# Modern Voting Tools, Social Incentives and Voter Turnout: Theory and Evidence

Patricia Funk[1]
Universitat Pompeu Fabra

This Draft: June 2006

## Abstract

This paper investigates social norms and voting behavior. I argue that social norms create incentives for signaling, i.e. voting for the purpose of being *seen* at the voting act. Empirical evidence on signaling can be gained by looking at the introduction of optional postal voting in Switzerland. Even though the possibility of mail voting reduced voting costs substantially, it didn't increase turnout. Consistent with my model's predictions, voter turnout decreased more in the smaller communities, but in the meantime, the share of cooperators (i.e. interested voters) was more positively affected there. Therefore, modern voting tools may decrease average turnout, but nevertheless increase the quality of the voting outcome. Current models predict the opposite, but ignore the effect of different voting systems on the incentive for signaling.

*JEL* H0, Z13. Keywords: Signaling, social norms, voting, mail voting, modern voting tools.

[1]Patricia Funk, Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain. Tel.: + 34 93 5421942, E-Mail: patricia.funk@upf.edu.
A previous version of this paper has been circulated under the title "Theory and Evidence on the Role of Social Norms in Voting".
For comments and discussions, I thank Renee Adams, Iwan Barankay, Aleksander Berentsen, Tore Ellingsen, Christian Hilber, Steven D. Levitt, participants at the 2005 Public Choice Conference in New Orleans, the 2006 meeting of the Midwest Political Science Association in Chicago, the 2006 ESRC/HMRC Public Economics Conference in Bristol, seminar participants at Stockholm School of Economics, at Universidade Nova de Lisboa, at Pompeu Fabra, at University of Mannheim, Brunel Business School, University of Basel and at University of Bern.

# 1 Introduction

Modern voting tools are commonly seen as a panacea against decreasing voter participation. Since mail- or online voting lowers the costs from the voting act, standard economic reasoning suggests an increased mobilization of voters, for whom poll-voting was too costly. Indeed, in one of the very few economic analysis on this subject, Valenty & Brent (2000) posit: "If online voting becomes secure and convenient, voter turnout will increase. However, there may be a trade-off between quantity and quality ... If these [uninformed] voters [with no interest in politics] are encouraged to vote online, the results could be damaging."

In this paper, I challenge this view. Theoretically as well as empirically, I show that modern voting tools may decrease, rather than increase turnout. Secondly, despite a possible turnout decrease, the quality of the voting outcome need not be worse.

The different assessment of modern voting tools stems from a different assessment of the incentives to vote. While economic factors such as voting costs are one component of voting decisions, social factors are probably at least as important. Modern voting tools affect both, economic and social incentives to vote. Concerning the first, the obvious effect is a reduction in voting costs, with a positive effect on turnout. Secondly, mail- or internet voting renders the voting act unverifiable. If social pressure matters for voting decisions, the presence of modern voting tools provides an opportunity to escape. Therefore, the more signaling motives matter for voting decisions, the more distinctive the modern voting tools' trade-off between cost reduction and a reduction in social incentives.

While several authors have mentioned that social pressure might explain why certain people vote (see e.g. Schram & van Winden (1991)), a clean theoretical model with empir-

ically testable hypotheses is missing so far. The innovation of this paper is to present both a signaling type of voting model with testable predictions regarding social and economic drivers to vote and an empirical test thereof. The natural experiment, which allows me to shed light on individual's motivation to vote, is the introduction of optional mail voting in Switzerland.

Tailored to the empirical setup, I analyze signaling benefits. Those can be social esteem from showing up at the voting booth, the avoidance of social sanctions, or benefits from being perceived as a cooperator. Since poll-voting was never replaced by mail-voting (and neither were poll stations closed during the period of investigation), the main "social effect" was removal of social pressure, and not reduced benefits from chatting at the polls. I argue that the degree of social pressure depends on community size. A large number of anthropological studies have documented that social control is particularly strong in small and close-knit communities. People know each other and gossip about who does their civic duty and who does not.[1] In Switzerland, community size captures connectedness pretty well, with social control being particularly strong in small-sized communities.[2]

Formally, I analyze signaling incentives and the dependency on community size in a model with two types of individuals: cooperators and defectors. Given an exogenous social norm that a good citizen should vote,[3] cooperators feel guilty from not voting. Or, formulated in

[1] As the anthropologist Sally Engle Merry writes (1984, p. 272): "The role of gossip in achieving social control in stable, bounded, morally homogeneous, and close-knit communities where escape and avoidance are difficult differs markedly from its function in large, fluid, open and morally heterogeneous communities where escape and avoidance are realistic opportunities ... ".

[2] Apart from daily experience, data support this view. For instance, mobility is lower in smaller communities, which creates incentives to invest in social capital (Glaeser, Laibson & Sacerdote (2002), Zingg & Benz (2003)).

[3] Downs (1957) argued that democracy would fail if nobody voted and therefore, voting is a contribution to a public good.

Riker & Ordeshook's (1968) term, they derive a benefit from from fulfilling their civic duty ($D > 0$). Defectors, on the other hand, have a low sense of guilt and zero civic duty benefits ($D = 0$). Yet, both types may receive signaling benefits from showing up at the polls, e.g. through social esteem from people who are at the polls and from those who learn about the voting act through gossip. As community size increases, social connectedness decreases and signaling benefits as well. I assume that in the smallest community, signaling benefits exceed the voting costs. Given this setup, equilibrium turnout in two situations is derived.

In situation (A), voting is only possible at the polls. Proposition 1 establishes that there exists a unique and non-positive relationship between (PBE) voter turnout and community size.[4] The basic intuition is straightforward: since social rewards are very high in small communities, pooling (or partial-pooling) equilibria are sustainable where (a part of the) defectors pool with the cooperators in order to get social rewards. As communities get larger and the signaling benefits smaller, defectors don't vote anymore and only a separating equilibrium is sustainable. In situation (B), voting at the polls or by mail is possible. Comparing equilibrium turnout in situations (A) and (B), Proposition 2 states that: the introduction of postal voting has a non-positive impact on turnout in small communities and a non-negative impact on turnout in large communities.

Four key results emerge from the model: first, signaling can create a positive turnout in situations where nobody would vote. Second, there is a negative correlation between turnout and community size (for given instrumental benefits), third, turnout may decrease after the introduction of a low-cost voting tool and fourth, small and large communities react differently to it.

[4]For establishing uniqueness, I have to restrict the beliefs off-the-equilibrium path. Banks & Sobel's (1987) refinement "universal divinity" is applied for that purpose.

To empirically test the model, I analyze turnout at federal party elections ("Nationalratswahlen") between 1971 and 2003.[5] Federal elections bear the advantage that instrumental benefits are the same in every Canton and community, since votes are counted on a national level. In a difference-in-difference framework, I study the impact of mail voting on Cantonal turnout. The effect is very small (an increase in 2 percentage points), and not statistically different from zero.

This result is remarkable, since in Switzerland, every citizen (above 18) is automatically registered, and gets the election documents delivered at home. Since mail voting requires no more than dropping the documents into the nearest letter box, the costs from executing the voting act virtually drop to zero. As such, if voting costs were the main inhibitor to vote, one would expect a big increase in voter turnout. This did not happen. And, as the data suggest, the lack of a positive effect is not caused by endogeneity.[6]

However, signaling may counteract with the cost-reduction effect. To test this, I compare differently-structured Cantons and find indeed that there was a more negative turnout effect in Cantons with a high share of people living in small communities with less than 1000 inhabitants. These results are highly suggestive for signaling playing a role.

Several robustness tests are conducted to rule out competing explanations.

Since standard individual drivers to vote (instrumental benefits, expressive benefits, civic duty benefits) were not affected by the introduction of mail voting, the most plausible alter-

[5]Elections are held every 4 years. Since women received the right to vote in 1971, it serves as a natural starting point.

[6]If early adopters were the ones which had experienced a particularly strong decline in turnout, it may bias the the postal dummy downwards. However, there is no evidence in the data that past turnout mattered for the timing of introduction. Rather, the language area of a Canton and the prevalence of direct-democratic institutions are correlated with the timing of introduction. Since those time-invariant factors are captured by the Canton-Fixed Effects, they cannot lead to biased estimations.

native explanation is some heterogeneity in the poll-voting costs, which might be correlated with community size. To gain information on the communities' poll-voting costs, I had to conduct a survey. I chose the Canton Zuerich, because it is the biggest Canton in terms of population and has the best communal-level data. The Canton Zuerich has 171 communities and gave citizens the possibility to vote by mail in 1994. By sending out questionnaires to all the presidents if the communities, information about the number of poll-stations, the opening days, the opening hours of the polls and the share of mail-votes could be gathered.

Before turning to the poll-voting costs, I replicate the mail-voting effect with communal (instead of cantonal) turnout data. I find again that mail voting had no significant positive effect, and that there was even a decrease in small-sized communities with less than 1000 inhabitants. These results also hold if the neighbor Canton St. Gallen, which gave citizens the option of mail voting in 1978, is taken as a control group. As such, the lack of a turnout boost and the differential reaction between small and large communities seem to be very robust results.

As I then show, this differential reaction is not due to different poll-voting costs. In fact, community size is highly significant in explaining the pre/post-mail voting turnout drop, even when communities' poll-voting costs are controlled for. Also, the communities with the *shortest* (and most unfavorable) poll-station opening hours experienced the biggest drop after mail voting. While voting cost considerations predict exactly the opposite (i.e. the largest turnout increase in the communities where poll-voting was very inconvenient), signaling provides a rational. In the communities, where there is one poll-station, open for a very limited time (in one community, for instance, the poll is open only half an hour on the voting weekend's Saturday and Sunday), social pressure was maximal, and the escape effect

provided by mail voting as well.

As a last robustness check, I test the predictions of the model concerning the share of cooperators and the use of postal voting. While the first should be more positively affected in small rather than large communities, the use of postal voting should be higher in larger communities. While the survey provides information on the share of votes handed in by mail, I had to construct a proxy for the share of cooperators. The share of changed ("panaschierte") lists in the communities is used for that purpose. In Switzerland, a voter can either simply hand in the list of the favorite party, or change the list of the favorite party by deleting candidates and replacing them with candidates from other parties. Since a voter, who has changed the list, has (normally) spent more time on filling out the vote, I use the share of changed lists as a crude proxy for the share of cooperators. While there is weak evidence supporting the first prediction (maybe due to the noisy measure), evidence for the (ceteris paribus) higher use of mail voting in larger communities is strong.

With these robustness tests at hand, I am confident to summarize the main results of the study as follows:

(1): The addition of a low-cost voting mechanism to the traditional poll-voting did not significantly increase aggregate turnout in the Swiss Cantons. (2): The structure of the community matters for the reaction to mail voting. In small communities, turnout decreased after mail-voting was given, whereas in big cities, turnout increased. (3): The signaling explanation seems to be the most compelling in explaining the following empirical regularities: that turnout did not increase after the introduction of a low-cost voting scheme, that turnout relatively decreased in small communities, that communities with high poll-voting costs experienced a more negative effect on turnout than the others and that postal

voting is (ceteris paribus) more frequently used in larger communities.

In order to put the significance of these results in a broader context, I relate them to several strands of literature (Section 2). Section 3 outlines the theoretical model and Section 4 displays the basic regressions. Some robustness tests are conducted in Section 5 and Section 6 concludes.

# 2 Relation to the Literature

By arguing that social incentives (social pressure, social esteem) drive certain people to the polls, I shed light on the question "Why do people vote". Since ever, economists have perceived high turnouts levels as a puzzle, since a single individual is unlikely to have an impact on the electoral outcome (this so-called "Voter Paradox" goes back to Downs (1957)). The approach chosen in this paper falls into the category of models that aims at explaining voting decisions with a non-instrumental value, in particular a "social value".[7] Recent attempts along these lines have been group-based voter models, which see the individual as part of a social group. Group-utilitarian models take as given a share of ethical voters: these adopt a strategy that maximizes the group's welfare (if followed by all group members).[8] Other models emphasize the leader-follower relationship within a group. Even though it might not be instrumentally rational to vote for an individual, a group leader may have an effect if he can convince enough followers to vote. By exerting social pressure and appeals to the sense of civic duty, certain individuals might be convinced to cast a vote despite the non-pivotal

[7]Geys (2006) presents the latest review of the literature and points out to the difficulty of instrumental voting models to explain turnout levels. The first, who included a social component in form of a civic duty benefit were Riker & Ordeshook (1968).

[8]Group identity is created from opposing views on which candidates or proposals are good or bad (Feddersen & Sandroni (2002), Coate & Conlin (2004)).

role (see Uhlaner (1989), Shachar & Nalebuff (1999), Schram & van Winden (2001)). Alternatively, the desire to build up a reputation (Buffacchi (2001)), or the fear from social sanctions from co-citizens (see Knack (1992) and Opp (2001)) can motivate a citizen to fulfill his voting duty.[9] As Shachar & Nalebuff note (1999, p. 535): "We believe that social pressure is very important ... The more people in a social network that encourage a person to vote, the more likely that person is to vote and to encourage others to do the same."[10]

The key contribution of this paper is to present a model that actually allows to test the relevance of social incentives in voting decisions. The two features that allow me to construct such a model are (1): the observation long acknowledged by anthropologists that group-size and group-cohesion matter for the extent of social pressure, and (2): that observability of the voting act facilitates exertion of social pressure. Theoretically, this allows me to go beyond the contributions on social pressure by stating testable conditions under which social pressure generates higher turnout: namely, small community size and a poll-voting regime. Note further that the above-mentioned group-utilitarian and civic-duty models (as well as any calculus of voting equation including some sort of consumption benefit) predict an increase in turnout to the introduction of a low cost voting scheme and no different reaction depending on community size.[11] Due to these contrasting predictions from my signaling

[9]Empirically, Knack (1992) and Opp (2001) find that citizens are more likely to vote, if they have politically active friends or partners. However, this result could be due to sorting: citizens with a high interest in politics are more likely to choose friends and partners with similar interests.

[10]In a similar spirit, Amaro-de-Matos & Barros (2004) model individual voting decisions dependent on the decisions made by the individuals in his/her social network. Interdependent decision making can lead to very high, or very low turnout, depending on the initial feeling for or against voting. In contrast to Amaro-de-Matos & Barros (2004), I emphasize external enforcement of norms and highlight the role of information about and observability of individuals' voting acts.

[11]Since the sense of civic duty or the ethical preferences are not affected by a low cost voting tool, more people vote as voting costs go down. Similarly, the key mechanism in Shachar & Nalebuffs' paper is to look at leader effort, which is tied to the closeness of an election. Since the elections studied in this paper are proportional elections, the instrumental benefit per election and individual is roughly the same. The key novelty and difference to these models is that I relate the degree of social pressure by co-citizens to the

model, I believe to empirically demonstrate that social incentives matter, and to an extent that may even offset the cost-reduction effect of modern voting tools. As such, this is the first paper that presents a testable model on social incentives and voter turnout and an empirical assessment with a real-world experiment.[12]

Secondly, the paper is related to the "Law & Social Norms" literature, which tries to understand the interaction between Law and Social Norms. Several authors have mentioned that the desire for social esteem and signaling motives make people adhere to law (McAdams (1997), Cowen (2002) and Posner (2000)). This study supports the view that classical Law & Economics analysis can be flawed if it ignores the effects of laws or institutions on social norms.

Finally, the paper also sheds light on the effectiveness of modern voting tools. Most of the studies on this subject analyze a single election that has been carried out by mail or internet (see Southwell & Burchett (1997), Magleby (1978) and Solop (2000)). If only a single election is held by mail or by internet, certain voters may test this new voting tool out of curiosity. Long-run effects on turnout might be quite different, however. Since the Swiss Cantons introduced the (same) system of optional mail voting at different points of time, difference-in-difference-estimation (with well known econometric advantages)[13] is possible. However, since community structure seems to be crucial for the effectiveness of modern voting tools, the results may differ from country to country.

---

community size and the voting system, and subject it to an empirical test.

[12]Theoretically, Benabou & Tirole (2006) analyze the relationship between monetary and non-monetary incentives and pro-social behavior in a very elaborate and general way. My model is a bit more tailored to the voting case and differs in that financial incentives are irrelevant, esteem wants to be granted for the voting act of a certain type, and that community size matters for the degree of social pressure.

[13]Fixed effects do not only allow to control for heterogeneity between the Cantons/communities, but also to absorb common trends, which might influence turnout in a particular year.

# 3 A Theoretical Analysis

The standard calculus of voting model describes the payoff from voting as $P = I - C$. $I = p \cdot B$ denotes the instrumental benefit from voting (= probability of being decisive $p$ times relative gain from being decisive $B$) and $C$ the voting costs. Since $I$ is close to zero (and smaller than $C$), a positive payoff from voting can be achieved by introducing some consumption benefits.[14] Obviously, without further specifying where this consumption benefit comes from, the consumption benefit hypothesis remains tautological (see Geys (2006)).

My interest lies in the role of social norms, and therefore, I focus on the civic duty component $D$ of consumption benefits. Some heterogeneity is introduced in the simplest possible form: There are two types, which I label cooperators and defectors. Cooperators have internalized the social norm and feel a pleasure (or "warm glow") from fulfilling the civic voting duty ($D > 0$). The share of cooperators in the population is $\alpha$. Defectors denote the share $1 - \alpha$ of the citizens that do not get a warm glow from obeying the social voting norm ($D = 0$). The defector's only incentive to vote is being seen at the voting booth and potentially receiving esteem and other benefits from norm-adherence,[15] the sum of which I call *signaling benefits* $S$. With $I$ normalized to 0 (w.l.o.g.), the payoff from voting is $P = D + S - C$ for a cooperator, and $P = S - C$ for a defector.

[14] Those can be a utility from expressing one's opinion, the possibility of defining oneself over the actions (Schuessler (2000)), or a pleasure from satisfying the voting duty (Riker & Ordeshook (1968)).

[15] Being seen at the voting booth may not only bring immediate benefits through social esteem, but also future benefits if citizens are more interested in interacting with a cooperative type. Posner (2000) made the point that people who observe a voter tend to think of him as a cooperative and responsible type and may be more willing to engage in future interaction with him. See Bufacchi (2001) for a similar argument.

## 3.1 Description of the Signaling Game

Standard signaling games model the interactions between a sender (who moves first and sends a signal) and a receiver (who responds to the observed signal). In the case of voting, every citizen is a sender, who thinks about going to the polls or not, and simultaneously a receiver, who sees other people at the polls (or learns about voting acts by gossip) and has to decide about esteeming them or not. For modeling this situation in a standard way, I assume that there is a "representative receiver" instead of many, possibly different, receivers.

In order to do this, I have to assume that all the members of the communities have the same payoffs from granting esteem or no esteem, independent of their own type (cooperator, defector). It is likely that somebody who doesn't get this warm-glow feeling from contributing to a public good may still praise other peoples' contributions. Similarly, cooperators as well as defectors like to trade with a cooperator, and therefore, different types of receivers react in the same way to the signal of the sender.

I will think of the representative receiver as being sitting at the polls, observing voters and non-voters, and getting perfect information about who votes at the polls. The representative receiver spreads this information to other members of the communities. In smaller communities, a larger amount of people get the information about who votes and who doesn't due to increased gossip. Therefore, if the representative receiver decides to praise a voter (or non-voter), total signaling benefits are an increasing function of the value $V$, one attaches to being praised and a negative function of community size $s$.[16] To keep things simple, I define

[16]The number of people met at the polls is similar in the communities. Communities typically adapt the poll days and hours to the expected number of people handing in votes, so that the number of people met at the polls during a given time interval is unlikely to vary systematically with community size. If anything, one would expect a negative correlation between community size and the number of people at the polls, one knows. However, since this only reinforces the intuition that signaling matters more in small communities, it can be ignored w.l.o.g.

total signaling benefits (given esteem) as $S = \frac{V}{s}$.

As for the receiver's reaction, it strikes me as most plausible that one wants to reward a cooperator who votes, but not necessarily a defector who votes (see also Bufacchi (2001) for a similar view). Therefore, I model the case where esteem is granted for the voting act of a certain type (cooperator). However, similar results occur if the game is modelled as a pure signaling game (esteem for a cooperator), or a game, where the voting act is approved (irrespective of the type).[17]

## 3.2 A community where voting takes place at the voting booth

Figure 1 upper part describes the extensive-form representation of the game, where voting is possible at the polls. After having learnt his type (cooperator, defector), the sender can send the messages voting (V) or not voting (NV). The receiver in turn can either grant esteem (e) or no esteem (ne). The receiver wants to esteem the act of voting, but only if it is a cooperator who votes ($H > 1$).[18]

— insert Figure 1 here —

In the following, equilibria are derived under the assumption that signaling benefits in the smallest community exceed the voting costs $C$, i.e. there is a role for signaling. From the structure of the game, it is clear that if a defector votes, a cooperator votes as well, since he additionally has benefits $D$. Therefore, the only sustainable equilibria are pooling (both

[17] A lengthy discussion about the differences between these models is available upon request and also included in the earlier draft "Theory and Evidence of the Role of Social Norms in Voting". For the sake of brevity, I skip it here.

[18] Other specifications of the receiver's payoffs make sense as well. The main results do not hinge on this specific payoff-structure. Only a large enough payoff from esteeming a voting cooperator is necessary to make esteem for a voter (at least sometimes) a best response.

types voting or both types not voting), separating with cooperators voting and defectors not voting, semi-separating with defectors not voting and cooperators randomizing, and partial pooling where cooperators vote and defectors randomize. Furthermore, it seems intuitive that in the pooling equilibria, a cooperator is more likely to deviate from a non-voting equilibrium and a defector to deviate from the pooling equilibrium with both types voting. The equilibrium refinement of Banks and Sobel (1987) ("universally divine equilibrium") captures this intuition and restricts the out-of-equilibrium beliefs in the following way: a deviation from a pooling equilibrium with nobody voting occurs with probability one from a cooperator, and a deviation from a pooling equilibrium with both types voting occurs with probability one from a defector.[19]

Proposition 1 states that (given this refinement), there exists a unique relationship between community size and voter turnout.

**PROPOSITION 1:** With Banks & Sobel's (1987) equilibrium refinement "universal divinity", there exists a *unique* and *non-positive* relationship between community size and (PBE) voter turnout as well as between community size and equilibrium signaling benefits.

While the formal derivation of the equilibria is delegated to the Appendix, I will sketch the intuition. There are four cases to distinguish, depending on whether the civic duty benefit $D$ is bigger or smaller than the voting costs $C$, and whether the share of cooperators is high enough to sustain a pooling equilibrium ($\alpha > \frac{1}{1+2 \cdot H}$).[20]

[19]Formally, it suffices to show that e.g. in a pooling equilibrium with nobody voting, the set of mixed-best-responses, which induces a cooperator to defect, is strictly larger than the set of mixed-best-responses, which induces a defector to defect. The proof is straightforward and the same logic can be applied to the pooling equilibrium with both types voting.

[20]If the share of cooperators is high ($\alpha > \frac{1}{1+2 \cdot H}$), the best response for a receiver is to grant esteem in a pooling equilibrium where everybody votes. On the other hand, if there are enough defectors in the

Consider first the case where $C > D$, i.e. a situation where *nobody* would vote without signaling benefits. Figure 2 upper part depicts the equilibria of the game for $\alpha > \frac{1}{1+2\cdot H}$. As can be seen therefrom, the (only) equilibrium for small communities ($s < \frac{V}{C}$) is pooling: cooperators as well as defectors go to the polls and voter turnout is 100%. As community size increases (and $C > \frac{V}{s}$), only a separating equilibrium is sustainable, where cooperators vote and defectors abstain. While for the defectors, the voting costs $C$ exceed the signaling benefits $\frac{V}{s}$, the net-benefits from voting are positive for the cooperators as long as $\frac{V}{s} + D > C$, i.e. $s < \frac{V}{C-D}$. Therefore, as soon as the community size exceeds $s = \frac{V}{C-D}$, the only equilibrium is pooling with nobody going to the polls.

The important point is that in a setting, where nobody would vote without any signaling benefits, people go to the polls. Furthermore, since signaling benefits decline in the community size, the model predicts a non-positive relationship between voter turnout and community size.

— insert Figure 2 here —

If $\alpha < \frac{1}{1+2\cdot H}$, pooling equilibria are no longer sustainable for $s < \frac{V}{C}$, but partial-pooling equilibria instead. Cooperators vote, defectors randomize and receivers esteem with probability $\underline{p}$ (mixed best response).[21] Finally, the case with $D > C$ differs from the previous case in that a separating equilibria is sustainable for all community sizes, since cooperators always vote.

population ($\alpha < \frac{1}{1+2\cdot H}$), the best response in this type of equilibrium is to grant no esteem, which destroys it, since the defectors don't vote anymore.

[21] Since a mixed-best response is only optimal for a randomizing probability $x^* = \frac{\alpha}{1-\alpha} \cdot 2 \cdot H$ (see Appendix for details), the share of defectors voting remains constant between $\underline{s} < s < \frac{V}{C}$ ($\underline{s}$ denotes the smallest community). However, $\underline{p}$ increases in $s$ in order to make the defector indifferent between voting and not voting and reaches 1 at $s = \frac{V}{C}$.

## 3.3 A community with modern voting tools

Assume that in addition to poll-voting, citizens are given the means of postal or internet voting. Obviously, this brings a substantial reduction in the voting costs (compared to the transaction costs from going to the polls). Therefore, citizens have the choice between postal/mail voting (which causes voting costs $\underline{C}$) and poll voting, which causes voting costs $C$ ($\underline{C} << C$). In the following, I assume that $D > \underline{C}$, so that cooperators want to vote (for all $s$).

Figure 1 lower part presents the extensive form representation of the signaling game if postal voting is an option next to voting at the polls. While cooperators and defectors still can vote at the polls ($PV$ stands for poll voting), a certain part of the non-poll-voters ($NPV$) may choose to vote by mail ($MV$). As can be seen from the payoffs, mail-voting is strictly dominated by non-mail-voting ($NMV$) for the defectors,[22] and the opposite is true for the cooperators.

I assume that receivers have a (publicly known) prior $\delta^E$ about the share of cooperators voting by mail. Depending on the prior, expectations become self-fulfilling and may generate *multiple equilibria* in communities up to size $s = \frac{V}{C-\underline{C}}$ (see Appendix). For communities larger than that, the only equilibrium is separating, with all cooperators voting by mail and defectors not voting.

Figure 2 lower part depicts the R.E. equilibria in the case where the share of cooperators is high, i.e. $\alpha > \frac{1}{1+2 \cdot H}$. As can be seen therefrom, a rational expectations equilibrium, where cooperators and defectors pool, can *only* be sustained for the belief that a cooperator votes

[22]Since signaling benefits do not depend on the unobservable act mail voting/non-voting, mail voting causes costs of $\underline{C}$ to the defector, but no benefits.

at the polls for sure ($\delta^E = 0$). As soon as the receiver believes that a cooperator votes by mail with positive probability, the best response to a non-voter is to praise him, which destroys the defector's incentive to go to the polls. The only other sustainable equilibrium for small communities is then a separating equilibrium, with cooperators voting and defectors not voting (the dotted arrow accounts for the possibility that the pooling equilibrium remains).

Proposition 2 resumes the expected effects, the introduction of postal/internet voting has on voter turnout in small and bigger communities (the proof results in a straightforward way from comparing the possible equilibria in a given community size, see Appendix).

**PROPOSITION 2:** If the option of postal/internet voting is given next to poll-voting, it has a non-positive impact on (equilibrium) voter turnout in small communities and a non-negative impact on (equilibrium) voter turnout in large communities.

The intuition behind Proposition 2 is straightforward: For large communities $s > \frac{V}{C-D}$, all cooperators vote by mail, whereas defectors do not vote.[23] Therefore, the introduction of postal voting has either no impact on turnout (if cooperators did vote at the polls before ($D > C$)), or a positive effect, if cooperators preferred not to vote ($D < C$). Intuitively, the cheap voting mechanism can activate the cooperators which did not vote before due to high voting costs. In medium-sized communities ($\frac{V}{C} < s < \frac{V}{C-D}$), there is no effect on turnout, since a separating equilibrium remains (with possibly a share of cooperators newly voting by mail). In small communities ($s < \frac{V}{C}$), the pooling (or partial-pooling) equilibrium collapses as soon as the receiver believes that a share of cooperators votes by mail. Since the esteem

[23] In fact, cooperators already switch to mail voting for communities larger than $s = \frac{V}{C-\underline{C}}$. However, since I want to display the turnout effects, the interesting cases are communities with sizes larger than $\frac{V}{C-D}$ and smaller than $\frac{V}{C}$.

from going to the polls relative to not voting is reduced, voting doesn't pay for the defector anymore, and turnout decreases.

## 3.4 Possible extensions of the model and comparison with standard voting models

While the setup of the model seems very plausible to me, there are other reasonable specifications. For instance, one could model approval for the voting act (independent of the type), allow for a flexible amount of esteem (which increases in the senders probability of being a cooperator), or model informal sanctions for non-voting instead of esteem for voting. Since the main idea goes through all these models, a detailed comparison seems unnecessary at this point.[24]

Comparing the predictions of the signaling model with a classical calculus of voting equation $P = I + E - C \cong E - C$ (where $E$ denotes some consumption benefit), it is clear that in the latter case, the introduction of a modern voting tool increases turnout through its negative effect on $C$. Furthermore, community size shouldn't matter for the reaction. Similar predictions occur from the more recent group-based approaches. According to my best knowledge, the signaling model is not only different with respect to evaluating modern voting tools, it is also the first voting model that predicts a negative correlation between turnout and community size for given instrumental benefits. As such, comparing turnout and the turnout reaction to postal voting in different-sized communities sheds light on the role of signaling. Yet, the model also makes predictions concerning voter composition and the use of postal voting, which can be exploited as further evidence. In particular, the model

[24]A detailed analysis is available upon request, or included in the earlier draft "Theory and Evidence on the Role of Social Norms in Voting".

predicts a decrease in the share of defectors in small communities after mail voting, and a more frequent use of postal voting in larger communities.[25]

# 4 An Empirical Analysis of Mail Voting

## 4.1 Switzerland and the Mail Voting System

Switzerland is a small federalist country with roughly 7 million inhabitants. The country consists of 26 major districts (called "Cantons"), which are further divided into minor districts (called communities, "Gemeinden"). The 26 Cantons have their own constitution and legislative power and are free to pass laws, as long as they do not contradict with federal law.

As for the regulation of the voting process, the Swiss Cantons differ with respect to the use of modern voting tools (postal voting). While certain Cantons introduced the *option* of postal voting[26] already in the 80's, the majority gave citizens the possibility to vote by mail in the 90's (see table 1). In fact, a federal law was enacted in the end of 1994, which prescribed the Cantons to introduce the option of postal voting in order to facilitate voting for the citizens. From then on, there was only some variation left with respect to the time until the process of mail voting was organized.

— insert Table 1 here —

The Swiss voting procedure generally, but also the system of mail voting is very simple. In contrast to the United States, where voters have to register, every eligible citizen from

[25]The model predicts a complete switch to postal voting after a certain community size, while equilibria with poll- and postal voting are sustainable in smaller communities.

[26]Mail voting never replaced the polls, but was offered as a further option.

Switzerland *automatically* receives the (election) documents per mail. If mail voting is allowed, a return envelope is added to the election documents, so that the voter has only to put a stamp on the envelope and to drop it in the letter box. Since the alternative is to bring the filled-out documents to the polls, the transaction costs from postal voting are much lower.[27]

The goal of the empirical part is to analyze the effect of postal voting on turnout and to check for differences in different-sized communities (as predicted by the signaling model). The Cantonal variation in the timing of introduction provides an ideal setup, since it allows for difference-in-difference estimation.

## 4.2 Econometric Model and Results

The subject of investigation is voter participation at national parliamentary elections ("Nationalratswahlen") from 1971 to 2003 (elections are held every four years).[28] Similar to the American House of Representatives, the "Nationalrat" is the one (of the two legislative chambers), where the number of seats assigned to each region (Canton) corresponds to the population of the region (Canton). The election is for parties, with the different parties' weight being (roughly) determined by the proportion of votes received (proportional representation).

Analyzing turnout for this type of election bears several advantages. Firstly, the voting subject is unchanged over time. Secondly, since supply-side shocks affect turnout on a

[27]Certain Cantons had a system of mail voting upon request in earlier times. Since writing a letter to request a mail vote involves substantial time, the gain in transaction costs (compared to poll voting) is not evident. In the analysis and in table 1, I only count automatic mail-voting, where the gain in transaction costs is large.

[28]The year 1971 qualifies as a natural starting point, since it is the first year where women were allowed to vote.

national rather than Cantonal level, they can be absorbed by time-fixed-effects. For instance, if a party engages in harder competition by increasing advertising expenditures, it most probably affects the perception of this party (and turnout) in all Cantons. Thirdly, the voting day(s) are determined on a national level. Therefore, shocks on turnout due to weather conditions are similar in the Cantons as well.

Before turning to the econometric analysis, I display the raw data. Figure 3 depicts the development of turnout of the early introducers compared to a control group of similarly sized Cantons.[29] As can be seen from the raw data, the introduction of postal voting doesn't seem to have boosted turnout remarkably (the vertical line depicts the last election with poll-voting only).

— insert Figure 3 here —

However, since turnout depends on many more factors than the voting process, I want to control for population size, its communal structure, age structure, education, unemployment, income, and for whether Cantons have symbolic fines for non-voting (see footnote 29). The data appendix provides a detailed description of the variables and its sources.

Summary statistics are presented in table 2. As can be seen therefrom, early and late introducers differ most in terms of language. The fact that German and French-speaking Cantons attach different importance to mail voting is not surprising for a Swiss. For many political issues, German and French-speaking Cantons show consistently different voting

[29]Not depicted is the Canton St. Gallen. St. Gallen abolished the (minimal) voting fine at the same time as the introduction of postal voting. A negative reaction after mail voting could partly capture the abolishment of the voting duty. Even though voting fines in Switzerland have always been minimal (normally less than 1 $), they might nevertheless influence voting behavior (see Funk (2004)). Therefore, a simple graph for the Canton St. Gallen might be misleading.

patterns. There is even a term "Roestigraben", which refers to the different (voting) behavior in these two language-regions. Furthermore, large Cantons (in terms of population) and Cantons with more direct-democratic control were among the early introducers.[30]

— insert Table 2 here —

In the subsequent analysis, I want to estimate the impact of mail voting on turnout, controlling for the factors described above. To account for unobserved heterogeneity between the Cantons and general trends in voting behavior, I include Canton- and Time Fixed Effects. The model to be estimated is the following:

$$VT_{st} = \alpha_s + \gamma_t + b_1 \cdot Postal_{st} + b \cdot Z_{st} + u_{st} \tag{1}$$

$VT_{st}$ denotes voter turnout (in percentage) in Canton (state) $s$ in election $t$. $Postal$ is the mail voting dummy, $Z$ are the controls and $\alpha_s$ and $\gamma_t$ are the Canton- and Time-Fixed Effects.

Since the effect of mail voting is likely to depend on the Canton's community structure, I also estimate a model with an interaction term:

$$VT_{st} = \alpha_s + \gamma_t + b_1 \cdot Postal_{st} + b_2 \cdot Small \cdot Postal_{st} + b_3 \cdot Small + b \cdot Z_{st} + u_{st} \tag{2}$$

Small measures the proportion of cantonal residents living in small communities (i.e. communities with less than 1000 inhabitants). In Switzerland, there exist data about the number of people who live in communities with different sizes. The smallest size is "less than

[30]Interestingly, however, turnout did not differ between more and less direct-democratic Cantons.

1000 people", and the highest is "more than 100'000". All in all, 8 classes are distinguished. Obviously, the expectation is that $b_2$ is negative, since signaling is stronger in Cantons with a higher fraction of people living in small-knit communities.

Table 3 shows the regression results. Standard errors are depicted in parantheses and account for heterogeneity between as well as autocorrelation within Cantons.[31]

As can be seen from the first column, the introduction of postal voting did not significantly increase voter turnout. However, if we estimate a model with an interaction term (column 2), the coefficient before the postal dummy as well as the interaction term become statistically significant. Concerning economic significance, the Cantons, where no citizens live in small communities, had an increase in turnout of 6 percentage points after postal voting was introduced. In contrast, in the Canton with the highest share of people living in small communities (=36 %), the effect on turnout was minus 7 percentage points. As such, community structure matters. Columns 3 and 4 split the sample in Cantons with big cities and the ones without. While mail voting had a positive effect on Cantons with big cities, it was insignificant for the others. Again, a Canton's structure matters for the effect of mail voting.

The variable share small and most of the controls are quite correlated with the Canton-Fixed Effects, and therefore lose their significance. However, since our main interest lies in estimating the effect of postal voting, I prefer to include Canton-Fixed Effects, even though they capture part of the impact of the controls.

Summing up, we find clear evidence that the structure of the Canton matters for the

[31] Bertrand, Duflo and Mullainathan (2004) show that the failure to account for within-unit-autocorrelation can lead to an underestimation of standard errors in difference-in-difference estimations. As for the proposed solutions, clustering at the state-level performs quite well and is applied here.

overall effect of mail voting on turnout. Cantons with a higher proportion of people living in small communities had a more negative (or less positive) effect on turnout.

— insert Table 3 here —

## 4.3 Endogeneity

While this differential impact on differently-structured Cantons cannot be explained by endogeneity, endogeneity may bias the coefficient in column 1 downward. In particular, it would be worrisome if the cantons, who had experienced a large decline in turnout, gave postal voting first. To test this, I first check for different trends in turnout prior to adoption. As can be seen from the summary statistics, early and late adopters had no significantly different drop in turnout between 1971 and 1975.

To systematically test for longer-term effects of turnout on the introduction of postal voting, I estimate a probit model. The dependent variable is a dummy, which takes a value of 1, if the Canton switched to mail voting, and 0 otherwise. As can be seen from table 4, past turnout cannot explain the probability of introducing postal voting. Therefore, postal voting was not introduced after a particularly large drop in turnout. Rather, cultural differences (as manifested in the language barrier) and the weight, Cantons give to direct democracy seem to have driven adoption. Since these factors are time-invariant and captured by the Canton-Fixed Effects, they cannot bias the estimated postal dummy.

— insert Table 4 here —

# 5 Robustness

## 5.1 A replication with community level data

While the Cantonal data suggest that community structure matters, a more direct way to test the model is to analyze community level data. Since all votes have to be handed in at the community level, data exist for community turnout for exactly the same federal elections. Unfortunately, other community-level data are often unavailable. The Canton Zuerich though, who is the biggest in terms of population, has the most advanced statistical data and is taken as a starting point.

The Canton Zuerich has 171 communities, which are of very different sizes. More than twenty percent of the communities can be described as very small, with a population of less than 1000 inhabitants. Since the Canton Zuerich introduced postal voting in the year 1994, it seems interesting to compare turnout in the last election with poll voting only (conducted in 1991) and the first election with optional mail voting (in 1995).

Figure 4 upper part shows turnout at the 1991 parliamentary election. As can be seen therefrom, there is a strong negative correlation between community size and voter turnout. This negative correlation is compatible with the signaling idea that social pressure forced (or social benefits motivated) citizens to go to the polls in small communities. However, other explanations can account for this pattern as well (higher share of cooperators in small communities, higher consumption benefits in small communities).[32] The lower part of figure 4 depicts turnout at the next election, where postal voting was given as an option for the first time.

[32]Note that instrumental benefits are the same in the communities, since votes are counted on a Cantonal level. Therefore, the negative correlation *cannot* be caused by different instrumental benefits of communities.

— insert Figure 4 here —

From pure visual inspection from the graphs, one can see a drop in turnout in the small communities. The curve gets even flatter if turnout in the elections 1991 and 1999 is compared (not depicted).

However, since other factors could have caused this drop (e.g. a change in the age structure of population in small communities), I proceed with a more careful econometric analysis.

The panel-data study comprises a time-horizon from 1983-2003 (i.e. three elections without postal voting and three elections with postal voting). In a first set of equations, I analyze the impact of postal voting for the communities in the Canton Zuerich. Since all the communities were hit by postal voting at the same time, difference-in-difference estimation is not feasible. Therefore, I include the communities of the Canton St. Gallen as a control group. The Canton St. Gallen is a neighbor Canton of Zurich and has the same language, a similar party-structure,[33] and the Swiss-German culture. It is slightly smaller than Zuerich, with 86 communities, 10 percent of which have less than 1000 inhabitants. Since St. Gallen introduced the option of postal voting already in the year 1979 and has quite good community level statistics as well, it serves as a perfect control for difference-in-difference estimation.

The main regression equations are the following:[34]

$$VT_{ct} = \alpha_c + \gamma_t + b_1 \cdot Postal_{ct} + b \cdot Z_{ct} + u_{ct} \tag{3}$$

and:

[33]Both Cantons have a strong right-wing party (SVP).

[34]In the regressions with the Canton Zuerich only, a linear trend is used instead of time-fixed-effects.

$$VT_{ct} = \alpha_c + \gamma_t + b_1 \cdot Postal_{ct} + b_2 \cdot Small_c \cdot Postal_{ct} + b_3 \cdot Small_c + b \cdot Z_{ct} + u_{ct} \quad (4)$$

$VT_{ct}$ denotes voter turnout in community $c$ at election $t$. In equation (3), the coefficient of interest is $b_1$, measuring the average impact of postal voting on turnout. Equation (4) includes an interaction term. The dummy variable $Postal \cdot Small$ takes a value of 1, if postal voting is given and the community had (a minimal value[35] of) less than 1000 inhabitants, and 0 otherwise.

The control variables are similar to the analysis of Cantonal data and include the share of citizens in different age classes, as well as a measure for education.[36] Furthermore, I control for differences in the tax rates (see Table 5 for summary statistics).

— insert Table 5 here —

The regression results are presented in table 6. Column 1 depicts the effect of postal voting on turnout for the communities in the Canton Zuerich. The highly significant coefficient suggests that after the introduction of postal voting, turnout dropped by roughly 5 %.

— insert Table 6 here —

In column (3), the estimations are replicated with the Canton St. Gallen used as a control group. Since the results look similar, the Canton Zuerich's decline in turnout in the

[35]Minimal value stands for the lowest population in the six election years. Since this time-invariant measure is captured by the community-fixed effects, it is not displayed in the regression results.

[36]The age classes are sometimes a bit different from the age classes used in the Cantonal analysis. As for education, a proxy is built by taking the total number of people with a high-school degree and higher, in percentage of the community population older than 19.

mid-nineties seems to be truly caused by the introduction of postal voting, and not by an unknown tendency of citizens in small communities to vote less.[37]

Columns (2) and (4) show the dependency of the effect of postal voting on community size and confirm the prior finding that voter turnout dropped much more in smaller communities.

Not depicted for the sake of brevity are several variations of the basic estimations. Instead of defining small communities as communities with less than 1000 inhabitants, I re-estimated the models after defining small communities as communities with less than 1500 inhabitants. Again, the results are similar to before and highly significant. Also, since the population in each community at the elections is known, I re-estimated the model with an interaction term Postal Dummy times community population. Again, the results are confirmed: the coefficient is significantly positive, which indicates that larger communities had a smaller decrease in turnout after mail voting was given.

One final note concerns the different results for Cantonal and communal data. In the Cantonal regressions, postal voting was insignificant, while it was negative in the communal regressions. Consider though that a large negative effect in an averaged-sized community and a small positive effect in a large community can easily cancel each other out when it comes to Cantonal turnout (since the large communities make up a larger share of voters). The main purpose of these communal estimations was to check for different effects in different sized communities. The overall turnout effects can be more rigorously estimated with the Cantonal data.

[37]If the Canton St. Gallen had experienced the same development of turnout as the Canton Zuerich, the dummies *postal* and $postal \cdot small$ would become insignificant.

## 5.2 Accounting for communities' different Poll-Voting Costs

The different response of small and large communities to the introduction of mail voting seems to be a robust empirical regularity. Before I attribute this to social incentives, I want to be sure that it is not caused by different poll-voting costs of small and large communities. Theoretically, the observed response to mail voting could also stem from larger communities having higher poll voting costs. In that case, the introduction of convenient postal voting would cause a more positive/less negative effect on turnout in large compared to small communities. The previously estimated size-effect could then simply capture the reaction to different poll-voting costs rather than to different social incentives.

To test whether this was the case, I need measures for the communities' poll voting costs. Since no such data were available, I conducted a survey in the 171 communities of the Canton Zuerich. By E-Mail, the presidents of the communities were contacted and asked about several cost factors as well as the use of postal voting.[38] Overall, 110 responses were obtained. Since I conducted the survey in summer 2003, the elections 1983-1999 are covered.

With this information three cost variables were built: The number of poll stations per populated acres, the average number of days, the polls are open, and the average number of hours, the poll stations are open per day. The goal is to compare, which factors can better explain the cross-sectional turnout drop after mail voting: community size or poll-voting costs. The dependent variable is turnout 1991 - turnout 1999, which is mostly positive and therefore measures the drop in turnout.

[38]The precise questions for extracting this information were the following: How many poll stations do you have in your community? How many days are the different stations open and how many hours on each day? Were there any changes between the 1983 and 1999 elections, and if so, which ones? What share of votes was handed in by mail in the different elections?

Table 7 first column estimates the effect of community size for the 110 communities which filled out the survey. As can be seen therefrom, small communities had a 5 percent larger turnout drop between 1991 and 1999. The variable small community can account for 14 % of the variation in the turnout-drop between the communities. Column 2 analyzes the explanatory power of the measures for the poll-voting costs. Only opening hours seem to matter, but with an unexpected sign. Communities which had polls open for fewer hours, had experienced a larger drop than the others. This is against the expectation that in communities with short opening hours, the reaction to mail voting should be more positive. However, if poll-hours were very narrow, signaling might have been stronger and the escape effect larger. Column 3 adds measures for the opportunity costs of time from voting, but none of them are significant. In the last column, all controls are added. As can be seen therefrom, community size is highly significant and has roughly the same size as in column 1. Therefore, the larger turnout drop in smaller communities is not due to different poll-voting costs in these communities. Note further that community size is the strongest single predictor for explaining the turnout drop.

— insert Table 7 here —

## 5.3 Testing additional implications of the model

The signaling model does not only make predictions about voter turnout, but also about the composition of the voters and the use of postal voting. Concerning the first, mail voting should reduce the share of defectors if signaling matters. Furthermore, the share of

cooperators increases relatively more in small communities. As for the use of postal voting, the model predicts a higher use in large communities where social incentives don't matter.

As a proxy for the share of cooperators, I use data on the number of votes that have been modified. In the Swiss system of parliamentary elections, the voter gets a list of each party with the eligible party members on it. If the citizen just wants to vote for the favorite party, he does this by simply putting the list in the envelope/poll station. However, there is the possibility of replacing candidates from the favorite party with candidates from other parties ("Panaschieren"). Obviously, changing the lists by deleting and replacing names takes time and is probably done by voters only, who have a real interest and care about the issue. For the communities in Zuerich and the election years 1987 - 2003, data exist for the percentage of the lists that have been changed. Unfortunately, no such data have been collected for the communities in the Canton St. Gallen. Therefore, difference-in-difference estimation is not feasible and the effect of postal voting on the share of cooperators cannot be rigorously tested. The problem is that factors specific to a certain election might generate incentives to change the lists or not. Without a control group, no (election) time fixed-effects can be estimated, which could absorb such influences. However, the data allow to test whether small communities experienced a more positive effect than larger communities.

— insert Table 8 here —

Table 8 first column shows the regression results, with the dependent variable no longer being voter turnout, but the percentage of lists that have been changed. Consistent with the signaling model's prediction, the share of "cooperators" increased relatively more in small

communities, as soon as postal voting was introduced. Yet, the coefficient looses significance, if enough controls are added.

Columns 3 and 4 explain the use of mail voting in small and large communities. As predicted, small communities use postal voting to a lesser extent than large communities. Therefore, not only the predictions concerning turnout, but also concerning the share of cooperators and the use of postal voting seem to be confirmed by the data.

## 5.4 Alternative Measures of Signaling Benefits

The extreme scenario for signaling would be a community with one polling place, which is open on one day during a very limited period of time. Even though this extreme case does not exist, there are a couple of communities which come close to this description. For instance, in the community "Dorf", there is one poll-station, which is open for half an hour on both days of the voting-weekend. As such, there is a total of 1 hour time, where one can cast the vote. Obviously, due to the flexibility brought by mail voting, economic factors would suggest the biggest turnout increase in this type of community. However, signaling suggests exactly the opposite.

Subsequently, I would like to contrast the communities with very restricted poll-hours to the others. The main measure I consider is the total number of poll-hours, during which one can cast a vote. This measure is computed as the sum of opening hours over the different poll-stations. While the mean lies at 7 opening hours, there is substantial variation, with the minimum being 1 hour and the maximum 49 hours. Ceteris paribus, the more opening hours, the less strong signaling motives, but the lower the transaction costs from voting.

— insert Table 9 here —

Table 9, first column shows that communities with long opening hours reacted differently to mail voting than the ones with short opening hours. As can be seen from the highly significant coefficient, communities with long (short) opening hours had a more positive (negative) reaction to mail voting. This supports the view that social incentives matter. From a voting costs perspective, one would have expected a more positive turnout effect in communities with short opening hours, and therefore a negative sign. Columns (3), (4) and (5) decompose the indicator "total opening hours" into its components average number of poll stations, average number of opening days and average number of opening hours. Only poll hours has a significant sign. Accordingly, communities with short opening hours showed a larger decline in turnout. Again, this goes with the signaling intuition, but not with the cost saving argument.

# 6 Conclusions

In Switzerland, postal voting was introduced with the hope of slowing down the steady decrease in voter participation. Since the costs from mail voting are much lower than the costs from poll voting, the (economic) expectation was a substantial increase in voter turnout.

The empirical analysis of this paper shows that this did not happen. The introduction of optional postal voting increased aggregate Cantonal turnout at parliamentary elections by 2 percentage points, but the effect is statistically not different from zero. In contrast to this zero average effect, substantial effects in differently-structured Cantons were found. I calculate that the effect of postal voting on turnout was 6 percentage points in the Cantons where nobody lives in small communities with less than 1000 inhabitants. In contrast, the turnout effect was minus 7 percentage points in the Canton with the highest share

(i.e. 36%) of citizens living in small communities. A replication of the same procedure with community-level data confirms that the turnout decrease was particularly a "small-community"-phenomenon.

My explanation for this pattern is a change in social incentives, most likely in the external benefits of norm-adherence. In Switzerland, there exists a fairly strong social norm that a good citizen should go to the polls.[39] As long as poll-voting was the only option, there was an incentive (or pressure) to go to the polls only to be *seen* handing in the vote. The motivation could be hope for social esteem, benefits from being perceived as a cooperator, or just the avoidance of informal sanctions. Since in small communities, people know each other better and gossip about who fulfills civic duties and who doesn't, the benefits of norm-adherence were particularly high in this type of community.

With the introduction of postal voting, the signal from going to the polls got weakened. While before, a citizen who did not show up at the polls could be identified as a shirker, it may be a mail voter now. Therefore, in small communities where social pressure forced a substantial share of people to go to the polls, turnout decreased as soon as mail voting and the possibility of cheating was given.

My paper sets up a theory of voting, which for the first time, includes external benefits of norm-adherence ("signaling benefits"). The introduction of postal voting in the Swiss Cantons serves as a perfect experiment for testing this theory, since predictions from traditional voting theory and the signaling model differ. The empirical results are certainly more compatible with the signaling model than with traditional models of voting.

Even though social pressure seems to be the most plausible explanation for the empirical

[39]In many Cantons, it also was a legal norm, even though only five Cantons had very minimal fines for non-voting. See Funk (2004).

regularities, other social incentives might have played some role as well. For instance, if citizens enjoy meeting known people at the polls, the introduction of mail voting might cause a decline in turnout. In the first election with mail voting, however, a person who wants to socialize is likely to stop by and see, who else is there. As such, the immediate drop in turnout, I observe in the data, is unlikely due to reduced benefits from meeting at the polls. In the long run, however, as more and more people vote by mail, certain citizens, who showed up at the polls to socialize, might be turned off to cast a vote. On the other hand, new voters are recruited for whom poll-voting was too costly.

As for future research, I think that similar real-world experiments would create great value-added. While in the voting-case, it is not entirely clear that one wants to encourage voters, who vote as a by-product of signaling/socializing activities, there are public goods (like blood-donations), where an increase in the number of donors seems socially desirable per se. More evidence on which types of institutions foster pro-social behavior could serve as important guidelines for solving collective action problems. So far, it seems that providing space for signaling may achieve a certain effect.

# 7 References

**Amaro-de-Matos, Joao & Barros, Pedro P.** (2004). Social Norms and the paradox of elections' turnout. Public Choice, 121: 239-55.

**Banks, Jeffrey & Sobel, Joel** (1987). Equilibrium Selection in Signaling Games. Econometrica, 55(3): 647-661.

**Benabou, Roland & Tirole, Jean** (2006). Incentives and Pro-Social Behavior. Forthcoming: American Economic Review.

**Bertrand, Marianne, Duflo, Esther & Mullainathan, Sendil** (2004). How much should we trust Differences-in-Differences Estimates? Quarterly Journal of Economics, 119(1): 249-275.

**Bufacchi, Vittorio** (2001). Voting, Rationality and Reputation. Political Studies, 49: 714-729.

**Coat, Stephen & Conlin, Michael** (2004). A group-Utilitarian Approach to Voter Turnout: Theory and Evidence. American Economic Review, 94(5): 1476-1504.

**Cowen, Tyler** (2002). The esteem theory of norms. Public Choice, 113: 211-224.

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

**Feddersen, Timothy J. & Sandroni Alvaro** (2002). A Theory of Participation in Elections. Unpublished Manuscript.

**Funk, Patricia** (2004). Is there an expressive function of Law? An empirical analysis of voting laws with symbolic fines. Forthcoming: American Law and Economics Review.

**Geys, Benny** (2006). Explaining voter turnout: A review of aggregate-level research. Forthcoming: Electoral Studies

**Glaeser, Edward L., Laibson, David & Sacerdote, Bruce** (2002). An Economic Approach to Social Capital. Economic Journal 112(483): 437-457.

**Knack, Stephen** (1992). Civic Norms, Social Sanctions and Voter Turnout. Rationality and Society, 4: 133-156.

**Magleby, David B.** (1987). Participation in Mail Ballot Elections. The Western Political Quarterly, 40: 79-91.

**McAdams, Richard** (1997). The Origin, Development and Regulation of Norms. Michigan Law Review, 96: 338-433.

**Merry, Sally Engle** (1984). Rethinking Gossip and Scandal. In Black, Donald: Toward a General Theory of Social Control. Orlando: Academic Press.

**Opp, Karl-Dieter** (2001). Why Do People Vote? The Cognitive-Illusion Proposition and its Test. Kyklos, 54: 355-378.

**Posner, Eric A.** (2000). Law and Social Norms. Cambridge, Massachusetts: Harvard University Press.

**Riker, William H. & Ordeshook, Peter C.** (1968). A Theory of the Calculus of Voting. American Political Science Review, 62: 25-43.

**Schram, Arthur & van Winden, Frans** (1991). Why people vote. Free riding and the production and consumption of social pressure. Journal of Economic Psychology, 12: 575-620.

**Schuessler, Alexander A.** (2000). Expressive Voting. Rationality and Society, 12(1): 87-119.

**Shachar, Ron & Nalebuff, Barry** (1999). Follow the Leader: Theory and Evidence on Political Participation. American Economic Review, 89(3): 525-547.

**Solop, Frederic I.** (2001). Digital Democracy comes at age: Internet Voting and the 2000 Arizona Democratic Primary Election. PS: Political Science and Politics, 34(2): 289-293.

**Southwell, Priscilla L. & Burchett, Justin** (1997). Survey of Vote-by-Mail Senate Election in the State of Oregon. PS: Political Science and Politics, 30: 53-57.

**Uhlaner, Carole J. (1989)**. Rational Turnout: The neglected Role of Groups. American Journal of Political Science, 33(2): 390-422.

**Valenty Linda O. & Brent, James C. (2000)**. Online Voting: Calculating Risks and Benefits to the Community and the Individual. In Vogelsang, Ingo & Compaine, Benjamin M.: The Internet Upheaval. Cambridge, Massachusetts: The MIT Press.

**Zingg, Elvira & Benz, Matthias (2003)**. Mobilitaet, Wahlbeteiligung und Sozialkapital in Schweizer Gemeinden. Swiss Political Science Review, 9(2): 59-87.

# A Formal Appendix

## A.1 Voting at the booth

**PROPOSITION A**: Separating equilibria. Separating equilibria exist only with cooperators voting and defectors not voting. They are sustainable for the following community sizes: $s > \frac{V}{C}$ if $D \geq C$ and $\frac{V}{C} < s < \frac{V}{C-D}$ if $D < C$. **PROOF PROPOSITION A**: *A. Existence of separating equilibria with cooperators voting and defectors not voting.* Given correct beliefs $p = 1, q = 0$, the best-responses for the receiver are to grant esteem to a voter, and to grant no esteem to a non-voter.[40] Given these responses, a cooperator with $D \geq C$ always votes[41] and a cooperator with $D < C$ votes if $\frac{V}{s} + D - C > 0$ or $s < \frac{V}{C-D}$.[42] Given these best responses, a defector votes if $\frac{V}{s} - C > 0$ or $s < \frac{V}{C}$. Therefore, with $D < C$, a separating equilibrium exists for all $s_1 < s < s_2$, with $s_1 = \frac{V}{C}$ and $s_2 = \frac{V}{C-D}$. For $D \geq C$, a separating equilibrium exists for $s > \frac{V}{C}$. *B. Non-Existence of separating equilibria with defectors voting and cooperators not voting.* With correct beliefs $p = 0, q = 1$, the best responses of the receiver are are not to esteem a voter as well as a non-voter. Given these best responses, a defector never votes, which destroys this separating equilibrium.

**PROPOSITION B**: Pooling Equilibrium with both types voting. A pooling equilibrium with both types voting exists iff $\alpha \geq \frac{1}{1+2 \cdot H}$. For $\alpha > \frac{1}{1+2 \cdot H}$, it is sustainable for the following community sizes: $\underline{s} \leq s < s_1$, with $s_1 = \frac{V}{C}$. For $\alpha = \frac{1}{1+2 \cdot H}$, it is sustainable in $\underline{s} \leq s < \frac{\underline{p} \cdot V}{C}$. **PROOF PROPOSITION B**: 1. For $\alpha > \frac{1}{1+2 \cdot H}$, the best response of the receiver is to esteem a voter[43] and not to esteem a non-voter. Given these best responses, a defector votes if $\frac{V}{s} - C > 0$ or $s < \frac{V}{C}$. Therefore, a positive $s_1$ and hence a pooling equilibrium with both types voting exists since $\frac{V}{\underline{s}} > C$ by assumption (cf. p. 13). 2. For $\alpha = \frac{1}{1+2 \cdot H}$, the receiver is indifferent between granting esteem and granting no esteem to a voter. Denoting $\underline{p}$ the probability of esteem, a pooling equilibrium exists for $\underline{s} \leq s < \frac{\underline{p} \cdot V}{C}$. 3. For $\alpha < \frac{1}{1+2 \cdot H}$, the best response of the receiver is to grant no esteem, for a voter as well as a non-voter. This destroys the pooling equilibrium with both types voting, since a defector prefers not to vote.

**PROPOSITION C**: Pooling Equilibrium with nobody voting. A pooling equilibrium with nobody voting exists iff $C > D$. Given $C > D$, a universally divine pooling equilibrium with nobody voting is sustainable for communities of sizes: $s > s_2$ (with $s_2 = \frac{V}{C-D}$). **PROOF**

[40] For a voter, the payoff from granting esteem is $H$, which is greater than the payoff from granting no esteem $-H$. For a non-voter, the payoff from granting esteem is 0 and from no esteem 1.

[41] Since signaling benefits are strictly positive in this equilibrium, a cooperator with $C = D$ votes.

[42] Since $\frac{V}{\underline{s}} > C$, there exists a positive $s = \frac{V}{C-D}$.

[43] The expected payoff from esteem ($\alpha \cdot H$) is bigger than the expected payoff from no esteem ($(1-\alpha) - \alpha \cdot H$) iff $\alpha > \frac{1}{1+2 \cdot H}$.

**PROPOSITION C**: If $C \leq D$, cooperators vote, which destroys the pooling equilibrium with nobody voting. With $C > D$ and (refined) beliefs $q = \alpha, p = 1$[44], the best response of the receiver is to esteem a voter and not to esteem a non-voter. Given these best responses, a cooperator goes to the polls if $\frac{V}{s} + D - C > 0$, i.e. $s < \frac{V}{C-D}$. Therefore, the no-voting equilibrium is only sustainable for $s > \frac{V}{C-D}$.

**PROPOSITION D**: Hybrid Equilibria. 1. A partial pooling equilibrium where cooperators vote and defectors randomize exists in the community with size $s = \frac{V}{C}$ iff $\alpha \geq \frac{1}{1+2 \cdot H}$ or $\alpha < \frac{1}{1+2 \cdot H}$ and $x \leq \frac{\alpha}{1-\alpha} \cdot 2 \cdot H (= x^*)$; $x$ denotes the randomizing probability. A partial pooling equilibrium where cooperators vote and defectors randomize exists in all communities with sizes $\underline{s} \leq s \leq \frac{V}{C}$ iff $\alpha < \frac{1}{1+2 \cdot H}$ and $x = x^*$. 2. A semi-separating equilibrium where cooperators randomize and defectors do not vote exists for the community of size $s = \frac{V}{C-D}$ iff $C > D$. 3. Other hybrid equilibria do not exist. **PROOF PROPOSITION D**: *A. Exclusion of the existence of other hybrid equilibria:* The two other possible hybrid equilibria "defectors voting, cooperators randomizing", and "defectors randomizing, cooperators not voting" can be excluded, since cooperators always vote if defectors vote. *B. Existence of Equilibria.* In a partial-pooling equilibrium with cooperators voting and defectors randomizing, the beliefs of the receiver are $p = \frac{\alpha}{\alpha + x \cdot (1-\alpha)}, q = 0$, where $x$ denotes the defector's probability of voting. 1. For $\alpha \geq \frac{1}{1+2 \cdot H}$, the best response of the receiver is to esteem a voter and not to esteem a non-voter. A defector is indifferent between voting and non-voting iff $\frac{V}{s} - C = 0$ or $s = \frac{V}{C}$. Therefore, a partial-pooling equilibrium exists for $s = \frac{V}{C}$. 2. For $\alpha < \frac{1}{1+2 \cdot H}$ and $x < \frac{a}{1-\alpha} \cdot 2 \cdot H (= x^*)$, the best response for voting is to grant esteem, which leaves the defector indifferent at $\frac{V}{s} = C$. 3. For $\alpha < \frac{1}{1+2 \cdot H}$ and $x = x^* = \frac{a}{1-\alpha} \cdot 2 \cdot H$, the receiver is indifferent between granting esteem and no-esteem to a voter. A defector is indifferent between voting and not voting iff $\underline{p} \cdot \frac{V}{s} = C$. Therefore, for $\underline{s} \leq s \leq \frac{V}{C}$, there exists a partial pooling equilibrium with defectors randomizing with (constant) probability $x^*$ and receivers esteeming with probability $\underline{p} = \frac{C}{V} \cdot s$. 4. For $\alpha < \frac{1}{1+2 \cdot H}$ and $x > x^* = \frac{a}{1-\alpha} \cdot 2 \cdot H$, no partial pooling equilibrium exists. The best response for voting is to grant no esteem, hence, a defector never votes. In a semi-separating equilibrium where defectors do not vote and cooperators randomize, the beliefs of the receiver are: $p = 1, q = \frac{\alpha - \alpha \cdot x}{1 - \alpha \cdot x}$. Therefore, the best response is to esteem a voter and not to esteem a non-voter. Given these responses, a cooperator (with $C > D$) is indifferent between voting and not voting iff $\frac{V}{s} + D - C = 0$ or $s = \frac{V}{C-D}$. For $D \geq C$, he cannot be indifferent, but goes to the polls.

[44]Since a cooperator has a greater incentive to deviate than a defector (the set of mixed best responses, which make a deviation optimal, is strictly larger than for a defector), Banks & Sobel (1987) suggest to set the out-of equilibrium beliefs to $p = 1$ (the resulting equilibria are called "universally divine equilibria").

## A.2 Voting at the booth and postal voting

**PROPOSITION E**: Separating equilibria. Separating equilibria only exist with cooperators voting and defectors not voting. With $\delta^*$ denoting the equilibrium share of cooperators voting by mail, three types of (rational expectations) separating equilibria exist: 1. $\delta^* = 0$ for $\frac{V}{C} < s < \frac{V}{C-\underline{C}}$, 2. $\delta^* = 1$ for all $s$ if $\alpha > \frac{1}{1+2\cdot H}$, for $s > \frac{V}{C-\underline{C}}$ if $\alpha < \frac{1}{1+2\cdot H}$ and for $s > \frac{V}{C-\underline{C}} \cdot (1-\underline{q})$ if $\alpha = \frac{1}{1+2\cdot H}$. 3. $0 < \delta^* < \frac{1-\alpha}{\alpha\cdot 2\cdot H}$ at $s = \frac{V}{C-\underline{C}}$ (or at $s = \frac{V}{C-\underline{C}} \cdot (1-\underline{q})$ for $\delta^* = \frac{1-\alpha}{\alpha\cdot 2\cdot H}$). **PROOF PROPOSITION E**: *A. Existence of separating equilibria with cooperators voting and defectors not voting.* In a rational expectations equilibrium, the receivers prior $\delta^E$ about the share of cooperators voting by mail induces a share $\delta$ of the cooperators to vote by mail ($\delta^*$: $\delta^E = \delta$). Three types of rational expectations equilibria may exist: $\delta^* = 0, \delta^* = 1, 0 < \delta^* < 1$. 1. With $\delta^E = 0$, the (Bayesian updated) beliefs of the receiver are $p = 1, q = 0$. The best responses are to esteem a poll voter and not to esteem a non-poll-voter. A cooperator votes at the polls if $s < \frac{V}{C-\underline{C}}$, and a defector if $s < \frac{V}{C}$. Therefore, a separating equilibrium with $\delta^* = 0$ is sustainable for $\frac{V}{C} < s < \frac{V}{C-\underline{C}}$. 2. $\delta^* = 1$. With $\delta^E = 1$, the (Bayesian updated) beliefs of the receiver are $p, q = \alpha$. For these beliefs, the best responses are granting esteem to a non-poll voter if $\alpha > \frac{1}{1+2\cdot H}$, granting no esteem if $\alpha < \frac{1}{1+2\cdot H}$ and esteeming with probability $\underline{q}$ iff $\alpha = \frac{1}{1+2\cdot H}$. For a poll-voter, esteem is the best response if $p > \frac{1}{1+2\cdot H}$, no esteem if $p < \frac{1}{1+2\cdot H}$ and a mixed-best response for $p = \frac{1}{1+2\cdot H}$. For $\alpha > \frac{1}{1+2\cdot H}$ and all $p$, the separating equilibrium is sustainable in all community sizes, since cooperators as well as defectors have a larger payoff from not going to the polls. For $\alpha < \frac{1}{1+2\cdot H}$, the best response to a non-poll-voter is to grant no esteem. A separating equilibrium with $\delta^* = 1$ is sustainable for all $s$, if $p < \frac{1}{1+2\cdot H}$,[45] and for $s > \frac{V}{C-\underline{C}}$ otherwise. For $\alpha = \frac{1}{1+2\cdot H}$, the separating equilibrium is sustainable for $s > \frac{V}{C-\underline{C}} \cdot (1-\underline{q})$. 3. $0 < \delta^* < 1$. A cooperator can only be indifferent between poll-voting and non-poll-voting if he gets no esteem for non-poll-voting (or esteem with probability $\underline{q} < 1$). With beliefs $p = 1, q = \frac{\alpha\cdot\delta^E}{\alpha\cdot\delta^E + (1-\alpha)}$, granting esteem to a poll-voter and no esteem to a non-poll voter are best responses if $\delta^E < \frac{1-\alpha}{\alpha\cdot 2\cdot H}$. At $s = \frac{V}{C-\underline{C}}$, a separating equilibrium with $\delta^* < \frac{1-\alpha}{\alpha\cdot 2\cdot H}$ is therefore sustainable. Iff $\delta^E = \frac{1-\alpha}{\alpha\cdot 2\cdot H}$, a separating equilibrium is sustainable for $s = \frac{V}{C-\underline{C}} \cdot (1-\underline{q})$.

**PROPOSITION F**: Pooling Equilibrium with both types voting. The only type of R.E. pooling equilibrium with both types voting exists for $\delta^* = 0$, i.e. both types vote at the polls. For $\alpha > \frac{1}{1+2\cdot H}$, it is sustainable for the following community sizes: $\underline{s} \leq s < s_1$, with $s_1 = \frac{V}{C}$. For $\alpha = \frac{1}{1+2\cdot H}$, it is sustainable in $\underline{s} \leq s < \frac{\underline{p}\cdot V}{C}$. **PROOF PROPOSITION F**: 1. With $\delta^E = 0$, the beliefs of the receiver are $p = \alpha$, $q$. Poll-voting can occur if esteem is granted to a poll-voter and no esteem to a non-poll-voter. These best responses are optimal

[45] However, since the cooperator's incentive to deviate is greater than the defector's incentive to deviate, universal divinity restricts $p = 1$ and destroys this equilibrium.

if $\alpha > \frac{1}{1+2\cdot H}$ and $q < \frac{1}{2\cdot H+1}$[46]. Given these best responses, a defector votes if $\frac{V}{s} - C > 0$, or $s < \frac{V}{C}$. A cooperator votes at the polls if $\frac{V}{s} + D - C > D - \underline{C}$, which is always the case for $\frac{V}{s} - C > 0$. Therefore, a R.E. pooling equilibrium with both types voting at the polls ($\delta^* = 0$) exists for $\underline{s} \leq s < \frac{V}{C}$. For $\alpha = \frac{1}{1+2\cdot H}$, a poll-voter is esteemed with probability $\underline{p}$ and a R.E. pooling equilibrium with both types voting is sustainable for $\underline{s} \leq s < \frac{\underline{p}\cdot V}{C}$. 2. $\delta^E = 1$. Since no esteem is the best response to a poll voter, a defector never votes. Hence, no R.E. pooling equilibrium with $\delta^* = 1$ exists. 3. With $0 < \delta^E < 1$, the best response to a non-poll-voter is to esteem (recall that a defector never votes by mail). This destroys the pooling equilibrium, since defectors don't vote anymore.

**PROPOSITION G**: Hybrid Equilibria. Note first that no semi-separating equilibria exist, since a cooperator always votes ($\underline{C} < D$). A partial pooling equilibrium where cooperators vote and defectors randomize exists only for $\delta^* = 0$. It exists for $s = \frac{V}{C}$ iff $\alpha \geq \frac{1}{1+2\cdot H}$ or $\alpha < \frac{1}{1+2\cdot H}$ and $x \leq \frac{\alpha}{1-\alpha} \cdot 2 \cdot H (= x^*)$. It exists in all communities with sizes $\underline{s} \leq s \leq \frac{V}{C}$ iff $\alpha < \frac{1}{1+2\cdot H}$ and $x = x^*$. **PROOF PROPOSITION G**: 1. For $\delta^E = 1$: In this case, only defectors vote at the polls and hence, granting no esteem to a poll-voter is the best response. Therefore, a defector never votes, which destroys the PPE. 2. For $0 < \delta^* < 1$. In a PPE, a defector has to be indifferent between Poll-Voting and Non-Voting. Therefore, the net signaling benefits from going to the polls (compared to not voting) have to be equal to the voting costs. Since a cooperator prefers to go to the polls in this case,[47] a PPE with $0 < \delta^* < 1$ cannot exist. 3. For $\delta^* = 0$. With $\delta = 0$, the Bayesian updated beliefs are: $p = \frac{\alpha}{\alpha+(1-\alpha)\cdot x}, q = 0$. 1. For $\alpha \geq \frac{1}{1+2\cdot H}$, or $\alpha < \frac{1}{1+2\cdot H}$ and $x < \frac{a}{1-\alpha} \cdot 2 \cdot H (= x^*)$, the best response of the receiver is to esteem a poll-voter and not to esteem a non-poll-voter. Since a defector is indifferent between voting and non-voting iff $\frac{V}{s} - C = 0$, a partial-pooling equilibrium exists for $s = \frac{V}{C}$. 2. For $\alpha < \frac{1}{1+2\cdot H}$ and $x = x^* = \frac{a}{1-\alpha} \cdot 2 \cdot H$, the receiver is indifferent between esteeming and not esteeming a poll-voter. A defector is indifferent between voting and not voting if $\underline{p} \cdot \frac{V}{s} = C$. Therefore, for $\underline{s} \leq s \leq \frac{V}{C}$, there exists a partial pooling equilibrium with defectors randomizing with (constant) probability $x^*$ and receivers esteeming with probability $\underline{p} = \frac{C}{V} \cdot s$. 3. For $\alpha < \frac{1}{1+2\cdot H}$ and $x > x^* = \frac{a}{1-\alpha} \cdot 2 \cdot H$, the best response for poll-voting is to grant no esteem. Since a defector prefers not to vote in this case, no partial pooling equilibrium with $x > x^*$ can exist.

[46]Note that universal divinity sets $q = 0$, so that the equilibrium is universally divine.

[47]For a cooperator, the net signaling benefits must only cover the *surplus* in voting costs ($C - \underline{C}$).

# B Cantonal Data

Population: number of inhabitants (per Canton and year). Cantonal data on population were collected in the population census, which was conducted roughly every ten years. Intermediary values were obtained by linear interpolation.

Age: percentage of inhabitants in different age classes (per Canton and year). The following age classes are considered: 0-19, 20-39, 40-59, 60-64, 65-74. The data stem from the population census as well. Missing data were obtained by linear interpolation.

Higher Education: Number of "high-school degrees" per number of 15 to 19 year old people. High-school is put in quotation marks, because the Swiss school system is different from the American one. After six years of primary school (commonly attended from 6 to 12 years), there are three options: the "Realschule" (lowest level), the "Sekundarschule" (intermediate level) and the "Gymnasium" (highest level, denoted as "high-school"). While completion of the first two types of education takes between two and three years, "high-school" lasts six years. Therefore, "high-school" is commonly completed at age 18 and the number of "high-school"-degrees per number of 15 to 19 year old teenagers represents an adequate indicator for the frequency of attendance of higher education. Data source: Statistical Yearbooks of Switzerland.

Unemployment Rate: Number of unemployed persons per active population, in percentage. The active population consists of individuals working more than 6 hours per week. Unemployment Rates in Switzerland are measured in relation to the active population. Data Source: State Secretariat for Economic Affairs (seco).

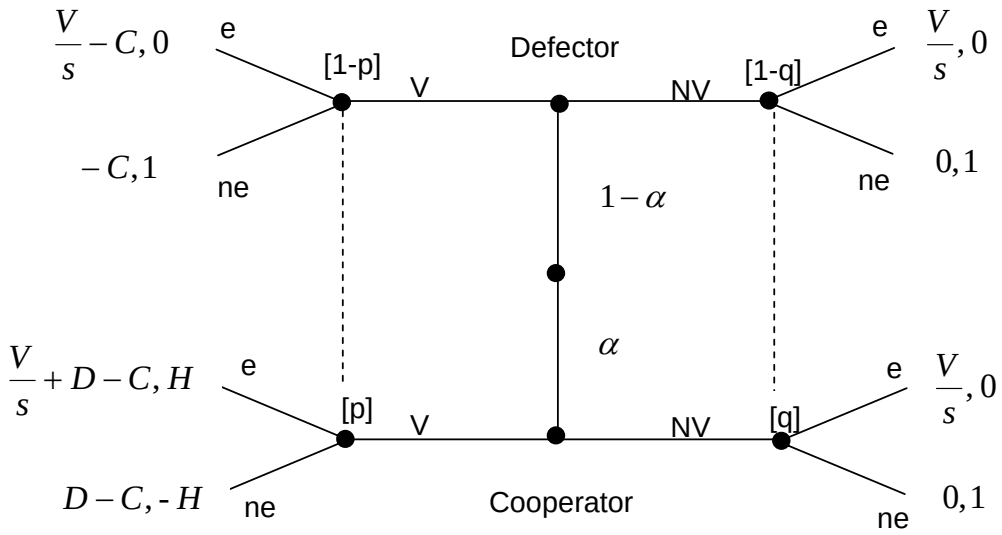
Income: Average Cantonal per capita income. The data stem from the bureau of statistics. One missing year (1971) was obtained by linear interpolation.

Close-Knit Community Structure: Percentage of people living in communities with less than 1000 inhabitants. The data stem from the population census as well. Missing data were obtained by linear interpolation.

Voting Fine: A dummy variable which takes a value of 1, if the Canton has a fine for non-voting. The fines are minimal and bear a symbolic rather than economic character.

**Figure 1: The Signalling Games**

**(1) Poll-Voting**




**(2) Poll-Voting and Mail-Voting**

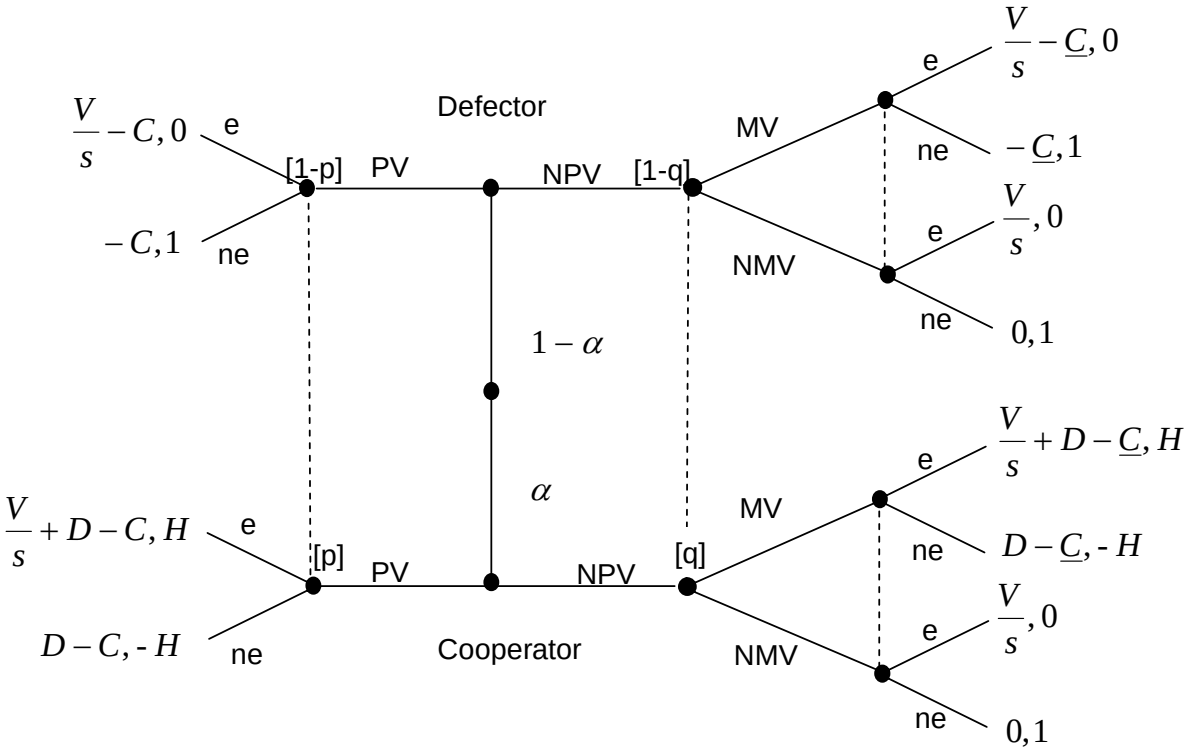
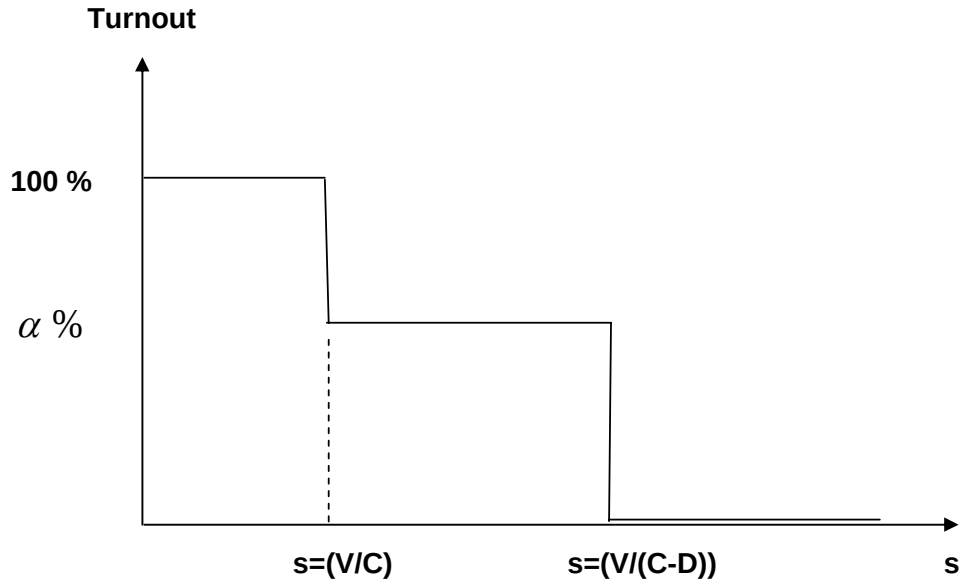


**Figure 2: Equilibria, C>D, share cooperators "high", i.e.** $\alpha > 1/(1+2\cdot H)$

**(1) Poll-Voting**




**(2) The Effect of Mail Voting on Equilibrium Turnout**

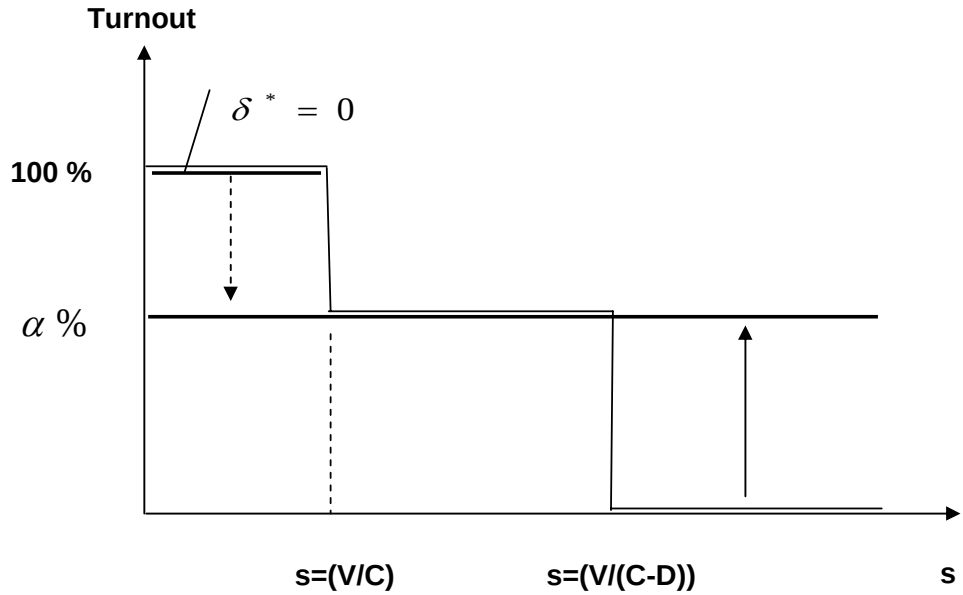
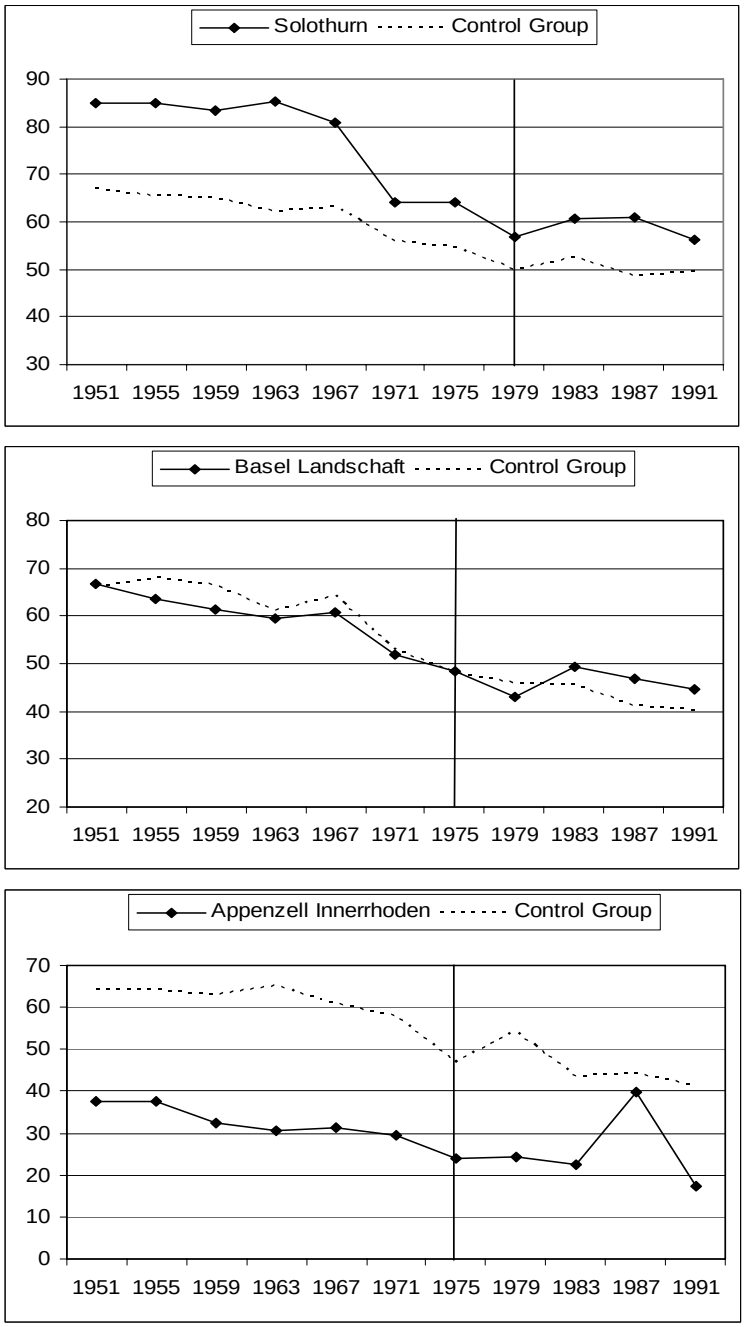
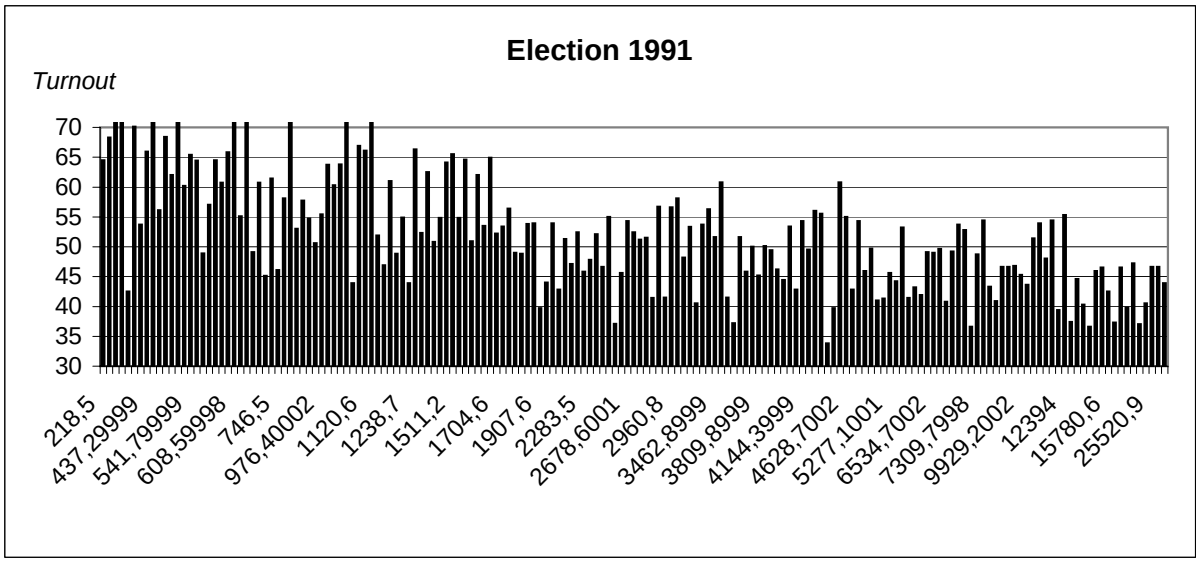


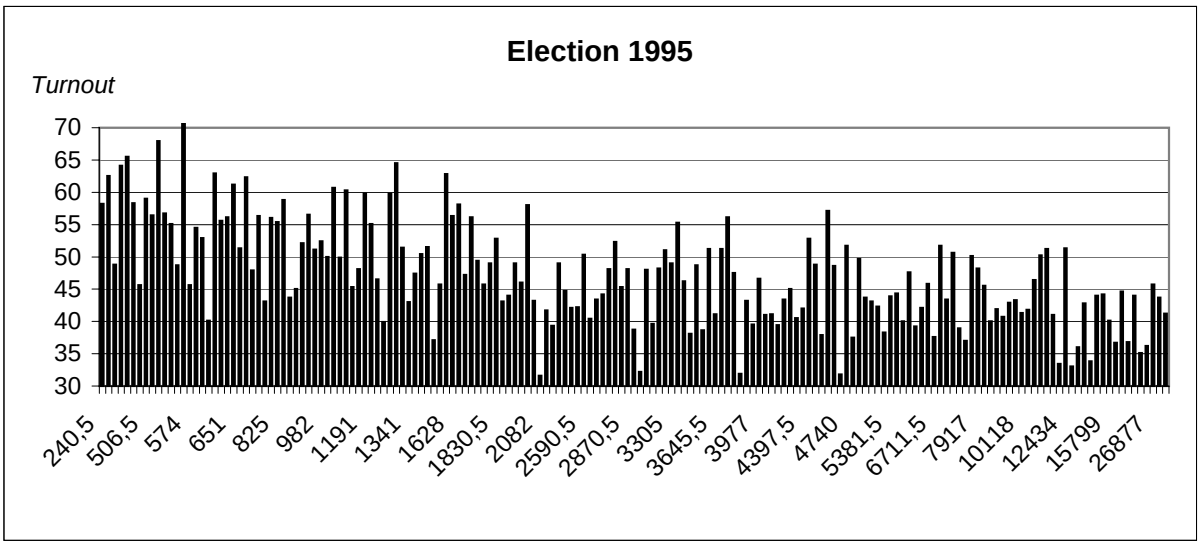
**Figure 3: The impact of postal voting on early introducer's turnout**




*Notes:* The control groups are similar sized Cantons, i.e. Cantons with more than 200'000 people in the Canton Solothurn, Cantons with more than 100'000 inhabitants in the Canton Basel Landschaft and Cantons with less than 100'000 inhabitants in the Canton Appenzell Innerrhoden. For the relevant Canton size, the average over the period 1951-1999 is taken. The vertical line depicts the last election where postal voting was not possible.

**Figure 4: Postal Voting and Turnout in the Communities of the Canton Zuerich**




*Note:* The graph displays turnout in the 171 communities of the Canton Zuerich. 1991 was the last election with poll-voting ony. 1995 was the first federal election, where postal voting was granted as an option next to poll voting.

**Table 1: Postal Voting in the different Cantons**

| Canton | Introduction Postal Voting | Canton | Introduction Postal Voting | Canton | Introduction Postal Voting |
|---|---|---|---|---|---|
| Basel-Land (BL) | 1978 | Basel-Stadt (BS) | 1995 | Wadt (VD) | 2002 |
| St. Gallen (SG) | 1979 | Zug (ZG) | 1995 | Neuenburg (NE) | 2003 |
| Appenzell Innerrhoden (AI) | 1979 | Freiburg (FR) | 1995 | Tessin (TI) | - |
| Solothum (SO) | 1980 | Schaffhausen (SH) | 1995 | Wallis (VS) | - |
| Thurgau (TG) | 1985 | Glarus (GL) | 1995 | | |
| Appenzell Aussenrhoden (AR) | 1988 | Uri (UR) | 1995 | | |
| Bern (BE) | 1991 | Graubuenden (GR) | 1995 | | |
| Aargau (AG) | 1993 | Obwalden (OW) | 1995 | | |
| Zuerich (ZH) | 1994 | Geneve (GE) | 1995 | | |
| Luzern (LU) | 1994 | Jura (JU) | 1999 | | |
| Nidwalden (NW) | 1994 | Schwyz (SZ) | 2000 | | |

*Notes:* The table depicts the year, postal voting was introduced (in addition to poll voting). A federal law was enacted in March 1994 and put into force in December 1994, which prescribed the Cantons to introduce a system of postal voting. Since there was no time limit specified, until when the Cantons had to change the system, there is some variation in the timing of introduction. "Early introducers" are the Cantons, which implemented a system of postal voting before 1995 (first column), and "late introducers" the others (second and third column).

**Table 2: Summary Statistics for Canton-Level Data**

| | 1971-2003 | | | | |
|---|---|---|---|---|---|
| | Early Adopters | | Late Adopters | | T Statistic |
| | Mean | Std. Dev | Mean | Std. Dev | Difference |
| **Turnout** | | | | | |
| Overall | 46.8 | 9.8 | 47.0 | 11.6 | 0.2 |
| Drop Turnout 1971-1975 | -5.6 | 4.0 | -6.4 | 11.6 | -0.2 |
| **Population Structure** | | | | | |
| Population (in 1000) | 370.8 | 354.2 | 176.4 | 144.6 | -5.8 |
| Citizens living in communities with less than 1000 inhabitants (%) | 6.7 | 4.7 | 14.7 | 11.5 | 6.5 |
| Citizens living in communities with less than 1500 inhabitants (%) | 23.3 | 10.3 | 26.3 | 17.5 | 1.4 |
| **Control Variables** | | | | | |
| Age 0 to 19 (%) | 27.4 | 3.8 | 26.6 | 4.7 | -1.4 |
| Age 20 to 39 (%) | 29.7 | 2.3 | 29.9 | 2.0 | 0.6 |
| Age 40 to 64 (%) | 29.1 | 2.8 | 29.3 | 2.6 | 0.7 |
| Age 65 to 74 (%) | 7.9 | 0.9 | 8.1 | 1.0 | 1.2 |
| 75 and Older (%) | 5.9 | 1.4 | 6.1 | 1.6 | 1.1 |
| Education (%) | 2.0 | 0.9 | 3.0 | 1.5 | 5.9 |
| Unemployment Rate | 1.1 | 1.3 | 1.5 | 1.7 | 1.9 |
| Pc. Income (in 1000 Sfr.) | 31.5 | 13.2 | 34.0 | 16.9 | 1.3 |
| Voting Fine | 0.1 | 0.3 | 0.1 | 0.3 | -1.2 |
| **Language** | | | | | |
| Non-German | 0 | 0 | 0.3 | 0.5 | 7.0 |
| **Direct-Democracy** | | | | | |
| Mandatory Law-Referendum | 0.7 | 0.5 | 0.4 | 0.5 | -4.1 |

*Notes:* The table reports summary statistics. Early adopters are the Cantons that implemented postal voting before the federal law prescribed it. Late adopters are the others. Turnout is measured in percentage of the eligible population, population is measured in 1000, community structure is captured by the share of cantonal recidents living in communities with less than 1000/1500 inhabitants, rural is the percentage of Cantonal residents living in rural areas. The share of residents living in age classes 0-19, 20-39, 40-64, 65-74, >75 is measured in percentage. Education measures the percentage of high-school degrees of the people aged between 15 and 19. The unemployment rate is defined as the percentage of unemployed persons of the active population (i.e. working 6 hours per week or more). Income is in annual 1000 Sfr., and Voting Fine a dummy for whether the Canton has a fine for non-voting. Language Non-German is a dummy variable taking values of 1 if the language is French or Italian, and Mandatory Law-Referendum takes a value of 1 if the Canton has such an institution for ordinary expenses.

**Table 3: Postal Voting, Turnout and Structure of the Canton**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dummy Postal | 2.3 | 6.5** | 3.2* | 2.8 |
| | (1.6) | (2.6) | (1.4) | (2.2) |
| Dummy Postal* Share Small | | -0.4** | | |
| | | (0.1) | | |
| Share Small | | 0.1 | | |
| | | (0.5) | | |
| Population | -0.05 | -0.04 | -0.03* | -0.05 |
| | (0.03) | (0.03) | (0.01) | (0.04) |
| Unemployment Rate | -0.7 | -0.9 | 0.5 | -0.8 |
| | (0.8) | (0.7) | (0.5) | (0.8) |
| Education | 0.6 | 1.3** | -0.3 | 0.7 |
| | (0.7) | (0.6) | (0.6) | (1.0) |
| Income | 0.1 | 0.0 | -0.1** | -0.1 |
| | (0.1) | (0.1) | (0.0) | (0.2) |
| Dummy_Fine | 3.0 | 4.6* | -6.4*** | 4.6** |
| | (2.2) | (2.5) | (0.4) | (1.9) |
| Age-Classes | YES | YES | YES | YES |
| Canton-Fixed Effects | YES | YES | YES | YES |
| Time-Fixed Effects | YES | YES | YES | YES |
| Split Sample | NO | NO | YES | YES |
| Subsample w. Big Cities | | | YES | NO |
| Subsample wt. Big Cities | | | NO | YES |
| Observations | 228 | 228 | 45 | 183 |
| R-squared | 0.99 | 0.99 | 0.99 | 0.99 |

*Notes:* Dependent variable is turnout at the federal elections 1971, 1975, 1979, 1983, 1987, 1991, 1995, 1999 and 2003. Dummy Postal measures the availability of Postal Voting and Small is the Cantonal share of citizens living in small communities with less than 1000 inhabitants. The Cantons with Big Cities (>100'000 inhabitants) are Zuerich, Basel Stadt, Bern, Geneva and Vaud. All estimations include Fixed-Effects. Robust standard errors clustered at the Cantonal level in parantheses. ***: significant at 1 percent, **: significant at 5 percent, *: significant at 10 percent.

## Table 4: Endogeneity of Adoption

| | Pr(Switch to postal) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Turnout last election | -0.003 | -0.004 | -0.001 | -0.002 |
| | (0.002) | (0.002)* | (0.002) | (0.002) |
| Turnout second-last election | -0.000 | -0.000 | 0.000 | 0.000 |
| | (0.002) | (0.003) | (0.001) | (0.002) |
| Turnout third-last election | | 0.001 | -0.001 | 0.001 |
| | | (0.002) | (0.002) | (0.002) |
| Election-Fixed Effects | NO | NO | NO | YES |
| Controls | NO | NO | YES | YES |
| Observations | 221 | 216 | 215 | 170 |
| Pseudo R-squared | 0.02 | 0.02 | 0.2 | 0.3 |

*Notes:* Dependent variable is the switch to postal voting, which takes a value of 1 in the year of the switching, and 0 otherwise. The model is probit, with marginal coefficients being reported. The controls in column (4) are the same as in table 3. Robust standard errors reported. ***: significant at 1 percent, **: significant at 5 percent, *: significant at 10 percent.

**Table 5: Summary Statistics for Community-Level Data**

| | 1983-2003 | | | |
|---|---|---|---|---|
| | Mean | Std. Dev | Min | Max |
| **Zuerich** | | | | |
| **Turnout** | | | | |
| Overall | 50.7 | 9.4 | 31.1 | 99.3 |
| **Population Structure** | | | | |
| Community Size | 7060 | 28290 | 213 | 365043 |
| **Control Variables** | | | | |
| Age 20 to 39 (%) | 30.2 | 3.8 | 18.7 | 44.8 |
| Age 40 to 64 (%) | 32.5 | 3.8 | 19.5 | 45.4 |
| Age 65 to 80 (%) | 9.0 | 2.4 | 2.6 | 18.3 |
| Education (%) | 12 | 6 | 2.1 | 42.1 |
| Tax Rate | 114.8 | 12.4 | 69 | 133 |
| Median Income (in 1000 Sfr.) | 45.3 | 5.9 | 29 | 69.3 |
| Percentage changed lists | 58.7 | 7.2 | 35.4 | 97.4 |
| **Survey-Data (110 communities)** | | | | |
| Number of poll-stations | 3.5 | 2.9 | 1 | 21 |
| Av. opening days (per poll-station) | 1.7 | 0.4 | 1 | 3 |
| Av. opening hours (p. station and day) | 1.2 | 0.4 | 0.3 | 3.5 |
| Total opening hours (=Av. Stations*Av. Days * Av. Hours) | 7.1 | 7.1 | 1 | 48.9 |
| Share votes handed in by mail | 9.5 | 17.9 | 0 | 79 |
| **St. Gallen** | | | | |
| **Turnout** | | | | |
| Overall | 44.7 | 6.7 | 31 | 68.1 |
| **Population Structure** | | | | |
| Community Size | 4741 | 8021 | 248 | 75664 |
| **Control Variables** | | | | |
| Age 20 to 39 (%) | 28.9 | 2.7 | 19.8 | 36.4 |
| Age 40 to 64 (%) | 28.2 | 3.2 | 21.1 | 38.6 |
| Age 65 to 80 (%) | 9.5 | 2.3 | 2.6 | 18.3 |
| Education (%) | 7.3 | 3.3 | 0.3 | 24.6 |
| Tax Rate | 148 | 16 | 90 | 175 |

*Notes:* The table reports summary statistics for the communities of the Cantons Zuerich and St. Gallen. Age structure is measured by the share of people in the different age classes. Education is defined as the share of citizens older than 19 holding a high-school degree and higher. The community tax rate is defined in percentage to the federal tax rate. Income is the yearly median income in 1000 Sfr. (only available for the Canton Zuerich). Data on Poll-Stations stem from the survey conducted in the Canton Zuerich. Poll stations are the number of polling places in the community. Poll days measures the average number of days, the polling stations are open and polling hours the average number of hours, the poll stations are open per day.

**Table 6: Postal Voting, Turnout and Community Size**

| | (1) ZH | (2) ZH | (3) ZH, SG | (4) ZH, SG |
|---|---|---|---|---|
| Dummy Postal | -5.3*** | -3.2*** | -6.4*** | -3.4*** |
| | (0.5) | (0.6) | (0.7) | (0.7) |
| Dummy Postal* Small | | -5.2*** | | -5.6*** |
| | | (0.9) | | (0.9) |
| Population (in 1000) | 0.1 | 0.0 | 0.1 | 0.0 |
| | (0.3) | (0.1) | (0.2) | (0.1) |
| Education | 0.4*** | 0.3** | 0.4*** | 0.3** |
| | (0.1) | (0.1) | (0.1) | (0.1) |
| Tax Rate | 0.0 | 0.0 | 0.0 | 0.0 |
| | (0.02) | (0.02) | (0.02) | (0.02) |
| Trend | -0.4*** | -0.4*** | | |
| | (0.1) | (0.1) | | |
| Control for Age Structure | YES | YES | YES | YES |
| Community-Fixed Effects | YES | YES | YES | YES |
| Time-Fixed Effects | NO | NO | YES | YES |
| Observations | 1564 | 1564 | 1564 | 1564 |
| R-squared | 0.99 | 0.99 | 0.99 | 0.99 |

*Notes:* Dependent variable is communal turnout at the federal elections 1983, 1987, 1991, 1995, 1999 and 2003. Dummy Postal measures the availability of Postal Voting and Small is a dummy measuring whether the community has less than 1000 inhabitants or not. Age structure is measured by the share of people in different age classes. Education is defined as the share of people older than 19 holding a high-school degree and higher. The community tax rate is defined in percentage to the federal tax rate. Standard errors are clustered at the community level. ***: significant at 1 percent, **: significant at 5 percent, *: significant at 10 percent.

**Table 7: Poll Voting Costs and the Response to Postal Voting**

| | (1) Drop Turnout 1991-1999 | (2) Drop Turnout 1991-1999 | (3) Drop Turnout 1991-1999 | (4) Drop Turnout 1991-1999 |
|---|---|---|---|---|
| Social Incentives | | | | |
| Small Community | 4.8*** | | | 4.7** |
| | (1.7) | | | (1.9) |
| Voting Cost Incentives | | | | |
| # Poll-Stations per 100 ha | | 0.1 | -0.3 | -0.6 |
| | | (0.3) | (0.5) | (0.5) |
| Av. # Poll Days | | 0.7 | 0.7 | -0.4 |
| | | (1.0) | (1.0) | (0.8) |
| Av. # Poll-Opening Hours p. Day | | -3.7*** | -3.8*** | -4.2*** |
| | | (0.9) | (0.9) | (1.2) |
| Opportunity Costs | | | | |
| Education | | | 0.3 | 0.3 |
| | | | (0.2) | (0.4) |
| Income | | | -0.3 | -0.1 |
| | | | (0.2) | (0.3) |
| Other Controls | | | | |
| Age Class 20-39 | | | | 0.6 |
| | | | | (0.8) |
| Age Class 40-64 | | | | 0.4 |
| | | | | (0.6) |
| Age Class 65-80 | | | | 0.9 |
| | | | | (0.8) |
| Population (in 1000) | | | | -0.1 |
| | | | | (0.1) |
| Tax Rate | | | | 0.1 |
| | | | | (0.1) |
| Observations | 110 | 110 | 110 | 110 |
| R-squared | 0.14 | 0.09 | 0.13 | 0.28 |

*Notes:* The dependent variable is the communities' turnout drop between 1991 and 1999. Small is a dummy measuring whether the community has less than 1000 inhabitants or not. Poll stations are per 100 hectares populated area. Poll days measures the average number of days, the polling stations are open. Polling hours are the average number of polling hours, the poll stations are open per day. Income is the median income in 1000 Swiss Franks. All other controls are defined as in Table 6. The control variables are taken for the year 1991. Robust standard errors in parantheses. ***: significant at 1 percent, **: significant at 5 percent, *: significant at 10 percent.

**Table 8: Voter Composition and Use of Postal Voting**

| | (1) % Changed Lists | (2) % Changed Lists | (3) % Postal Votes | (4) % Postal Votes |
|---|---|---|---|---|
| Dummy Postal*Small | 1.7** | 1.1 | -14.7*** | -10.1*** |
| | (0.8) | (0.8) | (2.8) | (2.8) |
| Education | | -0.2** | | 1.4** |
| | | (0.1) | | (0.5) |
| Age class 20-39 | | -0.3 | | -0.7 |
| | | (0.2) | | (0.7) |
| Age class 40-64 | | -0.2 | | -1.3* |
| | | (0.2) | | (0.7) |
| Age class 65-80 | | -0.4* | | 1.1* |
| | | (0.2) | | (0.6) |
| # Poll-Stations per 100 ha | | | -0.2 | 1.2 |
| | | | (1.8) | (1.7) |
| Av. # Poll Days | | | -22.6*** | -22.7*** |
| | | | (8.3) | (3.8) |
| Av. # Poll-Opening Hours p. Day | | | -11.7*** | 3.0 |
| | | | (3.4) | (1.9) |
| Community Fixed Effects | YES | YES | YES | YES |
| Time Fixed Effects | YES | YES | YES | YES |
| Observations | 853 | 853 | 504 | 504 |
| R-squared | 0.99 | 0.99 | 0.90 | 0.92 |

*Notes:* The dependent variable in the regressions (1) and (2) is the share of changed lists in the communities of the Canton Zuerich, elections 1987-2003. The dependent variable in regressions (3) and (4) is the share of votes handed in by mail. The sample contains the communities of the Canton Zuerich, which filled out the survey, and the elections are 1983-1999. The controls are defined as in Table 6. All estimations include Time- and Community Fixed Effects. Standard errors clustered at the community level in parantheses. ***: significant at 1 percent, **: significant at 5 percent, *: significant at 10 percent.

**Table 9: Alternative Proxys for Signalling**

| | (1) ZH | (2) ZH | (3) ZH | (4) ZH |
|---|---|---|---|---|
| Dummy Postal*Total Hours | 0.1*** | | | |
| | (0.1) | | | |
| Dummy Postal*Poll Stations | | -0.4 | | |
| | | (0.3) | | |
| Dummy Postal*Poll Days | | | -1.1 | |
| | | | (0.9) | |
| Dummy Postal*Poll Hours | | | | 2.2* |
| | | | | (1.2) |
| Education | 0.3 | 0.3 | 0.4 | 0.3 |
| | (0.2) | (0.2) | (0.2) | (0.2) |
| Age class 20-39 | -0.2 | -0.2 | -0.2 | -0.1 |
| | (0.3) | (0.2) | (0.3) | (0.3) |
| Age class 40-64 | 0.2 | 0.1 | 0.1 | 0.2 |
| | (0.2) | (0.2) | (0.2) | (0.2) |
| Age class 65-80 | 0.8*** | 0.7*** | 0.8*** | 0.8*** |
| | (0.2) | (0.2) | (0.2) | (0.2) |
| Community Fixed Effects | YES | YES | YES | NO |
| Time Fixed Effects | YES | YES | YES | YES |
| Observations | 550 | 570 | 555 | 550 |
| R-squared | 0.90 | 0.90 | 0.90 | 0.89 |

*Notes:* Dependent Variable is communal turnout at the federal elections 1983, 1987, 1991, 1995 and 1999 for the communities in the Canton Zurich. Total Hours is the sum of the total opening hours over the different poll stations in a community. Poll stations measures the number of poll stations in a community. Poll days stands for the average number of days, the poll-stations are open and poll-hours the average opening hours per station and day. All estimations include Community-Fixed Effects and Time-Fixed Effects. Standard errors clustered at the community level in parantheses. ***: significant at 1 percent, **: significant at 5 percent, *: significant at 10 percent.