institutions to develop digital services and for individuals to reach out to the state.

In addition to citizen-state interactions, X-Road is particularly suitable for queries involving multiple agencies and information sources. For example, checking vehicle registration data requires data retrieval from the population registry and vehicle registry, i.e. two otherwise unconnected data repositories. According to the State Information Authority, the conventional offline approach would require three police officers working on the request for about 20 minutes. With the X-Road, the entire information retrieval is conducted by one police officer within seconds. At the same time, citizens are not even required to carry their driving license or the car's registry documents around, as the information system that the police use, displays the status of these documents based on the driver's ID-card or licence plate number in real time.

Similarly, but in more complex terms, Figure 2.6 shows how many interfaces and datasets are currently jointly working to provide parents with options to apply for parental benefits.[15]

As one can see, the citizen interacts with the government through the Citizens' portal (which is the access point to most governmental e-services); the civil servant (if at all) works through the mini information system portal (MISP) and the X-Road provides access to all the relevant data repositories

---

[15] Retrieved from: https://www.ria.ee/public/x_tee/Xroad-technical-factsheet-2014.pdf

needed for one goal, i.e. processing a parental benefit application.

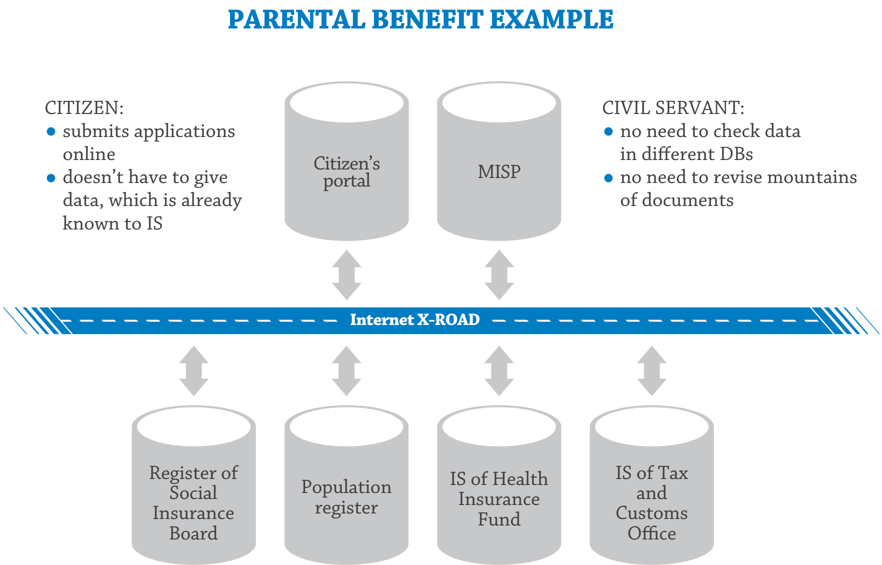

Figure 2.6: The parental benefit application system

Naturally, this open design is accompanied by rigid security measures: authentication, multilevel authorization, high-level log processing and monitoring, encrypted and time stamped data traffic (i.e. the basic functionalities that are covered within the very structure of X-Road).

As a result, Estonian state institutions are structurally incentivized to join the X-Road, simply because they can design services that would not be as efficient or convenient to develop and maintain individually.

2.5 The impact of the X-Road

Our next task was to understand the potential impact of X-Road. In other words, to understand how many institutions have actually joined the X-Road to render their services online, and more importantly, how many actual services are provided through X-Road to how many end-users?

In order to respond to these questions, we acquired the relevant time-variant data on X-Road usage numbers from the Estonian Information System Authority. We began with the question of how many services are offered via X-Road and from how many institutions. Figure 2.7 displays relevant growth rates for services and institutions using X-Road since 2003. As a reference, we also plotted the number of data repositories upon which those services run, but we will return to this feature later on.

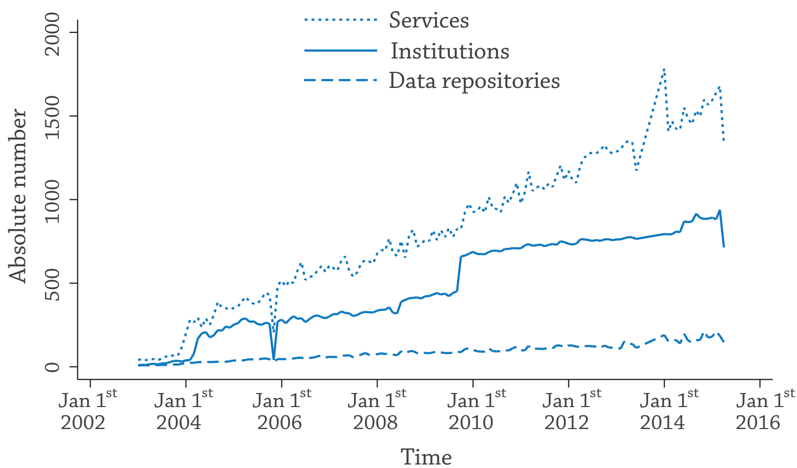

Figure 2.7: Number of institutions services and data repositories connected to X-Road

Let us being by considering the curve showing the growth of services over time. Services include both the actual end-products that citizens can use, but also those services that, for example, public servants use in their day-to-day work. The latter may include a policy officer retrieving information from the population registry to check one's vehicle registration, or a law office clerk checking your property ownership. Thus, the number of services does not reflect just the actual number of those services that we think of as part of e-government, but all the interactions between systems that exchange information via X-Road in order to run certain applications (the end-product).

We would prefer, naturally, to report the amount of actual applications available to citizens, but those are unfortunately not logged by the system, so we have to rely on a proxy measure: services. According to Figure 2.7 the total number of services grew from around 40 in 2003 to more than 1600 in 2015. The average growth of services amounts to about 123 services per annum.[16]

The growth in institutions joining X-Road follows a similar trend at a slightly lower level. X-Road had only about 10 institutions offering their services in 2003, however the number grew quickly to almost 900 by 2015. Notice that the number of institutions includes both public institutions at central and local government level and private companies. The share of public institutions was about 71% in 2014.

Also notice the two periods where the number of joining institutions grew very quickly (in 2004 and in 2009). These were the outcomes of two extensive communication campaigns directed to public institutions and citizens to encourage the use of the Estonian e-governance system. With every advancing year since 2003, on average 74 institutions joined the X-Road to render their services online.

Finally, we can also observe the number of data repositories that are connected to the X-Road. This is a very interesting characteristic to look at, because data repositories are the key building blocks for final products (the applications). To be sure, the number of data repositories does not equal the number of services, because one can build several services with one data repository and even more when interacting with multiple repositories.

Figure 2.7 demonstrates that the growth in the number of data repositories was quite modest when compared to the growth of institutions and services. Indeed, in 2003 only eight databases were connected to the X-Road. From there onward the growth has been stable but quite small, only about

[16]Linear estimation obtained by regressing the time on the number of services.

13 databases per year. The highest peak was in 2013, when 212 data repositories were connected to the X-Road. However, what makes the growth of databases more interesting than the other components of the X-Road, is their enabling nature, i.e. in themselves the dynamics of data repositories are not particularly interesting, but because they provide the foundation for application development, their association with actual usage of X-Road services is extremely relevant.

The next step in understanding how much time and money X-Road helps to save is to look at the number of actual end-users accessing services provided through the X-Road. Figure 2.8 illustrates the growth in end-users over time.

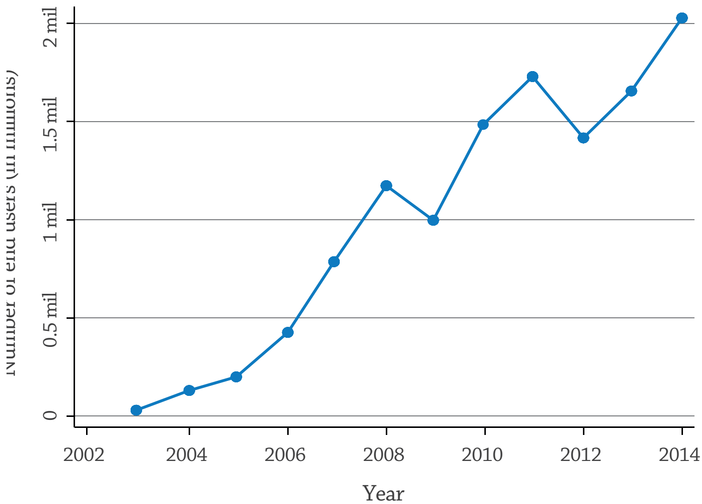

Figure 2.8: The growth in end-users of X-Road over time (from 2003 until 2014)

We observed a growth in end-users from around 30 000 in 2003 to slightly over 2 million in 2014. The annual estimated growth rate of end-users was about 150 000, which translates into an extra 1 300 end-users per month over the course of the last twelve years. The fluctuations in user numbers (sharp peaks in absolute numbers and waves fo smoothed averages) reflects the impact of public information campaigns that encouraged citizens to use specific e-services at specific time periods.

2.6 Economic impact of the X-Road

Our next task was to understand the economic impact of the X-Road. The best proxy measure to this end was the number of queries made within the X-Road system. Queries between different systems demonstrates the actual bloodstream of the X-Road, because every query is a point of interaction that substitutes a conventional service and thus has the potential to save time and money. Figure 2.9 demonstrates the growth in queries over time.

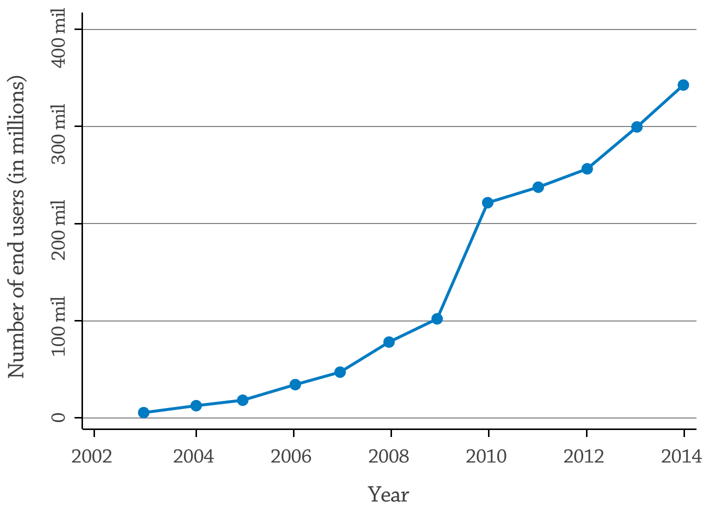

Figure 2.9: X-Road's bloodstream: the number of queries within X-road (personal and system queries combined)

There are two important inferences that can be made on the basis of this. First, as shown with the growth of ID-cards in earlier sections of this report, the number of queries did not increase much immediately after the kick-off of X-Road. Quite the contrary, it took several years before the number of queries started to visibly increase.

However, the growth, albeit with its slow start at the beginning, reached a point in early 2010 when the number of queries rapidly doubled that of the previous year. Mathematically, this is often referred to as a form of non-linear growth rate termed a "step function".

Next, an interesting association appeared between the number of queries and the number of databases connected to the X-Road. As mentioned earlier, databases are the nuclear entities of any effective e-government and allow for the building of e-services in the first place. The more data repositories connected to the X-Road, the more services that can be built. The question is how many actual queries one gets per connected database and how was the growth of data repositories related to the growth of queries?

These questions are relevant for any country looking at ways to digitize their public services with the aim of attracting the widest possible audience. Figure 2.10 provides the answer to this question. First, it shows that the number of data repositories was initially only modestly associated with growth of queries. Until the 50^{th} database the change in queries remained marginal. However, since the 50^{th} repository onwards, the growth became exponential.

Effectively, this is evidence that building an effective e-government does not happen overnight. Instead, it takes time and above all, hinges on the number of separate databases that are incorporated within a system such as the X-Road upon which institutions can subsequently start building their end-products for users. The discrete threshold for growth in Estonia appeared with the 50^{th} data repository.

Coming back to the nominal growth of queries made within the X-Road system, we can make inferences about the potential economic impact of Estonia's e-government ecosystem. According to the Estonian Information Authority's system, about one third of all queries are those conducted between individuals, i.e. citizens and public officials. We take this approximation as a point of departure and argue that every query (i.e. interaction within the Estonia e-government ecosystem) replaces a physical citizen-state interaction with a virtual one. We further argue that every interaction replaced with a virtual one saves time, both from the citizen's as well as the public official's perspective.

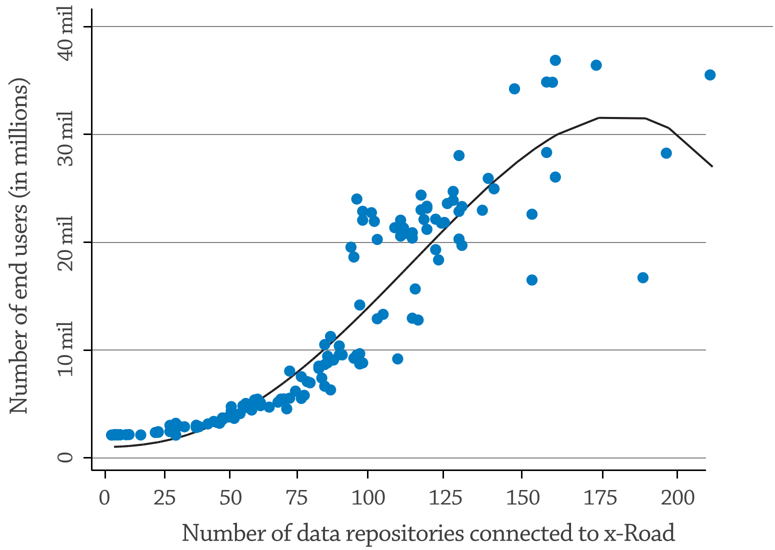

Figure 2.10: The relationship between the number of queries in X-Road and the number of data repositories connected to it

In order to calculate the plausible economic impact of the X-Road, we assumed that every query saved 15 minutes of a citizen's time. Considering that the time cost involves coming and going to the state institution, locating the relevant public official, booking a time and queuing or returning another day (if necessary), we are confident that the 15-minute time-saved period is a conservative measure. Next, using a time-saving of 15 minutes, we can further extrapolate the total time saved based on the number of X-Road queries for any given year. For example, in 2014 the annual number of human-to-human queries was approximately 113 million. This means that if every digital query saved 15 minutes, a total of 2.8 million hours or 3 225 years was saved. A comprehensive and intuitive interpretation of this number would posit that the time saved by X-Road in 2014 amounts to 3 225 people working continuously 24/7 for one year. Figure 2.11 displays the corresponding values for each year since 2002 and further illustrates the time-saving curves for alternative time saving scenarios, i.e., assuming that instead of 15 minutes X-Road only saves 5 minutes, or conversely as much as 30 or 60 minutes. Because this is the only assumption in computing the potential economic impact of X-Road and there is no opportunity for an empirical validation we believe that calculating alternative scenarios best describe the potential of Estonia's e-government ecosystem.

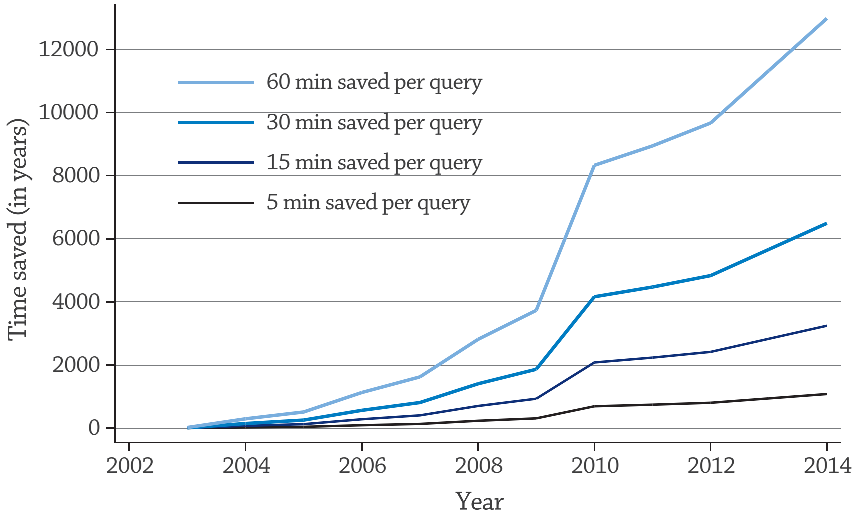

Figure 2.11: Time saved (in years) per year due to the availability of public e-services

Although there is no empirical validation available, we argue that it is reasonable to expect that the time saved by X-Road is somewhere between 5 and 60 minutes per query. In such a case, the representative mean value is in all likelihood close to 30 minutes and the corresponding economic impact can be interpreted accordingly. Note that the state Information Authority's (that administers the X-Road) entire annual budget for 2014 was about 14.2 million Euros, i.e. a fraction of what was actually saved in time for citizens in 2014.

Naturally, one should be cautious of taking this impact estimation at face value. Several factors for which no data are available can influence time saved in both directions. That said, we believe that it is a simple, powerful and fairly conservative assessment of the tangible impact of e-governance.

In the next chapters we depart from the X-Road as an enabling technical infrastructure for application development and explore one of the most tangible forms of citizen-state interaction in the realm of e-governance: internet voting.

Chapter 3

Historical development and legal aspects

Priit Vinkel

This chapter examines three interconnected topics. First, the discussions that took place regarding the constitutionality and legality of the Estonian e-voting system. Second, how the system has been developed over the years and third, what confidence building measures have been implemented over the course of this development in order to boost public trust in the system and ensure its continued viability.

3.1 The discussion over constitutionality

Before a first implementation of electronic voting, including remote internet voting, it is common to ask whether there is a need to change a country's constitution (Braun, 2006; Heindl et al., 2003; Rüß, 2000). Similarly in

Edited and reprinted with permission from Priit Vinkel "Remote Electronic Voting in Estonia: Legality, Impact and Confidence". Tallinn University of Technology Press, 2015. PhD thesis.

Estonia, adding a new voting method in addition to paper voting and more so, a fully remote way of voting, raised several questions in constitutional law.

In 2005 the Constitutional Chamber of the Estonian Supreme Court reviewed whether the process of e-voting was in accordance with constitutional principles, mainly with the principle of equality (uniformity) (Supreme Court, 2015). The President launched the case in 2005, just before the first e-enabled elections and the adoption of the redefined stipulations in the electoral laws. The central argument lay in the question of whether the e-voters' ability to change their vote by voting repeatedly electronically or once on paper would give them an unconstitutional advantage compared to the traditional on-paper voter.

A possible lack of legitimacy of an election results could stem from either of the following situations: that the privacy of the e-voting procedure cannot be ensured by the authorities. In which case, large-scale buying and selling of votes, as well as exercising other influence or pressure on a voter, could be possible. The people themselves cannot verify the e-voting results, and people need to have absolute faith in the accuracy, honesty and security of the electoral system (its organizers, procedures, software and hardware). For people who did not take part in developing the system, the computer operations could be verified only by knowing the input and comparing the expected with the actual output (similar to a black box). In a secret ballot system, there is no known input, nor is there any expected output with which to compare the electoral results.

Additionally, guaranteeing the freedom and secrecy of voting in an uncontrolled environment was examined during the review process. Based on the remote nature, one of the cornerstones of free voting - mandatory privacy of the voting process - is not possible in internet-based remote voting. The two sub-principles of secrecy of voting were analysed by the Supreme Court: privacy of voting and the anonymity of the vote. The court explained that to be constitutional, e-voting should have a "virtual voting booth", i.e. the possibility to change one's e-vote during the voting process so that the voter could change a coerced vote to one of their own choice at a later -

private - time. It is important to emphasize that the constitutionality of the internet as a communication channel, together with possible threats regarding anonymity and secrecy, was not analysed during that particular review and indeed has not yet been analysed by the Estonian Supreme court.

The secrecy of voting has traditionally been understood in Estonia, and elsewhere, as the right to cast the vote alone in a voting booth. In the case of e-voting, it is impossible to ensure the privacy aspect of the voting procedure. The voter's right to anonymity during the tallying of the votes can be guaranteed, indeed to the extent to which this can be secured in the case of remote postal voting (Kersting, 2004). Therefore, remote electronic voting required a rethinking of the privacy principle.

The principle of privacy is there to protect a person from any pressure acting against their free expression of a political preference. Such a teleological approach to the principle of privacy was the basis of the e-voting provisions from the very beginning of the whole project. Consequently, the provisions enabling e-voting are based on the premise that the government has to trust the citizen and avoid, whenever possible, interference with decision-making at the individual level. The voter has to be aware of the risks, and he or she has to have the right to decide whether to use the opportunity of e-voting. Therefore, e-voting cannot, under the same conditions, replace traditional paper voting and should be considered a complementary solution (Council of Europe, 2004). The 2005 ruling of the Supreme Court agreed with this position.

A second broader category of discussions on e-voting have taken place in the Constitutional Chamber of the Supreme Court, following on from specific electoral complaints. Complaints in Estonian elections (both on paper voting and on e-voting) can be issued via a fast-track appeal system, where institutions only have a limited period within which to reach a verdict (for electoral committees five working days, for the Supreme Court seven working days). In addition to the Supreme Court, appeals have to be scrutinized first by two lower tiers (county- and national-level) of electoral committees. Altogether, there are three tiers, so the maximum duration of dealing with an electoral complaint in all instances is about one month (Heinsalu et al., 2012). The principles of equality, secrecy, technical uniformity, procedural soundness and the security of e-voting have been raised in various different complaints. The potential effect of possible shortcomings on the overall election result is the overarching question that has to be analysed based on these complaints. By 2015, all of the complaints concerning e-voting had been dismissed. However, the complaints issued after the 2011 parliamentary elections had a strong influence on the parliamentary debates of 2012. Additionally, an issue that has arisen in these complaint debates is how to obtain applicable and sufficient evidence, which is by concept difficult, due to the anonymity of the vote. So far the Supreme Court has been quite innovative and liberal with regard to the e-voting electoral complaint judgements, however, always stating that the election organizers did their best in avoiding any malpractice.

3.2 Assessment of the constitutionality

On a broader note, the question of whether remote internet voting with binding results in public political elections complies with the constitutional principles of sound and fair voting cannot be answered with a simple "yes" a "no". As such, two sub-questions were proposed. The first sub-question was whether the legal norms in an abstract sense comply with the constitutional provisions of the state, and the second whether the technical solutions used to conduct e-voting procedures in a certain election guarantees constitutionality. The first sub-question can be answered based on theoretical analysis and could be decided upon in a constitutional review process, but the second should be examined before and after the actual elections. The fact that it is not possible to fulfil all of the theoretical and conceptual requirements set for an (originally paper-based) voting system is not enough for declaring e-voting to be unconstitutional. The second sub-question can be answered with "yes" only if sufficient measures are in place to check whether the IT solutions work properly. This leads to the requirement that auditing, verification and evaluation of the results be stipulated in law and electoral regulations.

In the case of Estonia, the legal norms comply with the constitutional provisions, because electronic identification enables secure remote identification, the digital ID-card has complete penetration, all advance voters (both electronic and paper) are placed under the same conditions, and the "virtual voting booth" (the right to replace an e-vote with another e-vote or a paper ballot) and the virtual double-envelope system ensure the freedom of anonymous voting and the uniformity of elections. Therefore, the answer to the first sub-question is also "yes". Moreover, the system is justified by the target to guarantee universal (general) suffrage in an information society where e-services (including e-voting) are demanded by a significant proportion of the electorate. Whilst formal equality can be provided, the questions of material equality (the access to computers and the internet) and the issue of the digital divide remains. In addition, complying with the principle of secrecy poses new obstacles for many countries. According to the teleological interpretation of the principle of secrecy, the act of voting is seen not as an aim but as a measure to guarantee freedom of voting, and the anonymity aspect of the principle of secrecy can be guaranteed. The analysis of the compliance of the Estonian e-voting system with the (ICCPR, 1976) has given positive results as well, but also emphasized the importance of special procedures to facilitate auditing and observation of e-voting (Meagher, 2008).

The answer to the second sub-question is more complicated. Internet voting in given elections is constitutional if the provisions of the law are fulfilled in practice: only people who are entitled to vote can vote, e-votes cast over the internet are recorded and tallied properly, and only one vote per voter is counted (OSCE/ODIHR, 2013). Independent IT auditing that covers all aspects of the system can prove its soundness. The proper performance of an IT system should be verified and audited before, during and after voting. Personal computers and the internet remain the weakest links in the system. Additional changes in 2012 introduced the first steps for individual vote verification with the Estonian system (see also Chapter 9) and therefore opened new possibilities to minimize the threats to personal computers. Nevertheless, remote online voting as a concept is never absolutely secure (the same applies to any voting method). Constant development of the system needs to be maintained to stay ahead of possible risks and threats. To date, the courts' answer to the second sub-question is a tentative "yes". Nevertheless, confidence and trust are the most important factors in judging the reliability of the system and they should be built and maintained by effective practical measures.

In conclusion, the 2005 constitutional debate has maintained its position throughout the years of e-voting implementation in Estonia. The principle of the "virtual voting booth" as a guarantee of freedom and the understanding of teleological voting secrecy have become the cornerstones of the Estonian system and are also adopted in other e-voting systems. Electoral complaints play an important role in highlighting possible challenges with the use of e-voting. During the first ten years, complaints on equality, secrecy, technical uniformity, procedural soundness and security of the system have been raised. However, no violations have been found. The constitutionality of an internet voting system can be assessed on the levels of general compliance with the electoral principles and the soundness of the implementation of the system in actual elections. The first-level question in the Estonian case could be answered positively, the system is in general compliance with the constitutional provisions. The answer to the second-level question in Estonia could also be seen in a positive light, but also depends heavily on the processes of verification and auditing. In addition, the appropriate measures to ensure security need constant upgrading and development.

3.3 The historical development

Let us turn from legal issues to the general historical development of the Estonian e-voting system. The expansion of e-voting in Estonia can be divided into three periods: 1) the setup and implementation phase; 2) the years of increasing participatory numbers and additional legal debates; 3) the introduction of verifiability and a stabilised use of the method.

The year 2002 marked the start of the setup phase, when a very general principle for remote electronic voting was stipulated under electoral law (LGCEA, 2002), allowing the election authorities to start with the project preparations, find a vendor for the system and prepare for the 2005 local elections. Legal debates on the topic were restarted in 2005 to broaden the regulations in the law (LGCEA, 2015). This period also held the discussions about the constitutionality of the system in the Constitutional Chamber of the Estonian Supreme Court (as discussed in the previous section). To test the features of the system a limited pilot was held in Tallinn during January 2005. The first e-enabled elections (for local government councils) were held in October 2005.

The second phase entailed a steady rise in user numbers and diffusion of the solution to the wider electorate. The legal stipulations had not been changed between the years 2005 and 2011. However, the technical solution was constantly updated for every election; the Mobile-ID support system and a new voter-application interface were developed for the 2011 general elections (Heiberg et al., 2012). The end of this phase is marked by a report by OSCE/ODIHR (2011), where several key features of the Estonian e-voting system and the regulations were revised as a result of recommendations made. This process was the main engine to launch renewed discussions in parliament to review the e-voting regulations and amend the procedures to bring more transparency and introduce additional steps regarding the ability of a citizen to verify their vote was counted correctly.

After the 2011 general elections, where almost a quarter of all votes were cast electronically, parliament decided to specify the norms of e-voting under electoral law in order to improve the legitimacy and transparency of e-voting. Until 2011, the electronic voting procedures had only very brief legislative regulations (despite the discussions in 2005). The parliament established a special working group (Constitutional Committee, 2011) which in addition to detailing procedures, had to propose a solution for raising levels of transparency and accountability with the e-voting system.

At the same time the technical community involved by Estonian National Electoral Committee (NEC)[17] in the discussions about the security and transparency of e-voting, came to the conclusion that a mechanism for the voter to

[17]NEC is the central electoral management body (EMB) in Estonia.

verify their vote was counted correctly was needed (Draft law 186SE 2012). The perceived aim was to detect possible malicious attacks on the e-voting system. The NEC has a better chance to discover attacks and react to those if e-voters, even a relatively small amount of them, verify their vote. If by using the verification system somebody finds out and reports to the NEC that their vote was not stored correctly, measures can be taken immediately (Heiberg et al., 2012). In addition, a second channel for verification had to be found, because if voters use the same personal computers for voting and verification, it only adds a limited amount of additional information regarding the voting computers. Therefore, an independent channel, such as a mobile phone or other mobile device, was introduced for verification purposes (Heiberg and Willemson, 2014).

In 2012 parliament adopted several amendments (Draft law 186 SE 2012) to the electoral law, stating that a new electoral committee - the Electronic Voting Committee - was to be created for the technical overseeingn of e-voting.

The first elections where this committee was in charge were the 2013 local elections. The law also regulates that before every implementation, the e-voting system must be tested and audited. The most significant change of the law was the statement that, from 2015 onwards, voters had to have the possibility to verify that their vote was received, stored at the central server of the elections and reflected the choice of the voter correctly.

The main lesson that can be learnt from this period is that together with the development of the technical environment, legal regulations also had to kept up. As Drechsler and Kostakis (2015) argue, technology is constantly evolving, but the law is generally not updated immediately. This allows for a process of consideration as to which technologies are desirable and sustainable for implementation. Verifiability was not implemented when it was available (years before the actual introduction), but only when there was a concrete need owing to the recent discussions that took place in the country. Moreover, only the quiet period between elections allowed these discussions to take place, so that a reasonable system could be selected and implemented. Additionally, widely accepted reports and input from the specialist community were strong initiators of the 2011-2012 legal processes. Moreover, the timing of possible reforms has to be taken into account, as the election-free period from 2011 to 2013 came after a long period of back-to-back elections and was the only time when the EMB and parliament had the time to oversee large reforms to the system.

The third phase of development could be defined as during the last three elections from 2013–2015, where the share of e-voters among all voters remained relatively high and additional steps for individual verification - recorded as cast and cast as intended- were implemented. The number of e-voters who verified their vote has grown throughout the years, reaching 4.7% for the 2015 elections. Despite the relatively small number of verifiers, mathematical models based on the data have shown the absence of any large-scale attacks or manipulations(Heiberg and Willemson, 2014).

The discussion about transparency and verifiability in a remote electronic voting system has clearly defined the general discussion on e-voting in the past (Krimmer, 2012; Spycher et al., 2012; Volkamer et al., 2011) and will continue to do so in the nearer future. The same is true for the situation in Estonia, despite the introduction of the first stages of verification (Springall et al., 2014). The OSCE/ODIHR election specialists' report (OSCE/ODIHR, 2015) emphasizes the need for added verifiability, and the electronic voting committee is actively seeking contributions from the ICT community (Electronic Voting Committee, 2015) to suggest new solutions; the fact that the next elections are in 2017 offers enough time for bolder development.

3.4 Experience from Switzerland and Norway

The e-voting landscape worldwide has been quite active (Stein and Wenda, 2014; Drechsler and Kostakis, 2014; Barrat et al., 2012; Krimmer and Kripp, 2009). Remote electronic voting has been utilized at some level in more than twenty countries, and several countries have analysde possible implementation (Faraon et al., 2015). The largest steps in Europe and maybe even worldwide, have been made (besides in Estonia) in Switzerland and Norway. Therefore, we will briefly examine the experiences in these two countries next.

As a confederation, the election units in Switzerland are usually made up of the cantons. With postal voting being a long-time favourite in a country where elections and referendums are held often, the step to online solutions was efficacious. Different cantons have held pilots since the early 2000s. Currently three different technical voting systems are in use, and more than half of the Swiss cantons use e-voting at some level of their electoral activity. Identification is based on unique passwords, and individual verification is offered. Since 2008, voting is also offered for Swiss expatriates. Similar to Estonia, the Swiss reached a stable user experience at the beginning of the 2010s and are today looking for possibilities to enhance their (different) systems by making them more transparent, observable and verifiable. For more details on the Swiss experience see (Schweizer Bundesrat, 2002, 2006, 2013; Gerlach and Gasser, 2009; Driza Maurer et al., 2012; Drechsler and Kostakis, 2014; Serdült et al., 2015).

Norway started its e-voting project with two pilots, the first during the 2011 local elections and the second during the 2013 general elections. Both pilots were held in a small number of local-government units. Norway implemented its system after rigorous constitutional analysis and an international public tender (Ansper et al., 2009). From the beginning, recorded as cast vote verifiability was implemented, and a large effort was made to ensure public trust by using the very latest security solutions for the system. Technically and from a public perspective, both pilots were perceived as successful. However, after some evaluation, the Norwegian government officially decided to discontinue e-voting pilots due to the possible risks contained in the system's security, though the underlying reason was actually a change in political leadership and the lack of trust this party's politicians held in the system. The Norwegian pilots are discussed in detail by OSCE/ODIHR (2012); Stenerud and Bull (2012); Barrat et al. (2012); Markussen et al. (2014).

Table 3.1 compares the central elements of the Estonian, Swiss and Norwegian e-voting systems. It can be seen that no single characteristic makes up a working system, but that verifiability and trustworthiness are features

Table 3.1: Legal norms providing the institutional background for e-voting in different ent European countries

	Estonia	Switzerland	Norway
Authentication method	electronic ID	Passwords through postal system	Unique ID tied with mobile phones
Implementation style	Snap implementation, nationally	Step-by-step, canton-based	Step-by-step, only limited pilots
Verifiability	Individual	Individual	Individual and universal
Multiple vote casting	Yes	No	Yes

that all implementers have/are investing in. Estonia is the only country with digital identification, while the two other countries use(d) either existing physical IDs or wholly new unique identification methods valid only for each election. What stands out is the different implementation methods. Estonia allowed e-voting for the whole population at once, while Switzerland and Norway rolled it out in steps. Norway has now completely rolled back e-voting in all the municipalities where it was piloted, while e-voting opportunities in Switzerland depends on the canton, with some having introduced and stuck to it continuously, while others introduced e-voting and then halted it for a period before re-implementing it. In any case, based on the table we can simply state that each e-voting system has been developed in line with the needs of the actual context it was implemented in. There is clearly no basis for generalizing based on individual features; it is the complete e-voting solution that needs to be looked at. What can be learnt from Norway, however, is that the ways of implementation are irrelevant if the politicians are not convinced that the election results would remain the same regardless of the new voting channels.

3.5 Building voter confidence in e-voting

Finally, lets us examine in more detail the various confidence building measures implemented over the years for the Estonian e-voting system. Trust has been shown to be the key determinant of e-voting use. Therefore, we looked at the factors that enhance the belief of the user that the solution at hand is trustworthy. The e-voter, as the user of the system, needs to be confident that the system cannot be manipulated and that the election organizers follow the prescribed rules and operate the system correctly, so that the systems' results reflect the actual will of the voters. We will examine in detail the factors that taken together form the pillars of trust in the e-voting process: 1) confidence in the overall e-government system; 2) confidence in the token of identification; 3) confidence in the electoral management body (EMB).

Confidence in e-government

The first factor takes into account whether there is an open and receptive society. With independence regained in 1992, Estonia started many processes anew, forcing society to adapt to rapid changes. This gave Estonian also a slight advantage in adopting new solutions (Kalvet, 2012).

According to the latest Global Information Technology Report (World Economic Forum, 2015), the overall ranking of Estonia in the Networked Readiness Index is 21st; in the category of government success in ICT promotion Estonia ranks 13^{ts}, ahead of such IT giants as the USA, Finland, Korea and Japan. In the category of quality of governmental e-services, Estonia reached a high fifth place. Since 2010, the official publication of Estonian Legal Acts, the State Gazette, is electronic, which means that Legal Acts are published only on the Internet. In addition, personal income tax declarations in Estonia are issued electronically in up to 95% of cases (Estonian Tax and Customs Board, 2015), and online banking has taken precedence over traditional banking for the vast majority of people. All these are signs of acceptance of e-services by society.

An important factor regarding the possibility to launch wholly new solutions, such as the official digital ID-card or e-voting, is the relative smallness of the population. Lennart Meri, the former president of Estonia, compared Estonia to a small boat in one of his speeches: "A super tanker needs sixteen nautical miles to change her course. Estonia, on the contrary, is like an Eskimo kayak, able to change her course on the spot." (Meri, 2000). Therefore, as the number of eligible voters is around 1 million, and there is generally a positive attitude towards innovation, such ideas as e-voting could be addressed more easily. In addition, the use of online ICT solutions in alternative democratic measures (e.g. participatory budget initiatives) further enhances citizens' commitment and confidence in using e-methods in general (Peixoto, 2009; Raudla and Krenjova, 2013). In the context of this model, this first factor could be summarized as there being a broad confidence in the general governmental environment where the e-voting solution is implemented.

Confidence in the token of authentication

The second factor of confidence is formed by having secure online authentication methods. As elaborated in Chapter 2, the cornerstone of Estonian e-services, public as well as private, is electronic identification. The ID card (together with other electronic ID tokens) has been the primary identification document since 2002.

The number of digital-IDs issued has exceeded 1 million, providing almost all Estonians with the possibility to use online services securely. Approximately half of the cardholders (507 606 persons in May 2015) actively (during January-May 2015) used the electronic identification functionality of their ID card (Certification Centre 2015).

Parliamentary debate over digital-ID cards raised several privacy and security questions, but the Parties supporting compulsory electronic identification commanded the majority of votes. The most controversial questions regarded the possible risks of identity theft and overall IT security. To prevent the use of an ID-card being used by an unauthorised person, respective provisions were added to the legislation. According to the law, fraudulent use of an ID-card is punishable with a fine (Penal Code, 2001). Therefore, confidence in the token of identification and in the authorities and services connected with the token are crucial for any successful remote electronic system.

Confidence in the electoral principles and the EMB

The third and arguably most important factor, can be understood as the effective measures to guarantee compliance to and similarity with traditional electoral principles, as well as confidence that the election organizers (in the Estonian case the National Electoral Committee) are able to guarantee these principles. The e-voting procedure has been adapted to the schematic rules of traditional voting. The double-envelope system (see Chapter 1), used in many voting systems (in particular postal voting) around the world, has been implemented as the logical structure for electronic voting. Its similar nature to the postal system allows the voter to relate to the e-system, helping build trust in an otherwise novel idea (Maaten and Hall, 2008).

The methods that have been used in Estonia to increase voter understanding of and confidence in the e-voting system, in an attempt to overcome any concerns about lack of transparency and complexity, are diverse. Eight particularly important features can be differentiated.

The first feature is validating the electronic voting system, with certification or verification procedures, and testing and auditing all considered (Council of Europe, 2004). The development and importance of internet voting verifiability have been discussed earlier. In 2013, the first steps for verifiability were added to the system, and have been used for three consecutive elections. Additional measures of verifiability are likely to be added to the system in the future. Verifiability, especially individual verifiability, where the voter can personally get information about the correct acceptance of their vote, helps the voter to understand the inner procedures of the voting system and allows for the EMB to claim widespread soundness regarding the conducting of an election and the results (Heiberg and Willemson, 2014). However, as the risk of receiving false-positive malignant claims of unsuccessful verifications might occur, the EMB has to have a procedure at hand to take appropriate measures.

Second, in most of the e-enabled elections in Estonia, the EMB has allowed all voters to test out the e-voting system prior to the voting period in order to encourage people to see how the system works, terming them mock or demo elections. This has helped voters to detect any problems they might encounter before the real e-voting period. For Estonia, the primary concerns among the country's election officials, outside observers, political parties and citizens relates to the acquisition of the hardware and software needed to use an ID-card on a personal computer, update an expired ID-card or Mobile-ID certificate, and renew the PIN codes needed for the electronic use of their ID-card or Mobile-ID. System-testing prior to elections by contracted testers, auditors, observers and the public is also an important factor in order to control functionality and accuracy.

Third, the Estonian e-voting system was developed with the principle that all components of the system should be transparent for auditing purposes: procedures are fully documented, with critical procedures logged, audited, observed and videotaped (since 2013 also published on YouTube) as they are conducted. A separate procedural audit by Certified Information Systems (CISA) auditors is procured by the EMB for every election. The scope of the audit is to ensure the validity of performed procedures in terms of the guidelines contained in the handbooks and technical documentation for e-voting. Additionally, auditors review and monitor security sensitive aspects of the process, such as updating the voter list, preparation of hardware and its installation, loading of election data, maintenance and renewal of election data, and the process of counting the votes.

Fourth, it is a common requirement that the source code of an information system is available for public audit (Council of Europe, 2004). In Estonia however, the source code of the e-voting solution was not universally available until 2013, but one could access it by signing a non-disclosure agreement with the EMB. However, after the legal debates of 2012, the source code of all central servers of the voting system, as well as the software of the vote verification application, have been made freely available on the internet (Electronic Voting Committee, 2013).

Fifth, according to Estonian electoral law, all procedures related to elections are public. Observers have access to the meetings of all election committees and can follow all electoral activities, including the voting procedures, counting and tallying of results. Internet voting has been no different. All significant documents describing the e-voting system have been made available to the public (National Election Committee, 2015), including observers. In order to enhance the observers' knowledge of the system they are invited to take part in a training course before each election. Besides the political parties, auditors and other persons interested in the e-voting system can take part in the training. Observers are also invited to participate in test elections during the set-up phase.

Sixth, it is important that observers be deployed for a long enough length of time to allow for meaningful observation. If some important stages that may influence the correctness of the final results have not been observed, conclusions about the integrity of the system cannot be made. For foreign observers especially, the length of the observation period appears to be a challenge. The OSCE reported on Estonian e-voting in 2007, 2011 and 2015 (OSCE/ODIHR, 2011, 2015, 2007) and in the 2011 report stated "The OSCE in general found widespread trust in the conduct of the Internet Voting by the NEC [National Electoral Committee]. However, [...] more detailed and formal control of software installation and reporting on testing of the internet voting system could further increase transparency and verifiability of the process." (OSCE/ODIHR, 2011). As a direct result, in 2012 a process to add transparency was created. Therefore, international observation is an influential and important source for receiving feedback and peer review from the international community, which helps to build general confidence in the EMB and the used voting methods.

Seventh, add an additional element of transparency, the number of e-voters was regularly published on the e-voting website (www.valimised.ee). This very simple process allowed a wider audience, as well as political parties and the media to follow how many people had e-voted and to determine if the number of voters casting e-ballots seemed reasonable.

Last, in order to convince voters that their votes had been correctly registered, they were provided with the option to check whether their e-vote had been reflected on the polling list on Election Day. In addition to the verification itself, a second option for confirming the arrival of an e-vote has been made possible during the e-voting period. If the voter decided to replace the e-vote with a new one, they were notified in the voting app of the previously recorded e-vote being stored in the central system.

As we have seen there are many different possibilities to give the wider public additional confidence in the procedures and organization of remote electronic elections. In summary, eight important features can be distinguished:

I Technical features

(1) Introducing stages of verifiability (both individual and universal)

(2) Introducing procedural auditing measures

(3) Publishing the source code of the system

II User experience features

(4) Providing mock elections for the public to test for glitches

(5) Providing safeguarding procedures for the voter so they can check their vote was counted

(6) Publishing the number of e-voters during the advance voting process

III Observation related features

(7) Inviting and training domestic e-voting observers

(8) Inviting and accepting international observers

3.6 Conclusion

This chapter discussed the central themes in the discussion of the constitutionality of e-voting. A major theme was the privacy and secrecy of the vote. These are arguably the most important issues in any electronic voting system, as they form the backbone of the secret ballot, which is central to democratic elections. E-voting adds a twist here, as these principles are partially ensured by way of complicated technical systems involving encryption and communications between IT systems that are hard for any layman to understand. One cannot however, settle for a situation where the integrity of the system is for a large segment of the population essentially a question of faith. The constitutional debate and the subsequent discussions of electoral complaints filed once e-voting had already been implemented hence focused on the need for more procedural clarity, and crucially, more transparency and verifiability regarding e-voting. The technical development of e-voting in Estonia has subsequently followed this trajectory, with the establishment of a separate electronic voting committee to oversee the process, more detailed and open procedures for electronic voting and the subsequent vote counting, and finally and possibly most crucially, the introduction of e-vote verifiability in 2013.

The structure upon which confidence in e-voting is built is more complicated than the narrower voting system itself. It stands on three pillars where the first two - the general e-government environment and e-identity (digital-IDs) - are more underlying components, whereas the third - EMB and e-the voting system - forms the backbone of peoples' confidence in the concept of e-voting. The third pillar offers the most possibilities to enhance public confidence, e.g. by the smart procedural and system related choices listed above. It is important to reiterate the importance of each of the three sets of features, as each's correct functioning in such a complex structure is necessary for voter confidence.

Similarly, former OSCE/ODIHR Director Lenarčič has compared electoral processes to a house (Lenarčič, 2011). He discussed that if elections [electoral processes] are fraudulent, i.e. the foundation of the house is not solid, then no matter how well the house is built, it will crumble. Therefore, if any of the three pillars show signs of weakness and does not guarantee the confidence of the voter, the whole structure of confidence, supporting the nominal e-voting “roof” concept, could be in danger of collapse.

Chapter 4

Diffusion of e-voting in Estonia, 2005-2015

Kristjan Vassil and Mihkel Solvak

4.1 Introduction

Remote Internet voting has long been discussed as a means of increasing voter turnout in developed democracies, especially among younger people (Alvarez and Hall, 2008; Alvarez et al., 2009; Norris, 2001, 2003). However, technology can only have a significant impact on political participation when its usage becomes widely diffused. Technology can empower people who have faced participation hurdles in the past (Vassil and Weber, 2011). Socially excluded groups or people with reduced mobility especially should benefit from modes that make it easier to cast a vote (Alvarez and Hall, 2004). This increased empowerment might also increase voter confidence and the willingness of eligible citizens to participate in elections (Alvarez

Academic reference: ”Diffusion of Internet Voting: Usage Patterns of Internet Voting in Estonia Between 2005-2013” by Kristjan Vassil, Mihkel Solvak, Alexander H. Trechsel, R. Michael Alvarez, Priit Vinkel and Thad Hall, presented at the 72nd Annual Midwest Political Science Association Conference, 3-6. April 2014, Chicago, US.

and Hall, 2006). As participation is needed for effective representation, easily usable voting modes should, in theory, ensure a better correlation between elected representatives and society. However, technology can also present additional barriers to the already disadvantaged, in effect nullifying its theoretical promise (Norris, 2003; Berinsky, 2005).

We can list multiple factors why a technological solution could counteract the wider societal trend of decreased turnout. First, this technology has the potential to empower people who have faced participation hurdles in the past (Vassil and Weber, 2011), in particular, socially excluded groups or people with reduced mobility (Alvarez and Hall, 2004). It might bring new voters to the polls to partly compensate for the share of the ones who normally would stay away. Secondly, e-voting can also improve the general administration of elections and increase voter confidence in elections and the willingness of eligible citizens to participate (Alvarez and Hall, 2006; Alvarez et al., 2008). A centrally administered, properly audited and transparent e-voting system in countries where the organization of elections has encountered problems could encourage certain groups to participate at increased rates. Finally, given that participation is needed for effective representation, easy to use ways of voting should, in theory, ensure a better overlap between elected representatives and society. E-voting could therefore theoretically contribute towards higher satisfaction levels with the democratic process in general.

All these positive scenarios can however only be realized if the use of this technology becomes widely diffused. Or in simpler terms, a critical mass of diverse individuals uses the technology. When this is not the case, the precondition of a positive societal impact is clearly not met. Instead, the e-voting technology might actually present additional barriers to the already disadvantaged, i.e. in being expensive and complicated to use it effectively nullifies its multiple potential benefits (Norris, 2003; Berinsky, 2005) or even introduce additional structural biases into the political process.

The actual practice of e-voting has been implemented in a limited number of countries and the jury on its potential impact is still out. We know from general studies on technology usage that the most likely users and beneficiaries tend to be young, tech savvy, well-resourced, and -connected people (van Dijk, 2000, 2005). There is clear evidence that the same applies to the early adopters of e-voting (Alvarez et al., 2009; Trechsel and Vassil, 2011), which would indicate that some of the positive theoretical potential is not, at least initially met. What we do not really know however, is whether e-voting has the potential to diffuse beyond these particular groups of people to a broader and less homogeneous group of voters or whether it remains a tool for those with skills and resources. Discussions about how and why new modes of voting might improve participation or representation require empirical evidence about the conditions for and patterns of adopting new technologies. If the rate of adoption of this new voting technology is slow and its diffusion limited to specific groups of people, then it is unlikely that e-voting can have a positive impact on voter turnout and the quality of representation.

The next sections address precisely this problem using data from Estonia. The data shows who e-voters are and how their profile has changed over time using a unique survey data set that covers each of the eight elections in Estonia from 2005 to 2015 where e-voting was an option. The goal was to determine whether e-voting has indeed diffused among voters or whether it remains a technical solution providing more convenience for a group of people already engaged in politics and who face limited barriers to participation to begin with.

4.2 Political participation and technological diffusion

We can think of e-voting as a technologically advanced way of engaging in what is otherwise conventional political behaviour, used most likely by a particular group of people among the electorate. Our objective here was to investigate how this group of people is changing over time. We did so with the help of the classical account of the diffusion of innovations provided by Ryan and Gross (1943) and Rogers (1962), which have been used throughout the years to explain a wide variety of phenomena ranging from the spread of agricultural practices (e.g. Fliegel, 1993), political reforms and policies (e.g. Starr, 1991; Jahn, 2006), medical practices (e.g. Greenhalgh et al., 2004) to management (e.g. Abrahamson, 1991) and most crucially, technological applications in very different fields (e.g. MacVaugh and Schiavone, 2010).

Rogers's account sees the diffusion of technology as a sequence of steps in an innovation decision process (Rogers, 1962). This process includes gaining knowledge of the technology, being convinced of its usefulness, and deciding where to implement it. Adoption results if people's expectations about the technology are positively confirmed by the experience of using it. Once a distinct subgroup has reached the adoption stage, subsequent spread to other groups, i.e. diffusion, is reminiscent of a bank-run, where the number of people adopting the technology is dependent on the number of previous adopters (Rogers, 1962, p. 206). This sequence has been demonstrated to apply both to collective and individual actors (see Wejnert, 2002), meaning it can be used to explain why organizations adopt innovations and why individuals decide to try out new technologies.

The crucial part in using Rogers' account to explain use of e-voting is in the characteristics of the adopters of the technology at different stages of the process. The first adopters tend to be a small number of well-informed and innovative risk-takers (Rogers, 1962, p. 263). These are people who are open to new experiences and more willing to try out new technologies. Quite naturally, such people also tend to have a distinct socio-demographic profile and the types of attitudes corresponding to their status. These first adopters are followed by a larger groups where the traits associated with the first adopters are continually less prominent. At the end of the diffusion sequence even technological laggards might be motivated to adopt the technology, i.e. when the relative gains outweigh the costs of adopting (Rogers, 1962, pp. 263-265). Rogers therefore sees diffusion as a sequence of adoptions by distinct user groups until the whole relevant population is using the tehnology. Depending on the account the names of some of the groups differ, but in general Rogers' differentiates between innovators, early adopters, the early majority, the late majority and laggards; for details see Rogers (1962).

As with every new technology however, adoption involves considerable costs, such as the need to come to terms with the higher complexity of the new technology and the need to evaluate relative gains compared to the previous solution. This means that a "technological bank-run" in terms of adopting e-voting might depend on whether a certain specific subgroup of voters can overcome the barriers to using e-voting. It suggests that the expected positive effects on voter turnout might at first only be realised for those who do not face any substantive hurdles to effective participation to begin with, namely citizens who are well resourced – in terms of education and income – and who also have a good command of modern technologies. Therefore, diffusion might completely stop well before its potential maximum due to certain groups not seeing the relative benefits of adoption or judging the adoption costs to be prohibitively high.

In sum, this means that while old thresholds of participation might be lowered, new barriers might be erected. This suggests the two different diffusion expectations pictured in Figure 4.1.

First, from Rogers's theory of the diffusion of innovations, one can expect a gradual dispersion of the use of a specific technology, driven by early adopters who are clearly distinguishable in terms of their sociodemographic profile. The early adopters should be younger, better educated and more well off individuals. Presumably, they should also be more tech savvy, trust technology more, and be distinct in their ideological positions. However, over time, new e-voters should be less distinct from regular paper ballot voters in terms of sociodemographic characteristics, as well as other behavioural and attitudinal characteristics.

This would mean the final level of diffusion into user groups including technological laggards; visually represented as scenario A in Figure 4.1. This scenario can be identified by investigating the profiles of e-voters. Given that the early adopters are, as a rule, a quite distinct set of individuals, later adopters should be more similar to the general population, or in other words e-voters should become more similar to regular paper ballot voters over time. Thus, we expected that the characteristics associated with a person's probability of casting an e-vote during the first e-enabled elections should become less important and useful for distinguishing e-voters for later elections. In a nutshell, e-voters should become more similar to regular paper ballot voters over time if full diffusion is indeed occurring.

A contrary expectation follows from the possibility that a new technology can present barriers to some voters. Although technology might spread according to the initial sequence suggested by Rogers, and the total number of users may rise at first, barriers might remain in place for a significant segment of potential users. If this is the case, e-voters will remain a clearly distinguishable subgroup of all voters and the growth of e-voters will quickly reach a plateau, as the technology fails to reach the less tech-savvy voters. It also means that e-voters should remain clearly distinguishable from paper ballot voters, irrespective of time or a moderate growth in e-voters. In essence this means the stopping of diffusion at a distinct level; visually represented as scenario B in Figure 4.1.

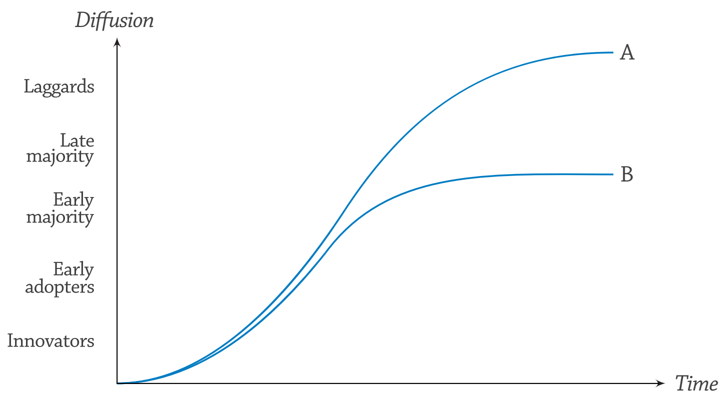

Figure 4.1: Levels of diffusion across time with two scenarios

Before examining the spread of e-voting, we must first solve the problem of choosing proper reference categories. If we expect e-voters to be different, then who are they different from? Comparing all e-voters to regular voters would likely be inaccurate, because the group of e-voters would contain not only the early adopters (like those who e-voted in the first e-enabled elections of 2005 or 2007), but also later adopters, who became e-voters only after it had been available for several elections. Because the motivations and characteristics of early and late adopters may differ considerably, we have distinguished them within the category of e-voters in general. To do so we conducted a detailed analysis of regular paper ballot voters compared to the set of e-voters who voted online for the first time in each e-enabled election. Those who voted online during the first e-enabled elections can be considered slightly different from those who voted online for the first time in the second e-enabled elections and so on. Focusing on the group of first time e-voters at each election provided us with more reliable estimations of how the characteristics of the people who adopted this technology at different points in time have changed.

Having outlined how we see diffusion happening and that we should observe first time e-voters in each election to determine if it is indeed going on in the case of Estonian e-voting, we also had to specify what precise voter characteristics to monitor to be able to answer our central question.

Prior studies on diffusion patterns in diverse settings have demonstrated that socio-demographic and economic factors, including ethnicity, are associated with diffusion rates (e.g. Berry and Berry, 1990; Hedström, 1994; Tolnay and Glynn, 1994). Such associations arise because the listed traits act as proxies for acceptance of technology and resources needed for effective usage. In addition to demographics there are also attitudinal dispositions that make people more conducive to voting online such as trusting online transactions in general and e-voting specifically. Finally, besides proxy measurements of resources and more direct measurement of attitudes, a very straightforward indicator in the form of self-reported computer skills should show a strong association with e-voting (Trechsel and Vassil, 2011), and provide an especially convincing test on whether true diffusion is taking place. Also, prior voting habits and frequency of engagement in political talk, both strongly associated with turnout in general (Cutts et al., 2009; Smets and van Ham, 2013), should show the engagement and political sophistication levels of first time e-voters, and test whether this way of voting stays the prerogative of political enthusiasts, analogous to Rogers' first adopters but in the political participation realm. Finally, we also investigated the ideological position of e-voters to account for the possibility that younger, more well off, tech-savvy individuals, that are the most likely first time adopters (Alvarez et al., 2009; Trechsel and Vassil, 2011), also occupy a recognisable ideological position in line with their socio-demographic traits.

With this clarified we utilized survey data over a 10-year period (2005–2015) to determine if we are dealing with a full-blown diffusion – scenario A in Figure 4.1 – or only partial diffusion – scenario B – of e-voting.

4.3 Diffusion patterns

The absolute and relative number of e-voters has grown in almost all elections over the years, with roughly one third of voters now casting their vote online (see Figure 1.1 in Chapter 1). These aggregate user numbers conform, however, with both the outlined diffusion scenarios, so let us proceed to look at the data on individuals to be able to determine the existence of full, partial or no diffusion. Our expectation is that in the case of full diffusion, first time e-voters will become less distinct from paper voters when it comes to the above outlined voter traits.

The specific voter characteristics presented visually here include age, ethnicity and computer literacy. The list of other characteristics that we also examined are presented in full, with the regression models, in the technical Appendix A.

The association with age is depicted in Figure 4.2[18]. The horizontal axis shows age and the vertical axis corresponds to the probability of e-voting. The figures show a very distinct pattern that gradually disappears as we go from one election to the next. Age had a positive association with choosing to e-vote for the first time up to the age of 40-50 years old – the higher the age up until then, the higher the probability to e-vote. After that point we saw a strong negative effect – the higher the person's age above about 50,

[18] For more information about the technicalities of the analysis and the results, see Appendix A.

the lower the probability of e-voting. This clearly shows that older people are less likely to initially choose e-voting over paper ballots. The association with age however, disappears over time and gradually flattens. In other words, age and e-voting are no longer associated and voters choosing to e-vote for the first time come equally from all age groups.

Figure 4.3 presents data for ethnicity over the course of the elections. The dots correspond to best guesses, based on the models, about the associations between the two variables and the vertical bars represent a range where the value of this association can be located with a 95% level of probability. The horizontal line which crosses the vertical axis at 0 indicates no association between the given variable and e-voting. If the bars from the dots cross this line, one cannot conclude that there is an association. The scale on the vertical axis shows how much the probability of e-voting is higher for those who are ethnically Estonian. For example, we can see that for the 2007 national elections, ethnic Estonians were about 20% to 50% more likely to e-vote than other ethnicities, with our best guess at about 35% more likely. The figure tells us that between the 2005 and 2009 elections, ethnic Estonians were close to 40 percentage points more likely to e-vote compared to other ethnicities. By the end of the period, however, ethnic Estonians are no longer more likely to vote online than other ethnicities, meaning e-voting is used regardless of ethnic background.

The association of e-voting with self-reported computer literacy is shown in Figure 4.4. A clear relationship that changed over time is again apparent. Specifically, first time e-voting was much more probable among people with good or average computer skills during the first e-enabled elections, but this difference gradually weakened over time. Though people with good PC skills are still more likely to vote online for the first time in comparison to people with poor skills, the same does not apply anymore to people rating their skills to be average. The Estonian data therefore suggests that computer literacy is no longer a clear driver of e-voting and thresholds set by modest skill level can over time be overcome with handily designed e-voting systems.

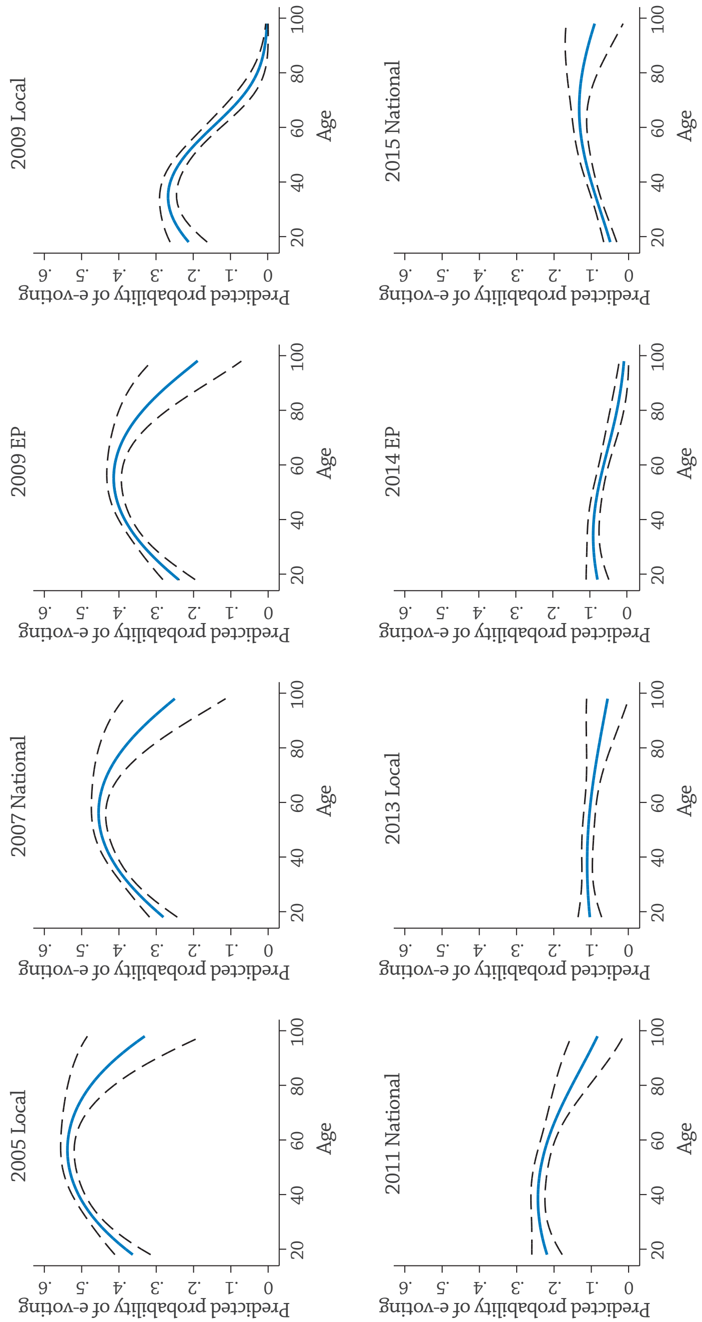

Figure 4.2: Effect of age on the probability of e-voting for the first time (2005-2015)

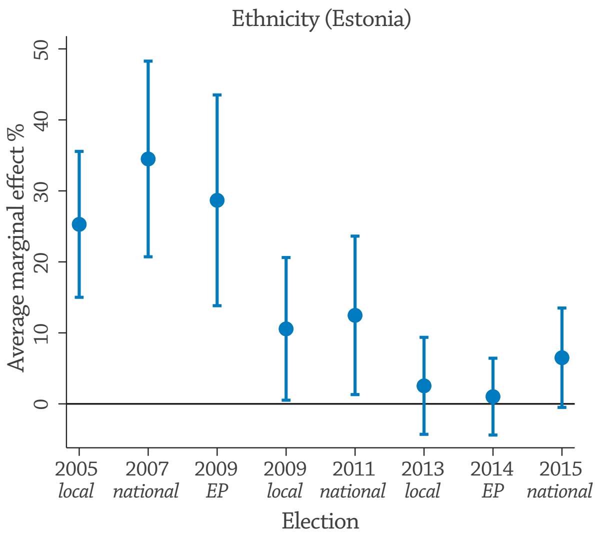

Figure 4.3: Effect of ethnicity on the probability of e-voting for the first time over time

DIFFUSION AFTER THREE ELECTIONS

For the first three elections multiple sociodemographic, attitudinal, and behavioural factors had a non-trivial association with being a first-time e-voter. However, from the fourth election onward, the importance of these factors gradually diminished, indicating the diffusion of e-voting among the Estonian electorate.

All–in–all our statistical models – using the aforementioned voter characteristics and a range of other factors (for details, see Appendix A) – did not distinguish between first time e-voters and regular voters after roughly three elections. To put it differently, whereas we could clearly say who was a "typical e-voter" in the first few e-enabled elections, we are no longer able to do this. Technological diffusion has traversed social boundaries and Estonian e-voters nowadays have become virtually indistinguishable from regular paper ballot voters. The evidence clearly suggests that full-blown diffusion of e-voting has taken place.

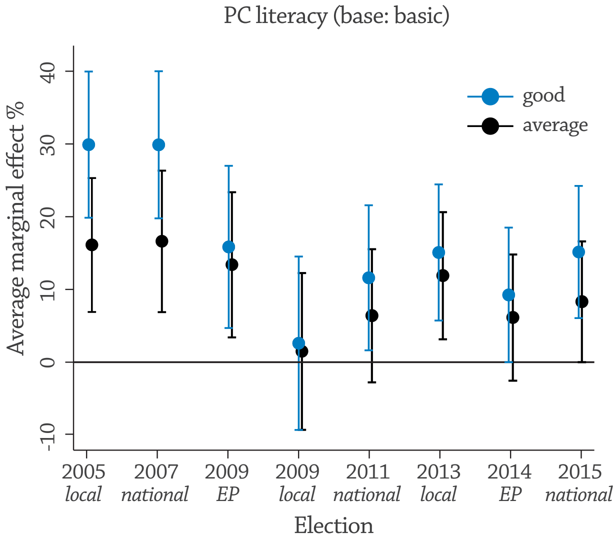

Figure 4.4: Effect of computer literacy on the probability of e-voting for the first time over time

4.4 Conclusion

We saw that e-voting has become widespread in Estonia, which *prima facie* already gave an affirmative answer to the question of diffusion. The crux of the matter, however, is that an increasing total number of e-voters does not necessarily mean a diverse set of voters is using remote internet voting. It would be problematic if this particular voting mode would spread only among a distinct subpopulation of voters who already enjoy relative ease of access to all modes of participation. True technological diffusion over time would mean that voters from a broad cross-section of the population, regardless of their social status or level of resources at their command, turn to e-voting. Following expectations derived from Rogers's account of the diffusion of innovation, we examined the profile of first time e-voters over the course of eight elections that spanned a period of ten years (2005–2015) to determine to what degree e-voting has been adopted by the wider general electorate.

Evidence shows that the characteristics of first time e-voters did indeed become more similar to the characteristics of paper ballot voters as time passed. Typical e-voters used to be ethnic Estonians from a distinct age group with good computer skills. From the fourth election onward however, one can clearly see that these characteristics no longer have a strong association with e-voting. As a result, the ability to predict who will take up or switch to e-voting using only sociodemographic and attitudinal data becomes increasingly difficult. In sum, the evidence shows that e-voting has diffused among the overall voter population and not just remained an activity of a privileged few.

We draw two main conclusions in light of this. First, technology has the potential to bridge social divisions and ease political participation not only for those not facing any substantive obstacles, but also for the less privileged who command fewer resources. The experience of Estonia seems to show that technology should not be considered as a hurdle, but as an enabler of political participation. The caveat is that technology presents only an efficient mode for participation. Structural hurdles that inhibit participation in general, regardless of the mode, such as belonging to socially marginalized groups, will most likely stay unaffected. What we have demonstrated is simply that technology itself does not seem to exclude anybody, as some sceptics have suggested.

The second and absolutely crucial conclusion, is that the potential enabling effects did not surface immediately after the adoption of the new voting technology, but required a period of at least three elections to appear. Adoption among a specific subgroup did happen immediately, but this seems to have represented a private benefit for people already enjoying a better socio-economic status. Wider public benefit can only be realised once the usage has diffused and this took time. Policy makers are well advised not to expect immediate results from the introduction of new technologies, but to recognise that different social groups adopt and use new technologies at differential rates. On the upside, evidence from e-voting in Estonia showed the process to be astonishingly quick. Characteristics that initially distinguished e-voters in the case of Estonia lost their association not after three electoral cycles, but after three separate types of elections within a four-year period. What seemed to matter was not time as such, but the frequency of being exposed to the possibility of casting a vote over the internet.

Chapter 5

E-vote log files 2013-2015

Taavi Unt, Mihkel Solvak and
Kristjan Vassil

The study of Estonian e-voting has since the 2013 local elections benefited from an additional data source in the form of anonymized e-voting system log data. The importance of this data in understanding how e-voting works can hardly be understated.

First, it covers the whole population of e-voters, meaning we do not need to operate with samples prone to measurement error and non representativeness.

Second, this data covers three e-enabled elections – 2013, 2014 and 2015 – which gives us unique insights into the period of "matured e-voting", i.e. the period when, as the previous chapter showed, e-voting had become fully diffused among the electorate.

Third, it gives us unmediated measures on voter behaviour, in other words, we do not face the measurement uncertainties that tend to result from survey respondents recalling their behaviour.

On the downside, the data includes a quite limited number of indicators, which clearly narrows the set of questions that can be studied.

The data itself is automatically generated when people e-vote and is foremost used by the election authorities and e-voting system administrators to identify failures and to detect anomalies and possible attacks against the system. This chapter uses the anonymized log data to examine general usage of e-voting and reports some interesting behavioural patterns.[19] Specifically, we give a general description ofthe e-voter in these three elections, followed by a detailed examination of how long it actually takes people to cast a vote online. We will also show how frequent multiple voting was in one election, which is a particular feature of the Estonian system. Finally, we will show how difficult it is for people to vote online by scrutinizing cases where a potential e-voter failed to successfully cast a vote at some stage of the electronic voting process.

The analysis below excludes log data on users who were not allowed to vote, i.e. under-aged or otherwise non-eligible, as they could still try to vote through logging into the system using their electronic identification credentials and thereby leave a print in the logs. Thus, only sessions that could theoretically end with a successful vote were included.

5.1 General description

We can describe a typical e-voting session with the set of variables included in the logs, including each voter's age, gender, the country where the vote was cast, the operating system (Windows, Linux, iOS), the method of identification (ID-card, Digi-ID or mobile-ID), the time of voting and the session length.

Let us look at age first, Figure 5.1 shows the distribution of voters' ages for three consecutive elections, the 2013 local, 2014 European Parliamentary and 2015 national elections. Keep in mind that the figure does not describe e-voting turnout within each age group, but the proportion of users from a given age for all who cast their vote online. As already hinted at in the chapter on diffusion, e-voting is mostly used by middle-aged peo-

[19] For a detailed description of the log process and another look at log files see: http://cyber.ee/uploads/2015/03/estivote_willemson.pdf

ple, between 35–45 years old. What stands out in particular in the figure is the very small share of the youngest voting eligible age segment, people between 18 to 25 years old. Heiberg et al. (2015) have already shown that the e-voting turnout among this age group is comparable to 70–75 year olds. This puzzling picture results from two things. First, turnout in general among the youngest eligible voters is very low. Second, this low turnout is coupled with a surprisingly small share of e-vote usage among the active voters within the youngest age group. To put it another way, e-voters comprise a clearly smaller share out of voters within the 18–25 age group than they do among the 35–45 age group. The distribution of age among active e-voters presented in Figure 5.1 therefore mirrors once more the somewhat surprising finding that voting online in Estonia is not the most popular option among the youngest and presumably most tech savvy citizens.

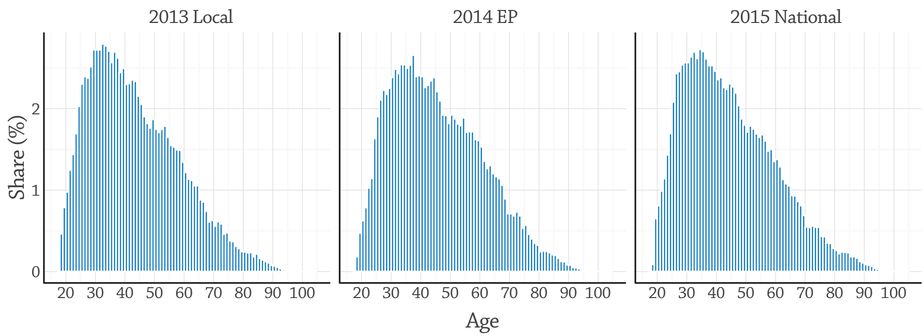

Figure 5.1: Age distribution of internet voters

Further descriptive statistics extracted from the log data are shown in Table 5.1. We saw that for each election the share of male e-voters was lower than for females. In addition to this, the data also shows how the share of women becomes progressively larger with age from around 45 years old onwards, a pattern which coincides with the age-gender structure of the Estonian population and tells us that e-voting is actually equally popular among males and females.

Examining the table further shows that votes from abroad hover around

5%. This means that in general every 20th e-vote was cast from outside of Estonia. In absolute terms the total number of e-votes submitted from outside Estonia in 2013, 2104 and 2015 were 5 640, 5 062 and 10 022 respectively. In contrast, only 1 251 people voted from abroad in 2015 at Estonian embassies/consulates or via post, meaning that 89% of votes from abroad in that year were e-votes. Such numbers tell us that traditional ways of voting from abroad have, thanks to e-voting, almost become a thing of the past. Given that more and more people travel, study and work in a foreign country, it is relevant to ensure that participation in elections is easy and sufficiently flexible for them, and e-voting seems to be a perfect tool for this.

Table 5.1: Descriptive statistics of the entire population of e-voters from system log data

	2013	2014	2015
Age (mean)	43.0	45.4	44.3
Male (%)	48.2	48.8	47.6
Voted abroad (%)	4.2	5.0	5.7
Operating system			
Windows (%)	93.8	93.3	92.6
iOS (%)	5.4	5.6	6.2
Linux (%)	0.8	1.2	1.1
Identification means			
ID-card (%)	90.2	87.5	86.4
Mobile-ID (%)	8.6	11.0	12.2
Digi-ID (%)	1.3	1.5	1.4
Mean session length (minutes)	2:52	2:21	2:36

Furthermore, we saw that the majority of voters used a version of Windows as their operating system and up to 90% of voters identified themselves with their ID-card. In each subsequent year the Mobile-ID, which is a rather new method for identification, was used by a larger number of people. Its share has grew from 8.6% in 2013 to 12.2% in 2015. The Digi-ID however, has remained a marginal mode of identification.

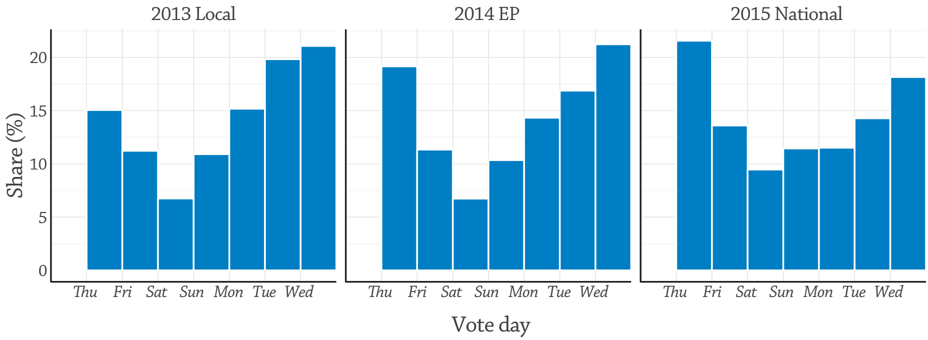

Figure 5.2: Distributions of e-votes cast per voting day

Finally, the table shows that on average an e-voting session lasted well below 3 minutes in duration. We will return to the time taken to e-vote in detail below.

Part of the general description of a typical e-voting session is also the distribution of e-votes over the e-voting period. In the three elections covered by the log data, each election started on a Thursday at 09:00 local time and lasted for seven days until 18:00 on Wednesday. The pattern of how the votes are distributed over the voting days is U-shaped, see Figure 5.2. Most e-votes were therefore given during the first and last day of the voting period and activity was lowest during the weekend falling in the middle of the voting period. By far the most popular hour for e-voting each year was the last hour of the e-voting period (not separately reported).

The portrait of the typical e-voter is therefore a 45 year old person, who votes within Estonia on a computer running Windows and does so within two and a half minutes. This is, however, only a general description and we will show a much more detailed picture in the next sections.

5.2 Session length

One of the main points of interest regarding Estonian e-voting is user convenience and the resulting considerable savings in terms of time spent on voting. The log data gives us precise measurements on how long it actually took individuals to cast a vote via the Internet.

Session length is a useful and interesting measure for multiple reasons. First, it shows how easily people managed to vote. Long average sessions might indicate system design problems. Second, session length can also be used to find patterns left by possible attacks, e.g. a very short session length might indicate a threat to the system. Finally, it gives a very precise point of reference for comparison with the time taken to cast a vote on paper and therefore shows us if and to what degree e-voting saves people time. We limited our interest to the speed it took people to e-vote and how the voting session lengths varied by user subgroups, and to a brief comparison with the estimated time taken to vote on paper.

Before proceeding to look at the numbers, we need to produce a definition of an e-voting session so as to be able to measure its length. One can argue about how to define the beginning and end of an e-voting session. Is it the total time spent sitting behind the computer while voting, including finding the relevant web page to load the necessary software and then commencing to vote? Do people actually specifically switch on a computer with the single purpose of voting and then perform the activity, or is it part of their natural workflow that simply happens while engaging in other online actions? These are all user specific questions which we could answer definitively and so needed to proceed with a definition that was quite strict and covered the essential steps necessary to cast a vote online.

We defined e-voting session length as the time between candidate list retrieval and vote submission, that is the time between when the e-voter was shown the list of candidates running in their district and the time when they digitally signed and submitted their electronic vote. We examine all sessions separately, so when a voter e-voted multiple times, each of their sessions was considered as a stand-alone session. The share of multiple sessions was however very low and will be examined in more detail in the next subsection. All sessions not ending with a vote, due to the voter deciding not to cast a vote after seeing the candidate list or other technical difficulties, were excluded.

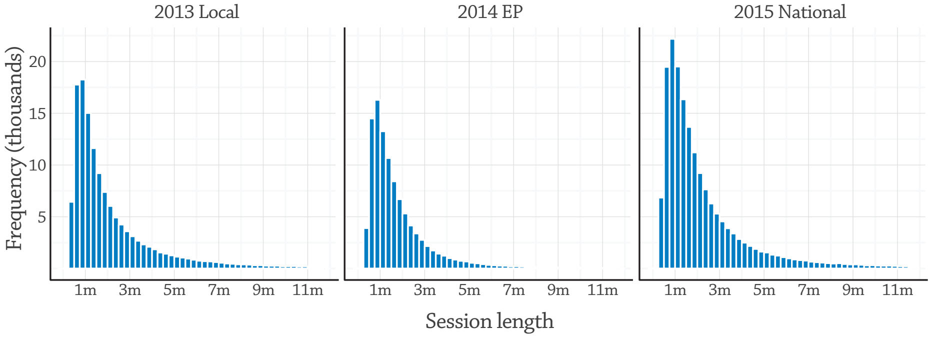

Figure 5.3: Distributions of e-voting session length

According to table 5.1, the average session length was 2 minutes 52 seconds in 2013, 2 minutes 21 seconds in 2014 and 2 minutes 36 seconds in 2015. Figure 5.3 shows the distributions of voting session lengths. One can see that the majority of voters were able to vote within a few minutes and the distributions actually peak around one minute. E-voting therefore is surprisingly fast. The mean values mentioned above are in fact somewhat conservative estimates of the speed of e-voting, as there were cases where a session lasted for many hours or even days due to voters leaving the application open without casting a vote straight away. The median session lengths, where half of the sessions were shorter and half longer, were 1:29, 1:21 and 1:36 minutes in 2013, 2014 and 2015 respectively. Notice that for the 2014 European Parliamentary election the sessions were noticeably shorter, both according to the mean and median value. This is most likely due to the type of election, as Estonia elects only six MEPs to the European Parliament and the whole country is one single voting district for such elections. Voters therefore have exactly the same ballot and it is also as a rule shorter in terms of the number of candidates than the district specific ballots of local and parliamentary elections. The voter theoretically therefore should face a simpler set of options from which to choose and hence also be able to e-vote faster.

E-VOTING SPEED

Median length of the e-voting session was 1:29, 1:21 and 1:36 minutes in 2013, 2014 and 2015. So in times of ”mature e-voting” people e-vote faster than it would take them to hard-boil an egg, read approximately two pages of text or listen to the 1969 release of ”(I Can’t Get No) Satisfaction” by the Rolling Stones.

To sum up so far, the log data shows that e-voting is astonishingly fast, taking a minute or less for a large segment of voters and well below three minutes on average. Such a quick process tells us that the e-voting system is entrenched and easily usable. We do not have comparable historical data for the first elections where e-voting was an option, so we cannot say if a learning mechanism occurred that led to a speeding-up of the e-voting process, but we can definitely say that in times of “mature e-voting”, people e-vote faster than it would take them to hard-boil an egg, read approximately two pages of text (Carver, 1990) or listen to the 1969 release of “(I Can’t Get No) Satisfaction” by the Rolling Stones.

Besides the typical voting speed and the time distribution among e-voters, we also took a look at whether voting speed changes with age, given that according to conventional wisdom older people might struggle with online services in general and tend to have lower computer skills. Figure 5.4 shows the average voting speed according to age separately for males and females. Each dot represents the mean session length for that particular age and a smoothing spline has been added to show the trend over age. The picture is striking and quite possibly the most interesting to emerge out of the log files. Older people e-voted faster and the increase in voting speed was almost linear, i.e. constantly faster the older the person was. For example, it took 25 years old males on average between 2.5 and 3.5 minutes to e-vote depending on the elections, while 75 years old males voted in less than 2 minutes in all three elections. We also saw that on average males voted more slowly than females over almost all age groups. Why did older people e-vote faster? We do not have any definitive answer to this question and can only specu-

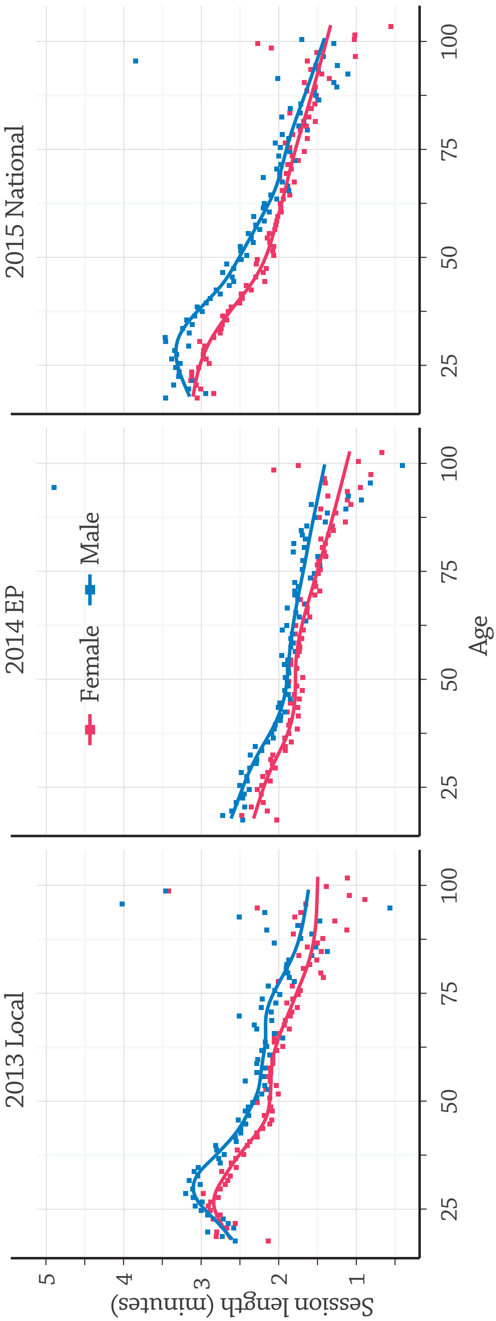

Figure 5.4: Average session length by age and gender (only sessions less than 30 minutes were included)

late. For one, older people might be more steadfast in their decision to cast their vote electronically. We know that increasing age correlates with fewer on-line activities, and it might therefore be that older people proceeded to cast a vote behind the computer with this single purpose in mind. In other words, going on-line and voting is the single purpose of the action, while younger people might e-vote in conjunction with doing something else on-line. It could also be that older people are helped in the e-voting process by someone, which might also increase voting speed. It is however unlikely that for example almost 40 000 e-voters who were 65 or older during the three elections under study were helped by someone to e-vote. Finally, given that age is a strong predictor of political participation, it might also be that the elderly were more determined when it came to their choice of party, i.e. they simply had made up their mind about their vote choice and did not need to spend additional time browsing through and reflecting upon the list of candidates presented to them.

To summarise, the detailed investigation of session length provided some interesting surprises, i.e. the process of e-voting is in general fast and faster still the older the user. One had expected it to be exactly the other way around, but the consistency of the pattern over three quite different elections suggests it to be a robust relationship. This indicates something quite interesting, for example one argument against e-voting is the potential entry barriers it sets against some voter groups not as well versed in technology, i.e. the elderly in particular. It might be an exclusive rather than inclusive technology for these groups when it comes to voting. The Estonian data suggests otherwise, i.e. we saw that the tens of thousands of e-voters who were nominally older than the retirement age were in fact voting faster than supposedly more computer literate younger people. We think that voting faster is a clear indicator of being able to do it with ease and the log data shows e-voting does act as a barrier for the participation of older age groups.This finding is the more striking as it comes from data that describes the entire e-voter population and does not suffer from possible bias or under-representation of certain e-voting age groups.

But session length patterns certainly do not exhaust the rich analytical potential of the log data. Next we will report upon another aspect of Estonian e-voting that might prove informative to the wider public: to what degree and who exactly e-votes multiple times?

5.3 Re-voting

E-voters can cast an unlimited number of e-votes during the seven-day e-voting period, with each new vote cancelling the previous e-vote by that person and they can also definitively vote on paper at the ballot station during advance voting if they so wish, which cancels the e-vote. Why is this? There are three main reasons.

First, if someone e-votes under duress, meaning they are coerced to vote in a specific manner, then the person can theoretically cast a new e-vote at a later time free from coercion.

Second, the knowledge that a vote can be changed should clearly lower the effectiveness of using coercion or buying votes, maybe even make it wholly pointless, as the potential vote manipulator has no guarantee that their machinations will deliver the desired results.

Third, if the e-voter is suspicious that their vote might have been compromised somehow (e.g. a malicious computer virus), then they can remove the potential threat and cast their vote anew from the same machine or another safe computer.

While the sensibleness of the re-voting option has not been disputed owing to these three reasons, it has been pointed out as being problematic from the perspective of a neutral voting environment. Campaigning is prohibited in Estonia on election day, but not during the advance voting period when e-votes are cast. This fact is seen as a risk by some political parties. It is argued that because people can re-vote as many times they want, they can also change their mind during those seven days, switch their vote between multiple parties and thus bring about uncertainty that political actors wish to reduce. Moreover, because people may be more susceptible to political campaigns during the election period, they also respond more promptly to party performance. When parties need to maximize their vote share and lock down their supporters, e-voting becomes an uncomfortable black box that makes it harder for parties to capture those votes simply because it allows for multiple vote switching.

We agree that there are valid arguments for and against the re-voting option, but our interest here is merely to demonstrate to what degree does multiple e-voting actually take place by mining the log data. Table 5.2 shows how widespread it is. As it turns out it is extremely rare; the overwhelming majority of e-voters, approximately 98% of those who cast a binding vote, e-voted exactly once. Between 1.5 and 2.1% e-vote twice and during each studied year only 0.14 to 0.17% (i.e. a few hundred) e-voters voted three or more times.

The data so far suggests re-voting to be a quite rare phenomenon and we also have to point out that re-voting did not necessarily mean that the vote choice was changed. There might have also been cases where the first vote was cast to test the system and the vote after that was the one initially intended and so on. Whether such cases occurred or what the proportion was cannot be known because the actual votes have been destroyed as stipulated by law and we are only operating with anonymized data on the so called outer envelops of the e-vote (see Chapter 1). So we can only speculate that most likely not all re-voters switched their party choice between multiple e-votes and the potential risk for political parties outlined above is even lower than the already mentioned nominal number of multiple e-voters suggests.

Table 5.2: Multiple e-voting

Times voted	2013 local	2014 EP	2015 national
1	131 222 (98.07%)	101 404 (98.31%)	172 457 (97.71%)
2	2 359 (1.76%)	1 603 (1.55%)	3 723 (2.11%)
3	186 (0.14%)	100 (0.10%)	254 (0.14%)
4 or more	41 (0.03%)	42 (0.04%)	57 (0.03%)
Total	133 808 (100%)	103 149 (100%)	176 491 (100%)

But let us not stop here and go on to examine multiple voting from one additional angle, namely: what happens to the session length with each new e-vote iteration? This is shown in Figure 5.5 As one can expect, each new session by the same e-voter was a bit quicker than the previous one. The solid light blue line in the figure represents the cumulative distribution function of session length for the first vote by all e-voters and the dashed light blue line represents the distribution of session length for the first vote by multiple e-voters. A comparison of these two light blue lines shows that in each of the three elections multiple e-voters voted somewhat slower during their first session than one-time e-voters. This is quite interesting, as it indicates that these multiple voters either had more trouble navigating the voting system or faced a greater vote choice dilemma when presented with the list of candidates.

We see further that the largest decrease in session length was between the first and second e-vote, while the decrease in length between the second and third e-vote was smaller and even more so for the third and fourth e-votes. So e-voting does become quicker the more a person does it, indicating a learning process, but the decrease in time spent on voting is not constant the more the person e-votes.

The fact that only roughly 2% of e-voters voted multiple times and hence re-voting was a rare event, does not address the underlining problem that this small share of voters might have been exposed to a (campaign) event during the e-voting period that led them to recast their vote for a different candidate. We have an indirect indicator for the likelihood of this happening: the time between the two e-votes by any particular voter. It would indeed be problematic if we see for example a pattern where the first e-vote was cast at the beginning of the seven-day e-voting period and the second e-vote at the end of it, i.e. with a whole week of potentially vote changing outside events happening in between. The cumulative distributions of the time between the first and second e-vote for people who voted at least twice is shown in Figure 5.6. It turns out for instance that in 30% of the cases the second vote was cast within 10 minutes of the first and in 50% of the cases within 12 hours.

Hence a remarkable proportion of re-voting took place within a relatively short time interval, indicating that any potential impact of an outside

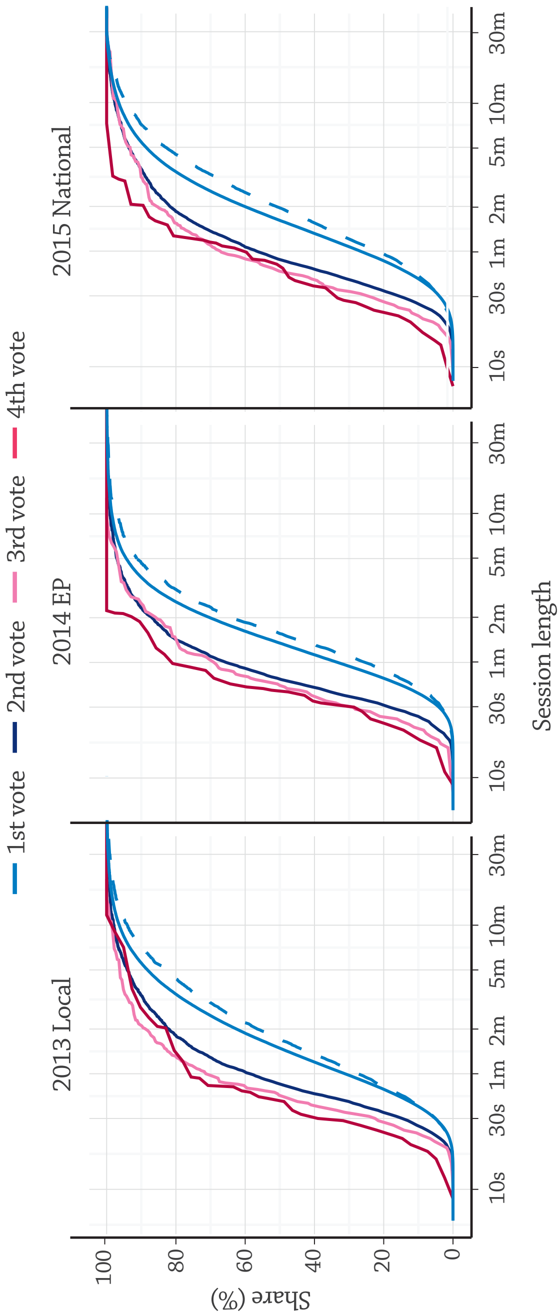

Figure 5.5: Cumulative distribution of session length by times voted

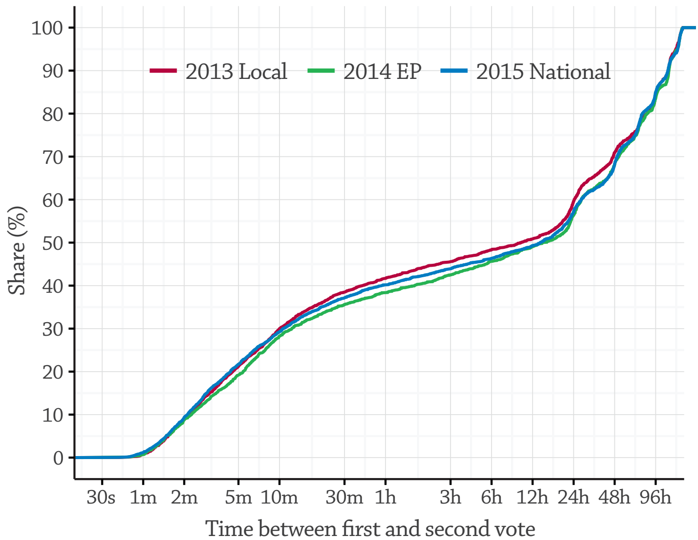

Figure 5.6: Cumulative distribution of time between the first and second e-vote

event is unlikely. In addition, we did not see any unexpected patterns when studying the timing of re-voting during the whole seven-day e-voting period. There were proportionally slightly more re-votes on the last voting day, but as reported before, the overall e-voting activity was also higher on the last election day. Generally, the second e-votes were distributed pretty uniformly over the voting period. We did not see a large chunk of second e-votes on specific days, rather multiple voters simply re-voted in close proximity to their first e-vote.

Figure 5.7 shows the re-voting probability according to age for males and females separately. Each dot on the figure represents the empirical proportion of re-voters according to a certain age and gender groups. The smoothing trend lines over age show the predicted re-voting probabilities estimated by logistic regression models with age as the single predictor, for both genders separately.

Two things stand out. First, we see that young men were clearly more

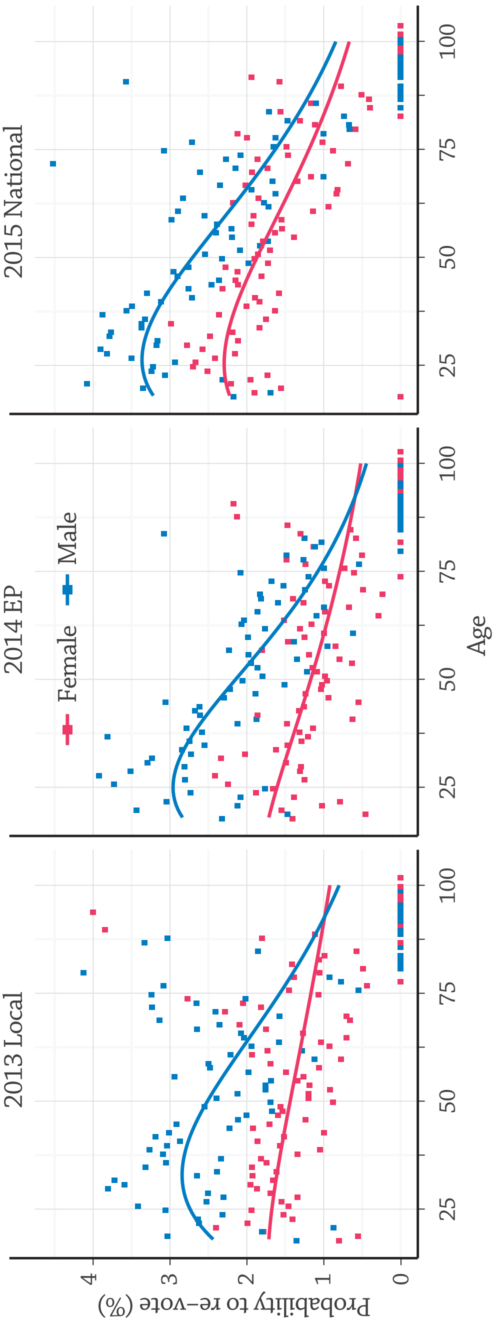

Figure 5.7: Probability to re-e-vote according to age for each gender

frequent re-voters than younger women. Second, it turns out that the proportion of re-voters declined as age increased. This makes sense when we think back to the associations seen with voting speed. Older people vote faster and seldom re-vote, hallmarks of a more certain party vote decision and possible single purpose usage of time behind a computer, i.e. e-vote and be done with it.

To sum up, re-voting was rare and people tended to re-vote in close proximity to their first vote. If in addition the assumption that not all re-voters actually change their party vote between different e-votes holds, then allowing people to e-vote an unlimited number of times does not seem to bring the feared risk of fluctuating preferences over the advance voting period owing to party campaigning. The patterns we see in multiple e-voting do not suggest it to be especially prone to being affected by specific vote influencing events, though we cannot fully rule it out with the given data.

5.4 Failure to e-vote

The unique nature of e-voting allows us to study one further interesting element, namely to what degree and who fails in completing the process, i.e. are there non-random patterns suggesting certain problems with the system for certain types of user. Any system needs to be sufficiently well designed and easy to use in order for successful diffusion to happen. We know already that this is the case in Estonia. All this nevertheless does not mean that the system is problem free and non-random user subgroups do not face problems. This section will take a closer look at to what degree this might be the case.

We showed in the session length section that e-voting happens very fast and that this indicates a relatively problem free e-voting process. The caveat is that the session length could only be computed for successful voting sessions, i.e. for sessions that resulted in a valid e-vote being cast. There are cases in the log data where that is not the case, because somewhere in the voting process a user or machine driven problem resulted in the session being terminated before a vote was submitted.

From a technical viewpoint, failures can occur in the candidate list retrieval and vote submission process. There are also cases where the candidate list is successfully received, but after this the session is closed for some unknown reason. One possible explanation for these cases is just a wish to observe the candidate list or to study the e-voting mechanism, without an intention to submit a vote. In this case it is technically not a problem of the system, the user simply decided not to cast a vote and closed the voting application. Another user specific reason might be that the user has forgotten their PIN2, which is required to confirm, i.e. digitally sign, their vote choice.

Table 5.3 shows the occurrences of different types of failures to vote. Nearly 2.7 % of all users in 2013, 1.1% in 2014 and 1.3 % in 2015 did not cast an e-vote by the end of the voting period. These are people who had problems with downloading the candidate list, signing the vote or did not even try to cast an e-vote.

Table 5.3: Percentage of users who failed at least once, failed every time, did not cast a vote

	2013	2014	2015
Failure to receive the candidate list at least once (%)	1.04	0.74	0.78
Failure to receiving the candidate list at every try (%)	0.06	0.04	0.03
Failure to cast a e-vote (%)	2.65	1.10	1.30

We also see that roughly 1% of users in 2013, and even less in 2014 and 2015, faced some difficulties in receiving the candidate list. But even if the first session was unsuccessful, the user can try until they succeed or fail in finally getting the complete list. We see that the latter share was maximized at 0.06% in 2013, which means 6 out of every 10 000 e-voters failed to get the candidate list, which was a very small share and was even smaller in 2014 and 2015. Another illustrative exercise should show how small that share really was. Given that we know 0.06% failed at the first crucial step of downloading the candidate list at every turn in 2013 and that 133 808 e-votes were cast during that year, which makes up the other 99.94%, then such a failure rate would mean 80 potential e-voters failed to cast their e-vote. This is a minuscule number in the given context.

More importantly, this small number is only potentially problematic if the failure is not random. In other words, if everybody has an equal chance to fail, the small amount of failures observed in the Estonian case, indicates that the system of e-voting functions very well.

Therefore, instead of looking at sheer failure ratios, we should focus at the potential non-randomness of failure and ask two questions in particular. Which age group has a greater chance of failing in general and does the rate of failure vary among different technical appliances used for voting? Ideally for actual randomness of failure to have occurred there should be no differences.

It turns out that the failure rate in candidate list retrieval process is highest within middle-aged users, especially among males, as shown in the upper row of Figure 5.8. It shows the rate of failure for each age level according to gender, the smoothing trend lines are again predicted failing probabilities extracted from a logit regression model with age as the predictor. The unexpectedly high failure rate by mid-aged men is however due to the fact that Mobile-ID usage for digital identification follows the same pattern, i.e. it is widely used by middle-aged men. It just so happens that Mobile-ID has more possibilities to fail technologically at different stages of the e-voting process than does the ID-card[20] given the additional steps built into that particular digital identification process. In addition the mobile network might be temporarily disrupted adding to more potential failure points.

A different picture emerges when taking a look with accounting for the separate identification modes, for example by looking at failure rates among only Mobile-ID users. This is done in the bottom row in Figure 5.8. The failure rate within middle-aged users is actually lower than among the youngest or oldest users. Notice also that failing to properly vote at least once is quite common within Mobile-ID users. We have point out however that only between 8 to 12% of e-voters used this identification mode during the elections between 2013-2015.

[20]Failure to receive the candidate list occurred only in a few sessions with ID-card usage.

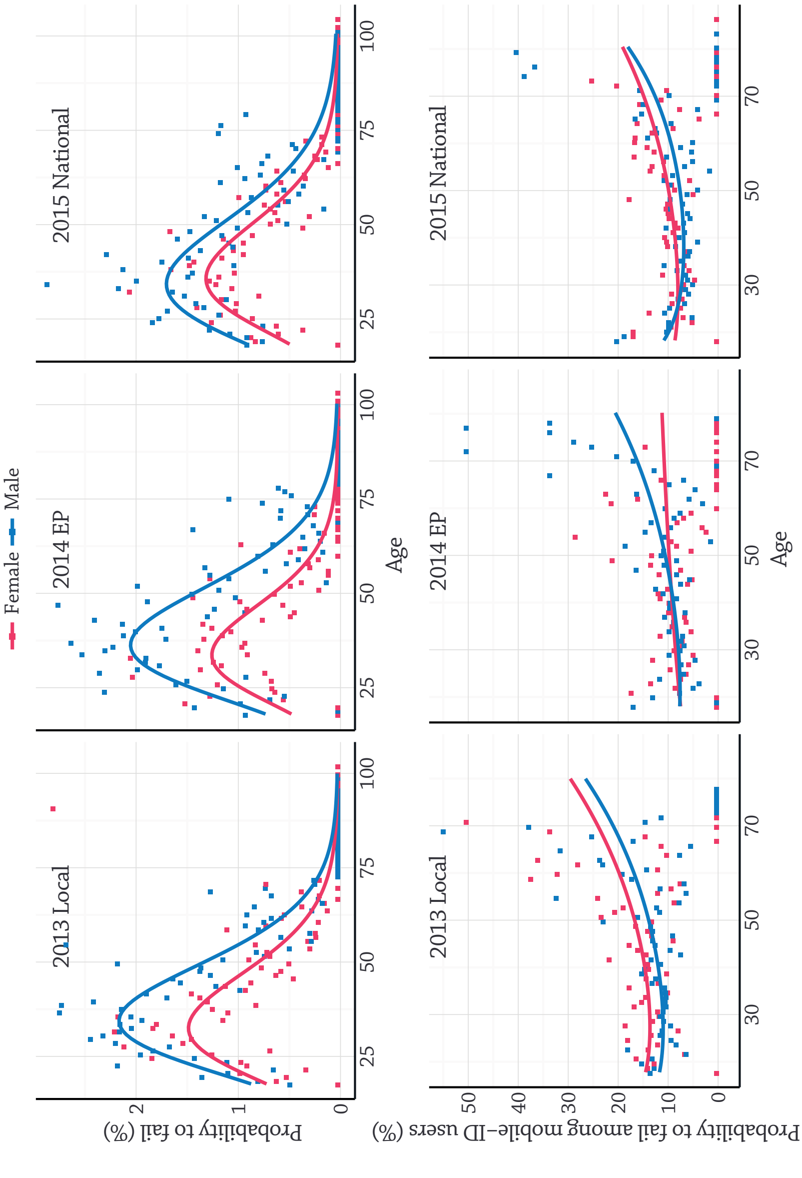

Figure 5.8: Probability to fail by age and gender

5.5 Conclusion

A detailed look at Estonian e-voting log files between 2013 and 2015 brought out a number of interesting findings. We list these here once more and discuss possible implications.

First, we saw that a typical e-voter is very much like the general Estonian population mid-aged. We knew from previous research that age as such does not differentiate e-voters from paper voters in Estonia any more. It is, however, still surprising to see the complete age distribution of e-voters peaking around 35-45 years and how small the share of the youngest voter segment is. This is surprising because it goes against the conventional wisdom of online participation being something that primarily younger people engage in. The log data shows once more that e-voting has the potential to diffuse widely and turn into a normal mode of voting.

A second major surprise is the speed with which people vote online. The whole transaction takes well below three minutes and the older the voter the less so. We put forward some potential explanations for the unexpected relationship between age and e-voting speed, but regardless of the exact reasons, it goes to show that the system is designed well enough not to place any difficulties in front of age groups who should in theory be less well versed in modern technologies. Given that survey data shows the average voter to have to take a 30-minute round trip to vote on paper at a polling station, the less than 3 minutes taken to e-vote in the comfort at your own home or workplace brings about a more than tenfold saving in time. The convenience and speed are probably the most important reasons why people chose to e-vote in the first place.

A third and no less important finding was the very small share of re-voters, which is one of the disputed aspects regarding Estonian e-voting. We saw that few people did so and the ones who did, re-voted within a relatively short time span, reducing even further the potential effects of vote relevant outside events. The log data therefore suggests that e-voting in Estonia does not introduce uncertainties into the advanced voting period, at least not to a degree that should worry anyone.

Finally, we also examined failure rates in e-voting and saw it to be low and mainly connected to a specific mode of digital identification, the Mobile-ID, which is more prone to problems due to its more complicated nature than the ID-card. Failure rates were therefore non-random, but low enough and related to external factors than to the voting itself.

In sum, the log data indicates Estonian e-voting to function remarkably well. All age groups e-vote, do it fast, generally e-vote only once and the overwhelming majority do not face technical difficulties in doing so.

Chapter 6

Mobilization

MIHKEL SOLVAK

6.1 Introduction

Multiple chapters in this book examine from different angles the potential impact of e-voting upon voter turnout. This chapter does so by re-examining the so called “bottleneck” model of e-voting proposed by Vassil and Weber (2011).

The bottleneck effect arises from two counteracting tendencies. Theoretical comparative accounts by various authors argue that use of e-voting is most likely among people who are typical voters in any case. However, the potential to mobilize someone into voting online can paradoxically only be largest among people who are very unlikely to participate in the first place. These two contrary developments might explain the rather small increase in aggregate voter turnout in Estonia that can be attributed to the introduction of e-voting. If a large share of citizens simply switches from paper voting to e-voting, while the non-participating minority does not start to participate in elections as a result of the new mode, then widespread usage of e-voting has negligible effects on voter turnout.

The crucial distinction here is therefore between the usage of e-voting and its impact (Vassil and Weber, 2011, p. 3). We will show that these two might not necessarily go hand-in-hand. Vassil and Weber noted that even though the usage of e-voting equals turning out to vote, it does not mean that the probability of turning out as such will be increased by the availability of this new mode. Impact, however, does mean that the individual has an increased probability of turning out due to this new voting mode. We will proceed to examine the difference between impact and usage and through that shed some light on whether and how e-voting might mobilize citizens to vote at higher rates.

6.2 Mobilization and bottleneck of e-voting

What constitutes usage of a voting mode is straightforward and does not need further explanation. Impact, however, is a somewhat less precise concept. Impact of e-voting on voter turnout according to my understanding constitutes an increase in the turnout rate resulting from the introduction of this new voting mode. Such an increase itself can however only come about when the introduced mode affects factors that are associated with a higher propensity to participate. Understanding this is crucial when discussing e-voting and turnout as the bottleneck model claims that usage does not equal impact, that is, mere usage might not be associated with a generally higher propensity to participate. Furthermore, it suggests that impact might actually decrease with an increase in the likelihood of usage (Vassil and Weber, 2011, p. 4).

What does that actually mean? Figure 6.1 outlines the crux of the matter. It shows three voting scenarios for two consecutive elections. In the first scenario, citizen A votes on paper in both consecutive elections; this voter exemplifies a situation with no usage of e-voting and no impact on turnout, because they voted in both elections. In the second scenario, citizen B votes on paper in the first and e-votes in the second election; this voter exemplifies a situation with usage of e-voting, but again no impact on turnout because they voted in both elections, albeit using different modes. In the third scenario, citizen C does not vote in the first election and e-votes in the second election; this voter exemplifies a situation with usage of e-voting and mobilization, because they switched from non-voting in the first to voting via e-voting in the second election.

We claim that the more likely scenario of e-voting is the one exemplified by citizen B and not C. Why?

A detailed look at e-voters in the first e-enabled elections showed them to be well resourced younger people. Given that such people are very likely to vote in the first place, the potential impact of e-voting might actually decrease with increasing usage potential. In other words, if the probability of voting as such increases together with the probability of e-voting then e-voting will subsequently have no real impact on behaviour. The ones already participating will now simply do so using a different voting mode, i.e. your regular voters will now e-vote. A true impact will only be achieved if non-voters are converted into voters. But given that typical non-voters tend to be the exact opposites of typical voters and hence also e-voters, the potential impact of e-voting should actually be greater among people who paradoxically have the lowest usage potential. This means that the impact of e-voting actually increases as the probability of voting and subsequently e-voting, decreases. This is precisely the heart of the aforementioned bottleneck effect: as the usage probability of e-voting increases, the probability of being mobilized by e-voting decreases the more similar someone is to a "typical" voter. This mechanism means that a substantial share of e-voters are most likely made up of voters like citizen B, while a small share is made up of voters like citizen C.

The given examples also point out why the hypothesis that e-voting will have a positive impact on turnout is if not outright misplaced, then it is at least overly optimistic. Turnout can by definition only increase when people who previously did not vote now do so. This means it can only come about by mobilizing people who for some particular reason are not your typical voters, it also means it is hard to get these people to vote. Accordingly, in order for e-voting to have an impact on turnout, it should reach the most difficult crowd in terms of mobilization. By implication it also means that e-voting should somehow address the factors that have kept these people

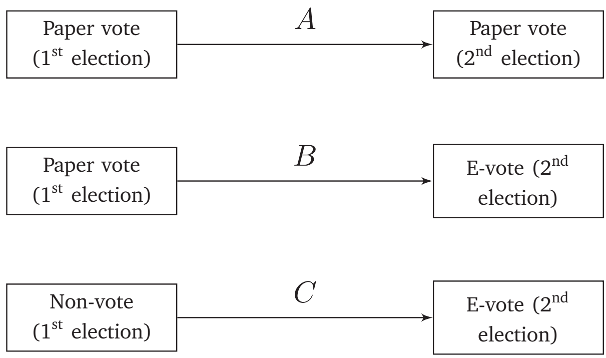

Figure 6.1: Schematic presentation of use cases: A no usage e-voting and no impact on turnout, B usage of e-voting, but no impact on turnout and C usage of e-voting with impact on turnout.

from participating in the past. Only then can we say that e-voting itself has an impact on voter turnout, meaning it has increased the probability of participation as such.

Similar "Catch-22" situations or bottleneck effects have been noted in other areas. A classical account on voting behaviour by Lazarsfeld et al. (1944), demonstrated how voters who follow politics more intensively have at the same time stable preferences, while people with more unstable preferences, and hence potential vote switchers, do not follow politics that much and are therefore paradoxically less exposed to information that might alter their preferences. The high potential for vote switching is hence not realized due to their low media exposure. Analogous effects have been pointed out in other accounts (see Zaller (1992)).

We will turn next to if and to what degree this empirically applies in the Estonian case, and by doing so hopefully untangle the apparent puzzle that whereas e-voting has become hugely popular, the impact on voter turnout has been negligible at best.

6.3 Analysis of mobilization

Figure 6.2 is already familiar from the introduction, but is reprinted here as it needs to be examined once more before turning to the mobilization mechanism. The figure shows how aggregate turnout has changed over the years in different types of elections in Estonia. Turnout after the introduction of e-voting has on average been slightly higher in parliamentary elections (by 0.12 percentage points) and clearly higher in local elections (by 3.5 percentage points). Although shown in the figure, a similar development for European Parliamentary elections is not reasonable, as the turnout was simply so low in the first election of this type that there was more room for increasing than decreasing turnout in subsequent elections.

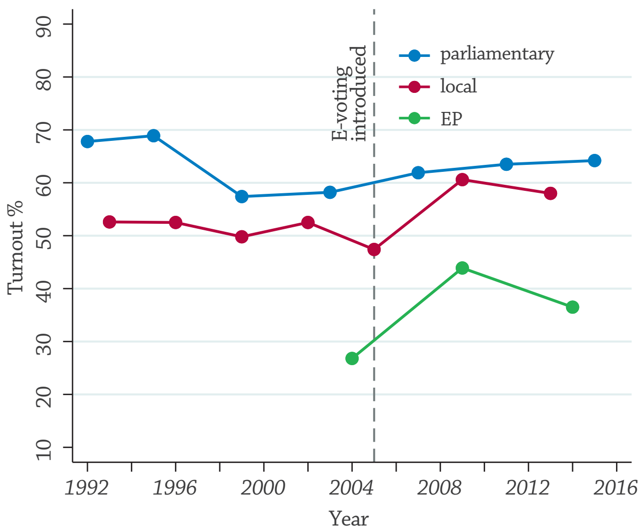

Figure 6.2: Turnout trends in different elections in Estonia

Such simple before-after comparisons do not yet suffice however to claim any positive impact of e-voting on turnout. We can simply state the fact that turnout is slightly higher now, but if it is higher due to e-voting is a much more difficult question to answer. For this we turn next to disentangle individual level associations.

Remember that we only see e-voting to be mobilizing when a person who has previously not voted does so now by using the e-voting mode (citizen C in Figure 6.1). Only then does e-voting have an impact on turnout. Ideally we would need an experimental setting to examine if the option to e-vote alone produces mobilization to vote. We do not have this luxury, as all vote eligible citizens in Estonia have since the first e-enabled election had the possibility to do so. We will therefore simply look at prior behaviour and see whether any change in this could be attributed to e-voting.

First, let us look at the share of people who behaved as the above outline citizen C, i.e. people who switched from non-voting to e-voting. We used survey data to do so and investigated self-reported behaviour for past elections. Table 6.1 shows: the share of people who reported to have e-voted in the current election, but not to have taken part in the previous two elections (*mobilized*); people who did not participate in any of the elections (*not mobilized*); and the other vote eligible people who have either always paper voted, e-voted or voted using a combination of these two over the examined elections (*others*).

Table 6.1: Share of e-voters mobilized by the option of e-voting (EP – European Parliament election)

	Mobilized (%)		Others (%)
	Yes	No	
2009 EP	3.1	8.8	88.1
2009 local	5.3	14.5	80.2
2011 national	4.1	16.1	79.8
2013 local	2.4	7.0	90.6
2014 EP	2.3	6.5	91.2
2015 national	2.4	9.3	88.3
Total	2.5	7.8	89.7

Though the shares fluctuate a bit over the years we see that roughly 2 to 5% of eligible voters were being mobilized into voting via e-voting.

Given that respondents tend to slightly over-report participation and the actual numbers are most likely a bit below these figures, we can say that the individual level picture is in line with the small increase in aggregate voter turnout observable after e-voting was introduced in Estonia as shown in Figure 6.2. There is hence likely some mobilization going on. A caveat is that we cannot be entirely sure that the ones we define to have been mobilized to vote by e-voting simply decided to participate in that election no matter what, i.e. would have voted also without e-voting. Such a counter-factual situation cannot be created, but we can examine what factors covary with the observed mobilization and juxtapose this with the usage of e-voting.

6.4 Bottlenecks

In order to demonstrate whether there was a bottleneck effect, we compiled a set of characteristics into a latent factor for e-literacy. This factor was made up of four separate items: 1) trust towards e-voting; 2) trust towards internet transactions; 3) internet usage frequency per week; 4) a self-evaluated level of computer skills. A high score for all these characteristics means the individual is a trusting and avid internet user with good computer skills, or in words an e-literate person. This e-literacy is captured in one score extracted from a factor analysis, the technical details of which are laid out in Appendix B. We used this e-literacy score in two separate models. First to statistically predict the likelihood of repeated usage of e-voting, by comparing people who have e-voted multiple times with people who have always paper voted. This showed whether and how e-literacy distinguishes e-vote users from non-users within the voting population. Second, westatistically predicted being mobilized by e-voting, juxtaposing people who had switched from non-voting to e-voting with people who had always e-voted. The aim of this comparison was to show how do typical e-voters differ from mobilized e-voters in terms of e-literacy. The first model will therefore tell us about usage of e-voting and the second about mobilization owing to e-voting.

The results are displayed in Figure 6.3, for the full results of the models, including the statistical controls, see Appendix B. Though merely a figura-

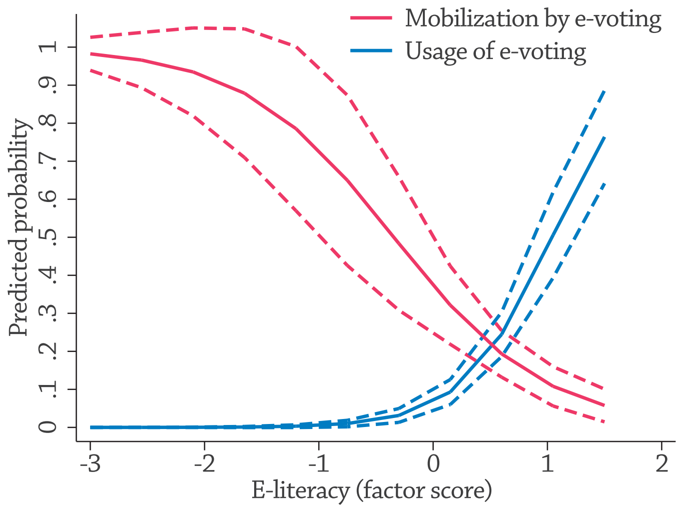

Figure 6.3: E-literacy and e-vote usage compared to mobilization

tive image, a particularly strikingly bottleneck emerges. We see the higher someone scores on e-literacy, the more likely they are a repeated user of e-voting. For those scoring lowest on e-literacy, the probability of being a repeated e-voter is in fact virtually indistinguishable from zero. The picture is completely the other way around for mobilization however. Mobilization is actually most likely among people who scored lowest on e-literacy. In other words, switching from non-voting to e-voting is most likely among people who are the complete opposites of repeated e-voters in terms of trusting online transactions and e-voting itself, and internet usage frequency and computer skills.

So in general, we can say that using e-voting is most likely among the usual suspects, but being mobilized into voting by e-voting is actually most likely among very unlikely e-voters. The reader might be puzzled by this empirical finding, but we emphasize again the distinction between usage and impact. High impact often goes hand-in-hand with low usage potential.

This also means that e-voting has indeed increased the probability of turning out to vote, but only for a small subset of the voting eligible population. Among a much larger group, this probability was already high so the introduction of e-voting simply led to them switching from paper voting to e-voting. The result is a puzzling pattern of large voter numbers e-voting, but a minute difference in aggregate voter turnout.

It might of course be that the average associations observed here are very different among specific subgroups of potential voters. It has been argued for example that e-voting should primarily help to combat low turnout among the young. We do know that younger people in very many societies have a generally lower voter turnout level than people in their middle-age (Franklin, 2004).

Figure 6.4 shows the association between age for e-voting use and mobilization as a consequence of e-voting.

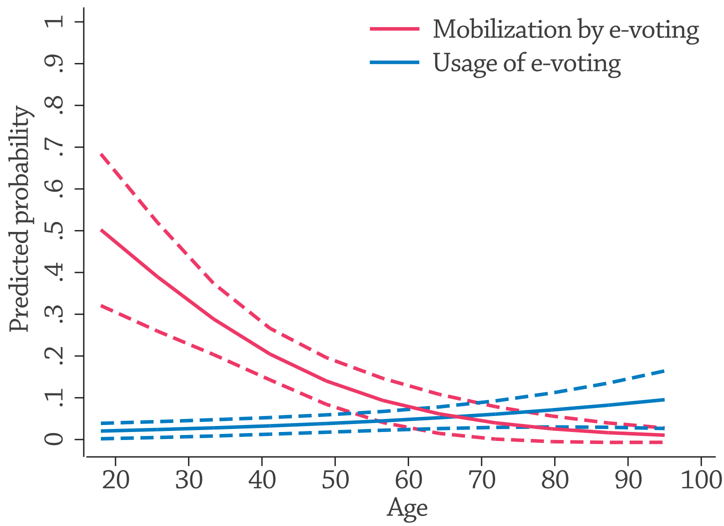

Figure 6.4: Age association with usage and mobilization by the availability of e-voting

We see how mobilization was most likely among the youngest eligible voters, but again the probability to e-vote was paradoxically also lowest amongst this group. This concurs with the evidence presented in Chapter 4 on diffusion, which showed that e-voting has never been most likely among the youngest segment of voters. Nevertheless, the figure clearly shows that mobilization is most likely among age groups that in general tend to have the lowest participation rates in elections and that e-voting does therefore fulfil its theoretical mobilization potential, at least partially, as intended.

Besides the general low turnout among younger people, often explained by the life-cycle, other subgroups in society also stand out as having below average voting rates. The resource theory of voter turnout for example predicts individuals with fewer cognitive and material resources to have lower participation rates (Smets and van Ham, 2013). This has been empirically verified in a number of countries with people from lower education and income categories demonstrating clearly lower turnout rates (Franklin, 2004; Gallego, 2010). Figures 6.5 and 6.6 examine this in relation to e-voting and mobilization in Estonia. Some of the education subgroups are simply too small to estimate with separate models, so we will examine the existence of a bottleneck with simple descriptive statistics. Figure 6.5 shows the associations with three educational categories.

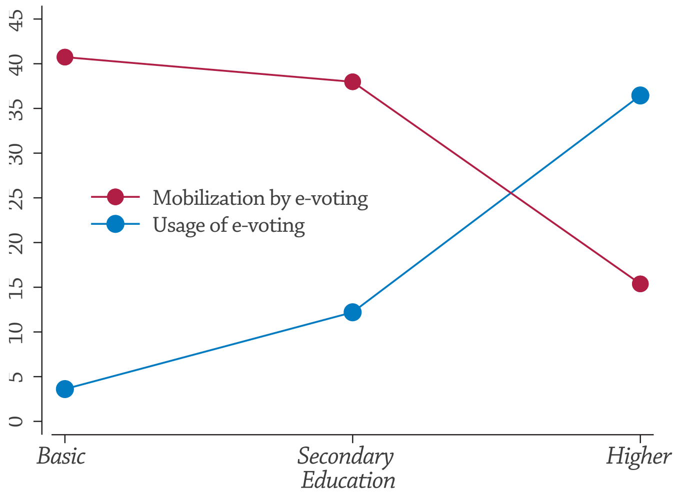

Figure 6.5: Usage of and mobilization as a consequence of e-voting according to level of education

Again the same pattern emerges. E-voting had the strongest impact within the lowest educational category, where at the same time the share of consecutive e-voters was the lowest. The mobilization potential is therefore indeed best realized in a group where participation rates are traditionally lowest.

The other voting resource proxy – income – is examined in Figure 6.6.

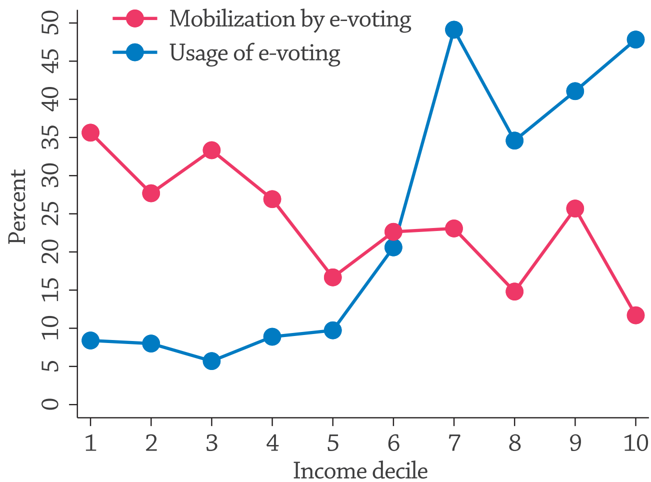

Figure 6.6: Usage of and mobilization as a consequence of e-voting according to income level

Though with slight fluctuations, the share of people mobilized by e-voting and the share of people engaging in repeated e-voting follows the same bottleneck pattern. The higher the income decile, the lower the share of mobilized people and the higher usage of e-voting. The pattern is positive in the sense that the lowest income segment is again mobilized by e-voting to a comparatively larger degree than people with high incomes. A downside is again the low usage share of e-voting by the less well off people.

All this means that internet voting has the most potential to affect those subgroups who traditionally have lower turnout levels. As such it clearly works as intended, but the question is how to incentivise people to a degree that it actually starts to increase aggregate turnout levels. One could of course settle with the current situation and claim that even modest mobilization levels among those groups least likely to participate is a good development. This is a valid argument given that it might bring direct positive consequences for the political representation of such societal groups. If that can be achieved by simply introducing a more convenient voting mode, then e-voting has fulfilled part of its positive potential.

However, given that the probability of voting among these groups is very low, the potential impact on turnout can only remain modest.

6.5 Conclusion

With this chapter we reported on the so called bottleneck effect that arises when e-voting usage is juxtaposed with e-voting impact. The more e-literate a person is, the more likely they are an avid user of e-voting, but the less likely they have been mobilized to vote by e-voting. This phenomenon explains the puzzle of large e-vote user numbers going hand-in-hand with very small changes in aggregate voter turnout numbers.

This contradiction comes about because an increase in turnout can only happen when people who previously did not participate had decided to do so. It just so happens that these very same non-voters are also quite unlikely to be sufficiently e-literate to e-vote in the first place. Paradoxically, to have a clear impact on voter turnout, e-voting should mobilize people who are very unlikely to be able to e-vote.

Nevertheless, given that the non-voting subgroups are exactly the target group for increasing turnout, even limited mobilization effects are an improvement.

The findings suggest quite simple recommendations. Given that the usage of e-voting is not a goal in itself, but a means of ensuring better representation and ease of participation, the most effective way to achieve this is to increase e-literacy and not strictly to promote casting one's vote remotely over the internet. The spread of any e-service, public or private, should by definition also increase the general ability of e-voting among those population segments who have tended to shun online participation. But when it comes to overall participation rates, we have to contend with a rather limited effect of e-voting.

On the bright side, the large overall numbers of e-voters shows that the intended ease of participation has indeed been realized. It was simply achieved by large numbers of habitual voters switching from paper to e-voting. All in all, we have to conclude that the hope of e-voting increasing turnout levels was somewhat misplaced. It presents a so called “technological fix” to a participation problem that cannot really be addressed by technology alone. Low turnout levels among certain social groups are down to multiple factors, with social marginalization and political discontent being often cited as the most prominent ones (Smets and van Ham, 2013). E-voting does not fix these issues, it is merely a new way to engage in voting.

Chapter 7

E-voting and the cost of electoral participation

Mihkel Solvak

7.1 Introduction

Chapter 4 demonstrated how e-voting has over time been picked up by a diverse set of voters. Having answered the question on whether the population will start using e-voting we can turn our attention to trying to disentangle the reasons behind usage. After all, when roughly a third of voters use it, two thirds still do not. Why is this?

E-voting should in theory impact upon participation rates. It saves time and should ease participation, especially among people who value ease and speed. In this context it is relevant to ask if e-voting is more prevalent among citizens for whom it brings the greatest relative benefit compared to regular voting. In other words, we investigated whether e-voting reduces the cost of

Academic reference: "I abstain if Voting Takes Me More than 30 minutes: The Impact of Internet Voting on Reducing the Cost of Electoral Participation" by Mihkel Solvak, Kristjan Vassil and Priit Vinkel, presented at the 8th ECPR General Conference, 3–6. September 2014, Glasgow, UK.

participation.

We might assume that the answer to this question is ”yes”, but previous research on voter turnout and the chapters above suggests that the answer might not be this straightforward. The previous chapter showed that the largest benefits from e-voting could potentially be realized for people who are very unlikely to vote in the first place. Instead of making participation easier for citizens for whom it is relatively more difficult, it seems to ease participation for those for whom it is relatively easy to begin with, i.e. for people who already participate regardless of e-voting. But for the careful reader it should be clear that we have not really provided a definitive answer to this questions as we have not looked at how the time spent on voting affects the likelihood of choosing to e-vote. We know from Chapter 5 how fast people e-vote, the next step is hence to present a detailed comparison with time taken to vote at the ballot station. How much of a relative gain does e-voting actually bring and does that compel some people to e-vote at higher rates than others. We turn again to survey data on Estonia to shed some light on this issue.

7.2 Cost and voter turnout

The core question is whether a reduction in costs associated with a behaviour will or will not make it more likely to happen. If e-voting makes voting easier, will more people cast their vote? The act of voting can thus be considered a result of a cost-benefit analysis. The decision to vote can be seen as a combination of three elements: the choices on the ballot; the probability of casting a decisive vote; the cost associated with voting (Downs, 1957, pp. 36-50). If the benefits outweigh the costs, going to the polls is likely.

It seems self-evident that e-voting provides a notable reduction in voting costs, as it makes going to a polling station unnecessary. The convenience and speed of e-voting should make it especially appealing, particularly for those whose decision to participate is influenced by such factors.

Naturally, whether people vote or not should depend on what they think about the selection of candidates, if they think their vote matters and the costs that are associated with going out to vote. Such a cost-benefit analysis happens against a backdrop of ideas about civic duty, and attitudes towards and engagement with the political system. In this context we are looking at how e-voting can reduce the costs of voting and affect this kind of participation. We can postulate that voting will take place if the benefits associated with the actual vote choice (the policy benefits a vote will bring) outweigh the costs, and if costs are lowered, all else being equal, the probability to vote should be higher.

Previous research that has looked at the effect of voting costs at a more general level (e.g. the day of voting or requirements of voting registration) has indeed shown that this is the case (Franklin, 2004). If voting rules stay the same, but voting costs are different between individual citizens, we can expect that those for whom the costs are lower are also more likely to vote.

However, it might also be the case that certain social and cognitive factors that decrease participation are related to higher costs of voting. Just think of the bottleneck phenomenon discussed previously; for example, elderly people with reduced mobility might benefit from e-voting, but paradoxically they might not command the resources needed to navigate in an online environment. As a result, the positive cost reducing potential of e-voting for these people will be nullified by them not being able to conduct their daily transactions online.

7.3 Analysis of voting costs

The voting cost is understood here as the amount of effort it takes to vote measured in the minutes it takes to make the round trip to the polling station. In other words, the physical proximity of the polling station. This measure is available for five elections, the 2009 local, 2009 and 2014 European Parliamentary, and 2011 and 2015 parliamentary elections.

A look at the distribution of this factor among our sample (see Figure 7.1) shows that for the majority of respondents, 75% in total, it should take no more than half an hour to make the trip to the station and back. There is, however, a non-negligible share of respondents (25% in total) for whom it takes in excess of half an hour, one or even two hours. It is apparent that there is considerable variance in the cost of voting among the respondents.

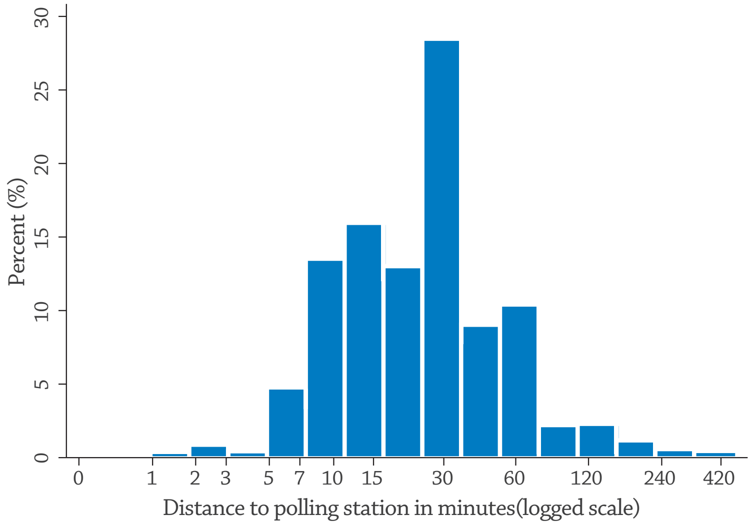

Figure 7.1: Distance to the polling station in minutes (logged scale)

We estimated a multivariate regression model that also includes other non-instrumental motivations and the perception of the choices on top of the measurement of voting cost as predictors. This should show whether association with voting cost is comparable to associations with other motivations behind the vote decision, see Appendix C for technical information on the model.

Analysis of the association between the cost of voting and the probability to either vote at a polling station or e-vote shows the expected relationships. *Id est* distance to the polling station had a small negative effect on the probability of voting, but a very strong positive effect on the probability of casting a vote online. These two effects of the voting cost on voting upon voting as such and e-voting specifically, extracted from a multivariate regression model, are compared in Figure 7.2. The contrasts are astounding. Voting costs seem to be less of an issue when deciding whether or not to, but very important regarding the choice of how to vote. While the association with cost is modest and negative for voting, meaning increased costs lead to a small decrease in voting probability, the association with e-voting is very strong and positive. One can see that e-voting already becomes more likely than voting on paper when the round trip to the polling station takes more than 30 minutes. Importantly, the cost saving effect kicks in immediately. For example, someone having to take a mere 10-minute trip to vote is already more likely to e-vote than someone who can do it in 5 minutes. This strong association spans the entire range of the voting cost measurement.

Appendix C lists in addition also the associations between non-instrumental voting motivations and the probability to vote or e-vote. It suffices to say there is a relationship with voting as such, but not with e-voting. This means the voting mode is indeed chosen based on the voting cost, while the decision to vote is primarily influenced by factors other than cost.

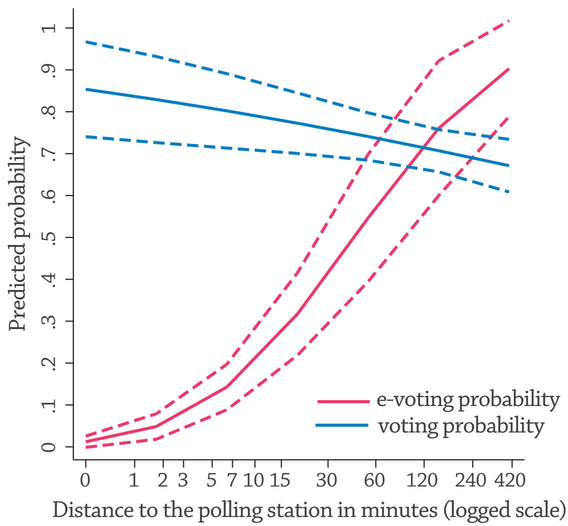

Figure 7.2: The association between distance to the polling station with voting and e-voting

There is however still the possibility that increasing costs might correlate with factors that at the same time suppress e-voting. For example, voting costs are higher in rural areas, but we also know that the rural population in Estonia is on average somewhat older, i.e. longer distances to polling stations correlate with age due to the non-random population distribution in Estonia. Similarly, the share of manual agricultural labourers with lower computer skills is higher among the rural population, creating again a non-random correlation between lower e-voting skills set and larger paper voting costs. The question is therefore if the overall pattern observable in Figure 7.2 also persists when we examine sub-populations with traits that might reduce their potential to e-vote. To do so we examined whether the association between voting cost and the probability to e-vote is related to socio-demographic factors (age, education and income) and technological skill (computer literacy). These associations are shown in Figures 7.3 to 7.6. The horizontal axes show distance to the polling station and the vertical axis the probability to e-vote. The lines show how the probability to e-vote changes as distance to the polling station increases. For each figure, the lines are respectively split between different levels of age, education, income and computer literary.

Figure 7.3 shows that for people living close to a polling station, differences in their computer skill levels do not translate into a tangible difference on the probability of e-voting. Only people with the poorest computer skills were somewhat less likely to e-vote in comparison to people with the highest skill level when they were within a distance of 15 to 60 minutes from a polling station. For all other cases we can say that regardless of their computer skill level, the further away from a polling station someone lived the more likely they were to e-vote. This is an important finding, for it partially refutes the claim that technological barriers will reduce the possible benefits that e-voting might bring. It seems that from a certain voting cost level upwards, e-voting brings such a reduction in costs that technological barriers will be overcome and eventually will stop functioning as prohibitively high thresholds to participation.

Turning to the interaction between distance to the polling station and age (Figure 7.4) also reveals a surprising picture. All age groups showed a similar increase in e-voting probability as their voting costs increase. Again, cost saving seems to override any age related problems in using online voting.

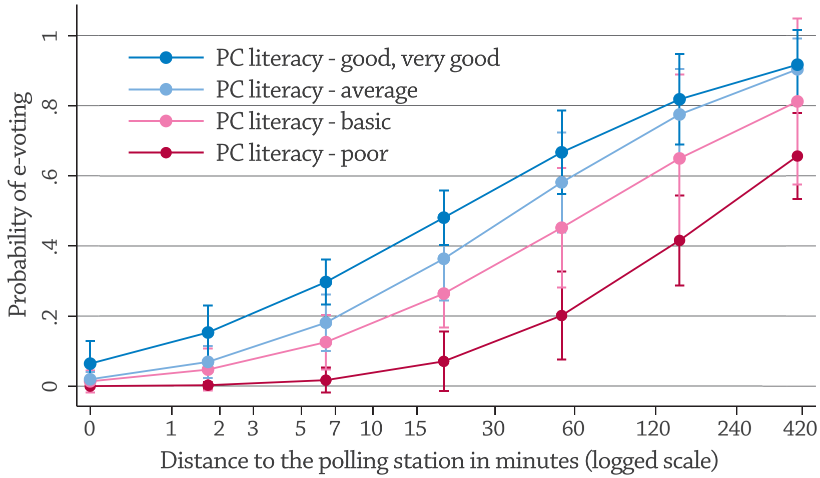

Figure 7.3: The effect of distance to the polling station on e-voting conditional on computer literacy

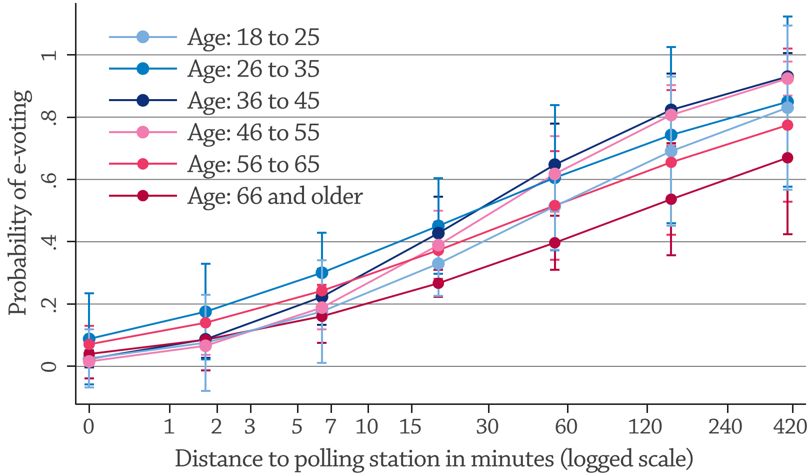

Figure 7.4: The effect of distance to polling station on e-voting conditional on age

Turning to the interaction between distance to the polling station and education (Figure 7.5), we saw no discernible difference. Regardless of the education level, the higher the voting costs, the more likely it was that someone opted to e-vote.

The same applied to income (Figure 7.6), i.e. regardless of income level, the further away the station was the more likely an e-vote became.

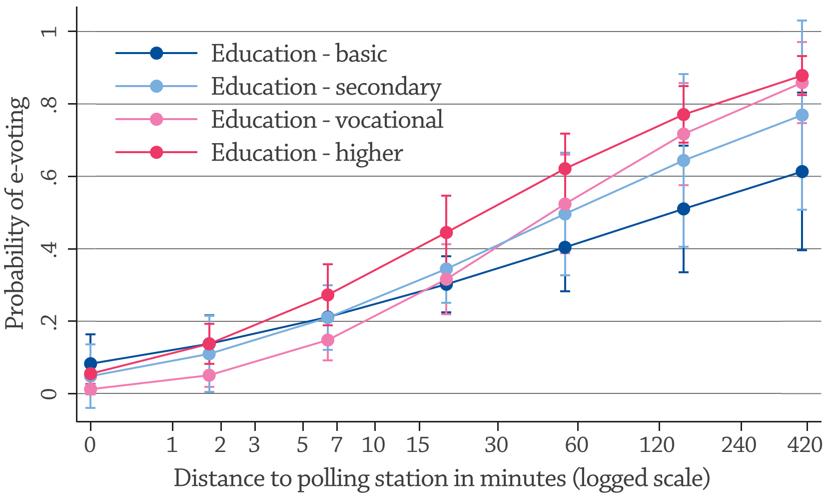

Figure 7.5: The effect of distance to the polling station on e-voting conditional on education

In sum, the results of these interactions were unexpected, as the traditional resource theory of voting suggests that those who have resources are more likely to participate. The much discussed digital divide problem is essentially a resource based one – those who have the resources to be able to command technological solutions should also be the main beneficiaries and users of new non-traditional voting technologies. This, however, is not the case here, as high physical voting costs significantly increased the probability of people participating through a new technology (e-voting), regardless of their resource levels.

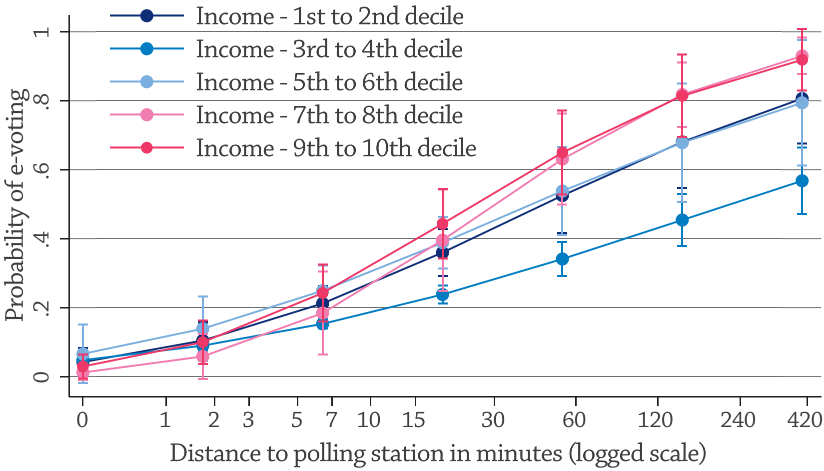

Figure 7.6: The effect of distance to polling station on e-voting conditional on income

7.4 Conclusion

We examined how e-voting functions as a cost saving mode of participating using post-election survey data from five Estonian elections. Proceeding from Downs (1957), we suspected that higher voting costs should reduce the probability of participation, but increase the probability of e-voting. The latter might however not hold depending on the technical skills and various resources commanded by the voter, i.e. cost saving might be merely theoretical and not realized when it happens to coincide with barriers to effective use of the e-voting technology.

The data clearly shows that e-voting is very likely when the cost of conventional participation increases. We saw that the further away the polling station, the lower the probability of voting, but this effect was quite modest. For e-voting however, there was a very strong positive effect. Citizens close to a polling station (a return trip of 0 to 30 minutes) were still more likely to opt for a paper ballot, but if the total distance was more than 30 minutes, e-voting was clearly the more probable option. The strong positive association between voting cost and e-voting also persisted after other socioeconomic factors were taken into account. Only a lack of computer skills played a small role in choosing whether to e-vote, i.e. for citizens the same distance away from a polling station, those with poorer skills were less likely to e-vote than very computer literate people. This difference disappeared however as the distance to the polling station became progressively larger. We did not detect an association between cost and e-voting related to education level, age or income. Regardless of these sociodemographic factors, which we take as indicators for cognitive and social resources, e-voting became more likely the further away from a polling station somebody lived. Lastly, we also found that factors clearly inherent to the voting decision, such as non-instrumental motivations and the perception of the choices on the ballot, do predict strongly the probability to vote, but do not predict the choice to e-vote, confirming it is really the reduction in participation costs that drives e-voting.

30 MINUTES TO VOTE ON PAPER

Does e-voting make participation easier for those for whom it is hard or for those for whom it is already easy? The data shows the longer the distance to the polling station, and subsequently the higher the cost of voting, the higher the probability of e-voting. The critical limit is a 30 minute round trip to the ballot station, anything above that makes e-voting already more probable than voting at the polling station.

All voters, regardless of their sociodemographic background, face some costs related to voting, most notably the actual time cost of getting to the polling station. I demonstrated that the cost reducing effect of e-voting seems especially apparent as the time costs related to voting increase. The attractiveness of a low-cost convenient voting mode seems to overcome social and technical barriers related to its usage. I found that e-voting does have the potential to influence participation by radically lowering the cost of voting. The usage of this mode became very likely once the actual physical cost of voting had reached a critical level, which in our case was a half hour or longer trip to the polling station.

These results have clear implications for future debates on the adoption of e-voting solutions in countries currently discussing this option. The fear that this technologically advanced voting mode might heighten already existing social inequalities, such as creating a voting mode specific high usage barriers for people for whom mobility is a problem, e.g. the elderly, does not seem to hold for Estonia. If cost saving is high enough, e-voting will become very likely regardless of the social resources somebody commands. This is an encouraging finding. However, the findings also once more at the limited success of e-voting in influencing turnout levels. While this mode of voting does indeed provide a easier option to participate, this effect will only be felt strongly among a small sub-population for whom time costs might be a problem. We saw that only 25% of people reported to having to spend more than 30 minutes on their trip to the polling station. Given that such a voter segment is a minority and the main problem of decreasing voter levels is actually down to low turnout among the younger generations, I remain sceptical about the prospect of e-voting being able to slow down or even reverse the malaise that has led to declining turnouts. It was simply demonstrated that technology has the ability to ease participation by lowering its cost. However, if the problems are beyond costs, technological solutions might not be able to counter a wider societal trend of decreasing turnout.

Chapter 8

"Stickiness" of e-voting

Mihkel Solvak

8.1 Introduction

Even though the discussion on the pros and cons of electronic voting is still ongoing, the growth of e-voting in Estonia stands in marked contrast to the scepticism regarding the inclusiveness of this new way of voting. We have shown above that there has been a notable level of diffusion of e-voting in Estonia, i.e. the people who have adopted this way of participation have become an increasingly indistinguishable group of voters. Also, the positive potential of e-voting, which lies in the reduction of physical voting costs and lowering participation hurdles for people who have mobility problems, seems to a certain level to have been realised in the example of Estonian.

We know that the characteristics of voters who were associated with e-voting during the first elections, no longer apply. We also know that e-voting becomes more probable than regular voting when the round trip to the polling station is more than 30 minutes, which applies to roughly 25% of

Academic reference: "Once an e-voter always an e-voter: "stickiness" of e-voting" by Mihkel Solvak, Kristjan Vassil and R. Michael Alvarez, presented at the 73rd Annual Midwest Political Science Association Conference, 16-19. April 2015, Chicago, US.

Estonian voters. This raises the question why a large share of people who do not stand out from the regular voting population and do not face high voting costs have kept on e-voting through multiple elections? There might be two possible explanations for this kind of behaviour. First, it might be that usage of e-voting among a certain population share has become essentially random, which is hard to believe. Second, it might be that usage of this mode was and still is non-random and in fact reinforced by voters' previous behaviour, akin to the practice of voting itself. Anecdotal evidence does suggest e-voting to be “sticky”, that is e-voting once seems to make a voter very likely to stay faithful to this mode of participation in subsequent elections. However, this has not been examined in detail, and that is something which we will try to remedy in the following sections.

We examined how persistent the patterns of e-voting are and whether it can be considered habitual. We again used survey data, this time from six e-enabled elections over a six-year period (2009 local, 2011 national, 2013 local, 2014 EP and 2015 national elections) and a modelling technique called path analysis to examine repeated e-voters. The results show e-voting to be strongly persistent, with clear evidence of habit formation. The findings indicate once more the potential of e-voting to be widely used by the voting population (Vassil et al., 2014), and due to its self-reinforcing character, being able to if not raise turnout, then at least to arrest its decline in Western democracies.

8.2 Habit and e-voting

E-voting is still not widely used worldwide and studies on whether it might be habit forming are therefore lacking. We will thus look at studies on voting habits in general, as well as online participation, to shape theoretical expectations with regard to e-voting as habit.

The classical definition of habit in social psychology sees it as “responding based on learned associations between context and responses without necessarily holding supporting intentions and attitudes” (Aldrich et al., 2011, p. 540). A habit is thus an *automated behaviour* and not *motivated decision*

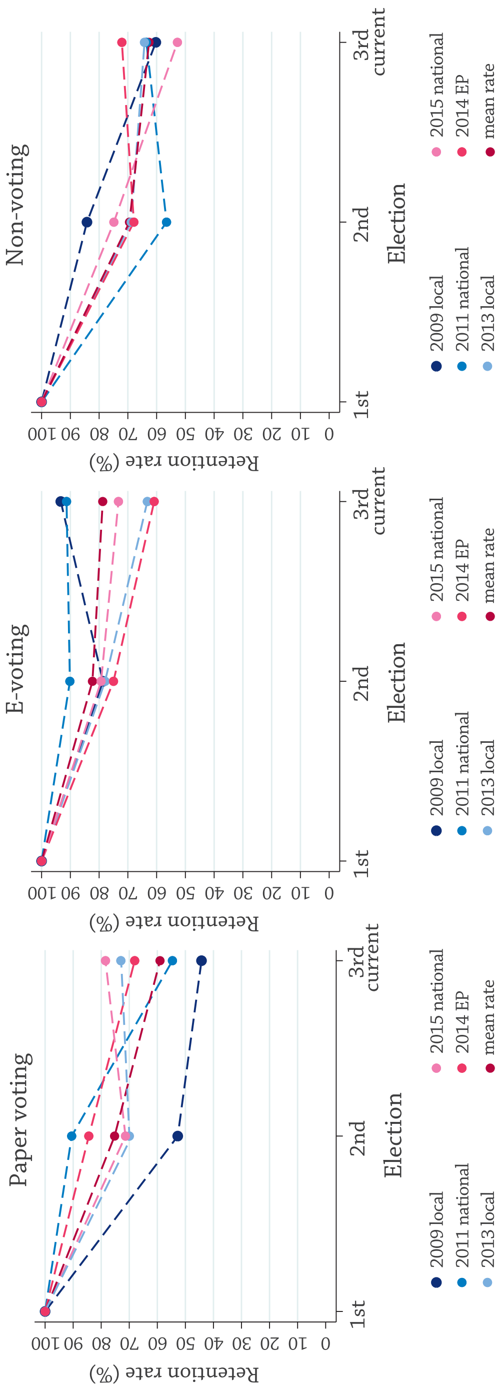

Figure 8.1: Share of people paper, e- or non-voting over three consecutive elections

making.

The easiest way to understand habit is to think of it as a behavioural strategy that simplifies decision making. Repetition equals a certain degree of automation. For choosing to e-vote, this implies a cognitive shortcut where the initial calculations that were made back in a previous election when a person e-voted for the first time are no longer apply when the voter e-votes in a subsequent election. So some sort of automation has taken over and one could say the voter e-votes simply because they have done so in the past.

E-voting can be seen as a habitual behaviour when we see the following:

- it is repetitive
- it is detached from the original motivations
- it is detached from supportive attitudes

We know from prior studies that people who vote in one election are very likely to do so in subsequent elections (Denny and Doyle, 2009; Aldrich et al., 2011; Gerber et al., 2003; Green and Shachar, 2000; Cutts et al., 2009; Nickerson, 2008; Plutzer, 2002; Dinas, 2012) and to a certain degree regardless of the type of the subsequent election (Gerber et al., 2003). Given that repetitive behaviour is the first precondition of a habit, we expect the same to hold for e-voting, i.e. past e-voting should be strongly associated with current e-voting.

We know already from Chapter 4 that a set of socio-demographic resources were initially strongly associated with e-voting, which fits well with findings from studies of online participation and engagement that showed specific resources act as enablers or preconditions of participation (Alvarez and Nagler, 2000; van Dijk, 2000, 2005; Margolis and Resnick, 2000; Putnam, 2001; Wilhelm, 2000). We also know that these resources clearly lost their associations with e-voting in Estonia over time. My expectation about meeting the second precondition of a habit is therefore necessarily similar, but with a twist. If e-voting is habit forming, then we should see the original resources associated with e-voting gradually replaced with a strong association with past e-voting.

E-VOTING AS A HABIT

We can consider e-voting to be a habit if it is a repeated behaviour that cannot be explained by either the social or demographic characteristics of a person, or by attitudes that a person might have about e-voting. E-voting is truly a habit if it is predicted only by previous e-voting.

Finally, besides commanding the resources needed to submit a vote over the internet, e-voters are also more likely to hold intentions and attitudes supportive of e-voting, such as trusting the system. This last factor is key to examining if and to what degree can e-voting be considered habit forming. Only if individuals engaging in repeated behaviour no longer hold supporting attitudes associated with that mode can we talk about a habit in any strict sense. This is the third precondition of a habit and also our final expectation.

To sum up, if e-voting is habit forming we should see the following. First, the behaviour needs to be repetitive, i.e. people who have e-voted once should also do so in subsequent elections. Second, their behaviour should not be associated with the traits usually associated with casting a vote online in the first place. And third, repeated e-voting should not have strong associations with attitudes in support of e-voting. Only then could we talk about this particular mode of voting being habitual.

8.3 Analysis of e-voting as a habit

To what degree e-voting is a repetitive behaviour can be inferred from Figure 8.1, which was constructed using surveys from six elections. In each of these surveys people were asked to recall their mode of participation in the current, previous and second to previous elections. The figure shows how the share of people who reported to have either paper voted, e-voted or non-voted in all three elections – the second to previous election (1^{st} election), previous (2^{nd} election) and current election (3^{rd} election) – changes. While there is some fluctuation depending on the election year it was obvious that e-voters keep on e-voting in subsequent elections to a larger degree than either paper voters or abstainers. For example, we saw that on average 80% of e-voters stayed faithful to this mode of voting in the subsequent two elections, while the corresponding share for paper voters and non-voters was roughly 60%. This confirms the first expectation, e-voting is repetitive – it “sticks” – and more so than voting on paper at a polling station or abstaining completely.

The second expectation will be tested by estimating a path model as outlined in Figure 8.2. Our focus was on how the current e-vote was influenced over time by e-voting in the second to previous election directly (path a) and indirectly by the second to previous election through an e-vote in the previous election (path $b \times c$). To keep things simple we will present the total association with prior e-voting, i.e. the direct and indirect effect path combined as $a \times b \times c$. These associations will be compared with the associations of the already familiar socio-demographic variables that act as proxies for the cognitive and material resources a voter commands (d_1 to d_4). If the second precondition of a habit is met, we should see how association $a \times b \times c$ becomes more manifest at the expense of associations d_1 to d_4.

We will present here an as non-technical interpretation of the path model as possible; the full results are outlined in detail in Appendix D. First, the data shows some associations in some years between social and demographic factors and e-voting at elections in different years, but these do not follow a distinct pattern of gradually diminishing importance. Only computer literacy eventually disappears as a predictor of e-voting.[21] Though the latter would be in line with the assumed logic of habit formation, we still have to look in more detail at the associations shown with paths a and $b \times c$.

[21] Keep in mind that these results cannot be directly compared with the findings from Chapter 4, as the current chapter does not examine first time e-voters in each election, does not employ all surveys used in Chapter 4, and the model specification also differs substantially.

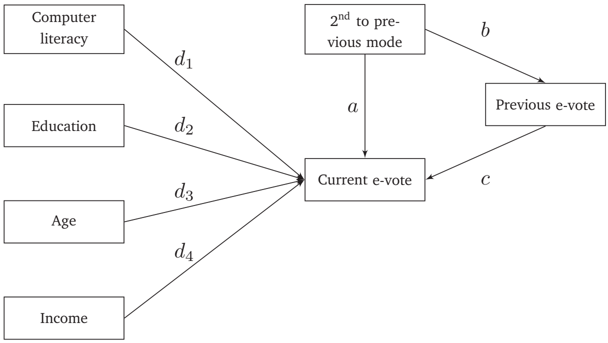

Figure 8.2: Path model of associations between prior e-voting and current e-voting (a – a direct effect of the second to previous mode on current e-voting; $b \times c$ – an indirect effect of the second to previous mode through the previous mode on current e-voting)

Table 8.1 displays so called odds ratios for e-voting in a given election when the voter has already used e-voting in the past. The total association, which is the combined association of paths a, b, and c ($a \times b \times c$) for example tells us that in 2009 the odds of e-voting in that election were 13 times higher for people who already e-voted rather than paper voted in the second to previous election or previous election. With the notable exception of 2011, we can say that e-voting in the past made a person 13 to 19 times more likely to also e-vote in the 2009, 2013, 2014 or 2015 elections. By and large, we can therefore say that prior e-voting has a very strong effect on current e-voting.

To properly address the third expectation with regard to e-voting as a habit, i.e. that repeated e-voting should not have a strong association with attitudes that are supportive of this mode of participation, we will finally examine those respondents who e-voted more than once. The aim was to see if supporting attitudes towards e-voting are associated with e-voting consecutively.

We expected to see no associations between supporting attitudes and repeated e-voting if e-voting was habitual in a strict sense. In detail, I looked

Table 8.1: Odds of e-voting in a given election when a person e-voted as opposed to paper voted during the second to previous election (paths a, b and c from Figure 8.2)

	2009	2011	2013	2014	2015
Total association ($a \times b \times c$)	13	34	13	19	15

For unrounded figures see D.2 in Appendix D

at whether repeated e-voters stand out as being more trusting towards e-voting and internet transactions in general; and whether they had higher confidence that their own e-vote and others' e-votes were counted as cast, while also statistically controlling for associations with socio-demographics. The detailed results are displayed in Table D.3 in Appendix D; they show that trust towards e-voting had a sizeable association with repeated e-voting. E-voters who generally trust e-voting are roughly 21% more likely to e-vote in consecutive elections. This did not however apply for other supportive attitudes. People e-voting in consecutive elections were not more trusting towards internet transactions, nor did they have higher confidence that their own and others' e-votes were counted as cast compared to paper voters and occasional e-voters. This clearly confirms the third expectation.

8.4 Conclusion

E-voting has been discussed as one possible remedy for the continuing turnout declines in Western democracies. Its potential has also been criticised, with fears citied that it will enhance rather than diminish existing social inequalities by setting up voting mode specific barriers for participation for voters who have less resources. The initial fears have been alleviated by the spread of e-voting among the wider electorate, at least in the Estonian example. As a side development to the diffusion process of e-voting in Estonia, anecdotal evidence of the “stickiness” of this voting mode have also been noted. If e-voting is indeed habit forming, it has the potential to arrest the continuing decline in voter turnout once a critical share of voters have experienced voting online.

In this chapter we examined the potential of e-voting to be habit forming using data from five elections and on how people reported their voting behaviour during three elections. The data shows e-voting to be very persistent, with large shares of e-voters clearly staying faithful to this way of participation in comparison to paper ballot or non-voters. E-voting once, makes one very likely to vote this way in subsequent elections. The crucial element of any habit, namely repetitive behaviour, is clearly present in e-voting.

We also examined if the sociodemographic resources associated by conventional wisdom with e-voting, such as age, higher education and wealth lose their effects at the expense of simply prior e-voting over the years. The results do not show that resources lose their association while prior behaviour gains one over time. We do see however, a consistent and strong effect of prior e-voting.

Lastly, for a habit in the strict sense to exist, the behaviour has to be not only repetitive and unrelated to factors that traditionally predict online participation, but also to be separated from the initial supporting attitudes for e-voting as a mode of participation. Only automated, not strictly reasoned repetitive behaviour is a habit. The evidence shows that consecutive e-voters do stand out as trusting the e-voting system more than non-consecutive e-voters or paper voters, but they do not harbour stronger attitudes in general in support of e-voting, such as trusting internet transactions or being more confident that theirs and others' e-votes were counted as cast. This is evidence in support for e-voting as habit forming.

In sum, keeping in mind the limitations of cross-sectional data in untangling habit formation, we believe that at the very least, there is clear evidence for the strong persistence of e-voting. Also, the fact that attitudes strongly supportive of e-voting are not significant in distinguishing "repeat offenders", suggests a certain automaticity, which is characteristic of a habit. E-voting seems to be self-reinforcing at a quite high level. We think these findings give again support for the potential of e-voting to be easily embraced by general electorate, if given a chance to cast their vote in such a manner. Scepticism about the positive potential of electronic voting to enhance representation was not borne out. To the contrary, if this participation mode is habit forming, then meaningful participation should receive if not a boost, then at least be stabilised at its current level.

Chapter 9

Verification and trust

MIHKEL SOLVAK

9.1 Introduction

In the local elections of 2013, Estonia introduced the feature of individual vote verification to the e-voting system. This gave individual voters the ability to verify whether their e-vote was: 1) cast-as-intended; 2) recorded-as-cast. This option has now been available in three consecutive elections, the 2013 local, 2014 European Parliamentary and 2015 national elections, warranting a closer look at its potential impact on e-voting and related matters.

Vote verifiability is a crucial element in ensuring a so called end-to-end (E2E) verifiable voting system. E2E verifiable systems add another layer of security and should ensure higher integrity of the voting process. The definition of an E2E verifiable voting system is quite strict (Popoveniuc et al., 2010) and the verification procedure introduced in Estonia as of 2013 does not yet meet that of a fully implemented E2E system; it does however cover a central element of it, namely giving individual voters the possibility to check if their vote was cast and counted as intended.

Though the main aim of introducing verifiability in Estonia was to detect possible large scale attacks on the system, it does so by encouraging individual voters to verify their e-vote with a separate smart device from the device used to cast their e-vote.

The interest in verifiability in this book arises from its potential to influence individual voter perceptions of the e-voting system's security. It is self-evident that conducting democratic free and fair elections is only possible when there is a baseline level of trust in the electoral procedures among the electorate. Building and ensuring that trust is key in legitimizing the outcome of the election. This is usually achieved through open and detailed regulation of election proceedings and the mutual oversight performed by national as well as international actors involved in the electoral process, i.e. election officials, party and candidate representatives, and independent outside observers.

For e-voting however, novel challenges in maintaining that level of trust in election proceedings arise due to the particular nature of the process (Mitrou and Gritzalis, 2002; Gritzalis, 2003). Given that people cannot physically observe how their e-vote is placed into a virtual ballot box, nor observe how these virtual e-votes are then "physically" counted by the election officials, a non-satisfactory answer to the question "what happens to my e-vote?" can discourage participation. In the absence of physical evidence in the form of paper ballots, it essentially becomes a question of trust. Building and maintaining trust levels is hence absolutely crucial for e-voting to be accepted among the electorate.

One ingredient for such trust is ensuring vote verifiability at the institutional level, but another is at the level of the individual voter. The former is needed to guarantee the integrity of the election process and to keep different actors from challenging the outcome, the latter is needed to encourage people to cast their vote online in the first place. Individual verifiability should in theory therefore ensure that the otherwise unobservable virtual voting process happened as intended. Though it does not resolve the non-observability problem, at the very least it should mitigate it somewhat by giving added insurance that due process occurred and we should see increased trust levels among users as a result.

The next sections give an overview of how individual verification works from the user side and examines in detail how this additional option in the e-voting experience has impacted upon trust in e-voting.

9.2 Verification procedure

E-vote verification is possible using a smart device that runs on Android, Windows Phone or iOS, has a camera to read a QR-code and internet connectivity.[22]

Verifying an individual e-vote is fairly straightforward. After casting an e-vote on a computer the voting application displays a note with a QR-code. The voter can simply close the application and be done with voting or they can take a separate smart device, download the verification app from Google Play, App Store or Windows Phone Store and use this app to read the displayed QR-code. Once the code is read and the smart device has communicated with the central server and received the encrypted vote, it will display a note on whether the e-vote cast from the computer was received by the server and upon request shows the candidate name and number for whom the vote was cast.[23] The verification app then closes automatically after 30 seconds. The app does not show any personal information of the voter and only one vote can be verified with it at a time. Each e-vote can be verified up to three times, but no later than 30 minutes after the vote was cast. If there is a conflict between the given vote and the information displayed on the separate smart device, either the computer used to cast the vote is compromised or there are more sophisticated problems with the e-voting system. In any case, the voter should report the conflict to the election authority and can revote with another device or vote on paper at the polling station.

---

[22] In 2013 verification was only possible using Android devices.

[23] For a detailed technical discussion of the system and the cryptography behind it see Heiberg and Willemson (2014)

9.3 Prerequisites

Given the way individual verifiability is implemented in Estonia, there are multiple prerequisites to its usage.

First, the e-voter needs to have access to a smart device. Though seemingly ubiquitous in today's world, not everybody has access to or user experience of smart devices. Furthermore, the usage pattern is most likely heavily non-random, meaning specific voter groups might systematically lack access to a device needed for e-vote verification. How access rates differ according to the age of eligible voters is shown in Figure 9.1, which was based on survey results from three given years.

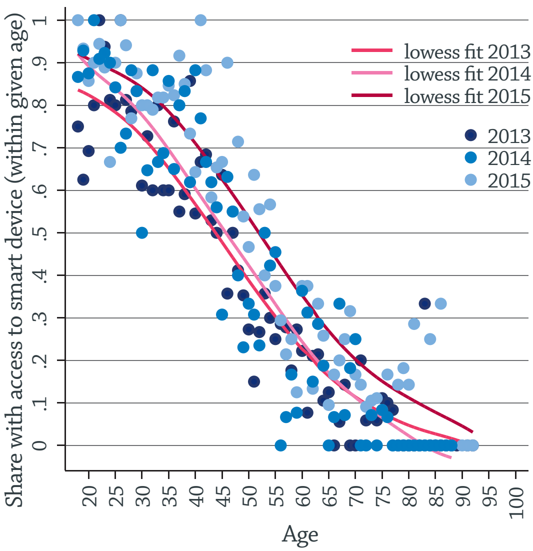

Figure 9.1: Access to a smart device according to age

As expected, access rates decreased with age. Note however that at almost all age levels, access increased even in the relatively short three-year window when vote verification has been possible in Estonia. Regardless of the latter the figure clearly suggests that verification can only have an impact among a certain subpopulation of e-voters, whose trust levels might paradoxically be somewhat higher to begin with, i.e. younger tech savvy people.

A second prerequisite is some familiarity with using a QR-code. Figure 9.2 shows the distribution of this according to age again based on survey results from the three years. Baseline familiarity with using QR-codes is clearly lower than the share who have access to smart devices and it is again highest among the youngest voter segment.

Looking at smart device access and QR-code familiarity in combination shows that only about 11% of eligible voters fulfil both of these prerequisites. A cursory look at the prerequisites therefore indicates that the possible positive effect upon trust towards e-voting as such is probably limited. The technical solution is simply somewhat excluding. We will turn to verification usage and impact next.

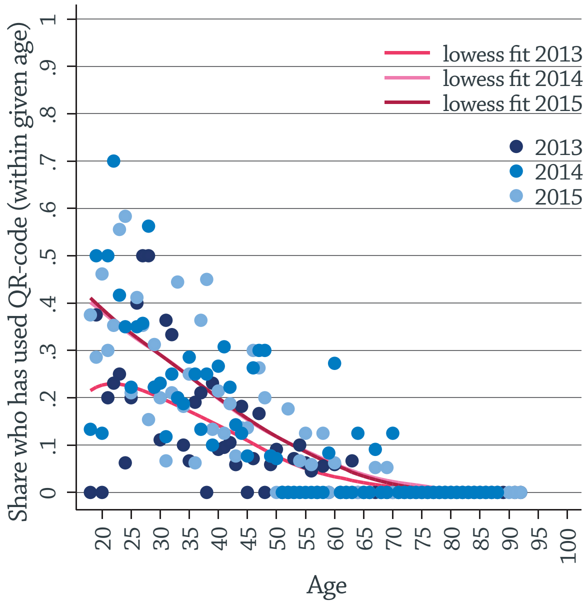

Figure 9.2: Familiarity with QR-codes according to age

9.4 Usage of verification

E-voting log file analysis gives us detailed information on the usage of e-vote verification; Table 9.1 shows how frequently it happened in the three elections where it was possible. Clearly a very small share of e-voters actually used this possibility, but we should point out that the absolute number of verifications in the 2015 elections was already close to the slightly more than 9 000 e-votes originally cast during the first e-enabled election of 2005. So even though verification was somewhat rare, roughly one out of 21 e-votes in 2015 was still verified, pointing to the wide potential usage of this technological solution in the future.

Table 9.1: E-vote verification frequency

	2013	2014	2015
Total e-votes	136 853	105 170	180 922
Verifications	5 024	4 924	8 439
Share verified	3.7%	4.7%	4.7%

Figures include multiple e-votes and verifications

At the same time, given that any technological impact takes time, as discussed in Chapter 4, we should be cautious regarding what impact it might have had on overall trust towards the e-voting system. These associations are examined in more detail in the next sections.

9.5 Verification and trust

There are two possible ways how verification could influence trust towards the system.

The first is quite obviously the users' direct experience. If an e-voter has verified their vote, they should consequently also display a higher level of trust towards the system, granted that the verification indeed returned a positive result, i.e. confirmed that the vote was received and counted as intended by the voter. We will juxtapose e-voters who verified their vote with those who did not do so to examine this possibility.

The second way emphasizes general knowledge instead of direct usage. Generalized as well as specific trust towards institutions has been noted as not necessarily connected to direct personal experiencee (Bachmann and Zaheer, 2006). Simply a good reputation might suffice to generate feelings of trust. It might therefore already be enough for a person to develop a sense of trust towards an institution by having heard of its good general standing; i.e. trust by way of osmosis and not from first hand user experience (Möllering, 2006). This possibility leads us to a different juxtaposition, namely comparing those who simply *know of* the e-vote verification possibility, but did not use it, with those who *did not know* at all of the verification option.

We used data from three post-election surveys from 2013, 2014 and 2015 and one panel study conducted before and after the 2014 EP election to examine these issues.

The general distribution of trust towards e-voting in elections where verification was possible is shown in Figure 9.3. The figure shows that e-voting polarizes. A substantial majority clearly trusts it, while a non-negligible part of the electorate does not trust it at all, and there is a large gap in between these two poles. This is somewhat unusual for questions of trust. As a rule, the majority tends to be in the middle between complete distrust and complete trust when it comes to trusting any institution. The same still applied here, we simply saw unusually large groups at the maximum distrust and maximum trust levels.

Regardless of the polarization, the overall trust level was still clearly high, higher than for other institutions of the Estonian state for example. But it does raise the question: if and how could trust in e-voting be improved? People who completely distrust the system are very unlikely to ever e-vote, so solutions that should produce additional confidence in e-voting, such as individual verification, will not really influence them. At the other end of the pole, among people with maximum trust, there is really no room for improvement, so verification will either change nothing or actually lower their trust when something goes wrong. The main room for improvement then, is with those with medium trust levels and that is where the impact of verification technology should be seen.

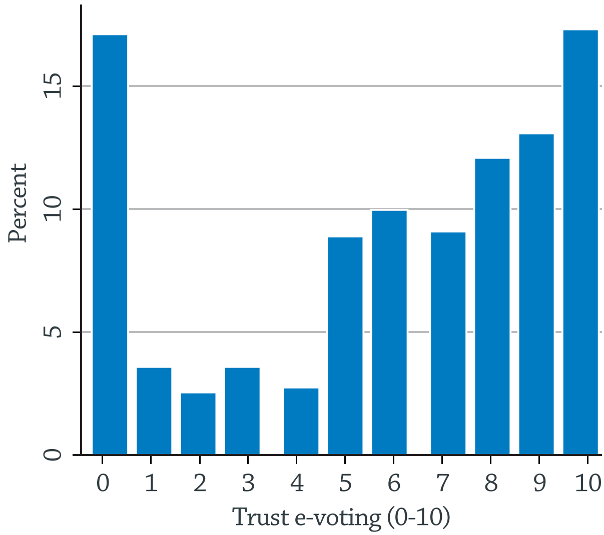

Figure 9.3: Distribution of trust in e-voting over 2013 to 2015 (0 - do not trust at all; 10 - complete trust)

We examined first if and to what degree might individual vote verification usage have an impact on trust towards e-voting in general. A panel study conducted during the 2014 EP elections measured trust before and after the election. A subsample of the surveyed voters reported to have heard of verification and a further subsample of those had verified their e-vote; Figure 9.4 presents the differences in mean trust towards e-voting before and after the election among those who verified their vote and those e-voters who knew about verification but did not do so.

We can see that the differences were quite small, there was a slight growth in mean trust level after the election, but this change was minute and was virtually the same among verifiers and non-verifiers. Further tests outlined in Appendix E confirmed that using verification does not bring a statistically significant increase in trust towards e-voting. We also tested for the possibility that the lack of change was simply down to an overall stability of trust perceptions in all possible public institutions before and after the election, i.e. that trust in e-voting is similar to trust in governance and

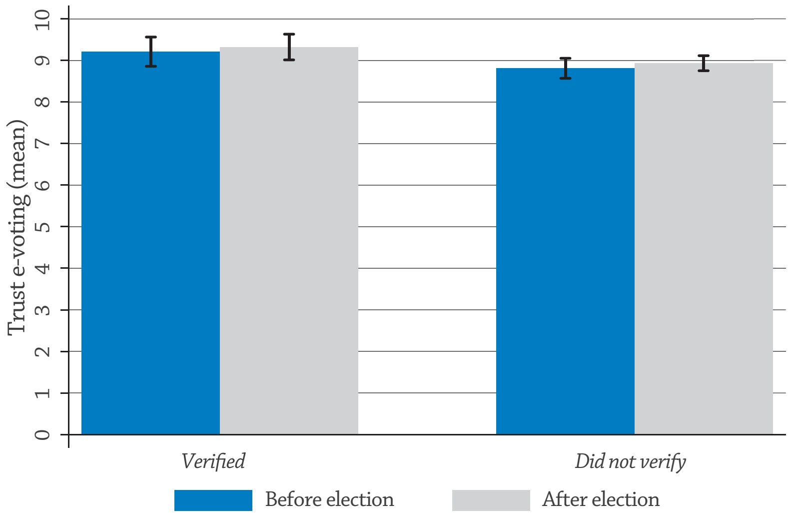

Figure 9.4: Mean trust of e-voting before and after the 2014 EP election for vote verifiers and non-verifiers (with standard error of the mean)

institutions in general. This however was not the case, with trust towards e-voting changing independently from trust towards other institutions; see Table E.3 in Appendix E.

Why did we see no difference in trust? One answer might come from the fact that those who verified their vote seemed to have a very high trust level to begin with, i.e. they trusted e-voting highly even before actually engaging in the act of verification. Figure 9.5 contrasts the change in trust levels before and after the election separately for e-voters and e-vote verifiers. We saw first of all that e-vote verifiers scored 8 or above on a trust scale of 0–10 before they verified. When the trust level is already so high there is effectively little room for improvement, so theoretically it is more likely to stay the same or even decrease. This was exemplified by people who gave the maximum trust score before the election and then verified their e-vote. Their trust had on average in fact decreased. E-voters who did not verify their vote exemplified an average increase in trust compared to levels before the election and the more so the lower their trust level was to begin with.

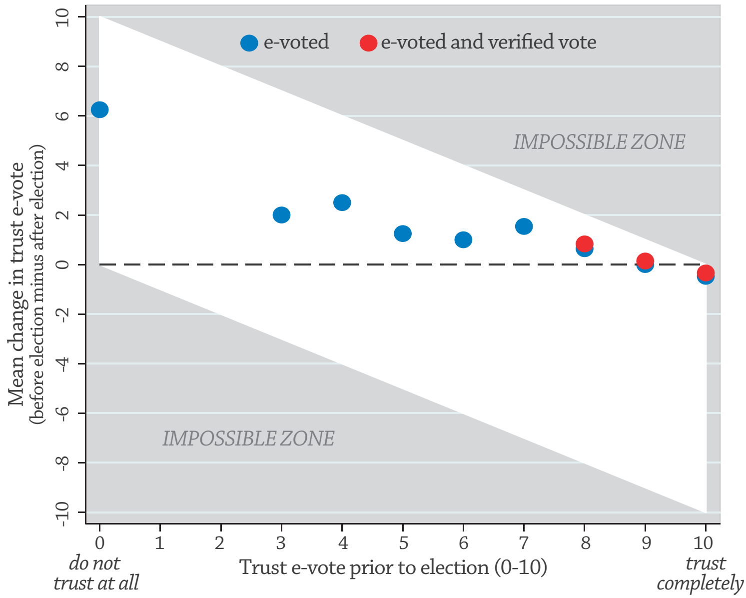

Figure 9.5: Prior trust levels and change in trust after the 2014 EP election for e-voters and e-vote verifiers

We emphasize especially this last fact, i.e. there are e-voters even among voter segments who show comparatively little trust in e-voting. But all-in-all, the picture suggests that a high frequency of verification might not indicate people not trusting the system overall. In Chapter 4 we already saw how first e-voters at the beginning of e-voting in Estonia used to be a distinct subgroup of the "usual suspects" of technology users, i.e. somewhat younger and computer literate voters. The same possibility suggests itself in case of verification. It seems the verification tool is mainly used among people who had no issue with trusting online voting in the first place. The subgroups among whom verification should show the comparatively largest impact, people with low trust in e-voting, simply did not use the technology in the take-off phase.

This brings us to the other possible mechanism of how verification might still impact upon voter trust in the e-voting system, namely whether simply *knowing about it* but not using it suffices for them to highly trust the system. To do so we compared the trust levels towards e-voting between three groups, those who used it, those who knew about it but did not use it, and those who did not even know that such an option existed. For this we examined post-election survey data from the 2013, 2014 and 2015 elections.

Figure 9.6 shows trust levels towards e-voting for these three groups. A rather interesting picture emerged, i.e. there was no real difference in the distribution of trust among users and knowers, simply the levels differed slightly, with users having a larger share of those who completely trusted e-voting. There was however, a crucial difference between the group knowing and the group not knowing about verification. Among the latter the single largest group were actually people with no trust at all (0 trust) in e-voting and there was also a roughly equal share of people for all trust levels above the scale midpoint. This picture tells us that not knowing about verification seemed to be associated with lower trust.

This possibility is further examined in Figure 9.7, which presents the associations extracted from a multivariate regression model detailed in Appendix E. The association is shown for two types: 1) e-vote verifiers with the reference point being non-verifiers who knew about verification; 2) for people knowledgeable of verification, with the reference point being people who did not know about it. The dot represents the mean controlled difference in trust in e-voting for the two respective groups and the whiskers the 95% confidence intervals. We saw that e-vote verifiers did not have a statistically different level of trust towards e-voting in any of the three years studied, as the whiskers intersect the zero association line.

But the picture was very different when contrasting knowing and not knowing about verification. We can say that after taking account of a range of other factors that might also influence trust, such as age, income and education, people knowing about verification still had on average roughly 2.5 points higher trust scores towards e-voting than people who did not know about verification in one of the three years. Given that the trust variable itself has a range of 0-10, the observed difference of more than two scale

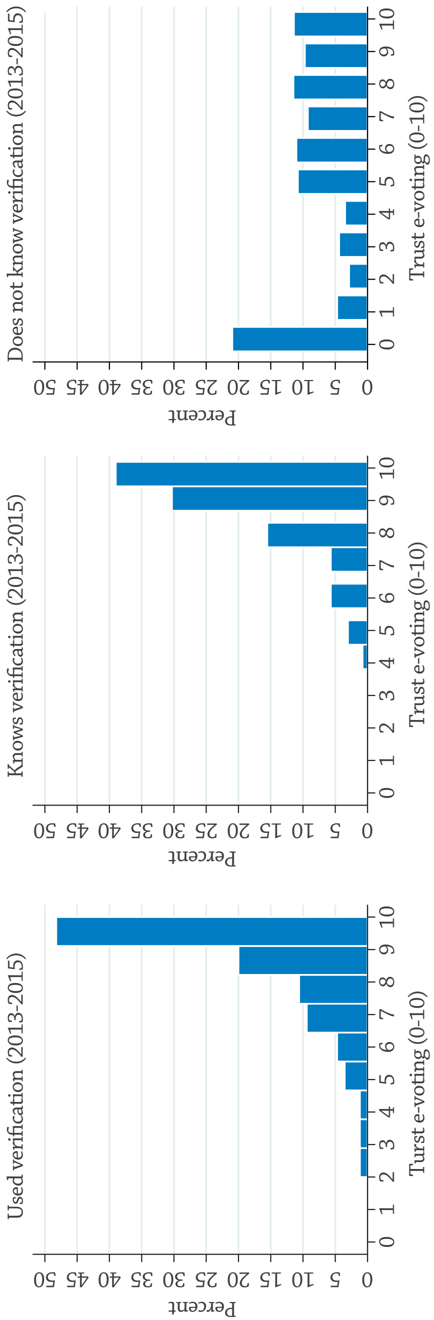

Figure 9.6: Distribution of trust in e-voting in 2013 to 2015 according to using, knowing and not knowing verification

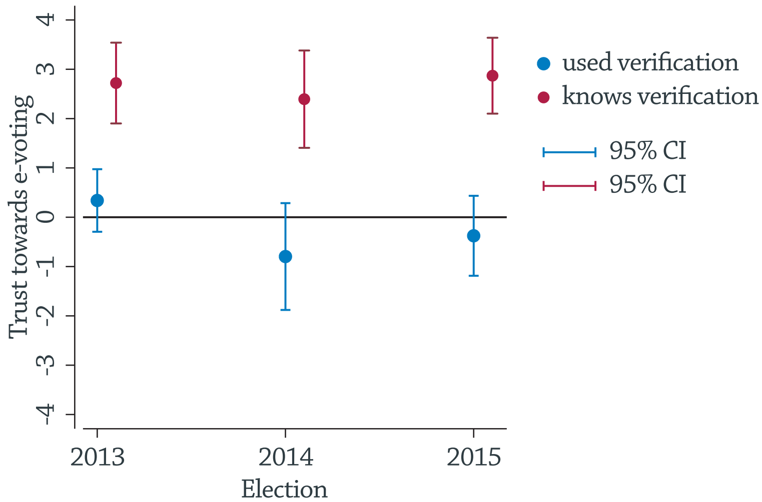

Figure 9.7: Associations between using verification or simply knowing about with trust in e-voting (OLS regression coefficients)

points is substantial (>20%).

We can also see that the associations did not really changed over the three years of elections where e-vote verification was a possibility. One can only speculate if a similar pattern as observed for e-voting in general could also be unfolding for vote verification, that is, if the actual impact of usage and not only knowledge will be seen when more regular e-voters start using it and not only the roughly 8 500 enthusiasts who did so in 2015.

9.6 Conclusion

In this chapter we examined if and to what degree might the latest technological innovation in Estonian e-voting, individual vote verifiability, have an impact on trust towards e-voting in general. Given that "what happens to my e-vote?" is a question often raised by voters, having the option to verify one's vote should foster higher trust in the whole process.

Investigating the distribution of trust towards e-voting amongst the Estonian voter population showed no problems with trust in general, with levels of trust high, but there is a considerable group who do not trust e-voting at all. We think the potential of e-vote verification to address the fears of this group is fairly limited however, as they are very unlikely to ever try e-voting. The trust span where the technology might have a potential impact is in the middle, i.e. between the two extremes of complete trust and complete distrust, and after all that is where the majority of the population lies.

The verification possibility was introduced in 2013 and was used to check 4.7% of all e-votes cast for the latest 2015 election. This share is small, but direct user experience is the most plausible mechanism of how a technological solution should have any impact, so one would expect e-vote verifiers to have elevated trust levels after using it, i.e. to gain additional confidence in an otherwise virtual e-voting process. Surprisingly, verifying one's vote did not result in a significant increase in trust. Instead we found that knowledge, not usage, is the real trust building mechanism. Contrasting those who knew about the verification option, but did not use it, with people who had no knowledge of this possibility, showed the former to have persistently higher trust levels.

The lack of an increase in trust for verification users can be explained by their very high trust levels to begin with. E-votes were somewhat paradoxically verified by people who had no problems trusting the system in the first place. This fact fits well with the patterns seen with the general diffusion of any technology, i.e. it first picked up by the more risk taking and experimenting individuals and only later are they followed by the rest of the population. We did see however, that e-voting as such, i.e. without verification, does lead to higher trust levels the lower the level was prior to voting online. This indicates the positive potential of user experience, absenting any problems in the process, to increase trust exactly among whom it is a problem, namely the people who remain sceptical of e-voting.

The main findings of this chapter raise a puzzling question. It is self-evident that building a technically complicated verification system that is at the same time easy for individual voters to use is difficult and expensive. Finding out that the fact of building it and not the usage of the tool seems to have a higher impact on trust in e-voting might be a somewhat sobering outcome. However it does make sense once seen in the context of the spread of e-voting in Estonia in general. The ones using verification in the first three elections where it was possible, are most likely part of the same non-random innovative technology users who were the first to pick up e-voting itself when it was first introduced. As with anything, it simply takes time for the technology to take hold and start to show its potential positive impact. The jury is still out on whether verification can achieve its intended aim of providing higher confidence in e-voting to the majority of the electorate. The user numbers are still too small and generated by very high-trusting individuals for any potential positive impacts of e-vote verification to emerge. We will have to wait until usage spreads to groups who currently have lower trust levels in order to see if it indeed does have an impact, and the signs so far clearly show that it might.

It is however already a positive outcome to see that simply knowing about the possibility to check whether one's e-vote was received and counted as intended suffices for higher trust in the system. Oddly enough, the national electoral authorities would simply need to spread the word about individual vote verification to increase trust levels but not necessarily to encourage more active usage of the tool.

Chapter 10

Political neutrality of e-voting

Kristjan Vassil

In this chapter we will present an analysis of whether remote internet voting in Estonia has the potential to bias aggregate election results by favouring some parties over others. First, we investigated the aggregate distribution of vote shares by parties and by the two modes of voting—traditional paper voting and e-voting—in order to demonstrate how large any discrepancy between them was.

Second, we explicated three individual level mechanisms that could potentially produce any bias in aggregate election results.

Third, each of these mechanisms was tested empirically using survey data from the national elections of 2007, 2011 and 2015.

Academic reference: ”Internet Voting and Politics: How e-voting influences turnout, political neutrality and trust” by Kristjan Vassil. Report for the Software Technology and Application Competence Center, 2013, Tartu.

10.1 What is political bias?

Aggregate election results from past elections in Estonia demonstrated that some parties consistently gained more electronic votes than others. For example, in the last three national elections the largest share of e-votes has been won by the current prime ministerial party, the Reform Party. The second-largest share of e-votes was gained by Pro Patria and the Res Publica Union, followed by the Social Democrats. One of the smallest shares of e-votes has been gained by the centre/left opposition party, the Centre Party, which at the same time tends to get the bulk of paper votes (see Table 10.1). The observable difference in the allocation of e-votes and paper votes has sparked a fierce political debate on whether e-voting produces a mode-specific political bias, and more importantly, whether it affects the outcome of elections. Those against e-voting rely on the aggregate evidence of the unequal vote share and posit that e-voting should be abolished because it structurally favours some parties over others and therefore violates the assumption of a voting mode having to be politically neutral. Conversely, those winning larger shares of e-votes advocate its convenience and potential to mobilize young voters, suggesting that the continuation of e-voting is pivotal for democratic participation. In this section I attempt to go beyond this normative debate and offer an evidence based empirical account by assessing the potential of e-voting to influence aggregate election results.

Before doing so however, we wish to conceptually clarify what is meant by *political bias* and how it relates to the vote shares gained by parties overall and as a result of specific modes of voting. In the context of e-voting we define political bias as the difference between the observed election outcome and the outcome that would have occurred in the absence of remote internet voting. The fundamental problem of causality in this case is that we only observe one of two counterfactual situations at any given election. We cannot turn back time, i.e. look at two identical elections in which one used and the other did not use e-voting and compare the results of the two elections. We also cannot directly compare elections before the 2005 local elections without e-voting to those in or after 2005 with e-voting, because any discernible difference might occur owing to other factors over time.

A way out of this inferential deadlock is to additionally define the aggregate political bias, i.e. at the level of election results, the bias within modes of voting (traditional election day voting, postal voting and internet voting). In the latter case, the bias is defined as the difference between the mode specific vote share distribution and the aggregate election results. It follows that the sheer distribution of vote shares according to different modes of voting may produce political biases within their modes, but it does not follow automatically that these mode specific biases translate into biases in the aggregate election results. It may so happen, but this needs to be (and can be) validated by empirical testing. This is precisely the research logic used the following analysis.

We proceeded by first investigating the aggregate distribution of vote share by party and by the two modes of voting: e-voting and conventional paper ballot voting. We then identified three mechanisms that could potentially induce political bias and empirically tested whether there was supporting evidence for any of these.

Aggregate distribution of vote shares

We began by looking at aggregate election results to understand how much the distribution of e-votes deviated from the distribution of traditional paper ballot votes for different political parties in Estonia. To achieve this, we relied on the results of the elections of 2007, 2011 and 2015, the three parliamentary elections where e-voting has been used. In order to evaluate whether the shares of e-votes were representative of traditional votes we calculated for each party the difference d between the traditional vote share t and the e-vote share e for each party ($d = t - e$). If both vote shares are equal, $d = 0$. If the traditional paper ballot vote share exceeds the share of e-votes for a particular party $d > 0$, and conversely if the share of e-votes is larger than the share of traditional paper ballot votes $d < 0$; Table 10.1 displays the outcome.

First, we saw that the largest discrepancy between the vote shares of the

Table 10.1: Vote share by mode of voting for the 2007, 2011 and 2015 national elections

Party	Vote share in 2007			Vote share in 2011			Vote share in 2015		
	Paper	E-vote	*d*	Paper	E-vote	*d*	Paper	E-vote	*d*
Reform Party	27.4	34.5	-7.1	25.8	36.9	-11.1	23.3	37.6	-14.3
Centre Party	27.1	9.1	18.0	27.7	10	17.7	32.4	7.6	24.8
Pro Patria and Res Publica	17.4	26.7	-9.3	18.9	25.4	-6.5	12.2	17.2	-5.0
Social Democrats	10.5	13.3	-2.8	16.8	18	-1.2	14.4	16.9	-2.5
Peoples Party	7.3	3.6	3.7	2.4	1.3	1.1	-	-	-
Greens	6.9	10.7	-3.8	3.6	4.3	-0.7	0.8	1.2	-0.4
Free Party	-	-	-	-	-	-	7.2	12.0	-4.8
Conservative People's Party	-	-	-	-	-	-	8.7	6.9	1.8
Others	3.4	2.1	1.3	4.7	4.1	0.6	1.0	0.6	-0.4

Raw data available at: http://www.vvk.ee two voting modes occurred for the Centre party, with a difference of close to 25 percentage points in three elections. The second largest discrepancy can be observed for the Pro Patria and Res Publica Union and also for the Reform Party. Both these parties win larger shares of e-votes than on-paper votes. For all other parties, the differences tend to be smaller.

Based on the aggregate results so far we can clearly see that e-votes have a very different party distribution than paper votes. The within voting-mode bias appeared to quite substantially favour the Reform Party in 2011 and especially in 2015, as well as the Pro Patria and Res Publica Union in 2007 and 2011. Yet, it is imperative to differentiate between mode-specific bias and bias of the overall election results, as the latter is not necessarily a function of the former. Imagine that some voters simply switch form paper voting to e-voting and that this switching is more probable among those who vote for the Reform Party. Although this non-random switching produces a mode specific bias, the aggregate election result will remain unaffected, as the same people still vote for the same party, see also the text-box.

MODE SWITCHING MISTAKEN FOR POLITICAL BIAS

Say we have 20 voters, out of whom 10 support party A and 10 support party B. The aggregate voting result would hence be that party A receives 50% of votes and party B 50% of votes. Now what if 9 out of 10 party A supporters voted online while only 1 out 10 party B supporters voted online? This would mean that we have in total 10 votes cast online, with 9 of them for party A and 1 for party B. If we would now examine online voting results separately, we would see that 90% (9 votes) of online votes went to party A while 10% (1 vote) went to party B. This seems like extreme bias, but the total election outcome in terms of votes is still 50%:50%, we simply have a larger share of party A supporters e-voting and a larger share of party B supporters voting with paper ballots.

It follows that making inferences about ”biasness” on the basis of aggregate election results is potentially spurious. However, bias can still occur due to a number of mechanisms associated with the non-random usage of e-voting. Our next task was then to explicate behavioural mechanisms that could potentially translate to mode-specific bias that could alter election results.

10.2 Mechanisms that produce bias

We explicated three mechanisms that can potentially skew aggregate election results due to availability of internet voting. In so doing we laid out the building blocks of each mechanism so that they could later be subjected to empirical testing in order to find evidence for whether any of the mechanisms are actually at play in Estonian elections.

Mobilization

The first mechanism is based on the idea that if internet voting can potentially bias election results, it may be due to the fact that it makes some citizens vote who would not have voted otherwise. We call this process e-voting's capacity to mobilize new voters. The basic logic of this process is that with e-voting the electorate is larger than it would have been in the absence of e-voting. Next, suppose that mobilization is non-random and that e-voting only mobilizes supporters of certain political preferences, thus structurally giving more voice to some parties while undermining others. Provided that these conditions hold, i.e. non-random mobilization based on political preferences happens, elections with e-voting would produce different election outcomes than those without it. Although we cannot compare an election with and without e-voting at the same time, because this counterfactual situation is never observed, we can test whether those who are mobilized by e-voting structurally represent some parties, but not represent others. We thus examined the following hypothesis: e-voting mobilizes new voters who structurally prefer some parties over others.

Normatively, we find actually no particular problems with mobilizing non-random new voters, because any means of voting that encourage political participation *decreases* any observed political bias instead of *increasing* it. This happens because more voters decide to participate in political decision making and thus one can argue that – on the contrary – elections without e-voting structurally exclude potential voters, and therefore some parties gain fewer votes than in elections with e-voting. This goes to show that seeing bias in a mode that opens up new ways of participation is problematic. A clearer case for bias exists when certain groups are intentionally kept from participating.

Vote switching

The second mechanism is based on the idea of vote switching. Voting research shows that some voters change their vote choice between elections, preferring some parties in some elections and other parties in other elections. It has been demonstrated that vote switching is particularly dominant between elections at different levels, i.e. national, European or local, with many voters using European Parliamentary elections as a means to signal their preferred party their discontent and warn them against what would happen in the next national elections. However, vote switching also occurs between same-level elections. The young in particular tend to have less stable party preferences and are more likely to change their vote choice between two parliamentary elections. It may also be possible that the sheer availability of e-voting induces vote switching for some particular reason. Although not really plausible in reality, this is a logically possible outcome that can be tested empirically.

To be sure, e-voting may bias election results if e-voters start voting differently than they would have voted in the absence of e-voting and that this change, again, is non-random. It follows that if e-voting induces switching, some parties may gain a competitive advantage over others (provided that switching is non-random). Granted, this is a very farfetched mechanism, but we wanted to test ways of how bias could occur, even when the mechanism seemed quite unlikely. Thus, we tested if the availability of e-voting induced non-random vote switching.

Mode-specific bias

Finally, we sought to test for within-mode political bias. As I have explained and showed above, a political skew can occur within specific modes of voting. For example, this was found to be the case within the mode of e-voting, where some parties consistently gained more votes than others. Thus, if e-voting is politically biased, it should be predictive of partisan choice, i.e. participation in online voting should be associated with higher propensities to vote for certain political parties. However, when testing for such associations we need to look beyond bi-variate relationships and include relevant controls to account for different propensities to use internet voting in the first place.

As the previous chapters have shown not all people are equally likely to vote online, because they for example lack the technical skills, physical access, time, experience or other perquisites for online voting. For example, let us assume that people of higher socio-economic status are more likely to vote online than others, even though the previous chapters showed this not to be case in Estonia anymore. We can further assume that the same set of characteristics that predicts internet voting at the individual level, might be predictive of partisan choice. In other words, people of higher socio-economic status could be more likely to lean to the right of the political spectrum and support the Reform Party; and they could also be more likely to vote online. Therefore, analysing the impact of e-voting on party choice contains two mechanisms: 1) the propensity to vote online as a function of party choice; 2) the propensity to vote online as a function of a set of socio-demographic characteristics that, at the same time, are predictive of partisan choice.

The key analytical objective for me was to disentangle the two mechanisms and demonstrate whether voting online is associated with political choice for either of the two mechanisms. The common strategy for separating the two is to control for relevant characteristics such as age, education and most importantly, the position on the left-right political spectrum. In so doing we are able to address the question: to what extent does political preference matter with respect to voting online and, more importantly, to what extent does baseline socio-demographic characteristics determine the choice to vote online. Thus, we tested for a relationship regarding whether the usage of e-voting is predictive of party choice even when a number of socio-demographic and attitudinal characteristics are controlled for.

Notice that we have formulated working hypotheses so that if confirmed, we find evidence to support the view that election outcomes with and without internet voting probably would yield different outcomes. In the following section we provide empirical tests for all three hypotheses.

10.3 Empirical tests

Mobilization

Given that we cannot emulate a situation without e-voting we rely on the same logic as elaborated in Chapter 6 vis-a-vie mobilization by e-voting and defined mobilized voters to be people who had switched from non-voting to voting via e-voting,i.e. the group we are going to compare with regular e-voters.

We specified a model on whether voter mobilization predicts partisan choice. If e-voting mobilizes voters with certain political preferences, then the mobilization should be associated with partisan choice. Conversely, if mobilized e-voters come from the entire political spectrum, its effect will be heterogeneous lending support for the argument that e-voting does not create political bias.

In order to assess any potential relationship between mobilization and party choice, we used data from the 2007, 2011 and 2015 parliamentary elections. For simplicity we focused only on the four largest parties in the Estonian parliament for all three elections, the centre right Reform Party, the centre left Centre Party, the center left Social Democrats and the conservative Pro Patria and Res Publica Union. We estimated separate logit regression models for all four in each election separately, with the outcome variable being whether the respondents voted for that particular party or not and the predictors being mobilization by e-voting and other vote relevant factors (for technical details of the regression analysis, see Appendix F).

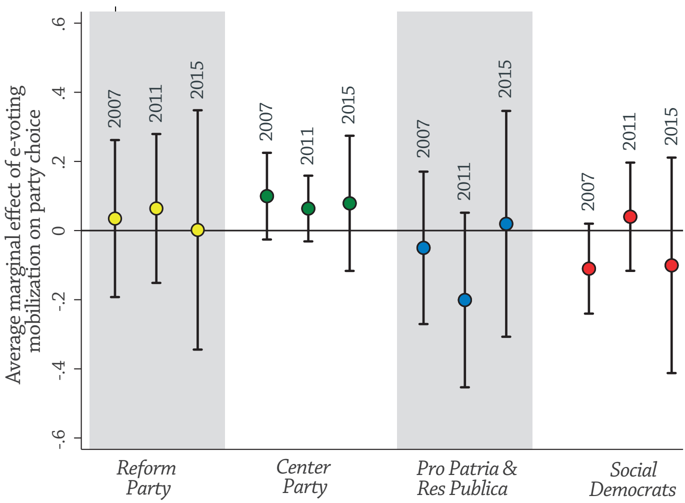

Figure 10.1: Associations between being mobilized by e-voting and party choice for the four main parties in the parliamentary elections of 2007, 2011 and 2015

Figure 10.1 presents the association between being mobilized by e-voting and party choice for all four parties over the three elections, as extracted from the regression model (for full model see again Appendix F). We saw no association for any party in any year, as the 95% confidence intervals represented by the whiskers of each point always crossed the zero effect reference line. What does this mean in simpler terms? It tells us that people being mobilized to vote by e-voting, i.e. people who previously did not vote but the e-voted, did not disproportionally support any one specific party. In other words, e-voting mobilized people who chose all kinds of parties and the aggregate level discrepancy between e-votes and paper votes cannot be explained with a mechanism where e-voting brings to the polls new voters who in disproportionate numbers support a specific party. Thus, we can rule out e-voting to be politically biased in this way.

We thus infer from this exercise that our first hypothesis of mode specific political bias due to mobilization does not hold, irrespective of the fact that as many as 14.6 percent of e-voters felt that the availability of e-voting induced them to vote. Our model shows that these voters, when controlled for their political preferences, were mobilized on a random basis with respect to party choice. They formed a heterogeneous group of voters who supported all four political parties with no structurally identifiable pattern.

Switching

Political bias occurring from the mode of voting may occur, if some voters non-randomly choose to switch their partisan preferences. In other words, if Reform Party voters choose to start voting for other parties because of the availability of e-voting, we might say, that this mode of voting induces people to switch their vote choice to another party. Although this is not very plausible in reality, we know that vote switching appears as a feature inherently linked to Estonian elections for a number of individual-level reasons (Solvak, 2011). If vote switching happens at random, no bias can be produced because switchers randomly change their vote choice to any of the political parties available. If, by contrast, switching has a pattern we might find that switching to some parties is more likely than others, potentially leading to a different election outcome.

Unfortunately, the only dataset that allows me to examine this arguably quite unlikely mechanism is the Estonian Election Study 2011 data (see: www.enes.ee). Using this post-election survey data, we coded a binary indicator on whether the voter voted differently in the 2011 national elections compared to the 2007 national elections. We call this variable “switching” and it takes the value 1 if a voter changed their vote choice between the two last national elections, and 0 if they did not. Next, we compared the frequency of switchers by the mode of voting, i.e. among e-voters and traditional voters in order to verify whether there is a statistically significant difference between the two groups. The data showed that 35 percent of on-paper voters and 28 percent of e-voters switched from one party to another between the two elections (see Table 10.2). This means that vote switching was by 7 percentage points less frequent among e-voters than among paper voters. Thus, even if the switching mechanism was hypothetically plausible, the lower switching rate among e-voters means it was unlikely to produce a large difference between aggregate voting tallies as a result of the voting mode.

Table 10.2: Party choice switching between the 2007 and 2011 national elections and voting modes

	Paper voted	E-voted
Switched party	117	28
	35.6%	28.3%
Did not switch party	212	71
	64.4%	71.7%
Total	329	99
	100.0%	100.0%

Nevertheless, let us examine the actual party switching patterns depending on the voting mode for the 2011 election in more detail. Only then can we say for sure if e-voting systematically brings about switching away from or towards certain parties. Table 10.3 shows how people who voted for any of the four large parties in 2007 voted in the next parliamentary election in 2011 depending on their voting mode for that election. We have highlighted in bold typeface the row percentages that show the share of people voting for the same party in both elections. There is indeed a systematic difference depending on the mode and it is that of the substantially larger share of e-voters staying faithful to their previous choice in comparison to paper voters. In any case, regardless of the mode, the largest group of supports for any party tended to vote for the same party in the next election.

We also point out that in the Estonian bilingual society, switching is predominantly a feature of native Estonians, while the Russian speaking minority is considerably less likely to change their party choice between elections (Solvak, 2011). In order to control for the potential confounding effect of

Table 10.3: Party choice switching between the 2007 and 2011 national elections and voting modes

Party vote in 2007	Vote mode in 2011	Party vote in 2011: Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats	Other	Total
Reform Party	E-voter	**72.9**	0.0	14.6	8.3	4.2	100.0
	Paper voter	**53.3**	4.4	17.8	21.1	3.3	100.0
Centre Party	E-voter	0.0	**93.3**	0.0	0.0	6.7	100.0
	Paper voter	0.0	**77.2**	4.4	10.5	7.9	100.0
Pro Patria & Res Publica	E-voter	7.7	0.0	**61.5**	26.9	3.8	100.0
	Paper voter	5.9	1.5	**55.9**	20.6	16.2	100.0
Social Democrats	E-voter	0.0	0.0	0.0	**100.0**	0.0	100.0
	Paper voter	0.0	2.9	23.5	**61.8**	11.8	100.0

language, we also looked at the relationship for the two ethnicities separately. Again, e-voters did not show abnormal switching rates away from or towards any particular parties.

Based on the two tables, there was no need to conduct additional statistical testing, we can safely say that e-voters are not inherently more unstable in their party choice and do not systematically switch away from certain parties.

From this brief exercise we find little support for the mechanism by which e-voting biases elections results, because voters who used this channel do not overwhelmingly switch their vote choice and therefore we can refute this second mechanism.

Mode-specific bias

As explicated above, my empirical strategy needed to account for two mechanisms that may potentially induce political bias due to the availability of e-voting. Let us start with a simple proposition. If e-voting is indeed politically biased, we should be able to observe, at the level of the individual, a relationship between the usage of e-voting and party choice; see association a between e-voting and party choice in Figure 10.2. In simple terms, it simply states that one's likelihood to use internet voting is associated with their party choice. For example, people who are likely to vote online are likely to vote for the Reform Party. We already know that this holds given the aggregate vote shares outlined in table 10.1.

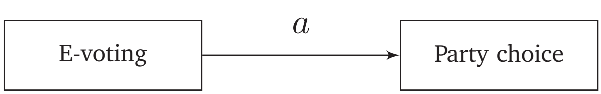

Figure 10.2: Relationship between e-voting and party choice.

At the same time, we know that party choice is first and foremost determined by one's socio-demographic profile. Young people with higher incomes and of Estonian ethnicity support the Reform Party more frequently than the Centre Party. However, the same set of characteristics could induce the probability of e-voting. In order to account for the potential relationship between voting mode and party choice, we expand my simple proposition by demographic determinants and linked them to both e-voting and party choice. Figure 10.3 links the usage of e-voting to party choice (link a), but it also links demographic profile to party choice (link b) and the probability of e-voting (link c). Our empirical goal is to verify whether a prevails irrespective of controlling for b and c. Consequently, we can extend the model by one's political preferences and argue that while they affect party choice, they might, through demographics, also induce the probability of e-voting. These confounding patterns are of particular interest for me, because e-voting can only produce a bias in election results when it is predictive of party choice irrespective of relevant confounders.

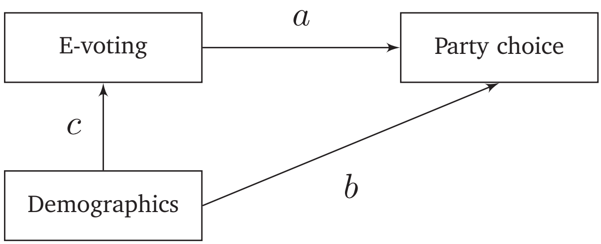

Figure 10.3: Relationship between e-voting and party choice.

A simple modelling strategy that accounts for the effect of demography is to specify a logit model where the outcome is the choice of party (e.g. a binary variable 1 for respondents who voted for the Reform Party and 0 otherwise) and the main predictor is a dichotomy between e-voters (coded 1) and paper voters (coded 0). Because we are interested in the effect of e-voting while controlling for demographics and other related covariates, we included a set of traditional determinants of party choice (such as age, gender, education, ethnicity, place of residence, income, left-right-placement and computer literacy). Subsequently we estimated the chances of voting for each of the four parliamentary parties; for technical details see section F.2 of

Appendix F. We present in Figure 10.4 the effect of e-voting extracted from these larger models, for full models again see Appendix F.

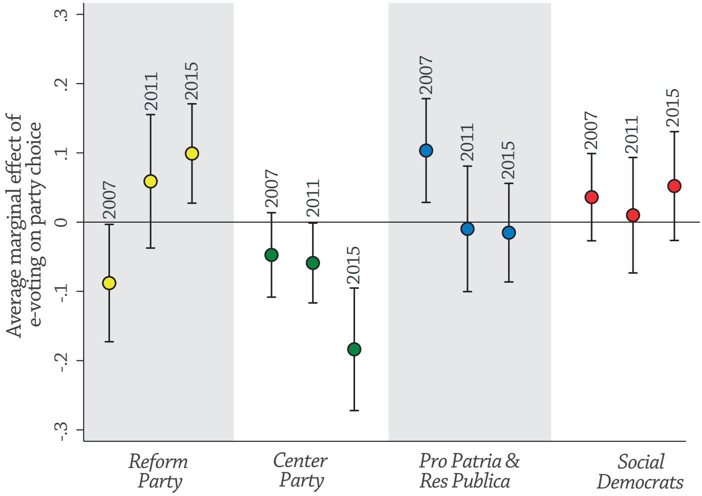

Figure 10.4: Associations between e-voting and party choice for four parties in the parliamentary elections of 2007, 2011 and 2015

The figures show associations and a specific pattern. E-voting is, as of 2015 positively associated with the probability of voting for the Reform Party, meaning people who e-vote are 9% more likely to cast their vote for this particular party rather than any of the other three largest parties. The reverse was the case for the Centre Party, i.e. people who e-vote are roughly 18% less likely to cast their vote for this party rather than any of the other three largest parties. We should point out here that the Centre Party and its leader have actively campaigned against e-voting and publicly called people not to vote online, so the association is hence also partly driven by this party's own actions and not only by a mode specific bias. Lastly, there was also a positive association with the probability of voting for the Pro Patria and Res Publica Union in 2007, but this was not the case in subsequent parliamentary elections.

The central question arising from this analysis is whether the statistical associations seen in the figure are a cause of concern when evaluating the individual level bias of internet voting? Think on our problem set in experimental research terms, where e-voting can be conceptualized as a treatment and traditional voting as the control, akin to a medical trial where certain patients are handed a medicine (e-voting) and the others remain in a control group (paper voting). In an experiment we would inquire whether a causal link between e-voting and party choice exists. In an experimental set-up, randomization of treatment assignment would ensure that the groups compared were similar to each other in all relevant characteristics except for treatment status (usage of e-voting and non-usage of e-voting). As we made use of observational data, people have self-selected themselves into modes of voting on the basis of their socio-demographic characteristics and other qualities relevant for making this decision (including political preference). Thus, our current sample is structurally skewed by a number of relevant variables that, if considerably different across two groups, would still show the effect of e-voting on party choice. If the treatment and control samples are not balanced, differences in voting behaviour could be a reflection these unbalances, rather than of the treatment itself (Lassen, 2005).

A conventional way to alleviate the problem of diverse subsamples is to use matching methods (Stuart and Rubin, 2008). Essentially, matching tries to recover a random research design from observational – or imperfectly randomized – data to provide a basis for causal interpretation of the estimates (Lassen, 2005, p. 109). The idea of matching is to match every respondent in the treatment condition (e-voters) and respondent in the control condition (traditional voting) on the basis of observed covariates. In so doing, matching methods emulate randomization by selecting a reference group that is statistically no different from the treatment group and thereby approximates randomization in an experimental design. Thus, the matching procedure ensures that inference is not based on treatment and control samples that are too different, which is not guaranteed by simply including these variables as controls in a regression in the standard fashion (Lassen, 2005, p. 109). We performed this matching and estimated the regression models again with the extracted results shown in Figure 10.5, for technical details, again see section F.2 of Appendix F.

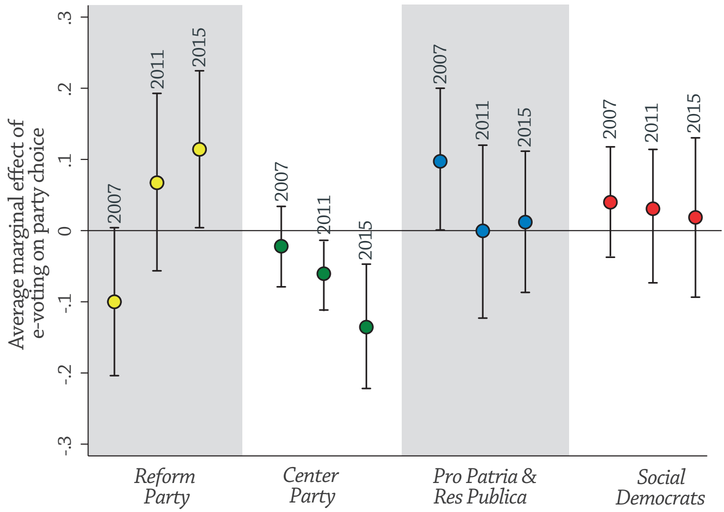

Figure 10.5: Associations between e-voting and party choice after matching for four parties in the parliamentary elections of 2007, 2011 and 2015.

What happened is best seen when the reader compares Figures 10.4 and 10.5. While not much changes with the Reform Party,i.e. e-voters in 2015 still had a 10% higher likelihood of voting for the party, two things are apparent for the Centre Party. First, we now see that in 2015 as well as in 2011 e-voters indeed had a lower likelihood of choosing that party. Second, the size of the effect in 2015 was considerably reduced after using the matching technique,i.e. it now stood at a 14% lower likelihood of e-voters choosing that party compared to paper voters. The differences observed in the aggregate levels and in the naive estimation models were hence partly due to inherent imbalances in terms of the observed characteristics between those voting on- and off-line.

10.4 Conclusion

In this chapter we presented tests for three individual level mechanisms that, if found true, would bias the aggregate election outcome. Through a series of empirical tests, we found supporting evidence for one mechanism. We saw that e-voters are indeed more likely to choose specific parties over others, even after controlling for the factors that supposedly affect party choice, such as age, income, education and political self-positioning. The associations were non-negligible and are therefore probably one reason behind the difference in aggregate level voting results based on different voting modes.

So why do some parties consistently gain more e-votes than others?

In order to respond to this concern, two outcomes needed to be considered that are relevant for voting behaviour. The first was the decision to vote for one of the parties competing in an election, i.e. vote choice. Although not exclusively, but often, voters act based on their political preferences. Whether short- or long-term, more or less responsive to policy changes, media influence, party leadership or political events, voters base their decision to vote for a party owing to their core political preferences. Even in the event of strategic voting, to some degree political preferences determine one's choice. To put it explicitly, political preferences determine people's vote choice. The second aspect of political behaviour is the decision to participate in elections, i.e. to turn out. The decision to vote, can and often is, a matter of principle, with some voters choosing to participate in elections at any cost, irrespective of unsatisfactory party performance or a rainy election day. Similarly, some voters may abstain as a matter of principle. Others may participate in elections occasionally, depending on the available choices, past party performance, weather conditions or something else.

We argue that the availability of voting modes – advance voting at the ballot station, postal voting, election day voting, as well as internet voting during the advanced voting period – is a feature of the electoral system that can first and foremost affect someone's decision to participate in elections. Consequently, I find it difficult to argue that the mode of voting, and e-voting in particular, can affect one's political preferences and therefore impact upon one's vote choice. It is logically possible, but empirically unlikely and the findings in the current chapter do not conclusively support it.

We suggest that the mechanism behind the unequal share of e-votes by different parties is, in fact, unrelated to political preference *per se*. Instead participation in e-voting requires its users to have certain skills and resources. These characteristics determine their frequency and habits regarding internet usage, and define a myriad of other aspects of their everyday behaviour. These characteristics simultaneously co-vary with their political preferences and consequently are predictive of the choices they make in political elections.When e-voting became available as a new means to cast a vote, these voters switched from traditional means of casting a ballot to online voting, because they have a higher propensity to use new technologies, not because they prefer certain political parties. The cause for e-voting, thus, is their propensity to use new technologies and not their political preferences *per se*.

However, the latter becomes accentuated when one only looks at the aggregate election results by mode of voting in terms of party shares, whereas the individual level mechanism driving the aggregate view remains uncovered. In this chapter we have approached and scrutinized these individual level mechanisms with a number of analytical techniques, across various datasets and elections, and found limited evidence that e-voting as a mode of voting induces political bias at the individual level. Therefore, we find it difficult to argue that it is biased at the aggregate level too.

One also needs to keep in mind the other findings outlined in this book. We know that e-voting has had a very limited impact on turnout and we know that current e-voters are already indistinguishable from paper voters. The only plausible explanation for the aggregate level differences in party vote tallies depending on the mode is therefore that a large share of typical voters have simply switched from paper voting to e-voting and this process is non-random, meaning a larger share of supporters of particular parties have done so. Should e-voting be discontinued, these people would simply switch back to paper voting.

In any case, voting modes are all used non-randomly, so the accusation of "bias" can be levelled at any mode, but a voting method that simply introduces one additional way of voting and functions in parallel with all other traditional modes can hardly be considered to bias results. What matters is that a maximum number of people are given the maximum number of options to participate and no one is hindered from participation.

Chapter 11

Conclusion and recommendations

This book represents a comprehensive study of the behavioural aspects of Estonian e-voting. In this concluding chapter we will briefly re-visit the central findings of the chapters, discuss their implications, and offer concluding remarks and recommendations.

The e-government ecosystem

An in-depth examination of the wider IT infrastructure surrounding the Estonian e-voting project goes to show that what really matters are the behind the scenes mechanics that normally remain unseen by your regular voter or outside observer. Without digital identification there could be no e-voting or at least no e-voting that did not rely on paper IDs or one-time identification methods implemented only for a particular election, and thus Estonian e-voting is in this sense fully online voting. It is however only a single e-service provided to citizens and wholly dependent on a wider infrastructure that makes a convenient voting mode possible. We showed how this infrastructure has been built and how widely it is nowadays used in providing government as well as private sector services in Estonia. It is important to consider e-voting as an integral part of such a system and not as a standalone phenomenon. Even though our focus here was on the behavioural aspects of e-voting, the fact that it is embedded in the wider e-service providing system ensures that citizens become familiar with such systems and develop the minimal level of trust needed for effective usage. Planting e-voting into a society that was still largely unaccustomed to interacting with public offices through online services would likely be a very expensive lesson in failure. E-voting was picked up in greater numbers in-line with the general growth of queries processed by the Estonian e-government ecosystem and the number of new e-services added to the system.

We saw that the growth in number of queries, which is an indicator of how well an e-government functions, was quite slow at the beginning, but started to grow exponentially once a critical number of services and data repositories had been added. This universal S-curve growth pattern in services, data repositories and processed queries in the Estonian e-governance system is one major factor that needs to be kept in mind. It tells us that any growth is bound to be slow in the beginning and patience is needed during the implementation phase. That patience needs to last many years before real-world positive effects in terms of working time saved become apparent.

System development

In parallel with the technical developments, a constitutional debate ensued that mainly focused on the privacy and secrecy of online voting. The debate on the abstract principles of a secret vote provided very clear input into the procedures that needed to surround e-voting and also highlighted the requirement to further develop the end-user side. As a result, Estonia decided to create a separate administrative level institution, the E-voting committee, to operate the e-voting system. This ensured clear responsibility and created an actor inherently tasked with safeguarding and developing the system. In addition, new procedures and technical solutions designed to enhance transparency and increase trust in the e-voting system were built after the first try-outs. It is again pivotal to point out that trust building measures related to e-voting alone where not effective when they are not complimented by similar efforts in related sectors. The strong public support and direct financial help of the banking sector, which was interested in widening electronic ID usage for secure online banking, should be especially noted. Encouraging citizens to use electronic IDs for online banking paved the way to their usage in other settings, such as e-voting. Creating such positive dynamics was the most effective way for increasing usage numbers, as citizens started to see cross-usage possibilities and usage of e-services, and by definition e-voting, as a normal way of going about their business.

This does not mean that the system of e-voting in Estonia was picked up by citizens by way of osmosis. The 2013 local elections saw one concrete technological solution being implemented that brought the system closer to end-to-end verifiability. To reduce the need to simply believe in the system's security in a situation where the average voter does not understand the workings of the sophisticated technology and encryption methods used in the Estonian e-voting system, individual e-vote verifiability was introduced. We devoted a whole chapter to its impact and saw it mainly to be negligible, a technological fix that does not really reach the people who might have potential trust problems, but it is nevertheless important to have. Any method that e-voters can use themselves to reduce potential uncertainties they might have over "what happens to my e-vote?" is worth implementing, as it helps to make the voting procedures appear as transparent as possible.

Finally, one also needs to keep in mind that the necessary legal framework and separate administrative unit were the result of an incremental process. Though the e-voting system was fully functional by 2005 and was launched nationwide in one go, the procedures have been consonantly improved. This goes to show that building a good functioning system cannot happen overnight and maybe it is not even advisable to try when one wants to get e-voting going. It is better to start small and simple and build on experience, which might also work better in terms of building user confidence. Showing that the system is being constantly improved, while not compromising end-user simplicity, helped to keep the solution visible. It might of course also lead to sceptics interpreting the need for improvement as a sign of dysfunctional system, but this is a necessary evil that goes together with developing any innovation in the politically charged environment of democratic elections.

Growth in usage

The central focus of this book has been on the e-voters themselves. When the first e-enabled election took place in 2005, only slightly over 9 000 voters decided to cast their vote online, which was close to 2% of all votes, meaning only every 50th voter e-voted. By the last parliamentary election in 2015, user numbers had grown to more than 170 000, which was more than one third of all votes cast during that election, so after exactly 10 years every third vote is now cast online. This is a remarkable development by any standard.

Nevertheless, such user numbers do not necessarily mean that usage of this voting mode has diffused among the whole voting population, which was the desired outcome, it might simply mean something akin to market saturation, with almost all younger tech savvy people using it and a substantial non-random segment not using it. The data showed that in the first three e-enabled elections, e-voters were indeed clearly distinct, they were younger, with better computer skills and mostly ethnic Estonians. From the third election onwards however, these differences started to disappear, meaning that e-voters became progressively less distinct from regular paper voters. By now we effectively cannot differentiate between e-voters and paper voters based on a list of socio-economic characteristics and can safely say that e-voting has become a tool of the masses, with all quite heterogeneous social groups engaging in this type of voting.

The second crucial finding was the speed of diffusion. Very much like the growth of e-services in general, the growth in user numbers was slow and the exclusive nature of e-voting persisted. As with anything, it simply takes time for regular voters to get comfortable with the idea of voting remotely over the internet. Politicians are well advised to be patient and not discontinue e-voting after a couple of trials when they do not see immediate wide usage, as was very much the case in Norway, for example. The Estonian experience shows e-voting to have the potential to spread among the masses and not stay a fancy technological tool for the already privileged and well connected citizens. Especially the fact that age as a factor does not differentiate between e-voters and paper voters in Estonia any more is encouraging. Elderly and other groups, who tend to have mobility problems, have taken to this technology quite effectively. E-voting could hence actually reduce participation thresholds for people with mobility issues and thereby improve their quality of representation.

Nature of usage

In addition to extensive survey data, we also made use of anonymized e-voting system log data from 2013 and 2015. This represents a period of massive e-voting and wide diffusion, which gives an especially good picture of the actual practice of e-voting among the general electorate.

The data showed a typical e-voter to be similar to a typical Estonian, i.e. middle-aged and the genders represented to the same degree as in the wider population. Furthermore, all age groups clearly e-vote with the exception of the youngest voter group, i.e. those who have just crossed the age of eligibility.

Perhaps the biggest surprise was the speed with which people e-vote. It takes on average less than three minutes and most surprisingly less so the older the voter is. This goes to show once more how much convenience e-voting brings. An action that took on average 30 minutes (the average time to vote by going to and from a polling station according to survey data) is reduced tenfold. Consider the economic impact such a time saving would have in countries that have voting during work days, as many old democracies still have. We should also not dismiss the time saving for individuals in weekend voting countries, such as Estonia. Being able to do your civic duty considerably faster will open up time to engage in other fruitful activities that might also bring macroeconomic benefits, not to mention valuable time spent with family.

We also examined the cost saving effect of e-voting separately using voter survey data. We saw that people were already more likely to vote online rather than on paper, when it took them a more than 30-minute round trip to vote at a polling station. This association persisted regardless of the age, education or computer literacy level of the voter. This means that the time saving potential is considered, so valuable in fact that even people who might shun online transactions due to their old age or low online skill took the leap and voted online. This is again a finding that needs to be emphasized. Think of countries that have a low population density and more difficult terrain; ensuring everyone in rural areas has the same easy access to voting points is quite expensive and difficult. We think the positive potential of e-voting could be realized to an even larger degree in such settings/countries, as it should be easier to provide better internet and communications cover than to build expensive physical infrastructure.

One must also not forget that e-voting makes it exceptionally simple to vote from abroad. Evidence shows that roughly 90% of votes from abroad in Estonian elections are now cast as e-votes; hardly anybody votes via post or in embassies anymore. Again, countries with large expatriate communities should see e-voting as an attractive way of keeping these people engaged with politics in their native country, given how widespread studying and working abroad is nowadays.

The log data also allowed us to pinpoint to what degree the feature of multiple e-voting (with only the last one counting) that was added to counter possible voting under duress is actually being used. We saw that a marginal share of e-voters actually voted multiple times and the ones who did so re-voted in close (time) proximity to their first or previous e-vote. With the risk of sounding overly euphoric, we have to say that Estonian e-voting functions remarkably well, with all age groups e-voting and doing so surprisingly quickly. Given that we can say this with actual system log data backing up each claim, should convince even the most sceptical that e-voting has great future potential to make the voting experience more convenient than it has ever been.

Unrealized potential for turnout

Amid the multiple successes of Estonian e-voting, there is also one key aspect where the hoped for potential has not been fulfilled.

When e-voting was first introduced in Estonia and wherever it is discussed worldwide, the potential to increase turnout or at least to arrest its continuing decline in Western democracies is usually mentioned. This has really not materialized in Estonia's case. The aggregate turnout did increased marginally after the introduction of e-voting, but we cannot really claim that this was only down to e-voting as such. This was all the more surprising as e-vote user numbers have grown immensely. This contradiction can be explained with a very simple fact. An increase in turnout can only happen when people who previously did not participate decided to participate because of the availability of e-voting. It just so happens that these very same non-voters are also quite unlikely to be sufficiently e-literate to e-vote in the first place. Paradoxically, to have a clear impact on voter turnout, e-voting should mobilize exactly those people who are very unlikely to e-vote.

We think this is one of the major implications of this book. E-voting does not address the underlining causes of turnout decline, such as disinterest, political disappointment and partisan de-alignment. E-voting simply makes voting easier for people, it will not necessary engage those for whom the problem lies in politics as such. The expectation of online voting increasing turnout is largely misplaced in light of the Estonian case. Low turnout is a problem, but online voting is not the solution to it.

This does not mean that e-voting will have no impact on turnout whatsoever, given that the non-voting subgroups are exactly the target group for increasing turnout, even limited mobilization effects provide an improvement. What we are simply saying is that in light of the limited impact of e-voting on turnout, making voting easier and more convenient should be a desired goal in itself.

In order to provide a positive note on turnout, we can only recommend to use more indirect ways to promote voting online and embed it into a wider emphasis on e-services. Given that the usage of e-voting is not a goal in itself, but a means of ensuring better representation and ease of participation, the most effective way to achieve this is to increase e-literacy and not strictly to promoting casting one's vote online. The spread of any e-services, public or private, should by definition also increase the general ability to e-vote among the population segments who tend to shun online participation. But when it comes to overall participation rates, we have to concede to the rather limited effect of e-voting.

The e-voting habit

Our detailed examination of e-voting development and practices raised the question of what drives people to vote online? We know now that the original distinctions between on- and off-line voters no longer apply. We also know that people for whom voting is physically difficult are e-voting more readily, though voting is already quite simple in Estonia for the majority of voters due to the high density of polling stations. We also know that e-voting has not mobilized many new voters to turn out. So why are people e-voting?

One rather simplistic, but quite possible explanation given the ten years of e-voting in Estonia, is that they do so because they have done so in the past. This means they have simply gotten used to this mode of voting and the original reason that drove them to e-vote in the first place no longer applies. This is a dictionary definition of a habit. It is automated and not strictly reasoned repetitive behaviour. Applied to a mode of voting, it would mean certain people started to e-vote a while back and have kept on voting in the same manner without thinking anymore why they do so. A closer examination of repeated e-voters clearly showed this is often the case.

Now one might consider it a bad thing when automation takes over and people no longer reflect upon why they do things the way they do. We argue that this is a good thing when it comes to the act of turning out. It is good when people participate because they have developed a habit of doing so, as it does not mean that they do not think about the political choices they make. It simply means turning out to vote is somewhat automated, the thought process that goes into the choice of political party itself is most likely not automatic as the electoral context is always a bit different when compared to the last.

The data also brought out something that is of wider relevance. Voting as a habit has been examined over longer periods in voting behaviour literature. The Estonian data showed e-voters to be more persistent voters than paper voters. this means once you make the switch to e-voting, you are more likely to participate in subsequent elections in comparison to when you have simply voted on paper. This is important, as it suggests e-voting to have the potential to condition people to participate more than paper voting. It also means that e-voting might be able to arrest continuing turnout declines, as long as a sufficiently larger share of voters opt to vote online. In other words, e-voting will not increase turnout, but it may keep citizens voting at higher levels than regular paper voting does. This a nuanced association, but something with great implications.

Trust and neutrality

In addition to e-voting behaviour, we also examined a crucial element or condition for it to happen in the first place: trust in e-voting. The Estonian voting authorities have made concerted efforts to increase trust and the e-vote verification solution was added in 2013 with this in mind. Giving voters a tool that can be used to check if their e-vote has been received by the vote counting server and counted in the intended manner should improve trust and increase the overall security of e-voting. Providing a satisfying answer to the ever present “what happens to my e-vote?” question is absolutely crucial for e-voting to succeed.

The data shows that the efforts of the electoral authorities have paid off, as the overall trust level in Estonian e-voting is very high. The great majority of people had high or very high trust levels to begin with. But we have to say that there is a clearly distinct group who does not trust the system at all. In other words, e-voting polarizes opinion, with a majority seeing no problem with it and a quite small, but nevertheless important group, having major trust problems.

The problem is that no technological solution is likely to address the non-trusting minority. We think the potential of e-vote verification to address the fears of this group is fairly limited, as they are very unlikely to ever try e-voting. So new technological solutions introducing end-to-end verifiable e-voting will never reach these people. Paradoxically, trust in e-voting has room to improve only among the subgroup who is the least likely to ever use the technology designed to improve trust. This is a classic case of a "technological fix" to an issue that technology is ill equipped to address. Trust building measures are important, but it is wise to approach different groups differently. For the ones already engaging in e-voting, trust is important to keep them e-voting; these people can be won over with continuous and incremental system improvements that show the e-voting system to be keeping up with changing times and technologies. Among such people technological fixes do indeed work.

As for the non-trusting group, a more indirect approach might be more efficient. A general growth in e-services and usage should go some way to address the trust issues towards e-voting, by showing them that online transactions are useful and commonplace in very diverse settings and voting online can be as normal and simple as online banking or shopping. Finally, e-voting is just a mode of voting and when it comes to elections as such, the integrity of the process becomes absolutely key in ensuring the legitimacy of the outcome. It is therefore probably wise to keep the procedures of electronic vote processing as similar to regular vote counting processes as possible, so the sceptics always see the equivalence of this or that step between the processes of on- and off-line voting. This is exactly what has been done in Estonia. We have an electronic voting committee as part of the national voting committee, use the digital equivalent of the double envelope system as used in postal voting worldwide, with the electronic votes are electronically ”opened” and ”counted” on election day with observers present, as is done in any polling station after polls close. All this should make the process as transparent as possible, even though one mode uses the latest in cryptography and the other a pen and a ballot paper in a voting booth, just as when the secret vote was introduced in the 19th century.

Connected to trust issues is the still ongoing political debates around the neutrality of e-voting. The aggregate levels results for e-votes and paper votes differ substantially, which is a puzzle to the average politician and voter. We examined various ways how this can come about and found the most plausible explanation to be that regular paper voters simply switched to e-voting in larger degrees depending on the party they support, which is supported by the fact that e-voting has not really brought new groups with distinct party sympathies to the polls. Should e-voting be discontinued, these people would simply return to paper voting.

In light of this we see the discussion on bias somewhat misplaced, as all voting modes are used non-randomly. By implication, no two modes produce the same outcome in terms of overall party vote shares. It is simply that different types of people of particular political preferences use different types of voting modes. What really matters is that nobody is actively kept from voting due to mode specifics. This is not the case for e-voting in a free and fair election, it simply gives yet another way to vote without limiting any of the other traditional voting channels. It is therefore a problem of perception, rather than a problem of substantive bias.

We see two possible solutions for the electoral authorities to address the misleading claims of political bias. One would simply be to publish the results without differentiating between voting modes. After all, not many countries publish voting results according to the voting modes. Estonia does so only for paper and e-voting, but not for the different types of paper votes, such as voting by ordering the ballot box to a home, voting from ships, voting via post and voting in polling stations during the advanced voting period. The problem of course is that not publishing party tallies based on e-votes could compromise the crucial transparency of the e-voting process.

The other solution would be another other extreme of publishing aggregate results according to all of the possible voting modes in the country for everyone to see that no two modes produce exactly the same aggregate outcome, and it is only the sum of all these modes that gives us the final election outcome, where all voters groups have had their say using very different voting modes. This might confuse some people, but it might also end the controversy around e-voting.

Future of Estonian e-voting [24]

We can safely say that internet voting is here to stay. The Estonian case is one outstanding success of e-voting, but that does not mean the system cannot be developed further.

During the past decade the e-voting system has been constantly evolving in line with developments of cyberspace in general. Given that elections are a subject of increased public attentiveness and scrutiny, the responsible authorities have been constantly introducing new security features to increase transparency and increase trust.

A clear plan for future developments is already in place. The upcoming local election in 2017 will make use of a internet voting system with end-to-end verifiability. This means that every step in the e-voting process will be available for scrutiny. Individual verification was introduced already in 2013, now the rest of the path – from recorded votes to final results – will be covered as well. The latest in advanced techniques – like homomorphic cryptography, mix-nets and zero-knowledge protocols – will serve to make this possible. The average voter with no mathematical training will of course not really notice this, but end-to-end verifiability will give anyone, if they wish so, the possibility to see the transparency and safety of the system.

Another development will allow for the online-part of the central e-voting system – the vote collection – to be fully outsourced to a third party, subject to strict conditions. This would allow the election administration to focus on their core work of administrating the elections, while still retaining control of the vote-collection run by the third party.

A third major development of the forthcoming system will be its universality and scalability. Previously used solutions were limited to Estonian elections only. There has however been wide international and also domes-

[24] A vision provided by Mr. Tarvi Martens, Chairman of the Estonian Electronic Voting Committee.

tic interest in using the technology, including for other kind of elections, like polls in large private organizations, stockholders meetings or company board meetings. The Estonian authorities are ready to provide engineering architecture and release the source code for e-voting, so it can be reused in various other elections.

All in all, though e-voting will not replace paper-based voting in the foreseeable future, it has clearly influenced our understanding of how election will be run in the future, given that e-voting provides a more controllable voting system that involves less of human factors in election administration.

The increased mobility of people will also increase demand for remote voting methods and the Estonian example so far should remind us to stay open minded and interested in innovating voting technologies.

Appendix A

Materials on diffusion

A.1 Variables and models

The dependent variable was dichotomous: respondents are either first time e-voters (coded 1) or paper ballot voters (coded 0). To ensure that the reference category does not include any occasional e-voters, we have excluded respondents who reported to have e-voted in previous elections, even if they voted with a paper ballot in the given election.

We examined the socio-demographic traits of voters and focused on: 1) age in years; 2) gender (male = 1; female = 0); 3) income (measured as income deciles); 4) education (a dummy variable for higher education and one for secondary education, with basic education as the reference category); 5) ethnicity (ethnic Estonian = 1; ethnic Russian =0); 6) the type of settlement in which the respondent resided (urban = 1; rural = 0). The behavioural and attitudinal variables of interest were: 7) prior voting habits (respondent had participated in all or most elections they were eligible for = 1; has participated rarely or never = 0); 8) self-reported computer literacy level (a dummy for good and one for average skill levels, with basic level as the reference); 9) trust towards the e-voting system (trust = 1; no trust = 0); 10) left-right self-position (10 point scale from left to right); 11) the frequency of engaging in political discussions with friends/family (often, sometimes = 1; never = 0).

To evaluate our expectations we estimate six separate logit models for each election in the following form:

$$ln\left\{\frac{Pr(evote_i = 1)}{1 - Pr(evote_i = 1)}\right\} = \beta_0 + \beta_i X_i \quad \text{(A.1)}$$

where X_i is vector of the above mentioned independent variables for an individual i. Our particular interest was in how β_i – the associations with the independent variables – changed over time.

A.2 Additional information on models and results

Table A.1 presents the results of the models that predicted first time e-voting.

Table A.1: Predicting first time e-voters (base: paper ballot voters only)

	2005 local	2007 national	2009 EP	2009 local	2011 national	2013 local	2014 EP	2015 national
Voting habit	−3.89	−1.29	16.33**	−0.44	2.82	−1.87	1.63	−7.97***
	(7.36)	(8.09)	(7.19)	(5.22)	(5.93)	(3.11)	(2.84)	(2.99)
PC literacy:good	29.99***	29.98***	15.89***	2.61	11.64**	15.13***	9.03	15.57**
(base: basic)	(5.14)	(5.18)	(5.70)	(6.10)	(5.10)	(4.79)	(4.73)	(4.64)
PC literacy:average	16.11***	16.68***	13.54***	1.48	6.42	12.02***	5.96	8.35*
	(4.74)	(4.97)	(5.05)	(5.51)	(4.67)	(4.42)	(4.42)	(4.25)
Trust e-voting	49.80***	33.47***	67.14***	35.66***	23.08***	13.51***	8.74**	13.59***
	(4.07)	(5.27)	(11.22)	(7.61)	(3.78)	(3.51)	(3.25)	(3.19)
Left-right position	0.05	1.08	0.05	−0.75	0.68	0.84	0.37	0.15
	(0.86)	(0.95)	(0.90)	(0.85)	(0.93)	(0.69)	(0.49)	(0.61)
Political talk	−2.2	7.82**	−7.14*	−6.78**	6.33**	1.04	−0.33	3.40
	(3.13)	(3.23)	(3.23)	(3.29)	(3.04)	(3.02)	(2.25)	(3.15)
Age	1.30**	1.38**	1.38**	1.25*	0.44	0.18	0.27	0.55
	(0.59)	(0.58)	(0.60)	(0.68)	(0.53)	(0.42)	(0.39)	(0.40)
Age^2	−0.01*	−0.01*	−0.01**	−0.02**	−0.01	0.00	−0.00	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Male	−0.94	3.37	−0.5	−1.03	6.84**	1.99	−1.60	4.59
	(3.09)	(3.11)	(3.18)	(3.37)	(2.88)	(2.34)	(2.11)	(2.37)

Continued on next page

Table A.1 continued from previous page

	2005 local	2007 national	2009 EP	2009 local	2011 national	2013 local	2014 EP	2015 national
Ethnic Estonian	25.29***	34.50***	28.67***	10.56**	12.46**	2.54	1.02	6.50
	(5.24)	(7.03)	(7.57)	(5.13)	(5.70)	(3.48)	(2.76)	(3.57)
Education: higher	7.34	5.92	11.95	4.09	3.32	8.42	−4.04	2.90
(base: basic)	(7.46)	(7.79)	(9.47)	(8.62)	(6.63)	(5.48)	(3.35)	(4.14)
Education: secondary	3.50	−0.03	4.75	−1.69	−4.78	4.90	−6.15*	−0.48
	(7.23)	(7.38)	(9.32)	(8.33)	(6.29)	(5.25)	(3.02)	(3.75)
Income decile	−0.18	1.38**	0.53	1.17	0.83*	−0.63	0.18	−0.63
	(0.58)	(0.59)	(0.68)	(0.71)	(0.48)	(0.52)	(0.42)	(0.52)
Urban	−0.94	2.88	−1.53	−4.06	−4.82	1.40	−0.24	5.99*
(base: rural)	(3.48)	(3.32)	(3.38)	(3.47)	(2.94)	(2.61)	(2.24)	(2.75)
Constant	−6.39***	−8.42***	−9.93***	−5.28**	−5.48**	−6.30***	−4.99***	−7.70***
	(1.20)	(1.11)	(1.37)	(1.45)	(1.21)	(1.55)	(1.56)	(1.52)
Observations	633	759	712	511	619	625	645	617
Nagelkerke Pseudo R^2	0.53	0.39	0.43	0.35	0.39	0.28	0.19	0.24
Correctly classified	0.80	0.75	0.73	0.80	0.82	0.90	0.93	0.89
Sensitivity	0.86	0.74	0.73	0.27	0.45	0.02	0.00	0.04
Specificity	0.74	0.77	0.73	0.95	0.92	0.99	1.00	0.99
Log Lik	−277.37	−382.29	−333.18	−197.52	−237.77	−162.71	−145.30	−163.73

Average marginal effects in percentages; standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A.1 presents various measures of model fit that supported our diffusion expectation. First, we saw that the general model fit shown by the pseudo-R^2 did indeed decrease over time, indicating an increase in the expected heterogeneity amongst e-voters. Second, our models drastically lost the power to correctly classify first-time e-voters (the true positives) from the fourth election onward. Notice that at the same time the share of correctly classified regular voters (the true negatives) remains constant or even increases slightly, providing the basis for a high level of overall correct classification. Taken together, the decline in model fit, as well as the significant drop in correct classification of first-time e-voters (the true positives), these findings clearly support the expectation that there has been diffusion of e-voting and not just an expansion of usage among a distinct subgroup of tech savvy voters.

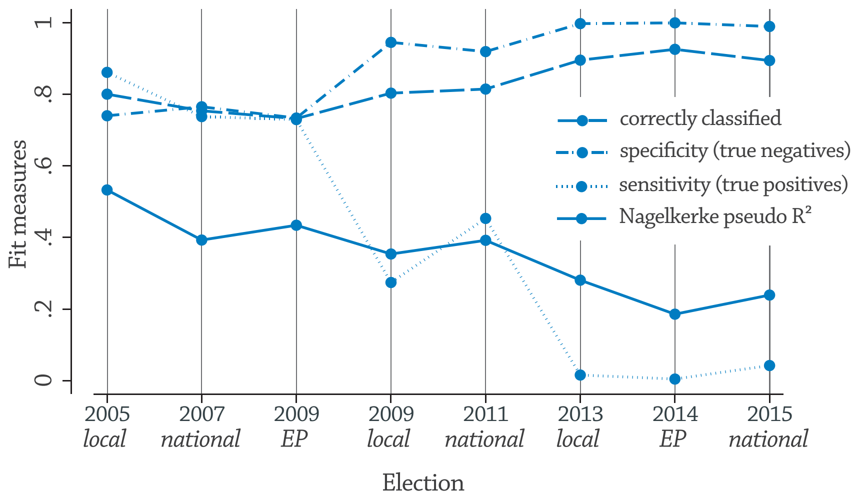

Figure A.1: Dynamics of the model fit for predicting first time e-voting

From 2013 onwards our basic model failed to distinguish between first-time e-voters and paper ballot voters completely, indicating that first time e-voters are virtually indistinguishable from paper voters for most of the characteristics included in the model.

Appendix B

Materials on mobilization

B.1 Factor analysis

Table B.1 shows the eigenvalues and proportion of variance accounted for by the factors. We saw that the first factor accounted for 53.4% of the total variance. Given that the other factors have clearly lower eigenvalues and explain less variance, we only used the first factor in the analysis.

Table B.1: Factor analysis: factor description

Factor	Eigenvalue	Diff	Proportion	Cumulative
Factor 1	2.14	1.20	0.53	0.53
Factor 2	0.94	0.35	0.24	0.77
Factor 3	0.59	0.26	0.15	0.92
Factor 4	0.33	–	0.08	1.00
LR test				2206.01^{***}

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The factor loadings are displayed in Table B.2. It shows that all variables were positively correlated with the factors and individuals who trust e-voting and internet transactions, use the internet frequently and have good PC skills scored high on the factor. The distribution of factor scores themselves are shown in Figure B.1

Table B.2: Factor analysis: factor loadings

Variable	Factor 1	Uniqueness
Trust in e-voting	0.58	0.66
Trust in internet transactions	0.70	0.51
Weekly internet usage	0.80	0.36
PC literacy	0.82	0.33

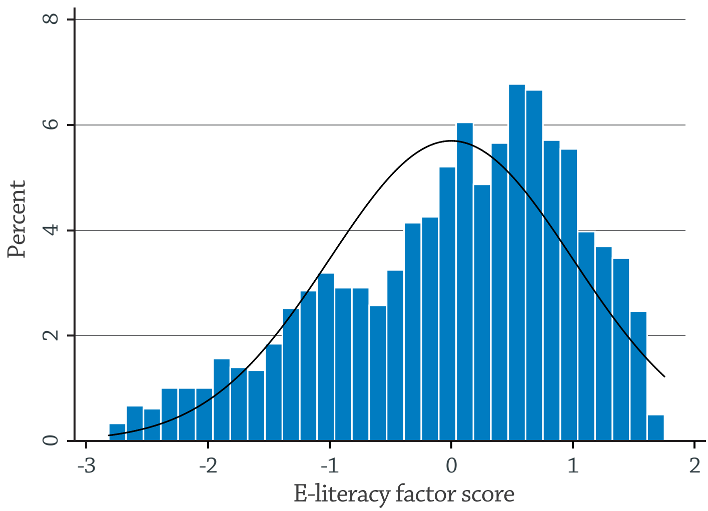

Figure B.1: Distribution of the e-literacy factor score

B.2 Additional information on models and results

Table B.3 shows the effect of e-literacy on two outcomes. First on the probability to vote online (usage) and second on the probability to switch from non-voting to e-voting (mobilization). We see it has a positive effect on usage while a negative effect on mobilization demonstrating the bottleneck effect.

Table B.3: Effects on the usage and the impact of e-voting

	Usage	Mobilization
E-literacy (factor score)	0.25^{***}	-0.21^{***}
	(0.24)	(0.04)
Age	0.00^{*}	-0.01^{***}
	(0.02)	(0.00)
Education: higher	0.00	−0.07
(*base: primary*)	(0.06)	(0.06)
Education: secondary	−0.02	−0.06
	(0.04)	(0.07)
Income decile	0.00	−0.03
	(0.00)	(0.01)
Ethnic Estonian	0.10^{*}	−0.08
	(0.05)	(0.01)
Male	0.04	0.00
	(0.02)	(0.05)
Constant	-4.89^{***}	4.50^{***}
	(0.78)	(1.15)
Sensitivity	39.13%	40.00%
Specificity	93.72%	96.53%
Observations	672	223
Nagelkerke pseudo-R^2	0.44	0.29

The table reports average marginal effects with standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix C

Materials on voting cost

C.1 Variables and models

We operated with two dependent variables. First, the reported participation of the respondent (coded 1), with non-voters as the reference group (0). Second, the reported mode of voting, contrasting e-voters (1) with paper ballot voters (0).

Three variables were included to capture the non-instrumental aspect of the act fo voting. First, a subjective evaluation of the frequency of prior participation in elections as a dummy comparing citizens who had participated in all or most prior elections (1) to the ones who had participated infrequently or not at all (0). Second, the frequency of engagement in political discussions with friends and family, comparing people who did so often or sometimes (1), with those who did it rarely or never (0). Third, the level of trust towards parliament, the government, politicians and the state, with all measured using a four category Likert-scale that was summed into an additive index where larger scores showed higher trust.

To express the utility of voting choice options on the ballot, we used a survey instrument that ranks parties by their perceived utility, the so called propensity to vote (PTV) measure. We follow van der Eijk and Oppenhuis (1991) in taking the voting propensity score of 8 as the cutoff and counted the number of parties assigned equal or higher scores by the respondent. The higher this number, asked for eight parties in total thus giving a range of 0-8 for the count, the more attractive the choices on the ballot were for the respondent.

Voting cost was measured with an item whereby respondents were asked to rate how many minutes (approximately) it took/would have taken them to go to the polling station and back. This gave us a tangible individual level interval measurement of the possible cost of voting, namely the physical proximity of the polling station.We also included age, education, income and gender as prominent predictors of participation (Matsusaka and Palda, 1999; Franklin, 2004; Smets and van Ham, 2013). Age was measured in years; education as three dummies for higher, secondary and vocational education, with basic education as the reference; income was measured as the income decile the respondent's household fell into; gender was measured with dummy contrasting males (1) to females (0). The technical skill levels of respondents were taken into account with an item on the subjective evaluation of computer literacy (1 - none-existent; 2 - basic; 3 - average; 4 - good to very good). We analysed the interaction of age, education, income and PC literacy with voting cost to see if the effect of cost on the usage of e-voting is dependent upon each citizen's cognitive ability and resources. Finally, as Estonia has a substantial ethnic Russian minority that tends to differ in their political behaviour, we include ethnicity as an additional control.

We estimated probit models to be able to compare estimations with the selection model explained below. The first two expectations were evaluated by estimating the following models:

$$Pr(vote = 1) = \varphi(\beta_0 + \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D) \tag{C.1}$$

$$Pr(evote = 1) = \varphi(\beta_0 + \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D) \tag{C.2}$$

where A is the vector of non-instrumental motivations, B the utility of the choices, C the cost of voting and D controls. These models were identical but for the dependent variable. We expected β_3 to have a negative sign in the first and a positive sign in the second model. With regard to β_1 and β_2 we expected to see significant positive effects in the first and non-significant effects in the second model.

Modelling voting and e-voting separately assumes that these two outcomes are generated by independent mechanisms, analogous to a two-part Tobit model. This is a strong supposition. An alternative is the partial interdependence of voting and e-voting mechanism, even after controlling for our regressors. This suggests a possible self-selection bias, where e-voters select to use this mode non-randomly even after the controls for the mechanism we proposed. To ensure our estimations of the central variables of interest, i.e. the cost of voting, were robust, we also specified a Heckman selection model, where the selection equation predicts voting and the outcome equation predicts e-voting.

The third expectation was evaluated graphically with estimates extracted from the following probit model:

$$Pr(evote = 1) = \varphi(\beta_0 + \beta_1 A + \beta_2 B + \beta_3 C \times R + \beta_4 D) \tag{C.3}$$

with R the vector of resources. We focus on β_3 i.e. how the effect of cost on e-voting differs depending on the level of resources and vice-versa. The data is multiply imputed and all models will be rendered using clustered standard errors by election type.

C.2 Additional information on models and results

The results of the multivariate models are reported in Table C.1, which lists the average marginal effects in percentages. The probit model coefficients were very similar with the Heckman selection model in the case of e-voting (see Table C.2) and the rho coefficient of it was not significant, indicating no self-selection problem[25] once all the controls are included.

Table C.1: Multivariate models of voting and e-voting

	Vote (base: non-vote)	E-vote (base: paper ballot)
Voting habit	42.23***	−0.82
	(2.37)	(10.01)
Political talk	5.80**	−3.81**
	(2.24)	(1.40)
Trust in institutions	1.88***	−0.14
	(0.35)	(0.41)
Utility of voting choices	2.77**	−1.11
	(1.14)	(1.14)
Distance to the polling	−1.62*	14.40***
station in minutes (log)	(1.14)	(2.19)
Age	0.99**	0.83***
	(0.43)	(0.27)
Age^2	−0.00*	−0.00***
	(0.00)	(0.00)
Male	−3.04	5.24
	(2.09)	(3.35)
Ethnic Estonian	−1.02	27.03***
	(3.04)	(3.34)
Education: high	11.32***	9.07**
(base: basic)	(3.54)	(3.83)
Education: secondary	4.86**	−0.72
	(2.41)	(6.37)
Education: vocational	3.57**	−2.63
	(1.51)	(6.25)
Income decile	0.36	1.35* * *

Continued on next page

[25]We used voting habit, political talk, trust of institutions and the utility of voting choices as exclusion restriction, as these predict voting strongly, but not e-voting, making them suitable for this purpose (Cameron and Trivedi, 2010, p. 558).

Table C.1 continued from previous page

	Vote (base: non-vote)	E-vote (base: paper ballot)
	(0.23)	(0.34)
PC literacy: good	12.80*	42.67***
(base: poor)	(7.71)	(7.30)
PC literacy: average	10.27	34.51***
	(6.63)	(6.54)
PC literacy: basic	9.40	25.05***
	(7.31)	(6.54)
Constant	−3.18**	−4.45***
	(0.74)	(1.25)
Observations	2999	2011
Nagelkerke Pseudo R^2	0.38	0.50
Correctly classified	78.97%	77.60%
Sensitivity	95.74%	84.44%
Specificity	44.82%	72.05%
Log Lik	−1,419.30	−915.55

Average marginal effects in percentages

Standard errors clustered by election in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.2: Robustness checks for e-voting (base: paper ballot voters)

	Probit	Heckman(outcome)
Voting habit	−0.82	
	(10.01)	
Political talk	−3.81**	
	(1.40)	
Trust in institutions	−0.14	
	(0.41)	
Utility of voting choices	−1.11	
	(1.14)	
Distance to the polling	14.40***	14.27***
station in minutes (log)	(2.19)	(2.18)
Age	0.83***	0.93***
	(0.27)	(0.18)
Age^2	−0.00***	−0.01***
	(0.00)	(0.00)
Male	5.24	5.26
	(3.35)	(3.35)
Ethnic Estonian	27.03***	27.10***
	(3.34)	(3.38)
Education: high	9.07**	9.76***
(base: basic)	(3.83)	(1.86)
Education: secondary	−0.72	−0.38
	(6.37)	(5.62)
Education: vocational	−2.63	−2.44
	(6.25)	(5.41)
Income decile	1.35***	1.39***
	(0.34)	(0.29)
PC literacy: good	42.67***	43.71***
(base: poor)	(7.30)	(6.97)
PC literacy: average	34.51***	35.34***
	(6.54)	(6.21)
PC literacy: basic	25.05***	26.00***
	(6.54)	(6.68)

Continued on next page

Table C.2 continued from previous page

	Probit	Heckman(outcome)
Constant	-4.45^{***}	-4.81^{**}
	(1.25)	(0.88)
Observations	2011	2999
Nagelkerke Pseudo R^2	0.50	
Correctly classified	77.60%	
Sensitivity	84.44%	
Specificity	72.05%	
Log Lik	−915.55	−13541.76
Rho		0.18

Average marginal effects in percentages

Standard errors clustered by election in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix D

Materials on e-voting habit

D.1 Variables and models

Our central variable of interest was a binary indicator of the mode of voting in a current election, coded 1 if the respondent e-voted and 0 if they voted on paper at a polling station. The path model itself is depicted in Figure D.1 and includes as predictors the main socio-demographic traits and proxies for resources, which are usually associated with a higher likelihood of online participation. First variable is the age of the respondents in its original and squared format (variables Age and Age^2). Second, the education level of the respondent in the form of dummies for higher and secondary education, with basic education as the reference category.

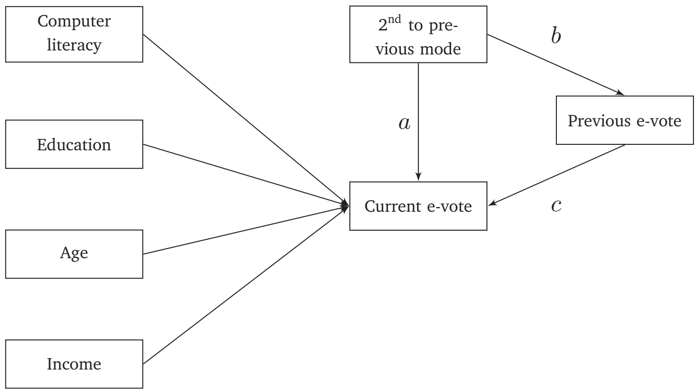

Figure D.1: Path model of associations between prior e-voting and current e-voting (a - direct effect of the second to previous mode on current e-voting; $b \times c$ - an indirect effect through the previous mode on current e-voting)

Third, the income level of the respondent in the form of the income decile the respondent's household fell into. Finally, three dummies on the self-reported PC literacy of the respondent, with poor computer skills serving as the reference category. We also included prior voting behaviour in the second to previous election and previous election, with two dummies each, one for e-voting and one for paper voting, with non-voting as the reference category. We allowed the second to previous voting mode to have a direct effect on the current e-vote, path a, as well as indirect effect through the previous voting mode, paths b and c.

We estimated this model for each year over the five-year period we had data for and were specifically interested if the associations between socio-demographics on the current voting mode choice decreased when the effects shown in paths a, b and c increased.

We will also examined the paths a, b and c in detail through a decomposition technique. Such decomposition of the total effect into direct and indirect effect works through counterfactuals. For example, we are interested in the effect of X on Y through Z, where X is voting behaviour in the 2$^{\text{nd}}$ to previous election (X=1 for an e-vote; X=0 for a paper vote), Y is the current voting behaviour (Y=1 for an e-vote; Y=0 for a paper vote) and Z is the behaviour in the previous election. First we computed the odds of e-voting in the current election for the X=1 and X=0 group, then we created a counterfactual scenario where the X=0 group had a distribution of Z of the X=1 group and computed the odds of e-voting of such a theoretical group. Finally, we contrasted the first odds of the original X=0 group with the counterfactual group, which gave the indirect effect (effect of Z on Y while X is fixed) and then the odds of the original X=1 group with the counterfactual group, which gave the direct effect (effect of X on Y while Z is fixed). This operation is formally expressed in the following manner:

$$\overbrace{\frac{O_{x=1,z|x=1}}{O_{x=0,z|x=0}}}^{\text{total}} = \overbrace{\frac{O_{x=0,z|x=1}}{O_{x=0,z|x=0}}}^{\text{indirect}} \times \overbrace{\frac{O_{x=1,z|x=1}}{O_{x=0,z|x=1}}}^{\text{direct}} \tag{D.1}$$

where the total effect, i.e. the odds of e-voting in the current election, is expressed as a product of the indirect and direct effect. For a thorough discussion of this method, see Buis (2010).

To address the third expectation that repeated e-voters did not have supportive attitudes for e-voting anymore we examined the subsample of respondents who had e-voted multiple times with a logit model where the dependent variable was whether the voter is a consecutive e-voter or not. The aim was to see if supporting attitudes towards e-voting are or are not associated with e-voting multiple times consecutively. This model was estimated for the combined dataset and took the following form:

$$\begin{split} ln\left\{\frac{Pr(evote_t=1)}{1-Pr(evote_t=1)}\right\} = \beta_0 + \beta_1 trustevote + \beta_2 trustinter + \\ + \beta_3 confown + \beta_4 confother + \beta_5 controls \end{split} \tag{D.2}$$

where $trustevote$ was a dummy on whether the respondent trusted or tended to trust e-voting in general, $trustinter$ was a dummy on whether the respondent trusted or tended to trust internet transactions, $confown$ was a dummy on whether the respondent was confident that her e-vote was counted as cast, $confother$ a dummy on whether the respondent was confident that e-votes by other voters were counted as cast and $controls$ a vector of the same socio-demographic variables as included in the path model in Figure 8.2 acting as controls.

D.2 Additional information on models and results

Table D.1 shows the associations between socio-demographics and e-voting in given elections, as exponentiated path model coefficients extracted from the model depicted in Figure D.1. Table D.2 shows the direct and indirect effect of prior e-voting (paths a, b and c), with all other covariates held at their mean.

Table D.3 shows supportive attitudes towards e-voting are not significantly associated with repeated consecutive e-voting.

Table D.1: Factors associated with e-voting in given elections. Exponentiated generalized SEM coefficients[a]

	2009	2011	2013	2014	2015
			$\mathbf{E-vote_t} <-$		
E-vote$_{t-1}$	47.91***	8.04***	9.42***	50.79***	18.9***
	(25.23)	(3.96)	(3.5)	(30.27)	(8.57)
Mode$_{t-2}$: paper vote	0.98	1.09	0.31*	0.20*	1.55
(base: non-vote)	(0.58)	(0.39)	(0.15)	(0.14)	(1.11)
Mode$_{t-2}$: e-vote	1.79	7.30**	0.82	0.35	2.73
	(1.69)	(4.73)	(0.53)	(0.31)	(2.05)
PC literacy: good	45.71*	5.97**	3.58	25.07*	16.36
(base: poor)	(78.62)	(4.1)	(2.94)	(37.94)	(705.56)
PC literacy: average	33.9*	4.07*	3.73	3.23	15.68
	(58.29)	(2.61)	(2.85)	(4.64)	(705.56)
PC literacy: basic	23.13	2.51	0.81	3.49	15.26
	(40.39)	(1.66)	(0.8)	(5.1)	(705.56)

Continued on next page

Table D.1 continued from previous page

	2009	2011	2013	2014	2015
Education: higher	0.25	0.51	2.51	1.28	1.14
(base: primary)	(0.25)	(0.31)	(2.0)	(0.72)	(0.4)
Education: secondary	0.16	0.24*	0.96	1.88	0.98
	(0.16)	(0.14)	(0.74)	(1.14)	(0.42)
Age	1.1	0.99	0.98	1.02	0.98
	(0.09)	(0.06)	(0.06)	(0.1)	(0.06)
Age^2	1.0	1.0	1.0	1.0	1.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Income decile	1.08	1.14**	1.01	1.05	0.91
	(0.07)	(0.05)	(0.06)	(0.09)	(0.06)
Constant	0.01	0.66	0.19	0.01	0.0
	(0.03)	(0.91)	(0.33)	(0.03)	(0.0)

Continued on next page

Table D.1 continued from previous page

	2009	2011	2013	2014	2015
			$\mathbf{E - vote_{t-1}} < -$		
$\text{Mode}_{\text{t-2}}$: paper vote	0.74	0.47^{*}	0.48^{*}	0.22^{***}	0.2^{***}
(base: non-vote)	(0.23)	(0.16)	(0.16)	(0.07)	(0.08)
$\text{Mode}_{\text{t-2}}$: e-vote	25.68^{***}	110.07^{***}	38.61^{***}	15.95^{***}	35.65^{***}
	(14.13)	(44.61)	(21.47)	(6.44)	(18.46)
Constant	0.46	0.12	0.28^{***}	0.31^{***}	0.32^{**}
	(0.13)	(0.03)	(0.08)	(0.09)	(0.12)
AIC	805.41	770.93	764.14	538.04	641.38
BIC	868.42	840.11	830.04	604.17	707.34
Pseudo R^2	0.57	0.56	0.44	0.62	0.47
Observations	493	744	598	607	600

Odds ratios with standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$,

*** $p < 0.001$

[a]Sub-index legend: given election is t, previous election $t - 1$,

2^{nd} to previous election $t - 2$

Table D.2: Direct and indirect effects of the second to previous voting mode on e-voting in given election

	2009	2011	2013	2014	2015
			E-vote vs. paper vote		
Path a	1.45	5.77^{**}	2.5	1.41	1.72
	(0.88)	(2.43)	(1.37)	(1.22)	(0.97)
Path $b \times c$	8.93^{***}	5.83^{***}	5.19^{***}	13.48^{***}	8.78^{***}
	(4.11)	(2.43)	(1.46)	(7.51)	(2.18)
Total effect	12.99^{***}	33.62^{***}	12.95^{***}	19.06^{***}	15.13^{***}
(path $a \times b \times c$)	(8.78)	(17.76)	(7.07)	(13.22)	(6.77)
Observations	373	531	447	315	451

The table reports odds ratios with bootstrapped standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D.3: Effects on consecutive e-voting

	Consecutive e-voting
Trust in e-voting	0.21**
	(0.06)
Trust in internet transactions	−0.05
	(0.04)
Confident own e-vote	−0.03
was counted as cast	(0.2)
Confident other e-votes	−0.13
were counted as cast	(0.09)
Controls included	
Sensitivity	99.8%
Specificity	8.8%
Observations	601
Nagelkerke pseudo-R^2	0.07

The table reports average marginal effects with standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix E

Materials on verification

E.1 Variables and models

Trust towards e-voting was measured with a question on whether the respondent trusted online voting using a scale of 0-10, where 0 meant not trust at all and 10 meant complete trust. Usage was captured with a dummy variable coded as 1 when the voter verified their e-vote and 0 when they knew about it but did not use it. Knowledge was measured with a variable coded as 1 when the responded knew that the verification option was available in the given election and 0 when they did not know. All other variables used in the models outlined below are the same as already used and explained in previous chapters.

The association between trust towards e-voting and usage of verification was examined with a linear regression model taking the following form:

$$trust = \beta_0 + \beta_1 used + \beta_i controls \tag{E.1}$$

where *trust* was the outcome variable and the dummy on usage, *used*, our main variable of interest and *controls* a vector of socio-demographic controls that were age, education, income, gender and ethnicity.

The association between trust towards e-voting and knowledge of the possibility to verify one's vote is examined with a linear regression model taking the following form:

$$trust = \beta_0 + \beta_1 knows + \beta_i controls \quad \text{(E.2)}$$

where *trust* was the outcome variable and the dummy on knowledge, *knows*, our main variable of interest and *controls* is a vector of socio-demographic controls that were age, education, income, gender and ethnicity. The results of the models are shown in Table E.2.

E.2 Additional information on models and results

Mean trust in e-voting on a 0 to 10 scale among e-vote verifiers before the election was 9.21 (standard deviation=0.83) and after the election 9.29 (standard deviation=0.75). We conducted a paired t-test of the mean difference and the results are shown in Table E.1.

Table E.1: Paired t-test of trust in e-voting before and after an election among e-vote verifiers

	t	df	p
Trust e-voting	-0.49	23	0.627

Table E.2 displays six OLS regressions where the outcome variable is trust towards e-voting measured on a 0-10 scale, where 0 means the respondent did not trust e-voting at all and 10 meant the respondent trusted e-voting completely.

Table E.3 displays seven OLS regressions where the outcome variable is change in trust before and after an election towards the given institution on a scale of -10 to 10, with the input variables change in trust levels for all the other given institutions. Regression diagnostics did not show any models to suffer from multicollinearity. Three things are worth pointing out: 1) only change in one trust item, trust towards the president, was significantly associated with change in trust towards e-voting; 2) change in trust towards e-voting itself also only predicted change in trust towards the president and no other outcomes; 3) the model fit shown by R^2 for the model with a change in trust towards e-voting as the outcome is lower by a factor of four to ten in comparison to the other models. This means the change in trust levels towards e-voting were indeed independent from other trust items.

Table E.2: Effects of usage and knowledge of verification on trust towards e-voting

	2013	2014	2015
		Using verification	
Used verification	0.34	−0.80	−0.38
	(0.32)	(0.54)	(0.41)
Age	0.01	0.02	−0.01
	(0.01)	(0.01)	(0.01)
Education: higher	−1.13	−0.41	0.91
(*base: primary*)	(1.34)	(0.79)	(0.66)
Education: secondary	−1.37	−1.84	0.65
	(1.33)	(0.82)	(0.67)
Income decile	0.01	0.02	0.04
	(0.06)	(0.08)	(0.07)
Ethnic Estonian	0.56	1.36^{*}	0.52
	(0.76)	(0.57)	(0.66)
Male	−0.03	0.25	−0.40
	(0.27)	(0.46)	(0.35)
Constant	9.15^{***}	7.70^{***}	7.66^{***}
	(1.56)	(1.16)	(1.21)
Observations	98	56	106
R^2	0.04	0.31	0.06

Continued on next page

Table E.2 continued from previous page

	2013	2014	2015
	Knowing about verification		
Knows verification	2.72***	2.40***	2.87***
	(0.42)	(0.50)	(0.39)
Age	−0.05***	−0.03	−0.06
	(0.01)	(0.01)	(0.01)
Education: higher	2.38***	1.57**	0.67
(*base: primary*)	(0.48)	(0.51)	(0.48)
Education: secondary	0.99*	0.55	0.16
	(0.41)	(0.44)	(0.43)
Income decile	0.05	0.07	0.11
	(0.06)	(0.06)	(0.06)
Ethnic Estonian	2.41***	1.86***	1.89***
	(0.0.32)	(0.38)	(0.36)
Male	−0.58*	−0.37	−0.39
	(0.29)	(0.33)	(0.31)
Constant	4.38***	4.91***	5.78***
	(0.74)	(0.83)	(0.78)
Observations	446	331	424
R^2	0.42	0.23	0.27

The table reports OLS coefficients with standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table E.3: Associations between trust items

	Trust in:						
	e-voting	government	parliament	president	politicians	parties	courts
Trust in e-voting		−0.02	0.01	−0.08**	−0.03	−0.04	−0.02
		(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Trust in government	−0.04		0.27***	0.22***	0.00	0.07*	0.07
	(0.05)		(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Trust in parliament	0.03	0.35***		0.20***	0.13***	0.06	0.20***
	(0.06)	(0.04)		(0.04)	(0.04)	(0.03)	(0.04)
Trust in president	−0.17**	0.26***	0.19***		0.12**	−0.03	0.07
	(0.06)	(0.04)	(0.04)		(0.04)	(0.04)	(0.04)
Trust in politicians	−0.07	0.00	0.12***	0.12**		0.39***	−0.05
	(0.06)	(0.04)	(0.03)	(0.03)		(0.03)	(0.04)
Trust in parties	−0.09	0.09*	0.07	−0.04	0.45***		0.19***
	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)		(0.04)
Trust in courts	−0.04	0.07	0.15***	0.06	−0.04	0.14***	
	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	
Constant	0.03	0.16**	0.07	−0.14*	0.00	−0.21***	−0.20***
	(0.08)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)
Observations	705	705	705	705	705	705	705
R^2	0.03	0.26	0.29	0.22	0.26	0.26	0.12

The table OLS coefficients with standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix F

Materials on political neutrality

F.1 Mobilization model

Mobilization associations with party choice was examined based on survey data from three parliamentary elections: 2007, 2011 and 2015. We estimated four separate logit regression models for each of the three elections.

Our outcome variables were discrete choices for the four largest parties in parliament during these three elections: respondents who voted for the Centre Party were coded 1 and the other parties 0 for the first dummy; respondents who voted for the Reform Party were coded 1 and the other parties 0 for the second dummy; respondents who voted for the Social Democrats were coded 1 and the other parties 0 for the third dummy; respondents who voted for the Pro Patria and Res Publica Union were coded 1 with the other parties 0 for the fourth dummy. In addition to our mobilization-dummy, we also included socio-demographic and attitudinal controls, such as age (ranging from 18 to 96), income decile (1 to 10) and left-right position (continuous from 0 to 10 where 0 meant left and 10 right). We had to keep the models concise, as they were only estimated on e-voters in a given election who were comprised of regular e-voters and new voters mobilized from not voting to e-voting for the current election. The generic form of our model took the following form:

$$ln \left\{ \frac{Pr(y=1)}{1-Pr(y=1)} \right\} = \beta_0 + \beta_1 mobiliz. + \gamma + \epsilon \tag{F.1}$$

The parameter of interest was β_1, which, if significant to at least at ¡0.05 level, would tell us that new voters were structurally mobilized from the supporters of these political parties for which the parameter was significant. The findings are presented in Table F.1.

The table shows average marginal effects with standard errors in parentheses. The main variable of interest is shown in the row mobilized by e-voting and the numbers show whether the person who was mobilized into voting via e-voting was more likely to choose a specific party over others, so 0.03 in the first row of the first column shows us that a mobilized voter would be 3% points more likely to vote for the Reform Party than any of the other three parties, but this effect was not statistically significant. In none of the years and for no party does the association reach any acceptable significance level, so we can only conclude that new e-voters being mobilized to vote did not support one specific party in overwhelming numbers.

Table F.1: Party choice and mobilization in national elections via e-voting

	Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats
		2007		
Mobilized by e-voting	0.03	0.09	−0.05	−0.11
	(0.12)	(0.06)	(0.11)	(0.07)
Age	−0.00	0.00	−0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Income decile	−0.02	−0.01	0.01	0.01
	(0.02)	(0.01)	(0.02)	(0.01)
Left-right position	0.03	−0.03*	0.05	−0.04

Continued on next page

Table F.1 continued from previous page

	Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats
	(0.02)	(0.02)	(0.03)	(0.02)
Constant	−0.62	−0.67	−2.19	2.28
	(1.29)	(2.65)	(1.38)	(2.66)
Observations	74	74	74	74
Pseudo R^2	0.04	0.28	0.07	0.20
		2011		
Mobilized by e-voting	0.06	0.06	−0.20	0.04
	(0.11)	(0.05)	(0.13)	(0.08)
Age	−0.01	−0.00	0.01	0.00
	(0.03)	(0.00)	(0.00)	(0.00)
Income decile	0.00	0.00	0.00	−0.01
	(0.01)	(0.00)	(0.01)	(0.01)
Left-right position	0.06^{***}	-0.02^{*}	0.02	−0.05
	(0.02)	(0.01)	(0.02)	(0.01)
Constant	−1.22	−0.36	-2.33^{*}	0.50^{***}
	(1.01)	(1.16)	(1.01)	(1.13)
Observations	123	123	123	123
Pseudo R^2	0.17	0.21	0.09	0.20
		2015		
Mobilized by e-voting	0.00	0.08	0.02	−0.10
	(0.18)	(0.05)	(0.17)	(0.16)
Age	−0.00	0.00	−0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Income decile	−0.03	−0.01	−0.00	0.04
	(0.02)	(0.02)	(0.02)	(0.02)
Left-right position	0.09^{**}	-0.05^{*}	0.03	-0.05^{**}
	(0.03)	(0.03)	(0.03)	(0.03)
Constant	−2.14	−1.02	-0.52^{*}	0.47^{***}
	(1.55)	(2.14)	(1.61)	(1.91)
Observations	59	59	59	59

Continued on next page

Table F.1 continued from previous page

	Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats
Pseudo R^2	0.17	0.20	0.03	0.29

The table reports marginal effects; standard errors are in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

F.2 Model of mode-specific bias

We investigated the possible effect of e-voting on party choice using a logit model, where the outcome was the self-reported voting choice for a particular party (coded 1) and the reference being a vote for any of the remaining three parties (coded 0). Such a model was estimated for each of the four parties separately for each of the elections under study. The main predictor was a dichotomy between e-voters (coded 1) and paper voters (coded 0). Because we were interested in the effect of e-voting while controlling for socio-demographics and other related covariates, we included a set of traditional determinants of party choice (age, gender, education, ethnicity, place of residence, income, left-right-placement and computer literacy). A generic form for our model took the following form:

$$ln\left\{\frac{Pr(y=1)}{1-Pr(y=1)}\right\} = \beta_0 + \beta_1 evoting + \gamma + \epsilon \tag{F.2}$$

where the primary parameter of interest is β_1, while controlling for the vector of socio-demographic, attitudinal and behavioural covariates (γ). If our primary parameter of interest was statistically significant, e-voting was associated with voting for this particular party. Results are displayed in Tables F.2, F.3 and F.4, where the columns contain four separate models estimated for each of the parties.

We saw that there was a positive association with e-voting and choosing the Reform Party in 2015 and with choosing the Pro Patria and Res Publica

Union in 2007. There was a negative association with e-voting and choosing the Centre Party in 2011 and 2015.

Table F.2: Party choice and e-voting in the 2007 national elections

	Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats
E-voted	-0.09^{*}	-0.04	0.10^{**}	0.04
	(0.04)	(0.03)	(0.04)	(0.03)
Age	-0.00	0.00^{**}	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Male	-0.07	0.04	0.04	-0.01
	(0.04)	(0.03)	(0.04)	(0.03)
Ethnic Estonian	0.61^{***}	-0.43^{**}	0.27^{*}	0.28^{*}
	(0.15)	(0.05)	(0.12)	(0.12)
Higher education	-0.06	-0.06	0.06	0.04
	(0.04)	(0.03)	(0.04)	(0.03)
PC literacy: good	0.06	-0.07	-0.01	0.01
	(0.05)	(0.04)	(0.04)	(0.04)
Income decile	0.02^{**}	-0.02^{**}	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Left-right position	0.02^{*}	-0.02^{***}	0.03^{***}	-0.03^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
Constant	-3.38^{*}	2.99^{***}	-4.45^{***}	-2.35
	(0.85)	(0.82)	(0.91)	(1.16)
Observations	570	570	570	570
Pseudo R^2	0.15	0.46	0.14	0.13

The table reports marginal effects; standard errors are in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

After preprocessing the data we continue by specifying exactly the same model as that in Equation F.2, only now we estimated the model on a balanced sample and thus, our results provide a basis for causal interpretation. The findings are reported in Table F.5. Because the observed covariates were

Table F.3: Party choice and e-voting in the 2011 national elections

	Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats
E-voted	0.06	−0.06*	−0.01	0.01
	(0.05)	(0.03)	(0.05)	(0.04)
Age	−0.00	0.00	0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Male	−0.06	0.02	0.05	−0.01
	(0.04)	(0.02)	(0.04)	(0.03)
Ethnic Estonian	0.72***	−0.32***	0.47**	0.41***
	(0.20)	(0.04)	(0.18)	(0.11)
Higher education	−0.11*	−0.02	0.14***	−0.02
	(0.04)	(0.03)	(0.04)	(0.04)
PC literacy: good	−0.03	0.00	0.01	0.02
	(0.05)	(0.03)	(0.05)	(0.05)
Income decile	0.00	0.00	0.00	−0.01
	(0.01)	(0.00)	(0.01)	(0.01)
Left-right position	0.04***	−0.01**	0.02*	−0.04***
	(0.01)	(0.00)	(0.01)	(0.01)
Constant	−4.29***	2.31**	−5.15***	−1.87*
	(1.12)	(1.04)	(1.14)	(0.89)
Observations	510	570	570	570
Pseudo R^2	0.19	0.61	0.14	0.15

The table reports marginal effects; standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

used as a basis of matching, they are not directly interpretable with respect to the outcome of interest. We have included them into the model as convention prescribes, but refrained from reporting them in the results table.

We made use of genetic matching that uses a search algorithm to iteratively check and improve covariate balance, and returned generalized propensity and Mahalanobis Distance matching scores (Diamond and Sekhon,

Table F.4: Party choice and e-voting in the 2015 national elections

	Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats
E-voted	0.09**	−0.18***	−0.02	0.05
	(0.04)	(0.05)	(0.04)	(0.04)
Age	−0.00***	0.01***	−0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Male	−0.05*	0.03	0.04	−0.01
	(0.03)	(0.03)	(0.03)	(0.03)
Ethnic Estonian	0.54***	−0.47***	0.32***	0.17***
	(0.11)	(0.02)	(0.08)	(0.05)
Higher education	−0.05	−0.05	0.06	0.04
	(0.04)	(0.04)	(0.03)	(0.03)
PC literacy: good	−0.00	0.06	0.03	−0.10*
	(0.04)	(0.04)	(0.04)	(0.04)
Income decile	−0.00	−0.00	−0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Left-right position	0.03***	−0.02***	0.01	−0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
Constant	−4.07***	0.77***	−4.50***	−1.17*
	(0.93)	(0.63)	(0.85)	(0.59)
Observations	592	592	592	592
Pseudo R^2	0.31	0.52	0.14	0.07

The table reports marginal effects; standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2005). In essence, it is a multivariate matching method that uses an evolutionary search algorithm developed by Sekhon and Mebane (1998) to maximize the balance of observed covariates across matched treated and control units (Diamond and Sekhon, 2005). Relevant observed variables used for matching were age, gender, education, urban residence, ethnicity, income, left-right self-position and computer literacy.

Applying the matching technique showed e-voting to indeed be positively associated with voting for the Reform Party in 2015 and negatively associated with choosing the Centre Party in 2011 and 2015. The strength of the negative association in 2015 was clearly weakened by using the more sophisticated matching technique. All-in-all, we saw therefore some support for a mode specific bias of e-voting, but it was not consistent over the years and did not exist to a degree that could explain the large differences in aggregate level election results.

Table F.5: Causal effect of e-voting on party choice in 2007, 2011 and 2015

	Reform Party	Centre Party	Pro Patria & Res Publica	Social Democrats
		2007		
Mobilized by e-voting	−0.10	−0.02	0.10	0.04
	(0.05)	(0.03)	(0.05)	(0.04)
		Controls included		
Constant	−0.03	3.26*	−2.55*	−0.84
	(0.57)	(1.53)	(0.65)	(1.34)
Observations	405	413	405	413
		2011		
Mobilized by e-voting	0.07	−0.06*	−0.00	0.03
	(0.06)	(0.02)	(0.06)	(0.05)
		Controls included		
Constant	−0.50	−4.55	−2.27***	0.55
	(0.67)	(1.59)	(0.71)	(0.78)
Observations	327	327	327	327
		2015		
Mobilized by e-voting	0.11*	−0.14**	0.01	0.02
	(0.06)	(0.04)	(0.05)	(0.06)
		Controls included		
Constant	−4.51***	0.95	−3.92**	0.21
	(1.38)	(1.12)	(1.27)	(0.95)
Observations	229	229	229	229

The table reports marginal effects; standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Bibliography

Abrahamson, E. (1991). Managerial fads and fashions: The diffusion and rejection of innovations. *Academy of Management Review 16*(3), 586–612.

Aldrich, J. H., J. M. Montgomery, and W. Wood (2011). Turnout as a habit. *Political Behavior 33*, 535–563.

Alvarez, R. and T. Hall (2008). *Electronic elections: The perils and promises of digital democracy*. Princeton University Press.

Alvarez, R., T. Hall, and A. Trechsel (2009). Internet voting in comparative perspective: The case of estonia. *PS: Political Science and Politics 42*(3), 497–505.

Alvarez, R. and J. Nagler (2000). Likely consequences of internet voting for political representation. *The Loyola Los Angeles Law Review 34*, 1115–1153.

Alvarez, R. M. and T. E. Hall (2004). *Point, Click, and Vote: The Future of Internet Voting*. Brookings Institution Press.

Alvarez, R. M. and T. E. Hall (2006). Controlling democracy: the principal–agent problems in election administration. *Policy Studies Journal 34*(4), 491–510.

Alvarez, R. M., T. E. Hall, and M. H. Llewellyn (2008). Are americans confident their ballots are counted? *The Journal of Politics 70*(03), 754–766.

Ansper, A., S. Heiberg, H. Lipmaa, T. Øverland, and F. Van Laenen (2009). Security and trust for the norwegian e-voting pilot project e-valg 2011. In M. Kripp, M. Volkamer, and R. Grimm (Eds.), *Identity and Privacy in the Internet Age*, pp. 207–222. Springer.

Bachmann, R. and A. Zaheer (2006). *Handbook of trust research*. Edward Elgar Cheltenham, UK.

Barrat, J., B. Goldsmith, and J. Turner (2012). *International Experience with E-Voting*. Stockholm: IFES Foundation.

Berinsky, A. J. (2005). The perverse consequences of electoral reform in the united states. *American Politics Research 33*(4), 471–491.

Berry, F. S. and W. D. Berry (1990). State lottery adoptions as policy innovations: An event history analysis. *American political science review 84*(02), 395–415.

Braun, N. (2006). *Stimmgeheimnis: Eine rechtsvergleichende und rechtshistorische Untersuchung unter Einbezug des geltenden Rechts*. Stämpfli Verlag.

Buis, M. (2010). Direct and indirect effects in a logit model. *Stata Journal 10*(1), 11–29.

Cameron, A. and P. Trivedi (2010). *Microeconomics Using Stata*. Stata Press.

Carver, R. P. (1990). *Reading Rate: A Review of Research and Theory*. Academic Pr.

Constitutional Committee (2011). *Minutes of the session of 9 June 2011 of the Constitutional Committee of the Estonian Parliament*. Estonian Parliament. `http://www.riigikogu.ee/download/6ca49b18-e6ec-b9f7-8e88-826d5d8ef5a4`.

Council of Europe (2004). *Recommendation Rec (2004) 11 "Legal, Operational and Technical Standards for I-voting" of the Council of Europe*. Council of Europe. `http://www.coe.int/t/dgap/democracy/activities/ggis/evoting/key_documents/Rec(2004)11_Eng_Evoting_and_Expl_Memo_en.pdf`.

Cutts, D., E. Fieldhouse, and P. John (2009). Is voting habit forming? the longitudinal impact of a gotv campaign in the uk. *Journal of Elections, Public Opinion and Parties 19*, 251–263.

Denny, K. and O. Doyle (2009). Does voting history matter? analyising persistence in turnout. *American Journal of Political Science 53*(1), 17–35.

Diamond, A. and J. S. Sekhon (2005). Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics*.

Dinas, E. (2012). The formation of voting habits. *Journal of Elections, Public Opinion and Parties 22*(4), 431–456.

Downs, A. (1957). *An Economic Theory of Democracy. New York: Harper.* New York: Harper.

Drechsler, W. and V. Kostakis (2014). Online-wahlen im internationalen vergleich. *Aus Politik und Zeitgeschichte 18*, 16–23.

Drechsler, W. and V. Kostakis (2015). Should law keep pace with technology? law as katechon. *Bulletin of Science, Technology and Society*.

Driza Maurer, A., O. Spycher, G. Taglioni, and A. Weber (2012). E-voting for swiss abroad: A joint project between the confederation and the cantons. In M. Kripp, M. Volkamer, and R. Grimm (Eds.), *Proceedings of the 5th International Conference on Electronic Voting (EVOTE2012)*, pp. 173–187. Bonn: Bonn: GI.

Electronic Voting Committee (2013). *The Source Code of the Estonian I-voting Solution*. Electronic Voting Committee. `https://github.com/vvk-ehk/evalimine`.

Electronic Voting Committee (2015). *The Electronic Voting Committee Call for Proposals for Amending the Internet Voting Solution by 2017 Elections*. Electronic Voting Committee. `http://www.vvk.ee/valimistekorraldamine/vvk-uudised/kutse-ideepaev-e-haaletamise-parenduseks/`.

Estonian Tax and Customs Board (2015). *Electronic Filing of Income Tax Returns is now Open*. Estonian Tax and Customs Board. `http://www.emta.ee/index.php?id=36597&tpl=1026`.

Faraon, M., G. Stenberg, J. Budurushi, and M. Kaipainen (2015). Positive but skeptical: A study of attitudes towards internet voting in sweden. In R. Krimmer and M. Volkamerm (Eds.), *CeDEM Asia 2014: Proceedings of the International Conference for E-Democracy and Open Government*, pp. 191–205. Krems: Edition Donau-Uni Krems.

Fliegel, F. (1993). *Diffusion Research in Rural Sociology: The Record for the Future.* Westport, CT: Greenwood.

Franklin, M. (2004). *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies Since 1945*. Cambridge University Press.

Gallego, A. (2010). Understanding unequal turnout: Education and voting in comparative perspective. *Electoral Studies 29*(2), 239–248.

Gerber, A., D. Green, and R. Shachar (2003). Voting may be habit-forming: Evidence from a randomized field experiment. *American Journal of Political Science 47*(3), 540–550.

Gerlach, J. and U. Gasser (2009). Three case studies from switzerland: E-voting. Technical report. Available at: `http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/Gerlach-Gasser_SwissCases_Evoting.pdf`.

Green, D. and R. Shachar (2000). Habit formation and political behaviour: Evidence of consuetude in voter turnout. *British Journal of Political Science 30*(4), 561–573.

Greenhalgh, T., G. Robert, F. Macfarlane, P. Bate, and O. Kyriakidou (2004). Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Quarterly 82*(4), 581–629.

Gritzalis, D. (2003). *Secure Electronic Voting*. Springer US.

Hedström, P. (1994). Contagious collectivities: On the spatial diffusion of swedish trade unions, 1890-1940. *American Journal of Sociology*, 1157–1179.

Heiberg, S., P. Laud, and J. Willemson (2012). The application of i-voting for estonian parliamentary elections of 2011. In A. Kiayias and H. Lipmaa (Eds.), *E-Voting and Identity: Third International Conference*, pp. 208–223. Berlin/Heidelberg: Springer.

Heiberg, S., A. Parsovs, and J. Willemson (2015). Log analysis of estonian internet voting 2013–2014.

Heiberg, S. and J. Willemson (2014). Verifiable internet voting in estonia. In R. Krimmer and M. Volkamerm (Eds.), *Proceedings of Electronic Voting 2014 (EVOTE2014)*, pp. 7–13. Tallinn: Tallinn: TUT Press.

Heindl, P., A. Prosser, and R. Krimmer (2003). Constitutional and technical requirements for democracy over the internet: E-democracy. In R. Traunmüller (Ed.), *Electronic Govenment*, pp. 417–420. Springer-Verlag.

Heinsalu, A., A. Koitmäe, M. Pilving, and P. Vinkel (2012). *Elections in Estonia 1992-2011*. National Electoral Committee.

ICCPR (1976). *United Nations International Covenant on Civil and Political Rights*. United Nations.

Jahn, D. (2006). Globalization as 'galton's problem': The missing link in the analysis of diffusion patterns in welfare state development. *International organization 60*(2), 401–431.

Kalkun, M. and T. Kalvet (2002). *Digitaalne lõhe Eestis ja selle ületamise võimalused*. Praxis.

Kalvet, T. (2012). Innovation: A factor explaining e-government success in estonia. *Electronic Government 9*(2), 142–157.

Kersting, N. (2004). Briefwahl im internationalen vergleich. *Österreichische Zeitschrift für Politikwissenschaft 33*(3), 341–351.

Krimmer, R. (2012). *The Evolution of E-Voting: Why Voting Technology is Used and How it Affects Democracy*. Tallinn: TUT Press.

Krimmer, R. and M. Kripp (2009). The use of electronic voting around the world. *Modern Democracy 2*, 8–9.

Lassen, D. D. (2005). The effect of information on voter turnout: Evidence from a natural experiment. *American Journal of Political Science 49*(1), 103–118.

Lazarsfeld, P., B. Berelson, and H. Gaudet (1944). *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. Columbia University Press.

Lenarčič, J. (2011). *Address by Ambassador Janez Lenarčič, Former Director of the OSCE/ODIHR, at the OSCE Chairmanship Expert Seminar on the "Present State and Prospects of Application of Electronic Voting in the OSCE Participating States", in Vienna, Austria on 16 September 2010*. OSCE.

LGCEA (2002). *Local Government Council Election Act*. passed on 27 March 2002.

LGCEA (2015). *Amendments to the Local Government Council Election Act*. passed on 28 June 2005.

Maaten, E. and T. Hall (2008). Improving the transparency of remote i-voting: The estonian experience. In R. Krimmer and R. Grimm (Eds.), *Electronic Voting 2008*. Bonn: Gesellschaft für Informatik.

MacVaugh, J. and F. Schiavone (2010). Limits to the diffusion of innovation: A literature review and integrative model. *European Journal of Innovation Management 13*(2), 197–221.

Margolis, M. and D. Resnick (2000). *Politics as usual: the cyberspace" revolution"*. Sage Publications, Inc.

Markussen, R., L. Ronquillo, and C. Schürmann (2014). Trust in internet election: Observing the norwegian decryption and counting ceremony. In

M. Kripp, M. Volkamer, and R. Grimm (Eds.), *Proceedings of Electronic Voting 2014 (EVOTE2014)*, pp. 173–187. Tallinn: Tallinn: TUT Press.

Matsusaka, J. and F. Palda (1999). Voter turnout: How much can we explain? *Public Choice 98*(3-4), 431–446.

Meagher, S. (2008). When personal computers are transformed into ballot boxes: How internet elections in estonia comply with the united nations international covenant on civil and political rights. *American University International Law Review 23*.

Meri, L. (2000). *Speech at the University of St Olaf Minnesota, U.S., 6 April 2000*.

Mitrou, L. and S. Gritzalis, D. Katsikas (2002). Revisiting legal and regulatory requirements for secure e-voting. In M. Ghonaimy, A. El-Hadidi, T. Mahmoud, and K. Heba (Eds.), *Security in the Information Society*. Springer.

Möllering, G. (2006). *Security in the Information Society*, Chapter Trust, institutions, agency: towards a neoinstitutional theory of trust, pp. 355–376. Edward Elgar Cheltenham, UK.

National Election Committee (2015). *Internet Voting Procedural Documents on the Estonian National Electoral Committee Webpage*. National Election Committee. `http://www.vvk.ee/valijale/e-haaletamine/edokumendid/`.

Nickerson, D. (2008). Is voting contagious? evidence from two field experiments. *American Political Science Review 102*(1), 49–57.

Norris, P. (2001). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge University Press.

Norris, P. (2003). Will new technology boost turnout? evaluating experiments in e-voting v. all-postal voting facilities in uk local elections. Technical report, Paper for the British Study Group Seminar Friday 31st October 2-4pm, Minda de Gunzberg Center for European Studies, Harvard University.

OSCE/ODIHR (2007). *SCE/ODIHR Election Assessment Mission Report. Estonia. Parliamentary Elections 4 March 2007*. OSCE. http://www.osce.org/odihr/elections/estonia/25925.

OSCE/ODIHR (2011). *OSCE/ODIHR Election Assessment Mission Report*. OSCE. http://www.osce.org/odihr/77557.

OSCE/ODIHR (2012). *OSCE/ODIHR Election Assessment Mission Report. Swiss Confederation. Federal Assembly Elections 23 October 2011*. OSCE.

OSCE/ODIHR (2013). *Handbook for the Observation of New Voting Technologies*. OSCE.

OSCE/ODIHR (2015). *OSCE/ODIHR Final Report on Estonian Parliamentary Elections 1 March 2015*. OSCE. http://www.osce.org/odihr/elections/estonia/160131.

Peixoto, T. (2009). Beyond theory: e-participatory budgeting and its promises for eparticipation. *European Journal of ePractice 7*(5), 1–9.

Penal Code (2001). *Chapter 19 Division 2 in the Penal Code*. Penal Code. https://www.riigiteataja.ee/en/eli/519032015003/consolide.

Plutzer, E. (2002). Becoming a habitual voter: Inertia, resources, and growth in young adulthood. *American Political Science Review 96*(1), 41–56.

Popoveniuc, S., J. Kelsey, A. Regenscheid, and P. Vora (2010). *Proceedings of the 2010 international conference on Electronic voting technology/workshop on trustworthy elections*, Chapter Performance requirements for end-to-end verifiable elections. EVT/WOTE'10.

Putnam, R. (2001). *Bowling alone: The collapse and revival of American community*. Simon and Schuster.

Rüß, O. (2000). Wahlen im internet: Wahlrechtsgrundsätze und einsatz von digitalen signaturen. *Multimedia und Recht 3*(2), 73–76.

Raudla, R. and J. Krenjova (2013). Participatory budgeting at the local level: Challenges and opportunities for new democracies. *European Journal of ePractice 14*(1), 18–46.

Rogers, Everett, M. (2003 [1962]). *Diffusion of Innovations. 5th edition.* Free Press, New York.

Ryan, B. and N. C. Gross (1943). The diffusion of hybrid seed corn in two iowa communities. *Rural sociology 8*(1), 15–24.

Schweizer Bundesrat (2002). *Bericht über den Vote électronique. Chancen, Risiken und Machbarkeit elektronischer Ausübung politischer Rechte*. Bundesblatt, Bern.

Schweizer Bundesrat (2006). *Bericht über die Pilotprojekte zum Vote électronique*. Bundesblatt, Bern.

Schweizer Bundesrat (2013). *Bericht des Bundesrates zu Vote électronique. Auswertung der Einführung von Vote électronique (2006–2012) und Grundlagen zur Weiterentwicklung*. Bundesblatt, Bern.

Sekhon, J. S. and W. R. Mebane (1998). Genetic optimization using derivatives. *Review of Economics and Statistics 7*(1), 187–210.

Serdült, U., M. Germann, F. Mendez, A. Portenier, and C. Wellig (2015). E-voting for swiss abroad: A joint project between the confederation and the cantons. In L. Teran and A. Meier (Eds.), *Proceedings of the Second International Conference on eDemocracy and eGovernment (ICEDEG)*, pp. 126–132. Quito: IEEE.

Smets, K. and C. van Ham (2013). The embarrassment of riches? a meta-analysis of individual-level research on voter turnout. *Electoral Studies 32*(2), 344–359.

Solvak, M. (2011). Parteivaliku vahetamine 2011. aasta valimistel. In R. Toomla (Ed.), *Riigikogu valimised 2011*, pp. 116–152. Tartu University Press.

Springall, D., T. Finkenauer, Z. Durumeric, J. Kitcat, H. Hursti, M. M., and J. A. Halderman (2014). Security analysis of the estonian internet voting system. In G.-J. Ahn (Ed.), *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pp. 703–715. New York: ACM.

Spycher, O., M. Volkamer, and R. Koenig (2012). Transparency and technical measures to establish trust in norwegian internet voting. In A. Kiayias and H. Lipmaa (Eds.), *The Application of I-voting for Estonian Parliamentary Elections of 2011*, pp. 19–35. Berlin/Heidelberg: Springer.

Starr, H. (1991). Democratic dominoes diffusion approaches to the spread of democracy in the international system. *Journal of Conflict Resolution 35*(2), 356–381.

Stein, R. and G. Wenda (2014). The council of europe and e-voting: History and impact of rec(2004)11. In R. Krimmer and M. Volkamerm (Eds.), *Proceedings of Electronic Voting 2014 (EVOTE2014)*, pp. 1–6. Tallinn: Tallinn: TUT Press.

Stenerud, I. and C. Bull (2012). When reality comes knocking: Norwegian experiences with verifiable electronic voting. In M. Kripp, M. Volkamer, and R. Grimm (Eds.), *Proceedings of the 5th International Conference on Electronic Voting (EVOTE2012)*, pp. 21–33. Bonn: GI.

Stuart, E. A. and D. B. Rubin (2008). *Best practices in quasi-experimental designs. Best practices in quantitative methods*.

Supreme Court (2015). *Judgment Number 3-4-1-13-05 of the Constitutional Review Chamber of the Supreme Court*.

Tolnay, S. E. and P. J. Glynn (1994). The persistence of high fertility in the american south on the eve of the baby boom. *Demography 31*(4), 615–631.

Trechsel, A. H. and K. Vassil (2011). Internet voting in estonia. a comparative analysis of five elections since 2005. report to the estonian national electoral committee. Technical report. Available at: http://www.vvk.ee/valijale/e-haaletamine/raportid/.

van der Eijk, C. and E. Oppenhuis (1991). European parties performance in electoral competition. *European Journal of Political Research 19*(1), 55–80.

van Dijk, J. (2000). *Digital democracy: Issues of theory and practice*, Chapter Widening information gaps and policies of prevention., pp. 166–183. Sage.

van Dijk, J. (2005). *The deepening divide: Inequality in the information society*. Sage.

Vassil, K., M. Solvak, P. Vinkel, A. Trechsel, R. Alvarez, and T. Hall (2014). Diffusion of internet voting: Usage patterns of internet voting in estonia between 2005-2013. *Paper presented at the 72nd MPSA Annual Conference, 3-6 April, 2014, Chicago*.

Vassil, K. and T. Weber (2011). A bottleneck model of e-voting. why technology fails to boost turnout. *New Media and Society 13*(8), 1336–1354.

Volkamer, M., O. Spycher, and E. Dubuis (2011). Measures to establish trust in internet voting. In E. Estevez and M. Janssen (Eds.), *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance (ICEGOV '11)*, pp. 1–10. New York: ACM.

Wejnert, B. (2002). Integrating models of diffusion of innovations: A conceptual framework. *Annual review of sociology*, 297–326.

Wilhelm, A. (2000). *Democracy in the digital age: Challenges to political life in cyberspace*. Routledge.

World Economic Forum (2015). *World Economic Forum Global IT Report 2015*, Chapter Global IT Report 2015. World Economic Forum. http://www3.weforum.org/docs/WEF_Global_IT_Report_2015.pdf.

Zaller, J. (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.