USENIX Journal of Election Technology and Systems (JETS)
Volume 2, Number 3 • July 2014
www.usenix.org/jets/issues/0203

Mitigating Coercion, Maximizing Confidence in Postal Elections

JACOB QUINN SHENKER, California Institute of Technology
R. MICHAEL ALVAREZ, California Institute of Technology

1. INTRODUCTION

Elections have traditionally depended on procedural safeguards and best practices to ensure integrity and instill trust. By making it difficult for individuals to manipulate ballots undetected, these policies electoral malfeasance. Even so, it is clearly preferable to move beyond this kind of best-effort security and instead provide strong guarantees of integrity and privacy.

An emerging literature on voting systems has identified two distinct approaches towards this end: build trustworthiness into the voting system, or audit the election after-the-fact to verify its integrity. The first strategy is embodied by end-to-end verifiable voting systems, which use cryptography to prove to the voter that their ballot was cast and tallied as intended (Chaum, 2004; Chaum, Ryan, & Schneider, 2005; Ryan, 2005; Adida & Rivest, 2006; Rivest, 2006). However, these systems are predicated on strong assumptions and use complicated, difficult-to-understand cryptography to deliver their security guarantees. Instead of attempting to provide these strict assurances, the auditing approach aims to output statistical evidence that an election was conducted properly (Stark, 2008, 2009; Aslam, Popa, & Rivest, 2007; Rivest & Shen, 2012).

Neither the literature on verifiable voting systems nor the one on post-election audits adequately addresses the problems specific to postal voting.[1] Indeed, the nature of postal voting makes an audit difficult. Any audit begins with a complete paper trail; in a postal election, where ballots can (and are) lost in the mail, it many be impossible to maintain a complete chain-of-custody regarding the postal ballots (Stewart, 2010). An audit can check the tally of ballots that were received, but this does not address postal voters' primary worry: that their ballots are being lost or tampered with in the mail.[2]

Since a key feature of end-to-end systems is that a voter may ascertain for themselves that their ballot was received unmodified, end-to-end verifiability should be a natural application to vote-by-mail. Yet previous work on end-to-end voting largely neglects voting by mail. This is lack of attention arises partly due to the difficulty of handling coercion in the postal voting.

Like any other remote voting protocol, postal voting allows much more pervasive coercion than is possible with in-person balloting. Researchers have designed many Internet-based end-to-end remote voting systems with coercion-mitigation techniques. In all of these systems, the voter is interacting with the system through their computer, which is capable of performing sophisticated

[1] We are not aware of any work on end-to-end auditing in postal voting (where nondelivery of ballots is detected by the audit.) We are aware of three voting system designs which apply cryptographic techniques to postal voting, but none address coercion: Popoveniuc and Lundin (2007) describe modifications to Punchscan and Prt Voter to make them suitable for use in postal elections; the Remotegrity (Zagrski et al., 2013) extension to Scantegrity II primarily targets electronic ballot return, but in principle could be used for mail-in Scantegrity ballots. Neither proposal is particularly attractive from a usability perspective: in the example one-race election given in the Remotegrity paper, the voter must use no fewer than six distinct authentication codes to cast and verify their ballot. A third proposal which looks promising, (Benaloh, Ryan, & Teague, 2013), is not strictly end-to-end verifiable but attains a high degree of verifiability with minimal cryptography, much in the spirit in the current work. It has been brought to the authors' attentions that Andrew Neff has commercialized a privacy-preserving postal ballot tracking product (Dategrity Corp., 2005).

[2] A survey of California postal voters indicates that many postal voters doubt that their ballots were delivered to the election authority: Bergman (2012) finds a full 18% of postal voters in California reported being either a little or not confident that their ballot was delivered safely, whereas 19% reported the same levels of confidence that their ballot was accurately counted and processed. Bergman notes that these two questions may measure the same underlying dimension, as there is a correlation $r = 0.8$ and a Cronbach's alpha of > 0.7. This suggests that voters' doubts about their ballot counting can be largely explained by doubts about ballot transport. Other surveys bolster this claim, finding postal voters have lower confidence in elections across-the-board compared to in-person voters (Alvarez, Ansolabehere, et al., 2009).

cryptography. These systems leverage this ability to provide coercion-resistant voting; a paper-based protocol, as is needed for vote-by-mail, has no such recourse to sophisticated cryptography, since all cryptographic operations must be performed by the voter without computational aids. As such, vote-by-mail shares the main difficulty of Internet voting but cannot use the same mitigation techniques.

In this paper, we make a first attempt to consider the problem of coercion in the postal voting setting. We demonstrate that the defining features of postal voting constrain the design of any postal voting protocol, and thus many established techniques for end-to-end voting simply cannot be used (section 2). Along the way, we propose a scheme for providing auditability to vote-by-mail (section 3.3). While our resulting system does not provide *coercion-resistance* as defined by Juels, Catalano, and Jakobsson (2005), it provides a seemingly-weaker property of *coercion-evidence* of Grewal et al. (2013) (section 3.6). We argue that far from being weaker, this second property is more valuable in practice for convincing the electorate of the fairness of an election (sections 3.5 and 3.6).

Our design builds upon previous techniques. Our contribution is to recognize that the protocol of Grewal et al. (2013) works even if the ballot encryption step is postponed until after vote casting. This allows our system to offer two novel features: vote casting without cryptography and privacy-preserving publication of plaintext ballots.

The protocol we describe is not fully verifiable. However, given the increasing importance of securing vote-by-mail elections and the weaknesses inherent in traditional postal voting systems, we think it is an important step to bring vote-by-mail. To our knowledge, it is the first proposal to do this in a way that specifically addresses the problem of coercion, which otherwise would be a significant deterrent to its implementation in real-world elections. It remains to be seen in future work whether the techniques we describe can be profitably incorporated into a fully end-to-end verifiable voting scheme for postal voting.

Whether or not our particular design is worthwhile, it is undeniable that the postal voting problem has received dispropotionately little attention, given its importance. In 1984, 4.5 million people voted by mail; in 2012, 21 million postal ballots were cast. In the intervening time, two states (Oregon and Washington) began conducting elections entirely by mail (Stewart, 2010). In November 2013, Colorado also began delivering postal ballots to all voters (Bland, 2014). Attracted by the promise of convenient voting, the electorate in these states strongly approves of vote-by-mail (Southwell, 2004; Alvarez, Ansolabehere, et al., 2009). Election administrators around the country are pushing for widespread implementation of mail-only elections as a way to curtail costs and increase turnout (Bergman, 2012; Gronke, Galanes-Rosenbaum, Miller, & Toffey, 2008). These efforts are bearing fruit: seventeen states already allow mail-only elections under special circumstances (NCSL, 2013).

This rapid growth will only exacerbate problems which are already being caused by postal voting. Many studies have shown that postal voting is less reliable than other methods by a number of metrics (Stewart, 2010; Alvarez, Stewart III, & Beckett, 2013) and that voters tend to trust it less than voting in-person (Bergman, 2012; Alvarez, Hall, & Llewellyn, 2008; Alvarez, Ansolabehere, et al., 2009).

Since passage of the Help America Vote Act, many California precincts have successfully improved their election infrastructure by replacing antiquated lever and punchcard machines by optical scanners, but these gains have been neutralized by a concurrent rise in no-excuse absentee and other forms of voting by mail; the rise of postal voting in California between 2000 and 2008 led to an additional 73,868 residual votes (Alvarez, Stewart III, & Beckett, 2013). Stewart (2010) estimates that in the 2008 election up to 3.9 million attempts to vote by mail did not result in a counted ballot; the resulting lost vote rate of 22% is more than five times the estimated overall lost vote rate. Given these statisitcs, the continuing and swift adoption of vote-by-mail poses important problems for election administration.

2. SYSTEM DESIGN

2.1. The Power of Plaintext

Any end-to-end voting system must prove to each voter that their ballot reached the ballot box unmodified. This proof must contain some information about how the voter has marked his ballot, otherwise the voter would have no assurance that the ballot received by the system was not altered. If this confirmation consisted simply of the voter's ballot choices, this would surely be adequate proof, but the voter could show it to a vote-buyer and use it to demonstrate complicity in a vote-selling arrangement. Proposals for electronic voting systems use cryptography during vote casting to get around this problem (e.g., designated verifier proofs), yet we are considering a paper-based protocol where these solutions do not apply (Jakobsson, Sako, & Impagliazzo, 1996; Hirt & Sako, 2000; Saeednia, Kremer, & Markowitch, 2004).

There is a solution suggested by JCJ and similar systems (Juels et al., 2005; Clarkson, Chong, & Myers, 2008; Bursuc, Grewal, & Ryan, 2012). During registration, the voter receives one true and a number of fake credentials. Whenever a voter submits a ballot, they include one of these credentials. All ballots are posted unencrypted, and this system may still be coercion-resistant provided we deny the coercer knowledge of which ballots have real credentials (which will affect the tally) and those which have fake credentials (which will not affect the tally).

In this way, we may reveal plaintext (unencrypted) ballots without compromising ballot secrecy nor coercion-resistance: a coerced voter may capitulate to the coercer's demand and vote according to their orders, but using a fake credential. They may then vote normally at another time using their true credential. As long as the adversary does not come to learn which credentials are true and which are fake, he cannot distinguish compliance from noncompliance. (We will argue in the next section that the requirements of the registration phase force us to abandon the distinction between true and fake credentials, and make modifications accordingly. But for now we consider a system that does have distinct true/fake credentials.)

Of course, verifiability requires that the system prove to the voter that their true vote was counted while their fake votes were not. By posting plaintext ballots, the system may demonstrate that ballots reached the ballot box unmodified without using cryptography; the system then uses cryptography to prove that only the true ballots in the ballot box were tallied. In the case where only encrypted ballots are posted, cryptographic proofs are needed for both steps. In a strict sense, showing the voter his plaintext ballot is no better than publishing encrypted ballots, since cryptography is required for full verifiability either way. However chief among the doubts of a postal voter, unlike a polling-place voter, is ballot transport: will his ballot make it to the ballot box unmodified? Seeing a publicly-posted image of his ballot, just as he marked and submitted it, would give him substantial confidence that his ballot was received unmodified.

The ability to post plaintext ballots provides additional advantages. Consider that election authorities are free to publish both the ballot scan itself as well as how the ballot was interpreted by the optical scanner or canvassing board. This allows anyone, not just government-approved auditors, to examine disputed ballots for themselves. Previous experience with election auditing suggests this capability would do much to increase election trustworthiness. Election transparency advocates in Humboldt County, California, with the cooperation of the County Clerk, scanned and publicly posted anonymous ballot scans after two 2008 elections, discovering 197 lost votes that were missed by ballot tabulation software (Greenson, 2009). The ability of anyone to audit gives a level of transparency otherwise unattainable, since they can audit every aspect of the election up until the final tally, which requires distinguishing between true and fake credentials so as to only count ballots with true credentials (using cryptography to do this in a verifiable way is discussed in section 3.3).

Notice that using a credential system that allows for overriding or cancelling votes, e.g., JCJ's true/fake credentials, is the only way to enable the public auditing of ballot scans in a coercion-resistant way, because any scheme that involves posting full ballot scans is susceptible to a pattern voting attack or the inclusion of intentional identifying marks on the ballot. Using true/fake creden-

tials, we can allow an adversary to trace a ballot back to a particular voter, because the adversary still will not know if that ballot was submitted with a true or fake credential.

So far we have assumed that we may communicate true and fake credentials to each voter without an adversary learning them. To do this, we must register voters over an *untappable* communication channel. Almost all previous voting systems have assumed the existence of such a channel, but recently it has become clear that technology has made this assumption more dubious, and that channels once thought to be effectively untappable in practice can in fact be tapped en-masse and at low cost (Benaloh, 2013; Nixon, 2013). In the next section, we discuss modifications to the true/fake credential scheme that allows it to mitigate coercion even without using an untappable channel. In particular, we find that in the absence of an untappable channel the system cannot make a distinction between true and fake credentials, so all issued credentials must be equally valid. Instead of true votes overriding fake votes, we adopt a scheme whereby any credential may cancel the vote of any other credential issued to the same voter.

2.2. Registration Phase

Registration is the process by which voters authenticate themselves to election authorities before an election. In the US, this usually involves the voter providing their address, social security number or driver's license number, and their signature. When it is time to submit a ballot, reproducing these personal data on their ballot serves as a voter's proof to the election authority that they are authorized to cast a vote.

The registration phase takes on special importance in the context of an end-to-end verifiable, coercion-resistant voting system because both the end-to-end verifiability and coercion-resistance mechanisms depend on the election authority and the voter sharing cryptographic secrets before the election. The verifiability and privacy guarantees of these systems are predicated on the perfect untappability of the communication channel between the voter and the registrar, and if that assumption is broken so too are those guarantees. That is to say, such a voting system is only as secure as its registration phase. Almost all previous work on end-to-end voting systems assumes that voters are able to register via mail or in-person in a perfectly secure way.

However, a perfectly secure channel is not necessary if intrusions or delivery failures can be detected and mitigated. Consider the election registrar who wishes to securely transmit a voting credential to each voter. If the registrar sends the credential in a tamper-proof sealed envelope, an adversary may intercept the credential mailing, but it will either be delivered with a broken seal or will not be delivered at all. In either case, the intrusion would be detected.

Since an adversary may intercept any given credential mailing and prevent the voter from receiving it, we assume that the registrar sends enough credentials so that it is highly likely that the voter receives at least one.[3] In doing so, the registrar has transmitted at least one credential to the voter that was not intercepted by an adversary. One might think that our task, of securely communicating a credential to the voter, is thus accomplished. However, in the process the adversary may have intercepted a number of credentials. One might respond that the voter could notify the registrar which credential mailing succeeded in reaching him unintercepted, but this is not possible, because an adversary in possession of a valid credential is indistinguishable from a voter in possession of a valid credential. The only way out would be to presuppose some secret information shared by the voter and registrar that the voter may use to authenticate himself, which is merely begging the question.

We have thus found a way to communicate a credential securely to a voter, but a number of equally-valid credentials may be intercepted by an adversary. To the election authority, these credentials are indistinguishable, so they may all be used to cast a ballot. Regrettably, this means an adversary may cast a ballot on behalf of a valid voter. Note, however, that we would expect most registered voters to cast ballots. If a voter was observed to have voted once with one credential, and again with another credential, we may suppose that one of those credentials was intercepted by an adversary since there would be no legitimate reason for a single voter to cast multiple ballots.

[3] In addition, one can allow voters to request additional credentials.

This observation inspired the notion of *coercion-evidence*, introduced by Grewal et al. (2013). In our implementation, the registrar assigns to each voter a set of credentials. If more than one ballot is submitted using credentials in the same credential set, these ballots are not counted; instead, they are set aside and marked as evidence of coercion. The system publicly outputs the tally (not including cancelled votes) as well as the number of cancelled votes without disclosing which or whose ballots have been cancelled.[4] In the following sections, we discuss how to use cryptographic techniques to do this in a verifiable way (section 3.3), how these cancelled votes can be used in a post-election audit (section 3.5), and how this approach to coercion mitigation compares with the more common notion of coercion-resistance (section 3.6).

3. SYSTEM DESCRIPTION

3.1. Preliminaries

credential. A human-readable password that a voter includes with their ballot in lieu of their name, signature, or any other identifying information.

bulletin board $\mathcal{BB}$. An append-only bulletin board listing public election data, presumably hosted on the election authority's website.

threshold encryption scheme. An encryption scheme wherein the secret key is distributed amongst a set of trustees, wherein a threshold fraction (e.g., a majority) of trustees need to cooperate to decrypt a ciphertext (Brandt, 2006).

plaintext equivalence test (PET). Given two ciphertexts encrypted with the same public key, a multiparty protocol may be performed by the trustees to prove whether their corresponding plaintexts are equal without revealing the decryption key or the underlying plaintexts. (Jakobsson & Juels, 2000).

verifiable reencryption mixnet. A mixnet which takes as input a list of ciphertexts and outputs a permuted list of reencryptions of those same ciphertexts; it outputs transcripts that are sufficient to verify that the shuffle has been performed correctly. (Chaum, 1981; Jakobsson, Juels, & Rivest, 2002).

trustees. A set of entities that execute a series of distributed protocols to process the election data.

registrar. An trusted entity responsible for maintaining the voter rolls and putting tamper-evident seals on credential mailings.

We adopt a randomized threshold encryption scheme with a plaintext equivalence test, such as distributed El Gamal (Elgamal, 1985; Brandt, 2006). We write $\{\text{plaintext}\}^r_{PK}$ to mean the ciphertext produced by encrypting *plaintext* with the public key PK and randomness r.

3.2. Assumptions

We make the following assumptions:

1. At majority of trustees are honest. A majority of trustees may generate arbitrarily malformed credential data, including additional credentials for vote-stuffing, or may decrypt any encrypted data. In particular, they may track ballots through the mixnet, and thus discover which ballots were cancelled.
2. The registrar is honest. This is not nearly as strong an assumption as it seems: in any voting system, there must be some entity which decides who is allowed to vote and maintains the voter rolls. Only the registrar knows the real-world voter identities (i.e., names and addresses).
3. The envelopes in which credentials are mailed satisfy two properties: a voter may ascertain that an envelope has not been opened, and that the provenance of the envelope can be perfectly authenticated. In this way, an adversary cannot intercept the contents of the envelope without being

[4] Again, implementation of this system would allow a process for voters to request additional credentials, as well as a process whereby the voter can identify herself in person to the election authority and cast a final ballot that could be included in the tally were all of the ballots associated with her previously-issued credentials used by coercers. These procedures would ensure that coerced voters do not lose their ability to vote.

detected. The former may be attained using a tamper-evident seal; the latter may be satisfied using any techniques for authenticating paper documents (e.g., the security features used in paper money).

4. Malware cannot spoof the bulletin board; standard Internet security techniques may be used to prevent this possibility.
5. Ballots are divisible into sections (B_k^l in the notation introduced below) such that no ballot is uniquely identified by a voting pattern on any given one ballot section. It is left to future work to adapt the protocol to handle write-in candidates or rich ballot types such as IRV.

3.3. Protocol

1. Trustees participate in a distributed key generation protocol, publishing a public key PK to $\mathcal{BB}$ in such a way as no minority of trustees can reconstruct the private key.
2. Trustees execute a multiparty oblivious printing protocol (Essex & Hengartner, 2012) to generate and print credential mailings in invisible ink.[5] This protocol outputs both cryptographic information (posted to $\mathcal{BB}$) and physical credential mailings.
3. As part of the oblivious priting protocol, trustees generate a list of credentials, $(\text{voter}_i, \{\text{cred}_{ij}\}_{PK}^{r_{ij}})$, where cred_{ij} is the j-th credential associated to voter i.
4. The oblivious printing protocol also outputs credential mailings, where the j-th credentials mailing for a voter i includes the plaintext credential cred_{ij} printed in invisible ink on both an adhesive label and a receipt slip, both enclosed in a tamper-evident sealed envelope.
5. The registrar is assumed to begin with a list of voter ID numbers and their mailing addresses, $(\text{voter}_i, \text{address}_i)$.
6. For each printed credential mailing for voter i, the registrar fingerprints (using, e.g., Sharma, Subramanian, and Brewer (2011)) a blank sheet of paper, delivers it to the first trustee to be printed and requests a credential for voter i; when it is returned by the last trustee, with the credential fully printed, the registrar verifies that the invisible ink has not been activated and that the returned sheet was the same one it delivered (using the fingerprint). This prevents the last trustee from revealing the invisible ink, copying down the credential, and printing an identical copy to mail offthis would allow silent interception of credentials, which the protocol must prevent.
7. The registrar then seals the credential mailing in an envelope with a tamper-evident seal and mails it to address_i.
8. Before the election, the trustees generate and print a large number of credentials, and the registrar mails each voter one of these credentials. The registrar sends each voter additional credentials from this set periodically during the election period, or at the request of the voter.
9. Trustees post on $\mathcal{BB}$ a committment to $(\{\text{cred}_{ij}\}_{PK}^{r_{ij}}, \{\text{voter}_i\}_{PK}^{p_{ij}})$, a list of encrypted credentials and the encryption of their associated voter identities.
10. During the balloting period, voters download a ballot form from $\mathcal{BB}$, print it, and fill it out. The voter then chooses any of the credential mailings they have received, opens it, and uses a special pen to activate the invisible ink on the mailing to reveal the credential. Verifying that the credential on the receipt slip matches the credential on the adhesive label, they place the label on the ballot and mail it back; they keep the receipt slip so that they may find their ballot on $\mathcal{BB}$ when its scan is posted.
11. After the balloting period has closed, trustees open the commitment to $(\{\text{cred}_{ij}\}_{PK}^{r_{ij}}, \{\text{voter}_i\}_{PK}^{p_{ij}})$. Furthermore, trustees jointly decrypt each $\{\text{cred}_{ij}\}_{PK}^{r_{ij}}$ and post proofs of correct decryption to $\mathcal{BB}$. By associating each decrypted plaintext credential with its encryption, the trustees then post

[5] The protocol is a generalization of the usual two-party visual cryptography scheme (Chaum, 2004), but extended to distribute trust amongst multiple printers. The printers each generate shares of a secret (in our case, a credential); each printer in turn prints its share in invisible ink, so that printers may not read previously-printed shares as they are printing their own. After all of the printers have printed their share, the invisible ink may be developed with a special pen to reveal the secret. The protocol guarantees the printing will be correct unless a majority of trustees conspire, and that none of the printers will know the secret (Essex & Hengartner, 2012).

$(\text{cred}_{ij}, \{\text{voter}_i\}_{PK}^{p_{ij}})$. Note that here it is crucially important that each voter identity $\{\text{voter}_i\}_{PK}^{p_{ij}}$ is encrypted with unique randomness; otherwise, by matching plaintext credentials with the same voter identity ciphertext, credentials belonging to the same voter could be linked.

12. The election authority scans all ballots, posting each ballot's credential, image, and textual representation on $\mathcal{BB}$. We write B_k^l for the ballot data corresponding to the l-th race on the k-th ballot and cred_k for the credential included on the k-th ballot. Using the output of step 5, an encrypted voter identity $\{\text{voter}_i\}_{PK}^{p_{ij}}$ may be associated to each ballot, where $\text{cred}_{ij} = \text{cred}_k$.
13. For each race l:
    i. Trustees post $(\{\text{voter}_i\}_{PK}^{p_{ij}}, \{B_k^l\}_{PK})$ to $\mathcal{BB}$.
    ii. Trustees execute a verifiable reencryption mixnet to shuffle $(\{\text{voter}_i\}_{PK}, \{B_k^l\}_{PK})$, posting the transcript and proofs to $\mathcal{BB}$.
    iii. Trustees execute plaintext equivalence tests between the encrypted voter identity for each pair of ballots, posting transcripts to $\mathcal{BB}$. The result is a list $(\{\text{voter}_i\}_{PK}, \{B_{n_1}^l\}_{PK}, \{B_{n_2}^l\}_{PK}, \ldots)$ where $B_{n_i}^l$ are the ballots whose encrypted voter identities have been shown to be plaintext equivalent.
    iv. Trustees jointly decrypt ballot information B_k^l in the case where only one ballot has been associated with a given voter identity, and post it to $\mathcal{BB}$ along with a proof of correct decryption. The tally is simply the sum of these.
    v. The final output is the tally, the decrypted ballot information $\{B_k^l\}$ for non-cancelled votes, and the number of cancelled votes (number of voter identities corresponding to cancelled ballots, *not* the number of cancelled ballots).

3.4. Attacks or Errors Prevented

— **Trustees adding or deanonymizing ballots.** No minority of the trustees can add a valid ballot to $\mathcal{BB}$ (since doing so would require generating a new credential, which requires the cooperation of a majority of trustees). Similarly, no minority of the trustees can associate a ballot with a voter (since doing so would require). Note that a majority of trustees still can do so, and this ability may be desirable for the purpose of investigating coercion after the fact.

— **Malformed credential mailings.** The oblivious printing protocol includes verifiability steps to ensure that credential mailings are printed correctly unless a majority of trustees conspire (Essex & Hengartner, 2012). We assume that a voter can distinguish valid credential mailings sent by the election authorities from spoofing attempts sent by an adversary.[6]

— **Removing or modifying ballots.** Any voter can look up the ballots corresponding to their credentials and verify that they match the ballots they submitted. The voter may make scans or copies of their ballots before submitting them if they wish, and they may use these as evidence of manipulation in case $\mathcal{BB}$ does not contain matching ballots.[7]

— **Deanonymization via bubble fingerprinting.** The coercion-mitigation property holds even if voters may voluntarily deanonymize their ballots, because voters will know to adhere to the vote-buyer's or coercer's demands in the deanonymized ballot but may submit a second, unidentifiable ballot to cancel it. It no longer holds if voters accidentally make their ballot identifiable, because the voter will not know to cancel their ballot. Calandrino, Clarkson, and Felten (2011) describe a machine-learning procedure that could be able to link ballots to individuals by examining the way in which they fill in the optical-scan bubbles. To combat this, ballot scans could be posted with the actual marks blurred or masked by solid black squares. To ensure that this masking is done correctly, a cut-and-choose-style protocol could be used: a limited number of bubbles could be unmasked, selected using a trusted random beacon, such as stock market data (cf. Clark, Essex,

[6] This can be done using well-known techniques, for example, those used in authenticating paper currency.

[7] Forensic techniques such as paper fingerprinting (Sharma et al., 2011) may be of use in proving that their ballot has been manipulated.

and Adams (2007), Clark and Hengartner (2010)). The number of unmasked bubbles per ballot would be chosen so that they would provide insufficient data for a Calandrino-style attack.

— **Misprinted ballots.** Ballot images are posted on $\mathcal{BB}$, so anyone may verify that the ballot was printed correctly and that the ballot design complies with election law.

— **Malicious optical scanner or canvassing board.** A textual representation of each ballot will be posted together with a high-resolution image of each ballot on $\mathcal{BB}$. Consistency between the two can be checked manually or with the assistance of ballot-auditing software (Kim et al., 2013). The textual representation is the output of the optical scanner, or in the case of a dispute, the interpretation of the canvassing board, and as such both may be verified. To our knowledge, this is only voting system which allows the publishing of ballot scans in a way insusceptible to coercion, and as such is the only system that allows canvassing board decisions to be audited by anyone.

— **Active coercion.** A voter can comply with any request the coercer makes, including pattern voting or abstention, and can turn over all of their credentials. The voter may then obtain a new credential and use it to submit another ballot, thus cancelling the coerced vote. The only way to successfully and undetectably coerce a voter is to intercept all of their communications from the beginning of the registration period (when the first credential is distributed) to the end of the balloting period; since this time period may be months or years, it would require enormous resources to coerce a significant number of votes.

— **Vote selling.** Again, a voter can reveal to a vote-buyer all their credentials and all of their submitted ballots, and the vote-buyer can indeed verify that these ballots appear on $\mathcal{BB}$, but the vote seller may at any time obtain a new credential and use it to submit another ballot, thus cancelling the sold vote and marking it as coercion-evidence. Note that voter-sellers are disincentivized from allowing sold votes to count, since if they sold their vote once, they could sell it again to additional vote-buyers; the multiple ballots submitted by these vote-buyers will all be cancelled. Thus, the price of a sold vote will be driven to zero.

— **Loss of privacy.** Because the registrar distributes credentials to the voter in a tamper-evident sealed envelope, a voter can trust that any credential mailing that arrives intact has not been intercepted. Thus, the only way an adversary can learn of a voter's true vote is if he discovers all of the voter's credential mailings after the voter has opened them. By hiding his credential mailings, a voter can make it arbitrarily difficult for his privacy to be violated. Note that the voter can always voluntarily give up privacy by revealing their credentials, but as we have seen above, this ability does not make them susceptible to coercion, since revealing credentials only reveals the ballots the voter has submitted using those credentials; there is no guarantee that any of those ballots would count.

— **Forced abstention or retribution.** Forcing abstention or exacting retribution require the adversary to learn at least one credential with which the voter has submitted a ballot. We have seen above that a voter can make this arbitrarily difficult. Additionally, a voter strongly afraid of coercion or retribution may implement the following strategy: he may obtain a number of credentials, submit blank ballots for all of them, and immediately afterwards destroy the credential mailings. Without the cooperation of a majority of trustees, the only way the adversary can learn the credentials the voter used is by intercepting the blank ballot mailings themselves. Furthermore, as long as two of the blank ballots are not intercepted, they will be marked as coercion-evidence.

— **Silent coercion.** A voter is said to be *silently coerced* if he is coerced without his knowledge (Grewal et al., 2013). A voter may be silently coerced if the adversary intercepts one of the voter's credentials and vote on his behalf without the voter's knowledge. These silently coerced votes will only count if the voter does not submit any ballots of their own. This makes our system, along with Caveat Coercitor (Grewal et al., 2013), one of the few systems that handle this kind of coercion.

— **Information leakage.** The above attack mitigations and privacy guarantees are predicated on the assumption that an adversary cannot learn which ballots are cancelled and which are not. The full set of ballots, the tally with cancelled votes removed, and the number of cancelled votes are

public information; in contrived cases, this information is sufficient to determine which ballots were cancelled and which were not. In Appendix A, we discuss this vulnerability and provide a simple remedy.

3.5. Error Recovery

Our protocol was designed to allow voters to see if their ballot was received intact by looking for it on $\mathcal{BB}$. If voters self-report missing or modified ballots, this information is included in an audit trail. If a significant number of complaints are received, the election authority or indpendent auditors may be prompted to investigate further. However, voters' self-reports cannot be assumed to be perfectly trustworthy or reliable. At the cost of a more complex procedure, we can do better, by allowing voters to correct these errors (resubmit their ballots until they are properly received) instead of merely declare them.

To do this, the system can post partial credentials[8] of the ballots as they are received; voters can check that their ballot was received, and can submit another if necessary. Note that this could lead voters to unintentionally cancel their own vote. Because of delays in postal service, a voter could see that their first ballot is missing from $\mathcal{BB}$ and proceed to submit a second one; if the first ballot is not lost, but merely delayed, the election authority will eventually receive both and cancel the vote. This can be prevented by instituting a policy of disqualifying ballots if they are received a certain amount of time (for example a week) after they were postmarked.[9] This way, a voter knows that he must resubmit a ballot if it does not appear on $\mathcal{BB}$ within a week of submission, and can be sure that this resubmission will not unintentionally cancel his vote. This protocol guarantees voters the ability to reliably cast a ballot even in the face of inconsistent postal service.[10]

Coercion will lead to ballots being cancelled; we now argue that this cancellation procedure prevents coercion from manipulating the outcome of an election. Consider the four regimes jointly characterized by the level of actual coercion (high or low) and the number of cancelled votes (fewer than the margin of victory, or in excess of the margin of victory).

In the low-coercion few-cancelled-votes regime, the cancelled votes would be due to a handful of instances of actual coercion or simply a few voters mistakenly submitting more than one ballot. These few cancelled votes would change the published tally slightly from the tally of voters' true preferences, but only by a small number of votes relative to the margin of victory, so would not come close to changing the election outcome or significantly modifying the margin of victory.

We now consider the case in which there are many cancelled votes, comparable to or exceeding the margin of victory, but little actual coercion. This means that there are many ballots being cancelled for reasons other than coercion: voters could be submitting multiple ballots themselves, or they could be publicly revealing their credentials so that others may cancel their vote for them. Based on existing research, we do not believe that many voters will intentionally cancel their ballots.[11]

---

[8] A partial credential is a truncated credential, where enough of the credential is posted so that it is uniquely identifiable but an adversary cannot efficiently brute-force guess the full credential. If full credentials are posted during balloting, then anyone can submit a ballot with any of these credentials, cancelling a vote.

[9] Disqualified ballots are still posted, but are marked as such, and are neither tallied nor can they cancel votes. To detect if the system is adversarially disqualifying ballots by falsely claiming they were received after the one-week deadline, scans of their enclosing envelopes (with the date they were postmarked) can be posted along with the ballots. It will then be evident if there are an abnormal number of such ballots. Additionally, the system can post the scans of these envelopes on $\mathcal{BB}$ before they are opened and the enclosed ballots scanned, so the system cannot preferentially disqualify ballots for a given candidate.

[10] Note that this has the undesirable feature of publishing a running tally of all ballots, including cancelled ballots, during the voting period. To prevent this, instead of posting partial credentials and ballot scans to $\mathcal{BB}$ during the voting period, one could instead publish partial credentials, a cryptographic commitment to the ballot scan, and the ballot scan itself encrypted with the full credential as encryption key. In this case, only those in possession of the full credential (by arguments above, only the voter) may examine his plaintext ballot scan. After the election, the commitments to all of the ballots are opened and anyone can examine (and audit) any ballot image, preserving the auditability properties discussed below.

[11] Some nations, including Sweden and Estonia, have procedures that allow voters to cast multiple ballots, with later ballots overridding earlier ones. Estonia's revoting process is a close analogue to what we propose here (ENEC, 2013a). Importantly, data from recent elections in Estonia have shown very low levels of revoting; for example in the 2011 Estonian parliamentary

In the opposite regime, where there are few cancelled ballots but high levels of coercion, many instances of coercion are not being detected. Either voters are knowingly being coerced and are simply choosing not to submit a second ballot to cancel the coerced votes, or coercers are intercepting many credentials from non-voters and using those to cast votes on their behalf without the non-voter's knowledge.[12] The former does not seem likely, so we consider the latter. The worst case arises if coercers are able to use demographic information and voter profiles to selectively target potential non-voters for credential interception. Even so, unless they are able to do this selection with close to perfect accuracy, there will be some fraction of suspected non-voters who will end up submitting a ballot themselves. These ballots will show up as cancelled votes, since both the coercer and the voter have submitted ballots for the same voter identity, so as long as the coercer is submitting a significant number of ballots on behalf of potential non-voters, this will arise in an anomolously-high cancellation rate, signaling election authorities or independent auditors to investigate the reason for these cancelled votes. Note that in low-turnout situations the margin of victory may exceed the number of cancelled votes; however, the number of cancelled votes would still have to be high in absolute numbers. Thus, the anomaly would be detected and further investigation would be prompted.

Similarly, in the high-coercion, high-cancellation regime, the system would announce a large number of cancelled votes, and election authorities would be prompted to investigate the detected coercion. The adversary *does* succeed in casting doubt on the integrity of the election. In this sense, our vote cancellation procedure does not seem sufficient to hold a fair election in this situation. However, in the presence of a high cancellation rate due to widespread coercion or suspected government corruption, cryptography would do little to dispel a lay voter's distrust of the outcome (especially since the government was likely involved with designing and implementing the voting system in the first place). Instead of cryptographic assurances that may be of little real value in convincing the public of a correct outcome, our system is highly transparent: it outputs a variety of information which will be useful in identifying coerced ballots and ensuring a correct election outcome. Publicly-accessible ballot scans, the physical ballots themselves, and and the number of cancelled votes (potential markers of coercion) comprise an extensive audit trail which may be used by auditors or in litigation addressing election impropriety. Furthermore, if it is desired, the protocol can allow auditors to deanonymize certain ballots. This would allow them to study ballots which have been cancelled and thus potentially coerced.[13]

In much the same spirit as a risk-limiting audit, a protocol may be agreed upon specifying how to determine an election outcome given this audit trail. In this way, elections under doubt would be handled in the courts, much the way they are now, but our system would provide direct information about coercion. A coerced voter could be assured that by submitting a second ballot to cancel his vote, he has announced his plight to the election authorities and they will follow this agreed-upon procedure for ensuring that this coercion does not manipulate the election outcome.

3.6. Beyond Coercion-Resistance

Coercion-resistant voting systems offer a mechanism which allows voters to pretend to acquiesce to a coercer's demand while actually voting how they please. The mathematical formulation of

---

elections, 4,384 multiple Internet votes were recorded, and only 82 Internet votes were cancelled by a later paper ballot (of a total of 140,846 Internet voters) (ENEC, 2013b). We have no reason to expect that there would be a greater incidence of revoting in our case, where revoting is not allowed (as it cancels the vote). Thus, attempts at intentional multiple voting in our system could be seen as protest voting, but again there is little evidence in the research literature that shows a great deal of protest voting in existing electoral systems, and we do not expect that protest voting would be more prevalent in our system. See Stiefbold (1965) for a classic discussion of protest voting and void ballots; or Sinclair and Alvarez (2004) for a more recent examination of intentional voiding of ballots.

[12] We call this *silent coercion*, following Grewal et al. (2013).

[13] There is precedent for this kind of deanonymization for the purpose of election forensics: in certain jurisdictions, such as the U.K., the government is legally obligated to deanonymize certain ballots at the request of an election judge (Smart & Ritter, 2009).

this property, introduced by Juels et al. (2005), states that in the course of the execution of such a protocol, an adversary is not able to learn any additional information beyond the tally itself. Put another way, the voting system does not allow the coercer to distingish between a coerced voter's compliance and non-compliance. As such, the voter need not heed the coercer's threat, and may vote according to their true preferences. The precision of this property is appealing; it purports to perfectly mitigate the threat of voter coercion. However, in practice a coercion-resistant voting system could fall significantly short of this goal.

Any coercion-resistant voting system presupposes an untappable channel, yet none of the channels over which remote elections are conductedmail, phone, and the Internetare perfectly untappable.[14] As recent descriptions of state-sponsored surveillance programs have illustrated, long-term, mass interception of mail is not a theoretical threat (Nixon, 2013). One might respond that instead of registering remotely, we could mandate in-person registration, which is surely more secure.

Benaloh (2013) argues that even this is not good enough, observing that the prevalence of cell phones and wearable cameras prevents even a polling booth from being truly private. Given that coercers or vote-buyers can instruct voters to surreptitiously record their registration or balloting sessions with these cameras, even the paradigmatic untappable channel (the private voting booth) is no more. Without any untappable channels, perfect coercion-resistance is impossible. Benaloh concludes that surrender is not an attractive option, but there seems to be little point to adding significant complexity to election protocols in an increasingly futile attempt to defeat pre-election coercion.

Instead of surrender, we advocate a strategic retreat. Our vote-cancellation procedure will still detect coercion perfectly even without an untappable channel. Once detected, an audit (and associated litigation) can use this information to neutralize coercion and ascertain the correct election outcome.

The messy process of a court case may seem far less appealing than the clean technical solution provided by coercion-resistance. However, we argue that coercion-resistance is only a partial solution to the problem of coercion: the goal of a voting system is not only to output the correct outcome, but also to convince voters that this outcome is indeed proper and correct. An audit may be messy, but voters are already familiar with its mechanics and understand how the adversarial legal system serves to arrive at a fair outcome; the lay voter is far less likely to understand why cryptography is able to guarantee the fairness of an election in the presence of coercion.

Furthermore, laws are the ultimate arbiter of election propriety, so far from a disadvantage, it is inevitable and beneficial that the courts be involved in adjudicating the election outcome. It is then the purpose of the voting system to provide extensive and clear evidence to guide the court. Cryptographic voting systems are designed to satisfy the mathematician that coercion has been mitigated in a given election, but this may not be the most useful evidence for the court's purposes. Our system offers a high degree of transparency: it outputs an audit trail that includes full ballot scans, physical ballots that may be subjected to forensic analysis, the number of cancelled (and possibly coerced) ballots, and possibly the voter identities corresponding to suspect ballots (if the protocol allows for their deanonymization). All of this data can be handled in a way analogous to that of a traditional audit. This represents a significant advantage over most cryptographic voting systems, which offer very little in the way of transparency or auditibility. As Benaloh (2008) mentions, audits are a complimentary approach to end-to-end verifiability and may better handle widespread attacks.

Abandoning the cryptography of coercion-resistance also allows for superior usability. In a coercion-resistant system, each voter must go through the rigamarole of a distributed registration protocol to construct a series of cryptographic credentials, must encrypt their ballots, and must submit appropriate proofs, and the voter must do this even if they are not being coerced. Our system

[14] In fact, the original paper by Juels et al. (2005) mentions that mail can be used as an untappable registration channel. This makes sense when designing an Internet voting system, when the goal may not be a system that is perfectly secure in an absolute sense, but rather a system that is no less secure than current election practice. We aim to design a system that is secure in an absolute sense, so we cannot assume mail to be untappable.

features a radically simpler ballot casting protocol completely free of cryptography. Furthermore, only coerced voters need to understand the details of the multiple-cast policy; most voters can simply cast a ballot using the first credential they receive and do not need to worry about the parts of the system that provide coercion-evidence and verifiability. In the limit where there are no coerced or malicious voters, and hence no cancelled votes, the tally is verifiable *without any cryptography*. In other words, the complexity of the coercion-evidence and verification mechanisms of the system only exhibits itself when it is necessary. In the absence of coercion, voters and administrators interact with the system in a way little different from current vote-by-mail practice.

4. DISCUSSION

The motivating feature of the voting system we have described is that it publishes ballot scans, bringing transparency to the voting process by allowing for public audits of the ballot box. This is far from a novel goal, however: volunteers in Humboldt County, California, scanned ballots from two 2008 elections, citizens in Colorado have sued for the right to access ballot scans, and similar efforts are underway in other states (Adler & Hall, 2013).

While posting ballot scans in the name of transparency may seem a beneficial development in election administration, Adler and Hall (2013) compellingly argue that doing so would do more harm than good to the integrity of the electoral process. This approach to transparency is plainly untenable if ballots could be associated with the voter who submitted them. Such a violation of privacy would be illegalthe constitutions of all fifty states guarantee ballot secrecyand furthermore would allow unrestricted vote-buying and coercion. Ballot publication seems possible, however, if one ensures that the ballots are not identifiable.

The problem is that it is impossible to guarantee that a ballot is truly anonymous. The most innocuous of stray marks is enough to distinguish a ballot. Moreover, ballots in the U.S. commonly include dozens of races; a coerced voter or vote-seller may uniquely sign their ballot by voting for an agreed-upon sequence of candidates. Many states have statutes that criminalize marking a ballot in an identifiable way or invalidate the vote therein; California law specifically prohibits the publication of ballots with identifiable marks (Adler & Hall, 2013). The trouble is, of course, that there is no way for an election official to reliably determine whether a stray mark or sequence of votes was made with the intent of making the ballot identifiable. Given the impossibility of such a task, one might reasonably conclude that ballot publication cannot be done without breaking the law and undermining anonymity.

Our proposed voting system is the first paper-based system to allow ballot publication while addressing the aforementioned concerns. A voter may choose to make any ballot they cast identifiable, but they can always cast another ballot to cancel the previous vote. As we have discussed previously, this is sufficient to neutralize vote-selling and coercion.

That said, publishing ballots may do harm to the electoral process even if voters have no rational basis on which to fear privacy loss. Gerber, Huber, Doherty, and Dowling (2012) have demonstrated that voter behavior is driven by their perception of privacy, which may be quite different than their actual level of privacy. In their survey, a quarter of respondents did not believe their ballot choices were kept secret. This surprisingly high fraction suggests that voters may be unfamiliar with the regulations and procedural safeguards in place to protect their privacy. These doubts are consequential: they lead to depressed turnout (Gerber, Huber, Doherty, Dowling, & Hill, 2013a) and in some cases may influence how a voter votes (Gerber, Huber, Doherty, & Dowling, 2012). Furthermore, Claassen, Magleby, Monson, and Patterson (2012) observe that voters' perceptions of privacy are correlated with their belief in a fair election outcome. In a later work, Gerber, Huber, Doherty, Dowling, and Hill (2013b) find that postal voters are more likely to doubt the secrecy of their ballot than in-person voters. In this survey, 43% of postal voters reported that it would be not difficult at all or not too difficult to find out who [they] voted for, and a similar number reported that they thought that election officials access [their] voting records to figure out who [they] voted for. Thus, ameliorating voters' privacy concerns should be a key goal of any vote-by-mail system.

Unfortunately, the posting of ballot scans runs the risk of inflaming these concerns. Every voter will know that anyone else could look at their ballot, and might believe that someone could identify which ballot was theirs, even if they did not have reason to believe this.

Unlike in existing vote-by-mail systems, the government cannot learn how they voted (except possibly with the authorization of an election judge). While voters may not understand or appreciate the cryptography that serves to protect their privacy, they do not need to: ballot secrecy under current election administration is assured not by mathematical proof but by procedural means. Gerber, Huber, Biggers, and Hendry (2013) find that mailings reminding voters of their rights to a secret ballot are effective in assauging voters' privacy doubts and yield a long-term increase in turnout.

Posting ballot scans also gives voters the ability to choose to give up their anonymity. If voters were to voluntarily give up their privacy in large numbers, it would further undermine confidence in the secret ballot. Moreover, if many voters denanonymized their ballot, it would create social pressure for others to follow suit. As such, laws prohibiting making ballots identifiable should be kept in place for the sake of upholding the *perception* of privacy even if they are not necessary to ensure to ensure actual privacy. Publicizing these regulations on ballots and other voting materials would go a long way toward ameliorating voter concerns.

We see that publishing ballot scans may have negative consequences for the perception of voter privacy, although reminders about secrecy regulations and procedures on election materials and through mailings may in large part effectively mitigate this. The purpose of publishing ballots, however, is to give voters confidence that their ballots were received; this is an unambiguous strength of our system. This approach is especially desirable in situations where ballot transport is highly unreliable. For example, our system could significantly improve the trustworthiness of UOCAVA voting, but would do so without voters to be coerced or their privacy violated.

However, our approach is useful more generally to combat the electorate's well-founded lack of confidence in postal voting. Alvarez, Hall, and Llewellyn (2008) have found that the fraction of absentee mail voters reporting that they are very confident that their vote counted was 16% lower than the corresponding fraction for optical scan voters. These doubts are not unfounded. (Alvarez, Hall, & Sinclair, 2008) find that absentee ballots cast by mail are much more likely to be challenged or not counted than ballots cast in person. In our system, since ballots and canvassing board decisions are posted publicly, voters can be directly verify that their ballot was received intact, before the deadline, and interpreted correctly. If their ballot was challenged or invalidated, they can see this too, and if they wish they may register a dispute with the election authorities.

Postal voters have been shown to have increased concern with privacy and decreased confidence in the integrity; both of these factors have been shown to depress turnout (Alvarez, Hall, & Llewellyn, 2008; Gerber et al., 2013a). By specifically reassuring the voter in both of these areas, our system may well fulfill one of the elusive promises of voting-by-mail: unambiguously increased turnout.

We have thus described a system which attempts to address exactly those concerns about which voters care most. The central verifiability mechanism of our designthe posting of plaintext ballotsis enabled by the multiple-voting with cancellation procedure of Grewal et al. (2013). In our system, however, we postpone the encryption of the votes until after casting. By doing so, our system allows voting-by-mail, and allows publication of ballot plaintext, two novel features in a voting system with coercion mitigation.

While the proposal allows highly transparent postal voting, it falls short of a fully-verifiable postal voting scheme. It remains a interesting, and highly relevant, goal for future work to construct a system that makes further progress in balancing verifiability with usability in the postal voting setting.

A. INFORMATION LEAKAGE FROM TALLY AND BALLOTS

We model a ballot with k binary options as a bit-vector $b \in \{0,1\}^k$ (a 0 represents no mark, a 1 represents a mark), the ballot data M for n ballots is a $k \times n$ matrix of bits, and the tally is given by a vector $t \in \mathbb{N}^k$. Recall that the tally is not the total number of marks for that ballot option, but the

total number of marks for that ballot option *not including cancelled ballots*. Note that the number of cancelled ballots, c, is evident from $\mathcal{BB}$. A solution vector is a vector $\chi \in \{0,1\}^n$ with entries $\chi_i, 1 \leq i \leq n$ such that $M\chi = t$ and $|\chi|_1 = n - c$. That is, a solution vector labels each ballot as either non-cancelled (contributes to the tally) or cancelled (does not contribute to the tally) in such a way as the tally of such non-cancelled ballots $M\chi$ equals the actual tally t, and furthermore since the number of cancelled ballots c is public knowledge, the solution vector must only label c ballots as cancelled. Let $\mathcal{S}$ be the set of solution vectors. The privacy guarantees we seek are negated when an adversary may learn with near-certainty that a particular ballot was cancelled. Thus, privacy loss occurs when an adversary may find a $P[\chi_i = 0]$ close to 1 for a ballot of interest i (where presumably *close* means $P[\chi_i = 0]$ considerably in excess of $\frac{c}{n}$, the probability one obtains knowing only the number of cancelled ballots and not the tally or ballot data). We may set $P[\chi_i = 0] = \frac{|\{\chi \in \mathcal{S} | \chi_i = 0\}|}{|\mathcal{S}|}$. To calculate this, one needs to find the solution set $\mathcal{S}$ given ballot data M and a tally t.

For typical election settingsmany ballots, many voterswe would not expect an adversary to be able to carry out this attack and violate privacy in this matter; it is left for future work to prove a privacy bound that makes such an argument rigorous. Alternatively, we can modify the voting protocol to prevent any possibility of such an attack. The notion of privacy we need is essentially that we want the output of our protocolthe tallyto be insensitive to which ballots we cancel. This is exactly the goal of *differential privacy*, a well-studied framework for privacy-preserving computation (Dwork, 2006). The usual method to implement a differentially-private algorithm is to add a small amount of noise to the output. While adding noise to an election seems untenable at first, notice that the amount of noise we would need to add (on the order of one vote) is negligible compared to other sources of noise in real-world elections. Furthermore, the probability that this noise would change the outcome of the election is exponentially small, and we can neglect it for all practical purposes.

We now sketch how one might add noise using a cut-and-choose protocol. Before the registration phase, the trustees generate one extra credential $\text{cred}_{\text{noise}}$ which will be used to inject noise, and a noise source (does not have to be trusted) generates N (with $N \sim 1000$) instances of random ballot data $B_i^l, 1 \leq i \leq N$ for each race l, and posts the encryption $\{B_i^l\}_{PK}$ of each of them to $\mathcal{BB}$. After the balloting phase, a trusted source of randomness (e.g., stock market data, cf. Clark, Essex, and Adams (2007), Clark and Hengartner (2010)) is used to select an $1 \leq k \leq N$. During tallying, $(\text{cred}_{\text{noise}}, B_k^l)$ is then included in the mixnet and processed like any other ballot. After the tallying, the trustees jointly decrypt the other $N-1$ instances of random ballot data; for large N, it can be verified that B_k^l was selected randomly with high probability.

References

Adida, B. & Rivest, R. L. (2006). Scratch & Vote: Self-contained Paper-based Cryptographic Voting. In *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society* (pp. 29–40). WPES '06. ACM.

Adler, E. S. & Hall, T. E. (2013). Ballots, Transparency, and Democracy. *Election Law Journal*, *12*(2), 146–161.

Alvarez, R. M., Ansolabehere, S., Berinsky, A. J., Lenz, G., Stewart, C., III, & Hall, T. E. (2009). *2008 Survey of the Performance of American Elections*. Caltech/MIT Voting Technology Project.

Alvarez, R. M., Hall, T. E., & Llewellyn, M. H. (2008, July). Are Americans Confident Their Ballots Are Counted? *The Journal of Politics*, *70*(03).

Alvarez, R. M., Hall, T. E., & Sinclair, B. (2008). Whose absentee votes are returned and counted: The variety and use of absentee ballots in California. *Electoral Studies*, *27*(4), 673–683.

Alvarez, R. M., Stewart III, C., & Beckett, D. (2013). Voting Technology, Vote-by-Mail, and Residual Votes in California, 1990-2010. *Political Research Quarterly*, *66*(3), 658–670.

Aslam, J. A., Popa, R. A., & Rivest, R. L. (2007). On estimating the size and confidence of a statistical audit. In *Proceedings of the 2007 USENIX/ACCURATE Electronic Voting Technology Workshop (EVT07)*.

Benaloh, J. (2008). Administrative and Public Verifiability: Can We Have Both? In *Proceedings of the conference on electronic voting technology*. EVT '08. Berkeley, CA: USENIX Association.

Benaloh, J. (2013, August). Rethinking Voter Coercion: The Realities Imposed by Technology. *Journal of Election Technology and Systems (JETS)*, *1*(1), 82–87.

Benaloh, J., Ryan, P. Y. A., & Teague, V. (2013). Verifiable postal voting. In B. Christianson, J. Malcolm, F. Stajano, J. Anderson, & J. Bonneau (Eds.), *Security Protocols XXI* (Vol. 8263, pp. 54–65). Lecture Notes in Computer Science.

Bergman, E. (2012, February). Administering Democracy: Public Opinion on Election Reform in California. *The California Journal of Politics & Policy*, 1–24.

Bland, S. (2014, February). Tracking Voters in Real Time in Colorado. Retrieved February 22, 2014, from http://www.nationaljournal.com/magazine/tracking-voters-in-real-time-in-colorado-20140224

Brandt, F. (2006). Efficient Cryptographic Protocol Design Based on Distributed El Gamal Encryption. In D. Won & S. Kim (Eds.), *Information Security and Cryptology (ICISC 2005)* (Vol. 3935, pp. 32–47). Lecture Notes in Computer Science. Springer Berlin Heidelberg.

Bursuc, S., Grewal, G. S., & Ryan, M. D. (2012). Trivitas: Voters Directly Verifying Votes. In *Proceedings of the Third international conference on E-Voting and Identity* (pp. 190–207). VoteID '11. Springer-Verlag.

Calandrino, J. A., Clarkson, W., & Felten, E. W. (2011). Bubble Trouble: Off-line De-anonymization of Bubble Forms. In *Proceedings of the 20th USENIX Conference on Security*. SEC '11. USENIX.

Chaum, D. (2004). Secret-ballot receipts: True voter-verifiable elections. *IEEE Security & Privacy*, *2*(1), 38–47.

Chaum, D. L. (1981, February). Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, *24*(2), 84–90.

Chaum, D., Ryan, P. Y., & Schneider, S. (2005). A practical voter-verifiable election scheme. In *Computer Security (ESORICS 2005)* (pp. 118–139). Lecture Notes in Computer Science. Springer.

Claassen, R. L., Magleby, D. B., Monson, J. Q., & Patterson, K. D. (2012, May). Voter Confidence and the Election-Day Voting Experience. *Political Behavior*, *35*(2), 215–235.

Clark, J., Essex, A., & Adams, C. (2007). Secure and observable auditing of electronic voting systems using stock indices. In *Canadian Conference on Electrical and Computer Engineering, 2007 (CCECE 2007)* (pp. 788–791). IEEE.

Clark, J. & Hengartner, U. (2010). On the Use of Financial Data as a Random Beacon. *IACR Cryptology ePrint Archive*, *2010*, 361.

Clarkson, M., Chong, S., & Myers, A. (2008). Civitas: toward a secure voting system. In *Security and privacy, 2008. sp 2008. ieee symposium on* (pp. 354–368). doi:10.1109/SP.2008.32

Dategrity Corp. (2005, May). VoteHere Announces Mail-in Ballot Tracker Audit Solution. Retrieved from http://www.marketwired.com/press-release/votehere-announces-mail-in-ballot-tracker-audit-solution-663738.htm

Dwork, C. (2006). Differential privacy. In *Automata, languages and programming* (pp. 1–12). Springer.

Elgamal, T. (1985). A public key cryptosystem and a signature scheme based on discrete logarithms. *Information Theory, IEEE Transactions on*, *31*(4), 469–472. doi:10.1109/TIT.1985.1057074

Essex, A. & Hengartner, U. (2012). Oblivious Printing of Secret Messages in a Multi-party Setting. In A. D. Keromytis (Ed.), *Financial Cryptography and Data Security* (Vol. 7397, pp. 359–373). Lecture Notes in Computer Science. Springer Berlin Heidelberg.

Estonian National Electoral Committee. (2013a). Internet Voting in Estonia. Retrieved December 3, 2013, from http://www.vvk.ee/voting-methods-in-estonia/engindex/

Estonian National Electoral Committee. (2013b). Statistics about Internet Voting in Estonia. Retrieved December 3, 2013, from http://www.vvk.ee/voting-methods-in-estonia/engindex/statistics

Gerber, A. S., Huber, G. A., Biggers, D. R., & Hendry, D. J. (2013). Ballot Secrecy Concerns and Voter Mobilization: New Experimental Evidence about Message Source, Context, and the Duration of Mobilization Effects. Manuscript, Yale University, Department of Political Science.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2012, July). Is There a Secret Ballot? Ballot Secrecy Perceptions and Their Implications for Voting Behaviour. *British Journal of Political Science*, *43*(01), 77–102.

Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., & Hill, S. J. (2013a). Do Perceptions of Ballot Secrecy Influence Turnout? Results from a Field Experiment. *American Journal of Political Science*, *57*(3), 537–551.

Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., & Hill, S. J. (2013b). The Voting Experience and Beliefs about Ballot Secrecy. Manuscript, Yale University, Department of Political Science.

Greenson, T. (2009, March). SOS report: Numerous deficiencies in elections software. Retrieved August 5, 2013, from http://www.times-standard.com/ci_11841759

Grewal, G. S., Ryan, M. D., Bursuc, S., & Ryan, P. Y. A. (2013). Caveat Coercitor: Coercion-Evidence in Electronic Voting. In *Proceedings of the 2013 IEEE Symposium on Security and Privacy* (pp. 367–381). SP '13. IEEE.

Gronke, P., Galanes-Rosenbaum, E., Miller, P. A., & Toffey, D. (2008, June). Convenience Voting. *Annual Review of Political Science*, *11*(1), 437–455.

Hirt, M. & Sako, K. (2000). Efficient receipt-free voting based on homomorphic encryption. In *Advances in Cryptology (EUROCRYPT 2000)* (pp. 539–556). Springer.

Jakobsson, M. & Juels, A. (2000). Mix and match: secure function evaluation via ciphertexts. In T. Okamoto (Ed.), *Advances in Cryptology (ASIACRYPT 2000)* (Vol. 1976, pp. 162–177). Lecture Notes in Computer Science. Springer Berlin Heidelberg.

Jakobsson, M., Juels, A., & Rivest, R. L. (2002). Making Mix Nets Robust for Electronic Voting by Randomized Partial Checking. In *Proceedings of the 11th USENIX Security Symposium* (pp. 339–353). USENIX Association.

Jakobsson, M., Sako, K., & Impagliazzo, R. (1996). Designated verifier proofs and their applications. In *Advances in Cryptology (EUROCRYPT 96)* (pp. 143–154). Springer.

Juels, A., Catalano, D., & Jakobsson, M. (2005). Coercion-resistant electronic elections. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society* (pp. 61–70). WPES '05. ACM.

Kim, E., Carlini, N., Chang, A., Yiu, G., Wang, K., & Wagner, D. (2013, August). Improved support for machine-assisted ballot-level audits. *Journal of Election Technology and Systems (JETS)*, *1*(1), 88–105.

National Conference of State Legislatures. (2013). Absentee and Early Voting. http://www.ncsl.org/legislatures-elections/elections/absentee-and-early-voting.aspx.

Nixon, R. (2013, July). U.S. Postal Service Logging All Mail for Law Enforcement. Retrieved July 3, 2013, from http://www.nytimes.com/2013/07/04/us/monitoring-of-snail-mail.html

Popoveniuc, S. & Lundin, D. (2007). A simple technique for safely using Punchscan and Prt Voter in mail-in elections. In *Proceedings of the 1st International Conference on E-voting and Identity (VOTE-ID'07)* (pp. 150–155).

Rivest, R. L. (2006). The ThreeBallot voting system. Retrieved from http://theory.csail.mit.edu/~rivest/Rivest-TheThreeBallotVotingSystem.pdf

Rivest, R. L. & Shen, E. (2012). A Bayesian method for auditing elections. In *Proceedings of the 2012 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE'12)*.

Ryan, P. Y. A. (2005). A Variant of the Chaum Voter-verifiable Scheme. In *Proceedings of the 2005 Workshop on Issues in the Theory of Security (WITS '05)* (pp. 81–88). ACM.

Saeednia, S., Kremer, S., & Markowitch, O. (2004). An efficient strong designated verifier signature scheme. In *Information security and cryptology (icisc 2003)* (pp. 40–54). Springer.

Sharma, A., Subramanian, L., & Brewer, E. A. (2011). Paperspeckle: microscopic fingerprinting of paper. In *Proceedings of the 18th ACM conference on Computer and communications security* (pp. 99–110). CCS '11. New York, NY, USA: ACM.

Sinclair, D. E. B. & Alvarez, R. M. (2004). Who Overvotes, Who Undervotes, Using Punchcards? Evidence from Los Angeles County. *Political Research Quarterly*, *57*(1), pages.

Smart, M. & Ritter, E. (2009). Remote Electronic Voting with Revocable Anonymity. In *Proceedings of the 5th international conference on information systems security* (pp. 39–54). ICISS '09. Berlin, Heidelberg: Springer-Verlag.

Southwell, P. L. (2004). Five Years Later: A Re-Assessment of Oregon's Vote by Mail Electoral Process. *PS: Political Science and Politics*, *37*(1), 89–94.

Stark, P. B. (2008). Conservative statistical post-election audits. *Annals of Applied Statistics*, *2*(2), 550–581.

Stark, P. B. (2009). CAST: Canvass audits by sampling and testing. *IEEE Transactions on Information Forensics and Security*, *4*(4), 708–717.

Stewart, C., III. (2010). Losing votes by mail. *NYU Journal of Legislation & Public Policy*, *13*, 573.

Stiefbold, R. P. (1965, June). The Significance of Void Ballots in West German Elections. *American Political Science Review*, *59*, 391–407.

Zagrski, F., Carback, R., Chaum, D., Clark, J., Essex, A., & Vora, P. L. (2013). Remotegrity: Design and Use of an End-to-End Verifiable Remote Voting System. *IACR Cryptology ePrint Archive*, *2013*, 214.

ACKNOWLEDGMENTS

Foremost, the authors would like to thank Kenneth Hung, as the ideas presented here largely arose in discussions with him. This work benefitted greatly from careful and detailed comments on earlier drafts by Katrina Ligett, Barath Raghavan, Ron Rivest, Jonathan Katz, Charles Stewart III, and the anonymous referees.

Received August 2013; revised June 2014; accepted July 2014