The impact of remote Internet voting on reducing the cost of electoral participation[1]

May 26, 2026

Abstract

Remote Internet voting is often seen as a way to address declining turnout and low youth participation by making voting more convenient and reducing participation costs. Using micro-level data from six Estonian post-election studies conducted between 2009 and 2023, it tests whether perceived voting costs correlate with reduced participation, but increased likelihood of voting online. The findings show that higher voting costs have a negative association with participation overall but a strong positive association with the probability of choosing Internet voting. If reaching a polling station takes more than 30 minutes, voters are already more likely to vote online. This effect is strong enough to offset the barriers often associated with digital participation, such as low computer skills, lower education, and older age. The paper concludes that, while Internet voting is unlikely to substantially raise turnout, it seems to reduce voting costs enough to possibly overcome barriers that might otherwise limit participation.

Internet voting, i-voting, voting cost, turnout

1 Introduction

One of the main reasons remote Internet voting began to be discussed in the late 1990s was its expected potential to counter declining turnout by lowering participation costs and by offering a convenient voting mode for young, tech-savvy voters (Alvarez and Nagler, 2000; Norris, 2001, 2003). Evidence suggests that this has occurred in some contexts. Goodman and Stokes (2020) find positive turnout effects in Canada, while Germann (2021) shows that expatriates - who arguably face the highest voting costs - appear to vote at higher rates due to Internet voting in Switzerland. Ciancio and K¨ampfen (2023) find that the availability of this voting mode increases the turnout among older Swiss voters. Solvak and Vassil (2018) show that Internet voting appears to help Estonian voters continue to vote at higher rates than regular paper-ballot voters. Although all of these studies attribute the observed effects

[1] This work was supported by the European Union Horizon 2020 grant number 857622. Data and scripts used for the analysis in this paper will be made available in a data repository such as ICPSR or Dataverse upon manuscript acceptance.

to reduced participation costs, none test this mechanism directly. This paper does so using individual-level data to examine whether remote Internet voting is more likely among citizens who would achieve a greater relative reduction in voting costs. The cost of voting itself is measured through perceived travel time to the polling station.

Drawing on Downs's *Economic Theory of Democracy* (Downs, 1957), we provide a direct test of the mechanism assumed in the turnout studies cited above. We outline a micro-level cost-reduction mechanism and estimate how reductions in voting costs affect the choice of Internet voting over the traditional paper ballot. We also examine whether this cost differential is strong enough to outweigh factors usually associated with lower rates of online participation and technology use.

To do so, we analyze post-election survey data from Estonia, the first country in the world to hold nationwide elections in which citizens could cast binding votes online. Since 2005, fourteen Internet-voting-enabled elections have been held, with Internet voting peaking at 51.1% of all votes cast in the 2023 national elections. Internet voting has thus become something of a tradition in Estonia (Ehin et al., 2022), and usage patterns should now be stable enough to allow a closer examination of the behavioral mechanisms behind its use.

The paper proceeds as follows. Section 2 presents the theoretical argument linking costs to voting and Internet voting. Section 3 describes the data, variables, and modeling strategy. Section 4 presents the findings, section 5 concludes.

2 Model of voting cost and Internet voting usage

To evaluate remote Internet voting as a cost-reducing mode of participation, we must first specify a model of turnout. Electoral costs need to be linked to participation before their effect on the choice of voting mode can be assessed.

Reviews of the turnout literature suggest that no clear "core model" of participation has yet emerged (Geys, 2006; Smets and van Ham, 2013). Because our question is whether a reduction in costs makes a behavior more likely, we adopt a Downsian perspective as a baseline. In this framework, the decision to vote depends on the utility or attractiveness of the options on the ballot and the costs of voting (Downs, 1957, pp. 36-50). When benefits exceed costs, turnout becomes more likely; when direct "shoe-leather costs" outweigh expected benefits, abstention becomes more likely (Woller et al., 2022). It also follows that a reduction in costs should increase the probability of participation, which has been examined in detail by turnout literature (Bhatti, 2012; Karp and Banducci, 2000; Brady and McNulty, 2011; Dyck and Gimpel, 2005).

Remote Internet voting is in theory convenient, fast, and involves minimal physical effort. It should therefore appeal most to voters whose participation is sensitive to cost.
Yet voting costs are only one part of the turnout calculus. Participation also depends on the utility voters attach to the available options. Because such utilities are difficult to observe directly, researchers typically rely on voting propensities or like/dislike scores, which are

assumed to capture the attractiveness of the ballot choices (van der Eijk et al., 2006). We follow the same logic further below.

Downs also acknowledged that additional non-instrumental motives, beyond utilities and costs, may enter the calculus of participation (Downs, 1957, p. 27). The most prominent of these is civic duty (Riker and Ordeshook, 1968, p. 28), since voting remains the most common conventional form of regime-supportive participation (Dalton, 1988, p. 36). This view is supported by studies showing that voting often functions as norm-satisfying behavior among engaged and trusting citizens. However, these partly social motives do not undermine the rationality of participation itself (Blais and Young, 1999; Olson, 1965; Santana and Aguilar, 2021; Fisher and Savani, 2022).

We therefore model participation as a function of three factors: 1) the cost of voting (C), 2) the utility of the choices on the ballot (U), and 3) non-instrumental motivations to participate (D). Voting occurs when cost is outweighed by utility and duty, that is, when $C < (U + D)$. From this it follows that physical voting costs should reduce the probability of participation:

H1: Probability of participation decreases as voting cost increases

The choice of voting mode, however, should vary inversely with the costs of in-person participation. All else equal, voters facing higher physical voting costs should be more likely to choose remote Internet voting over other modes:

H2: Probability of remote Internet voting increases as voting cost increases

Note also that if the voting cost form the central mechanism behind remote Internet voting usage, then ballot utility and non-instrumental motivations should only be related to participation, but not the choice of voting mode itself.

Such a parsimonious mechanism may nevertheless be somewhat incomplete. For some citizens, social and cognitive barriers may offset the benefits of lower physical costs. If digital voting systems are difficult to use, they may reproduce or even reinforce existing inequalities in participation (Alvarez and Nagler, 2000; van Dijk, 2000, 2005; Margolis and Resnick, 2000; Putnam, 2001; Wilhelm, 2000). Lower physical costs may not translate into greater use among citizens who lack the skills needed to use remote Internet voting effectively. Resource-based theories of participation similarly show that cognitive resources shape turnout (see Smets and van Ham, 2013). We therefore account for the possibility that Internet voting may be more common among wealthier, younger, and better-educated citizens who command better resources and face fewer participation barriers to begin with (Norris, 2001; Alvarez and Nagler, 2000; Solop, 2002; Mossberger et al., 2003).

Because these cognitive and social resources are unevenly distributed, we expect the effect of voting cost on Internet voting to depend on them. We therefore test the following proposition:

> *H3: Probability of remote Internet voting increases as voting cost increases, conditional on cognitive and social resources*

Since these resources cannot be observed directly, we rely on socio-demographic traits commonly associated with them, including age, education, income (Smets and van Ham, 2013; Galicki, 2018), and especially perceived technological skills. We expect respondents with higher education, income, and perceived technical skills to be more likely to vote online than otherwise similar respondents facing the same voting costs.

3 Data, variable selection and model estimation

3.1 Data

The main hypotheses are tested using pooled survey data from six Internet-voting-enabled elections in Estonia between 2009 and 2023. All surveys are post-election studies conducted after the 2009 and 2014 European Parliament elections, the 2009 local elections, and the 2011, 2015, and 2023 national elections. The 2014–2023 surveys used stratified random samples, whereas the 2009–2011 surveys used random sampling with quotas for participation mode, with paper voters, Internet voters, and non-voters each comprising one third of the sample. Each survey included roughly 1,000 respondents, and the samples are representative of the voting-eligible population. In the modeling below we use demographic survey weights (combined for age, gender, region and ethnicity) throughout all the models and cluster standard errors by election in the pooled sample. Because pooling surveys has consequences we also estimate the main models for all elections separately and report those as supplementary material. The robustness of the findings is also tested using survey data from elections outside the six main cases, details further below. A technical overview of the surveys is provided in Appendix A.

3.2 Variable selection

We use two dependent variables. The first is turnout, coded 1 for voters and 0 for nonvoters. The second is the use of Internet voting among those who turned out, coded 1 for Internet voters and 0 for paper-ballot voters.

To capture non-instrumental motivation to vote, we include two measures. The first is the frequency of political discussions with friends and family, contrasting those who discuss politics often or sometimes (1) with those who do so rarely or never (0). Previous research shows that political discussion is associated with higher participation (Pattie, 1999; Pattie and Johnston, 2001). The second is trust in parliament, government, politicians, and the state. Each item is measured on a four-category Likert scale, recoded into a binary indicator of trust (1) versus no trust (0), and combined into an additive index, with higher values indicating greater institutional trust. More trusting citizens have been shown to be more likely to vote (Belanger and Nadeau, 2005).

To measure the utility of ballot choices, we use the propensity-to-vote (PTV) instrument, which captures the perceived utility of parties (see van der Eijk and Niem¨oller, 1984; van der Eijk and Franklin, 2009; van der Eijk et al., 2006; van der Brug et al., 2007; van der Eijk and Franklin, 2009).[1] Following van der Eijk and Oppenhuis (1991), we use a cut-off of 8 on the 0–10 scale and count the number of parties to which respondents assign that score or higher. Respondents evaluated eight parties, producing a count from 0 to 8. Higher values indicate that the respondent perceives more attractive options on the ballot.

Our main independent variable is perceived voting cost. This is measured by asking respondents how many minutes it took, or would have taken, to travel to the polling station and back. This provides an individual-level interval measure of perceived voting cost based on the physical proximity of the polling station.

Following resource-based theories of participation, we also include age, education, income, and gender as predictors of participation and as variables to be interacted with voting cost (Matsusaka and Palda, 1999; Franklin, 2004; Smets and van Ham, 2013). Age is measured in years. Education is represented by three binary variables for higher, secondary, and vocational education, with basic education as the reference category. Income is measured by household income decile. Gender is coded as male (1) and female (0).

Technical skills are measured using respondents' self-assessed PC literacy (1 = nonexistent, 2 = basic, 3 = average, 4 = good/very good). In addition to the additive model, we estimate interactions between voting cost and age, education, income, and PC literacy to assess whether the effect of cost on remote Internet voting depends on citizens' cognitive abilities and skills, as suggested in earlier work (Alvarez and Nagler, 2000; van Dijk, 2000, 2005; Margolis and Resnick, 2000; Putnam, 2001; Wilhelm, 2000). Because Estonia also has a substantial ethnic Russian minority with distinct patterns of political behavior, we include ethnicity as an additional control, coded 1 for self-reported ethnic Estonians and 0 for others, most of whom are ethnic Russians. In addition, given that the voting cost measure might act as a proxy for the rural urban divide we also include a dummy on the type of settlement of the voter, coded 1 for a urban and 0 for a rural settlement. For limitations of the survey item based approach see section 3.4 below.

[1] The survey question is the following: "Some people are quite certain that they will always vote for the same party. Others reconsider in each case to which party they will give their vote. I shall mention a number of parties. Would you indicate for each party how probable it is that you will ever vote for that party?" The respondent is provided with the list of parties in the respective polity with a scale ranging from 0 to 10, where 0 means "Will certainly never vote for this party" and 10 means "Will certainly vote for this party at some time".

3.3 Model specification

We are going to estimate two probit models to be able to compare estimations with the selection model elaborated further below. The first and second hypothesis will be evaluated by estimating the following models:

$$Pr(vote = 1) = \varphi(\beta_0 + \beta_1 D + \beta_2 U + \beta_3 C + \beta_4 A) \quad (1)$$

$$Pr(evote = 1) = \varphi(\beta_0 + \beta_1 D + \beta_2 U + \beta_3 C + \beta_4 A) \quad (2)$$

, where D stands for a vector of non-instrumental motivations, U for the utility of the choices, C the cost of voting and A for controls. These models are identical but for the dependent variable. We expect β_3 to have a negative sign in the first and positive in the second model. With regard to β_1 and β_2 we expect to see significant positive effects in the first and non-significant effects in the second model. Modeling turnout and remote Internet voting separately assumes that these two outcomes are generated by independent mechanisms, similar to a two-part Tobit model. This is a strong assumption. An alternative is that the mechanisms behind turnout and remote Internet voting are partially interdependent, even after controlling for the regressors. This raises the possibility of self-selection bias, whereby Internet voters do not choose this mode randomly, even after accounting for the mechanisms we propose. To ensure that our estimates of the key variable of interest - voting cost - are robust, we therefore also specify a Heckman selection model, in which the selection equation predicts electoral participation and the outcome equation predicts the use of remote Internet voting. There are further possible modeling options of turnout and mode choice, for a brief discussion of those see the next section.

The third and final hypothesis will be evaluated using graphical presentation of marginal effects extracted from the following probit model:

$$Pr(evote = 1) = \varphi(\beta_0 + \beta_1 D + \beta_2 U + \beta_3 C \times R + \beta_4 A) \quad (3)$$

, with R standing for a vector of resources. We focus on β_3 i.e. how the effect of cost on remote Internet voting differs depending on the level of resources and vice-versa. All models will be rendered using robust standard errors clustered by election. To ensure that pooling of the data does not create compositional effects not really observed when behaviors are examined for each elections separately we also run the full analysis for all elections one by one and report these results in the appendices.

We have also access to four datasets on elections held between 2015 and 2023, that is outside of the the original six elections examined here, but due to questionnaire changes we could not specify the exact same theoretically substantiated regression model, the variables needed to operationalize utilities are missing. Nevertheless, to examine if our central variable of interest - voting cost - plays a role also for all elections in a country were voting via the

Internet has become a norm we will also estimate a model for Internet voting without the utilities component in the following from:

$$Pr(evote = 1) = \varphi(\beta_0 + \beta_1 D + \beta_2 C + \beta_3 A) \tag{4}$$

, and report the results in the appendix.

3.4 Limitations and additional robustness checks

A limitation of the analysis is that some key variables are based on self-evaluation. In particular, travel time to the polling station and PC literacy are self-reported rather than directly observed. This raises the possibility of reporting bias correlated with the outcomes of interest. Non-voters, for example, may overstate travel time in order to rationalize abstention, while respondents who did not use Internet voting may understate their digital skills to rationalize non-use. If such patterns are systematic, they may bias the estimated relationship between voting costs, technical skills, and participation mode. We therefore interpret these variables as perceived rather than purely objective measures and treat the results with appropriate caution. In addition, some measure, such as the urban–rural indicator of place of residence, may be overly general, causing the cost variable to capture unobserved factors beyond Downsian voting costs and thus raising the possibility of omitted-variable bias. Finally, the modeling strategy chosen has clear implications as the vote and mode decision can in theory be modeled separately, jointly, or as a two-stage model with all approaches taking on their own assumptions.

We address these concerns in several ways. First, we present detailed travel-time statistics by voting mode and abstention status in order to be transparent about potential issues. Second, we test models with extreme top and bottom travel times trimmed (by 1%, 5% and 10%) as well as with a four category coding of the cost variable and report the results. This will show if the estimates are dependent on extreme self-reported distances. Third, we interpret these variables as perceived rather than purely objective measures and treat the results with appropriate caution. At the same time, because perceived costs and perceived competence may themselves shape behavior, these measures remain substantively relevant even if they are not free from reporting bias. In any case, we avoid strong causal language when interpreting the results.

To account for the different ways turnout and mode choice can be modeled, we employ three approaches: two separate probit models and a Heckman selection model in the main body of the paper, as well as a multinomial logit model with three outcomes: abstention, paper-ballot voting, and Internet voting. If the effect of voting cost remains clear and substantively consistent across these modeling strategies, we can be more confident that the unavoidable limitations of the data and models do not unduly obscure the underlying relationship. The chosen approach adheres to what is typical in turnout and vote choice studies. A standard approach for turnout has been using the logit or probit model (Abrams et al., 2011; Smets and van Ham, 2013; Fieldhouse and Cutts, 2018; Bechtel and Schmid, 2021). A preferred approach for vote- and mode choice has been the multinomial logit or probit model (Alvarez and Nagler, 1998; Dow and Endersby, 2004; Bhatti et al., 2020; Shino and Smith, 2022). Some caution against the latter when non-voters are included as the baseline category due to the heterogeneous nature of non-voters (Bagozzi and Marchetti, 2017). Yet others suggest two-stage or joint turnout and vote choice models are overall better aligned with the rational of voter behavior (Sanders, 1998). We partially cover all those as the probit models assumes an independent process, the Heckman a two-stage and the multinomial a simplified joint choice process.

4 Results

4.1 Main findings

The distribution of our key variable of interest, the cost of voting, measured as travel time to the polling station in minutes, is shown in Figure 1. Election-specific distributions are reported in Appendix C, Figure 4. Descriptive statistics by voting behavior are presented in Table 1, with election-specific statistics in Appendix B, Table 4. Summary statistics for all other variables used in the analysis are listed in Appendix B, Table 5. To account for the possibility that each additional minute has an increasing or diminishing effect as travel time rises, we use the logarithm of reported minutes.

(FIG 1 HERE)

Fig. 1 Distribution of the distance to the polling station in minutes (logged)

Table 1 Descriptive statistics of distance to polling station in minutes by voting behavior (not log transformed)

Behavior	Mean	Standard deviation	Median	25th percentile	75th percentile	Min	Max
Abstained	37.4	40.8	30.0	15.0	30.0	1.0	480.0
Voted on paper	25.4	23.0	20.0	10.0	30.0	1.0	360.0
Voted via Internet	38.2	47.1	30.0	15.0	45.0	1.0	720.0

For most respondents, the round trip to the polling station takes no more than half an hour. However, a substantial share report travel times of more than one or even two hours, indicating considerable variation in voting costs. For non-voters and Internet voters, the central tendencies reported in Table 1 are remarkably similar, while the 75th percentile and maximum values suggest that Internet voters include more individuals with very long travel times. This pattern is consistent across survey years, except in 2014, when non-voters reported markedly longer travel times. Paper voters and Internet voters, by contrast, show substantially different travel times, as expected.

The results of the multivariate models are reported as average marginal effects in Table 2.

Table 2 Multivariate models for the probability of voting and of remote Internet voting

	Vote (base: non-vote)	Internet vote (base: paper ballot)	Heckman (outcome)
Distance of polling station	-0.05***	0.15***	0.16***
in minutes (log)	(0.01)	(0.01)	(0.02)
Political talk	0.10***	-0.09***	-0.11**
	(0.02)	(0.01)	(0.03)
Trust institutions	0.05***	0.04*	0.04
	(0.01)	(0.02)	(0.02)
Utility of voting choices	0.05***	0.02	
	(0.01)	(0.01)	
Age	0.01**	0.01	0.01
	(0.01)	(0.00)	(0.00)
Age^2	-0.00	-0.00*	-0.00
	(0.00)	(0.00)	(0.00)
Male	-0.03	0.05***	0.05***
	(0.02)	(0.02)	(0.02)
Estonian	0.02	0.21***	0.21***
	(0.02)	(0.05)	(0.05)
Education: high	0.13***	0.17***	0.14**
(base: basic)	(0.02)	(0.04)	(0.05)
Education: secondary	0.08***	0.10**	0.09**
(base: basic)	(0.02)	(0.03)	(0.03)
Education: vocational	0.06***	0.05	0.03
(base: basic)	(0.01)	(0.04)	(0.04)
Income decile	0.01***	0.01*	0.01
	(0.00)	(0.00)	(0.00)
PC literacy: v. good or good	0.08***	0.21***	0.19***
(ref: poor or basic)	(0.01)	(0.03)	(0.04)
PC literacy: average	0.14***	0.32***	0.328***
(ref: poor or basic)	(0.03)	(0.04)	(0.05)
Urban settlement	-0.02	-0.01	-0.01
	(0.02)	(0.01)	(0.01)
Sensitivity	95.68%	69.41%	
Specificity	21.33%	78.95%	
Correctly classified	77.70%	74.89%	
Pseudo R^2	0.14	0.29	
Observations	4400	3314	4378
Rho			-0.39

Average marginal effects with standard errors clustered by election in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first and second hypotheses are clearly supported. Distance to the polling station has a negative association with the probability of voting, but a very strong positive one with the probability of casting a vote online. This finding is stable over time and across different elections; see Tables 6 and 7 in Appendix C. Figure 2 compares these associations graphically by showing the marginal effects of the cost variable across its range, with all other covariates held at their average values. It demonstrates the strength of the perceived cost-saving association with use of remote Internet voting.

The predicted probability of remote Internet voting crosses the 0.5 threshold at a travel time of around 30 minutes to the polling station. Moreover, the probability of casting a remote Internet vote rather than voting by paper ballot at the polling station rises steadily across the full range of distance. The comparison of the two association in the figure also shows clearly that voting costs appear to matter less for the decision whether to vote, but much more for the choice of voting mode. These associations are again repeated across elections over time; see Figures 5 and 6 in Appendix C. A reduced model using data from 10 elections tells the same story consistently; perceived voting costs are strongly associated with the choice to vote via the Internet; see Table 8 and Figure 11 in Appendix C.

Returning to Table 2, one expectation is not fully met; non-instrumental motivations still help distinguish between paper-ballot voters and Internet voters. This does not weaken the clear support for the main hypotheses, however. The higher the physical costs of voting, the more likely citizens are to use Internet voting.

(FIG 2 HERE)

Fig. 2 The effect of distance of polling station on voting and remote Internet voting (marginal effects from probit models with 95% confidence intervals)

The associations with sociodemographic controls are examined in more detail when evaluating the third hypothesis. It is already clear, however, that the factors distinguishing voters from non-voters differ somewhat from those distinguishing paper voters from Internet voters. Internet voters tend to be younger, better educated, have higher incomes, and report stronger PC skills than paper-ballot voters.

The third hypothesis was evaluated by estimating interactions between voting cost and socio-demographic variables, as well as technological skills, while controlling for all other variables listed in Table 2. These results are shown in Figure 3. If the interacted variables serve as proxies for barriers to use, we would expect the effect line of distance on the probability of Internet voting to remain flat among low-skill respondents, but to have a positive gradient among high-skill respondents. By contrast, if the cost savings are sufficient to outweigh limited skills, we would expect the probability of Internet voting to correlate with an increase for all skill groups and to be indistinguishable.

We first examine the interaction between perceived PC literacy and distance to the polling station. The figure shows that differences in perceived computer skills do not affect the probability of remote Internet voting among respondents living very close to a polling station. There are differences where confidence intervals don't overlap between high and average skill levels only at distances of roughly 15 to 60 minutes. Outside this range, the confidence intervals overlap, and the probability of Internet voting increases with distance for all skill groups. This suggests that PC literacy is not a major barrier to use.

Overall, low-skill citizens living far from a polling station are nearly as likely to vote online as higher-skill citizens. This pattern is broadly consistent across elections over time, with the exception of the 2014 European Parliament election; see Figure 7 in Appendix C.

(FIG 3 HERE)

Fig. 3 The effect of distance to polling station on remote Internet voting conditional on: (1) PC literacy levels, (2) education; (3) age and (4) income (marginal effects from probit model with 95% confidence intervals).

Turning to the interaction between distance to the polling station and education, again no clear difference emerges. Regardless of education level, higher voting costs make Internet voting more likely. This pattern is also repeated across elections, with the exception of 2014; see Figure 8 in Appendix C.

The interaction between age and voting cost is also somewhat surprising. All age groups show a similar increase in the probability of remote Internet voting as voting costs rise. This pattern appears separately in three of the six elections; see Figure 9 in Appendix C. The interaction therefore suggests that cost savings can outweigh age-related barriers to remote Internet voting.

Finally, we examine the interaction between income and voting cost. To avoid overloading the graph, some income deciles are grouped together. Regardless of the grouping, higher voting costs make remote Internet voting more likely at all income levels. This pattern is again stable across elections and over time; see Figure 10 in Appendix C.

In sum, the interaction results are somewhat unexpected. Traditional resource theories of voting suggest that citizens with greater resources are more likely to participate. The widely discussed digital divide is essentially a resource problem; those with the skills and resources needed to use technological solutions should be the main beneficiaries and users of new, non-traditional voting methods, whereas those with lower skill levels should face greater barriers to use. Instead, we find that high perceived physical voting costs are positively associated with the probability of voting through new technologies across resource levels. The reduction in voting costs made possible by Internet voting appears substantial enough to make even those who might otherwise be less likely to use technology more likely to adopt it. This finding suggests that the technology may partially fulfill its purpose of making voting easier for those for whom voting would otherwise be comparatively more costly.

4.2 Robustness checks

We also conducted multiple robustness checks to be certain that the main probit model results stand. For remote Internet voting, the coefficients from the probit models are very similar to those from the Heckman selection model, and the ρ coefficient in the latter is not significant. This indicates that there is no general self-selection problem once all controls are included. We used the utility of voting choices as the exclusion restriction, as it strongly predicts voting but not remote Internet voting, making it suitable for this purpose (Cameron and Trivedi, 2010, p. 558). Although the Heckman model for one election - the 2009 local election; see Table 12 in Appendix D - indicates self-selection and corrects the distance effect downwards, the effect remains significant and substantial even after the correction.

The effects of the cost variable in probit models in which the extreme top and bottom 1%, 5%, and 10% values are trimmed are reported in Appendix D, Table 9. The effects remain significant and are even strongest in the model with 10% trimming. Effects using a four-category coding of the cost variable are reported in Appendix D, Table 10. Again, the effects of cost on Internet voting remain persistent and in the expected direction. Thus, neither the coding logic nor extreme values of perceived cost appear to affect the results. The estimated multinomial model, with non-voting, paper-ballot voting, and Internet voting specified as the dependent variable, is reported in Appendix D, Table 11, and Figure 12. It shows that, as distance to the polling station initially increases, the probabilities of both Internet voting and non-voting rise in a similar manner. At longer travel times, however, the effect becomes clearly stronger for Internet voting. Although these robustness checks cannot fully rule out endogeneity or omitted-variable bias, they provide reasonable confidence that the mechanism is specified adequately, as they all tell the same story and the substantive interpretation of the probit models in Table 2 is trustworthy.

5 Conclusion

We examined the extent to which remote Internet voting functions as a cost-saving mode of participation using post-election survey data from six elections in a country where remote Internet voting is available nationwide. Drawing on Downs (1957), we argued that voting costs should correlate with reduced probability of participation but increased probability of choosing remote Internet voting. We further expected that while ballot utility, noninstrumental motivations, and electoral costs all shape turnout, voting costs should play the central role in the choice of voting mode. Finally, we tested whether the association between voting costs and remote Internet voting depends on voters' technical skills and broader resources.

The results provide partial support for these expectations. The farther away the polling station is perceived to be, the lower the probability of voting, although this effect is modest. For remote Internet voting, by contrast, the association is strongly positive. Citizens who perceive the round trip to a polling station to take roughly 30 minutes or less are still more likely to vote by paper ballot, but beyond that point Internet voting becomes the more likely option. This positive association remains robust across controls and over time. We find that only perceived low computer skills are associated with a lower likelihood of Internet voting among otherwise similar voters, but this difference disappears as distance to the polling station increases. No significant interaction effects are found for education, age, or income. In other words, regardless of these socio-demographic characteristics, which we treat as proxies for cognitive and social resources, remote Internet voting becomes more likely as perceived physical voting costs rise.

All voters face some costs of participation, even if these are not equal in relative terms. What they share is the need to travel to the polling station. Our findings show that the expected cost-reducing effect of remote Internet voting becomes especially pronounced as this physical burden increases. The appeal of a low-cost and convenient voting mode appears strong enough to overcome select social and technical barriers to its use. This suggests that Internet voting can affect participation by lowering the costs of voting. Once physical voting costs reach a critical level — in our case, roughly a 30-minute trip — remote Internet voting becomes highly likely. These findings have implications for countries considering the adoption of remote Internet voting. Concerns that this technologically advanced mode may simply reproduce social inequalities, for example by creating barriers for citizens with lower mobility such as older voters, are not supported here. If the cost reduction is large enough, remote Internet voting becomes likely regardless of the resources a person commands.

At the same time, our findings also suggest that remote Internet voting has limited capacity to raise turnout overall. It provides a cheaper mode of participation, but its strongest effects are concentrated among a relatively small group for whom physical voting costs seem a real barrier. In this respect, Internet voting may resemble postal voting: it is taken up by voters who otherwise face higher costs of participation (Karp and Banducci, 2000; Nemcok and Peltoniemi, 2021), but is unlikely to substantially increase overall turnout (Wass et al., 2017) or voter composition (Southwell and Burchett, 2000). Because this group is likely to remain a minority, and because declining turnout is driven largely by low participation among younger generations, we remain skeptical that remote Internet voting can reverse broader turnout decline. Our results show that the use of a technology becomes more likely the greater the benefit it offers in reducing perceived costs. But when the obstacles to participation lie beyond costs, technological solutions alone are unlikely to counter wider social trends of electoral disengagement.

References

Abrams, S., Iversen, T., & Soskice, D. (2011). Informal social networks and rational voting. *British Journal of Political Science, 41*(2), 229–257.

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

Alvarez, R. M., & Nagler, J. (1998). When politics and models collide: Estimating models of multiparty elections. *American Journal of Political Science, 42*(1), 55–96.

Bagozzi, B. E., & Marchetti, K. (2017). Distinguishing occasional abstention from routine indifference in models of vote choice. *Political Science Research and Methods, 5*(2), 277–294.

Bechtel, M. M., & Schmid, L. (2021). Electoral reforms and the representativeness of turnout. *Political Science Research and Methods, 9*(3), 485–499.

Bélanger, É., & Nadeau, R. (2005). Political trust and the vote in multiparty elections: The Canadian case. *European Journal of Political Research, 44*(1), 121–146.

Bhatti, Y. (2012). Distance and voting: Evidence from Danish municipalities. *Scandinavian Political Studies, 35*(2), 141–158.

Bhatti, Y., Fieldhouse, E., & Hansen, K. M. (2020). It's a group thing: How voters go to the polls together. *Political Behavior, 42*, 1–34.

Blais, A., & Young, R. (1999). Why do people vote? An experiment in rationality. *Public Choice, 99*(1–2), 39–55.

Brady, H. E., & McNulty, J. E. (2011). Turning out to vote: The costs of finding and getting to the polling place. *American Political Science Review, 105*(1), 115–134.

Cameron, A., & Trivedi, P. (2010). *Microeconomics using Stata*. Stata Press.

Ciancio, A., & Kämpfen, F. (2023). The heterogeneous effects of internet voting. *European Journal of Political Economy, 79*, Article 102444.

Dalton, R. (1988). *Citizen politics in Western democracies*. Chatham House Publishers.

Dow, J. K., & Endersby, J. W. (2004). Multinomial probit and multinomial logit: A comparison of choice models for voting research. *Electoral Studies, 23*(1), 107–122.

Downs, A. (1957). *An economic theory of democracy*. Harper.

Dyck, J. J., & Gimpel, J. G. (2005). Distance, turnout, and the convenience of voting. *Social Science Quarterly, 86*(3), 531–548.

Ehin, P., Solvak, M., Willemson, J., & Vinkel, P. (2022). Internet voting in Estonia 2005–2019: Evidence from eleven elections. *Government Information Quarterly, 39*(4), Article 101718.

Fieldhouse, E., & Cutts, D. (2018). Shared partisanship, household norms and turnout: Testing a relational theory of electoral participation. *British Journal of Political Science, 48*(3), 807–823.

Fisher, J., & Savani, M. M. (2022). Who's in charge? The impact of delivery and perception of risk on the willingness to voting online. *The British Journal of Politics and International Relations*. Advance online publication.

Franklin, M. (2004). *Voter turnout and the dynamics of electoral competition in established democracies since 1945*. Cambridge University Press.

Galicki, C. (2018). Barriers to voting and the cost of voting among low socioeconomic, young and migrant voters in New Zealand. *Political Science, 70*(1), 41–57.

Germann, M. (2021). Internet voting increases expatriate voter turnout. *Government Information Quarterly, 38*(2), Article 101560.

Geys, B. (2006). Explaining voter turnout: A review of aggregate-level research. *Electoral Studies, 25*(4), 637–663.

Goodman, N., & Stokes, L. C. (2020). Reducing the cost of voting: An evaluation of internet voting's effect on turnout. *British Journal of Political Science, 50*(3), 1155–1176.

Karp, J. A., & Banducci, S. A. (2000). Going postal: How all-mail elections influence turnout. *Political Behavior, 22*(3), 223–239.

Margolis, M., & Resnick, D. (2000). *Politics as usual: The cyberspace "revolution"*. Sage.

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

Mossberger, K., Tolbert, C., & Stansbury, M. (2003). *Virtual inequality: Beyond the digital divide*. Georgetown University Press.

Nemčok, M., & Peltoniemi, J. (2021). An examination of the two opposing factors impacting adoption of postal voting among citizens living abroad. *Political Behavior*.

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* [Paper presentation]. British Study Group Seminar, Minda de Gunzburg Center for European Studies, Harvard University.

Olson, M. (1965). *The logic of collective action*. Harvard University Press.

Pattie, C., & Johnston, R. J. (1999). Context, conversation and conviction: Social networks and voting at the 1992 British general election. *Political Studies, 47*(5), 877–889.

Pattie, C., & Johnston, R. J. (2001). Talk as political context: Conversation and electoral change in British elections 1992–1997. *Electoral Studies, 20*(1), 17–40.

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

Riker, W., & Ordeshook, P. (1968). A theory of the calculus of voting. *American Political Science Review, 62*(1), 25–42.

Sanders, M. S. (1998). Unified models of turnout and vote choice for two-candidate and three-candidate elections. *Political Analysis, 7*, 89–115.

Santana, A., & Aguilar, S. (2021). How costly is voting? Explaining individual differences in the costs of voting. *Journal of Elections, Public Opinion and Parties, 31*(1), 119–139.

Shino, E., & Smith, D. (2022). Political knowledge and convenience voting. *Journal of Elections, Public Opinion and Parties, 32*(2), 408–428.

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

Solop, F. (2002). Digital democracy comes of age: Internet voting and the 2000 Arizona Democratic primary election. *PS: Political Science & Politics, 34*(2), 289–293.

Solvak, M., & Vassil, K. (2018). Could internet voting halt declining electoral turnout? New evidence that e-voting is habit forming. *Policy & Internet, 10*(1), 4–21.

Southwell, P. L., & Burchett, J. (2000). Does changing the rules change the players? The effect of all-mail elections on the composition of the electorate. *Social Science Quarterly, 81*(3), 837–845.

van der Brug, W., van der Eijk, C., & Franklin, M. (2007). *The economy and the vote: Economic conditions and elections in fifteen countries*. Cambridge University Press.

van der Eijk, C., & Franklin, M. (2009). *Elections and voters*. Palgrave Macmillan.

van der Eijk, C., & Niemöller, B. (1984). Het potentiële electoraat van de Nederlandse politieke partijen. *Beleid en Maatschappij, 11*, 192–204.

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

van der Eijk, C., van der Brug, W., Kroh, M., & Franklin, M. (2006). Rethinking the dependent variable in voting behavior: On the measurement and analysis of electoral utilities. *Electoral Studies, 25*(3), 424–447.

van Dijk, J. (2000). Widening information gaps and policies of prevention. In K. L. Hacker & J. van Dijk (Eds.), *Digital democracy: Issues of theory and practice* (pp. 166–183). Sage.

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

Wass, H., Mattila, M., Rapeli, L., & Söderlund, P. (2017). Voting while ailing? The effect of voter facilitation instruments on health-related differences in turnout. *Journal of Elections, Public Opinion and Parties, 27*(4), 503–522.

Wilhelm, A. (2000). *Democracy in the digital age: Challenges to political life in cyberspace*. Routledge.

Woller, A., Justesen, M. K., & Hariri, J. G. (2022). The cost of voting and the cost of votes. *The Journal of Politics*. Advance online publication.

A Survey descriptions

All surveys are post-election surveys with fieldwork conducted over 30 days after the election date. The 2009 – 2011 surveys used quota sampling according to voting mode due to the still relatively low number of remote Internet voters in the population. The samples are representative for eligible voters in terms of age, gender, ethnicity and region. The 2014 to 2023 surveys used stratified random samples and are representative for eligible voters in terms of age, gender, ethnicity, settlement type and region. Table 3 lists the interview methods and number of respondents for all the surveys used in this study.

Table 3 Description of post-election surveys used in the study

Post-election survey	Sampling method	Interview method	Number of respondents
2009 EP	quota sample (1/3 non-voters; 1/3 paper voters; 1/3 e-voters)	CATI	1000
2009 local	quota sample (1/3 non-voters; 1/3 paper voters; 1/3 e-voters)	CATI	1000
2011 national	quota sample (1/3 non-voters; 1/3 paper voters; 1/3 e-voters)	CATI	1007
2013 local	stratified random sample	CAPI	1042
2014 EP	stratified random sample	CAPI	1001
2015 national	stratified random sample	CAPI	1007
2017 local	stratified random sample	CATI	1000
2019 national	stratified random sample	CATI	1000
2019 EP	stratified random sample	CATI	1002
2021 local	stratified random sample	CATI(30%)/CAWI(70%)	1153
2023 national	stratified random sample	CATI(50%)/CAWI(50%)	1001

B Summary statistics

Table 4 Descriptive statistics of distance to polling station in minutes by voting behavior by survey

	Behavior	Mean	Standard deviation	Median	25th percentile	75th percentile	Min	Max
2009 EP	Abstained	32.4	30.7	30.0	15.0	30.0	2.0	260.0
	Voted on paper	23.9	27.0	15.0	10.0	30.0	1.0	240.0
	Internet voted	36.8	32.7	30.0	15.0	45.0	1.0	240.0
2009 local	Abstained	38.5	38.7	30.0	15.0	45.0	3.0	360.0
	Voted on paper	24.9	25.1	15.0	10.0	30.0	1.0	330.0
	Internet voted	45.5	53.1	30.0	20.0	60.0	2.0	480.0
2011 national	Abstained	40.4	36.8	30.0	15.0	60.0	2.0	250.0
	Voted on paper	27.8	27.1	20.0	15.0	30.0	2.0	360.0
	Internet voted	40.6	44.1	30.0	15.0	45.0	5.0	420.0
2014 EP	Abstained	39.0	53.9	25.0	15.0	40.0	1.0	480.0
	Voted on paper	27.3	24.6	20.0	15.0	30.0	2.0	360.0
	Internet voted	29.9	22.9	25.0	15.0	32.5	5.0	120.0
2015 national	Abstained	33.9	27.5	30.0	15.0	40.0	2.0	180.0
	Voted on paper	24.6	17.1	20.0	15.0	30.0	1.0	180.0
	Internet voted	29.8	24.3	30.0	15.0	33.5	1.0	180.0
2023 national	Abstained	38.5	44.6	30.0	15.0	40.0	1.0	250.0
	Voted on paper	23.8	20.2	20.0	10.0	30.0	2.0	180.0
	Internet voted	38.4	66.3	30.0	15.0	30.0	2.0	720.0

Table 5 Summary statistics (means for all and standard deviations in parentheses for interval variables)

	All	2009 EP	2009 local	2011 national	2014 EP	2015 national	2023 national
Voted	0.69	0.68	0.67	0.66	0.62	0.81	0.83
E-voted	0.41	0.59	0.49	0.50	0.23	0.24	0.51
Distance of polling station	3.22	3.19	3.29	3.31	3.18	3.13	3.12
in minutes (log)	(0.73)	(0.75)	(0.76)	(0.75)	(0.74)	(0.65)	(0.80)
Political talk	0.65	0.62	0.58	0.57	0.68	0.80	0.85
Trust	1.11	1.12	1.12	1.32	0.95	1.04	1.09
	(0.99)	(0.89)	(0.88)	(0.82)	(1.13)	(1.16)	(0.94)
Utility of voting choices	1.14	1.09	1.18	1.30	0.81	1.09	1.35
	(1.04)	(1.03)	(1.13)	(1.07)	(0.78)	(0.90)	(1.16)
Age	48.97	47.27	47.67	50.37	49.31	50.18	50.46
	(17.84)	(17.31)	(17.37)	(18.10)	(18.14)	(18.07)	(18.20)
Male	0.43	0.46	0.41	0.39	0.46	0.45	0.44
Estonian	0.84	0.90	0.84	0.88	0.80	0.80	0.76
Education: high	0.33	0.38	0.36	0.34	0.28	0.29	0.34
Education: secondary	0.22	0.20	0.21	0.21	0.24	0.23	0.54
Education: vocational	0.32	0.34	0.34	0.34	0.29	0.12	
Education: basic	0.13	0.08	0.09	011	0.19	0.18	
Income decile	4.57	4.77	4 44	3 85	4 95	4 96	4.96
	(3.09)	(2.97)	(2.95)	(3.38)	(3.12)	(2.87)	(3.03)
PC literacy: v. good or good	0.41	0.43	0 39	0 36	0 45	0 42	0.61
PC literacy: average	0.30	0.30	0.32	0.31	0.26	0.31	0.28
PC literacy: poor or basic	0.29	0.27	0.30	0.32	0.30	0.27	0.11
Urban settlement	0.65	0.65	0.67	0.63	0.65	0.65	0.64

C Results per year

(FIG 4 HERE)

Fig. 4 Distribution of the distance to the polling station in minutes (logged) for each election in the sample

Table 6 Multivariate models for the probability of voting per year (2009-2023) *(base: non-voting)*

	2009 EP	2009 local	2011 national	2014 EP	2015 national	2023 national
Distance of polling station in minutes (log)	-0.01	-0.07***	-0.04*	-0.04	-0.05**	-0.05**
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)
Political talk	0.01	0.08***	0.10***	0.12**	0.11***	0.09**
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
Trust institutions	0.05***	0.08***	0.07***	0.05**	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)	(0.01)
Utility of voting choice	0.03*	0.03*	0.07***	0.06*	0.07***	0.05***
	(0.01)	(0.01)	(0.07)	(0.02)	(0.02)	(0.01)
Age	0.03***	-0.00	0.01*	0.02**	0.01*	0.01*
	(0.00)	(0.01)	(0.06)	(0.01)	(0.01)	(0.00)
Age^2	-0.00***	0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.06)	(0.00)	(0.00)	(0.00)
Male	-0.12***	-0.03	-0.04	-0.02	0.01	0.02
	(0.03)	(0.03)	(0.01)	(0.04)	(0.04)	(0.02)
Estonian	-0.04	-0.02	0.08*	-0.07	-0.01	0.10***
	(0.05)	(0.04)	(0.01)	(0.04)	(0.03)	(0.03)
Education: high	0.11	0.15*	0.21***	0.11	0.13**	0.10**
(ref: basic)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)
Education: secondary	0.07	0.05	0.10	0.08	0.08*	0.05
(ref: basic)	(0.06)	(0.06)	(0.05)	(0.06)	(0.04)	(0.03)
Education: vocational	0.08	0.04	0.08	0.09	0.08**	0.06
(ref: basic)	(0.06)	(0.06)	(0.01)	(0.06)	(0.03)	(0.06)
Income decile	0.00	0.01	-0.00	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.07)	(0.00)
PC literacy: v. good or good	0.10*	0.10*	0.10*	0.00	0.10*	0.02
(ref: poor or basic)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)
PC literacy: average	0.13*	0.09	0.20***	0.07	0.13**	0.13*
(ref: poor or basic)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.05)
Urban settlement	0.02	-0.03	-0.03	-0.09*	0.01	-0.02
	(0.03)	(0.03)	(0.15)	(0.03)	(0.04)	(0.03)
Sensitivity	94.87%	92.97%	92.36%	88.53%	97.90%	98.01%
Specificity	24.14%	26.36%	36.79%	40.48%	16.95%	16.83%
Correctly classified	76.65%	74.06%	76.90%	71.28%	84.92%	87.14%

Pseudo R^2	0.20	0.18	0.28	0.21	0.27	0.31
Observations	788	775	762	585	736	754

Average marginal effects with standard errors in parentheses.
* p<0.05, ** p<0.01, *** p<0.001

Table 7 Multivariate models for the probability of Internet voting per year (2009-2023) *(base: paper ballot voting)*

	2009 EP	2009 local	2011 national	2014 EP	2015 national	2023 national
Distance of polling station	0.17***	0.20***	0.10***	0.07*	0.12***	0.11***
in minutes (log)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Political talk	-0.09**	-0.03	-0.06	-0.05	-0.01	-0.08
	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.06)
Trust institutions	-0.00	-0.01	0.01	0.04*	0.03*	0.14***
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Utility of voting choice	0.03	-0.03	-0.02	0.01	0.01	0.03*
	(0.02)	(0.01)	(0.07)	(0.01)	(0.02)	(0.02)
Age	0.01	0.01	0.00	0.00	0.01*	-0.00
	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
Age^2	-0.00	-0.00*	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.06)	(0.00)	(0.00)	(0.00)
Male	0.04	0.00	0.10**	0.03	0.07*	0.05
	(0.03)	(0.03)	(0.01)	(0.03)	(0.03)	(0.04)
Estonian	0.34***	0.22***	0.17*	-0.02	0.15**	0.11*
	(0.05)	(0.04)	(0.01)	(0.04)	(0.05)	(0.05)
Education: high	0.29***	0.16*	0.06	0.06	0.03	0.01
(ref: basic)	(0.08)	(0.06)	(0.07)	(0.06)	(0.06)	(0.08)
Education: secondary	0.26**	0.09	-0.11	0.02	0.01	0.03
(ref: basic)	(0.08)	(0.08)	(0.03)	(0.06)	(0.06)	(0.08)
Education: vocational	0.21**	0.01	-0.08	-0.04	-0.03	0.02
(ref: basic)	(0.06)	(0.08)	(0.01)	(0.05)	(0.06)	(0.11)
Income decile	0.01	0.01	0.02***	0.02***	0.01	0.00
	(0.01)	(0.01)	(0.05)	(0.00)	(0.01)	(0.01)
PC literacy: v. good or good	0.21**	0.20***	0.15*	0.07	0.13**	0.24***
(ref: poor or basic)	(0.06)	(0.05)	(0.06)	(0.04)	(0.03)	(0.06)
PC literacy: average	0.29***	0.32***	0.20**	0.34***	0.29***	0.34***
(ref: poor or basic)	(0.07)	(0.06)	(0.07)	(0.07)	(0.05)	(0.07)
Urban settlement	-0.03	-0.06	-0.02	-0.07	0.02	-0.04
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
Sensitivity	87.98	78.79	79.70	34.21	27.74	75.31
Specificity	72.43	78.62	78.14	93.88	94.96	68.11
Correctly classified	81.51%	78.70%	78.91%	81.62%	79.93%	71.70%

Pseudo R^2	0.55	0.51	0.51	0.37	0.28	0.31
Observations	584	554	550	370	613	643

Average marginal effects with standard errors in parentheses
*p<0.05, ** p<0.01, *** p<0.001

(FIG 5 HERE)

Fig. 5 The effect of distance of polling station on voting (2009-2023)

(FIG 6 HERE)

Fig. 6 The effect of distance of polling station on Internet voting (2009-2023)

(FIG 7 HERE)

Fig. 7 The effect of distance to polling station on remote Internet voting conditional on PC literacy levels

(2009-2023)

(FIG 8 HERE)

Fig. 8 The effect of distance to polling station on remote Internet voting conditional on education levels

(2009-2023)

(FIG 9 HERE)

Fig. 9 The effect of distance to polling station on remote Internet voting conditional on age (2009-2023)

(FIG 10 HERE)

Fig. 10 The effect of distance to polling station on remote Internet voting conditional on income decile

(2009-2023)

Table 8 Multivariate models without utilities for the probability of Internet voting per year (2009-2023) *(base: paper ballot voting)*

	2009 EP	2009 local	2011 national	2014 EP	2015 national	2017 local	2019 national	2019 EP	2021 local	2023 national
Distance of polling station	0.19***	0.20***	0.10***	0.07*	0.12***	0.15***	0.10***	0.14***	0.09***	0.11***
in minutes (log)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Political talk	-0.07*	-0.04	-0.04	-0.03	0.01	-0.10	0.04	0.04	-0.08	-0.09
	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
Trust institutions	-0.00	-0.01	0.01	0.03*	0.03*	0.04**	0.07***	-0.03	0.09***	0.15***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Age	0.01*	0.01	-0.00	0.01	0.01	0.01	-0.00	-0.00	0.00	-0.00
	(0.00)	(0.01)	(0.06)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Age^2	-0.00**	-0.00**	-0.00	-0.00	-0.00*	-0.00*	0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Male	0.04	0.01	0.11***	0.03	0.06	0.02	-0.06	0.15***	0.06	0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Estonian	0.34***	0.21***	0.15**	-0.01	0.15***	0.31***	0.20***	0.21**	0.17***	0.13*
	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.04)	(0.05)
Education: high	0.28***	0.16*	0.06	0.06	0.02	0.21	0.17	0.31***	0.19**	0.01
(ref: basic)	(0.08)	(0.08)	(0.08)	(0.09)	(0.06)	(0.11)	(0.10)	(0.10)	(0.07)	(0.08)
Education: secondary	0.23**	0.10	-0.11	0.02	-0.00	0.17	0.08	0.19*	0.09	0.03
(ref: basic)	(0.08)	(0.08)	(0.08)	(0.07)	(0.06)	(0.09)	(0.09)	(0.10)	(0.07)	(0.08)
Education: vocational	0.19*	0.02	-0.07	-0.05	-0.05		0.11	0.22*	0.22*	-0.00
(ref: basic)	(0.06)	(0.08)	(0.07)	(0.07)	(0.06)		(0.10)	(0.11)	(0.10)	(0.11)
Income decile	0.01*	0.01	0.02	0.01	0.01	0.01	0.02*	0.01	-0.02*	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
PC literacy: v. good or good	0.24***	0.21***	0.16**	0.07	0.15*	0.17*	0.22***	0.15*	0.17*	0.24***

	2009 EP	2009 local	2011 national	2014 EP	2015 national	2017 local	2019 national	2019 EP	2021 local	2023 national
(ref: poor or basic)	(0.06)	(0.05)	(0.06)	(0.04)	(0.03)	(0.06)	(0.05)	(0.07)	(0.07)	(0.06)
PC literacy: average	0.30***	0.33***	0.21**	0.35***	0.30***	0.31***	0.36***	0.28***	0.33***	0.34***
(ref: poor or basic)	(0.06)	(0.06)	(0.07)	(0.06)	(0.05)	(0.06)	(0.06)	(0.08)	(0.08)	(0.06)
Urban settlement	-0.03	-0.06	-0.02	-0.08*	0.01	0.03	-0.03	0.02	-0.02	-0.04
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)
Sensitivity	89.57%	77.61%	79.93%	32.47%	27.14%	67.83%	65.35%	72.96%	71.68%	74.84%
Specificity	71.37 %	79.59%	77.66%	94.41%	94.98%	80.82%	80.76%	67.20%	64.29%	68.28%
Correctly classified	81.96%	78.65%	78.76%	81.89%	79.61%	75.80%	74.48%	70.19%	68.04%	71.52%
Pseudo R^2	0.55	0.55	0.51	0.37	0.27	0.40	0.36	0.30	0.22	0.31
Observations	593	562	565	381	618	595	623	520	682	653

Average marginal effects with standard errors in parentheses.
* p<0.05, ** p<0.01, *** p<0.001

(FIG 11 HERE)

Fig. 11 The effect of distance of polling station on Internet voting (2009-2023)

D Robustness checks

Table 9 Multivariate models for the probability of voting and of remote Internet voting with trimmed extreme costs

	Vote (base: non-vote)	E-vote (base: paper ballot)
	Trimmed top and bottom 1%	
Distance of polling station	-0.05***	0.15***
in minutes (log)	(0.01)	(0.02)
Controls included	*effects not reported*	
Observations	4290	3241
	Trimmed top and bottom 5%	
Distance of polling station	-0.05**	0.15***
in minutes (log)	(0.02)	(0.03)
Controls included	*effects not reported*	
Observations	4113	3130
	Trimmed top and bottom 10%	
Distance of polling station	-0.05*	0.18***
in minutes (log)	(0.02)	(0.02)
Controls included	*effects not reported*	
Observations	3427	2634

Average marginal effects with standard errors clustered by election in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table 10 Multivariate models for the probability of voting and of remote Internet voting with four category cost variable

	Vote (base: non-vote)	E-vote (base: paper ballot)
Distance of polling station:	0.00	0.08***
> 15 to 25 minutes	(0.03)	(0.03)
base: up to 15 minutes)		
> 25 to 35 minutes	−0.06	0.18***
base: up to 15 minutes)	(0.02)	(0.01)
more than 35 minutes	−0.08***	0.26***
base: up to 15 minutes)	(0.02)	(0.04)
Controls included	*effects not reported*	
Observations	4400	3314

Average marginal effects with standard errors clustered by election in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table 11 Multinomial logit for the probability of voting and of remote Internet voting with non-voting as base category (relative risk ratios)

	Paper vote (base: non-vote)	E-vote base: non-vote
Distance of polling station in minutes (log)	0.50***	1.20*
	(0.05)	(0.11)
Trust institutions	1.20**	1.55***
	(0.08)	(0.09)
Utility of voting choices	1.34***	1.51**
	(0.09)	(0.19)
Age	1.08**	1.12***
	(0.02)	(0.04)
Age^2	0.99	0.99**
	(0.00)	(0.00)
Male	0.75*	0.95
	(0.10)	(0.15)
Estonian	0.72	2.56***
	(0.17)	(0.36)
Education: high *(base: basic)*	1.44*	3.96***
	(0.23)	(0.58)
Education: secondary *(base: basic)*	1.32***	2.43***
	(0.01)	(0.36)
Education: vocational *(base: basic)*	1.28*	1.74*
	(0.14)	(0.44)
Income decile	1 02	1.07***
	(0.02)	(0.01)
PC literacy: v. good or good *(ref: poor or basic)*	1.26**	4.30***
	(0.09)	(0.63)
PC literacy: average *(ref: poor or basic)*	1.36	7.66***
	(0.27)	(2.31)
Urban settlement	0.87	0.80*
	(0.07)	(0.07)
Pseudo R^2	0.18	
Observations	4378	

Relative risk ratios with standard errors clustered by election in parentheses
* p<0.05, ** p<0.01, *** p<0.001

(FIG 12 HERE)

Fig. 12 Effect of the distance to the polling station in minutes (logged) on non-voting, paper voting and Internet voting (predicted probabilities from the multinomial logit model)

Table 12 Robustness checks for remote Internet voting per election

	2009 EP	2009 local	2011 national	2014 EP	2015 national	2017 local	2019 national	2019 EP	2021 local	2023 national
Distance of polling station	0.19***	0.20***	0.10***	0.07*	0.12***	0.15***	0.10***	0.14***	0.09***	0.11***
in minutes (log)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Political talk	-0.07*	-0.04	-0.04	-0.03	0.01	-0.10	0.04	0.04	-0.08	-0.09
	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
Trust institutions	-0.00	-0.01	0.01	0.03*	0.03*	0.04**	0.07***	-0.03	0.09***	0.15***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Age	0.01*	0.01	-0.00	0.01	0.01	0.01	-0.00	-0.00	0.00	-0.00
	(0.00)	(0.01)	(0.06)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Age^2	-0.00**	-0.00**	-0.00	-0.00	-0.00*	-0.00*	0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Male	0.04	0.01	0.11***	0.03	0.06	0.02	-0.06	0.15***	0.06	0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Estonian	0.34***	0.21***	0.15**	-0.01	0.15***	0.31***	0.20***	0.21**	0.17***	0.13*
	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.04)	(0.05)
Education: high	0.28***	0.16*	0.06	0.06	0.02	0.21	0.17	0.31***	0.19**	0.01
(ref: basic)	(0.08)	(0.08)	(0.08)	(0.09)	(0.06)	(0.11)	(0.10)	(0.10)	(0.07)	(0.08)
Education: secondary	0.23**	0.10	-0.11	0.02	-0.00	0.17	0.08	0.19*	0.09	0.03
(ref: basic)	(0.08)	(0.08)	(0.08)	(0.07)	(0.06)	(0.09)	(0.09)	(0.10)	(0.07)	(0.08)
Education: vocational	0.19*	0.02	-0.07	-0.05	-0.05		0.11	0.22*	0.22*	-0.00
(ref: basic)	(0.06)	(0.08)	(0.07)	(0.07)	(0.06)		(0.10)	(0.11)	(0.10)	(0.11)
Income decile	0.01*	0.01	0.02	0.01	0.01	0.01	0.02*	0.01	-0.02*	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
PC literacy: v. good or good	0.24***	0.21***	0.16**	0.07	0.15*	0.17*	0.22***	0.15*	0.17*	0.24***
(ref: poor or basic)	(0.06)	(0.05)	(0.06)	(0.04)	(0.03)	(0.06)	(0.05)	(0.07)	(0.07)	(0.06)
PC literacy: average	0.30***	0.33***	0.21**	0.35***	0.30***	0.31***	0.36***	0.28***	0.33***	0.34***
(ref: poor or basic)	(0.06)	(0.06)	(0.07)	(0.06)	(0.05)	(0.06)	(0.06)	(0.08)	(0.08)	(0.06)
Urban settlement	-0.03	-0.06	-0.02	-0.08*	0.01	0.03	-0.03	0.02	-0.02	-0.04
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)
Sensitivity	89.5%	77.61%	79.93%	32.47%	27.14%	67.83%	65.35%	72.96%	71.68%	74.84%
Specificity	71.37%	79.59%	77.66%	94.41%	94.98%	80.82%	80.76%	67.20%	64.29%	68.28%
Correctly classified	81.96%	8.65%	78.76%	81.89%	79.61%	75.80%	74.48%	70.19%	68.04%	71.52%
Pseudo R^2	0.55	0.55	0.51	0.37	0.27	0.40	0.36	0.30	0.22	0.31
Observations	593	562	565	381	618	595	623	520	682	653

Average marginal effects with standard errors in parentheses.
* p<0.05, ** p<0.01, *** p<0.001