Random-Sample Voting

Far lower cost, better quality and more democratic

David Chaum

ABSTRACT: Random-sample voting can be used locally, nationally, regionally, or even globally, with results that are more irrefutable than with current elections but at less than one-thousandth of the cost.

As a new member of the democracy toolbox, such voting holds great promise, for instance making practical today: Petitions of government that prove majority support. Binding consultations of constituencies by officials or parties. Ladder competitions that elect the most important and clearly stated issues to be put to vote. Juries for public policy issues with unprecedented resistance to manipulation. Even, in the extreme, full Athenian-style direct democracy, practical at the scale and complexity of society today.

Voters may be better motivated and informed since each vote carries more weight and each voter can meaningfully investigate and study the single issue that voter is asked to help decide. Voter confidence in the increased integrity of the election process may also enhance participation. Anyone can verify online that neither randomness of voter selection nor integrity of outcome can have been manipulated, even by governments. The ballots are sent voters by paper mail, but vote buying is made ineffective by a novel technique.

The approach is compared with voting today in terms of ten attributes of election quality, the operational concept along with its properties sketched, historical context discussed, and various adoption scenarios laid out. An appendix gives implementation details, based on systems already proven in governmental elections, sufficient to show technical and practical feasibility.

It is accepted as democratic when public offices are allocated by lot;
and as oligarchic when they are filled by election.

—Aristotle (Politics 1301a28-35)

INTRODUCTION

A new way is introduced here for translating the “will of the people” into governance. This white paper aims to show that its new approach can already be deployed in a variety of significant ways, with or without participation of government—there is no catch. It also aims to explore the advantages and potential of the new approach. The aim, however, is neither values neutrality nor a manifesto for a particular ultimate system.

The number of voters sampled can be small, depending on how close the contest, yet give overwhelming confidence. For instance, if the margin is at least ten percent, then a thousand votes will likely yield a result that itself, without any assumption about the margin and with only a one-in-a-million chance of error, establishes that a majority are in favor—even with an electorate of millions or

billions. This dramatic reduction in the number of voters participating in each election compared to a conventional election today yields a substantially proportionate reduction in cost.

A general benefit of such low cost, and also taking advantage of the small number of voters needed for an election, is that many more elections can be conducted, both in parallel and more often. This in turn means that each voter need only be involved in an election relatively infrequently. It also means that each election can be on a single issue, allowing voters ample opportunity to research and consider that issue. The mechanics of supplying ballots would be through paper mail; nevertheless, with today's widely adopted information technology, such as the web, search, forums, electronic access to experts, and so on, voters will be able with the increased time and focus to investigate and deliberate even complex issues with unprecedented ease and depth.

Another benefit of the dramatic cost reduction is an opening up of who can conduct elections. Any interest group can create their own election initiatives, if they can bear the modest expense, without requiring permission or assistance from government. A new type of election infrastructure could, as another example, allow all interested parties to submit questions and would then conduct mini elections between proposed questions until one prevails and is put to a larger vote. Political parties or candidates for office could also conduct their own non-governmental elections and even bindingly agree in advance to be bound by the results. Every use helps establish efficacy, increases acceptance, and—even if national leaders are never selected this way—is potentially hugely beneficial to society.

Representative democracy, by no means the only type of democracy, is based on a right to vote towards representatives and that every vote should be counted. This hard-won mechanism of universal suffrage, from which political legitimacy of most governments is extracted, may unfortunately be failing.

Trends over the last several decades bear this out. Trust in elected governments has plummeted, at least for major democracies. The number and, according to the various published ratings, almost all metrics of democracies are also declining, the so-called "democracy deficit." Divergence of public policy and fundamental values of society is accelerating, for instance on issues such as income distribution, human rights, war/peace, and environment. All this in spite of huge technology advances, especially for information, with its unprecedented ability to provide transparency, tackle complexity and scale.

Common explanations for the failing of representative democracy are the low utility to the individual of actually casting a ballot in view of the choices available and the corrupting influence of economic power in such elections. An alternate explanation may be that the transparency brought by information technology is exposing the failings of current systems. The approach introduced here, in contrast, promises to obviate corrupting influences on governance and make each vote more meaningful, so that policy can converge with widely held values—while actually benefiting from the scale of information technology and transparency in a way that can meet the complexity of society today.

Furthermore, using the technique of democracy outside governments opens new possibilities. Issues cutting across national boundaries, including even election-related issues such as access to information, for instance in terms of anonymous/uncensored online access and media supported by voted quality, become something even a modestly-well-funded non-profit organization could put to global vote.

To whatever extent binding, random-sample voting can give powerful voice to issues while gaining acceptance and raising the bar on election system quality and performance. Gradually, as the techniques are proven out in practice, refined, and acceptance broadens, adoption in governmental elections is at least plausible. The techniques could be used for such things as filling more positions by election. They

are also a good fit for the various types of referenda and initiative, such as used in many countries and in the ten western-most US states, potentially used to make major policy decisions. They then hold promise for increasing participation, improving perception of fairness and efficacy of the process, and strengthening the foundations of democracy and arguably society more generally.

This new type of election, more fundamentally, allows continuous, effective and indisputable monitoring of the will of the people on a very wide range of issues. It thus offers an alternative to representative democracy—at least, to whatever extent it may be desirable, by solving a challenging "open problem"—Athenian democracy at the scale and complexity of governments today.

This new approach can also improve voter protection. Elections in which voters are able to verify convincing online evidence that their votes were correctly included in election outcomes have already been used to fill political offices in a United States city. This strongest known techniques for integrity and transparency of elections, as explained later and detailed in the appendix, is used to implement the new approach presented. An important property achieved here, and verifiable by anyone online, is that which voters are included is truly random and cannot be known in advance, preventing discrimination in selection or targeted influence of voters. An extension ensures that vote buying is ineffective. An optional extension can even hide the identity of voters in perpetuity.

Weather data or stock market prices on a future date-certain provide the randomness. Data is "committed" to in advance by being publishing in unchangeable encrypted form. It's like the encrypted data is the result of a potentially suspect coin flip still hidden by a hand on a wrist, with weather or stock data the random "call" made after the coin can no longer be changed but before the heads or tails is revealed—to ensure unmanipulatability of the combined outcome. Because the decryptions can be checked by anyone online, the process is fully transparent. Such outcomes determine what to decrypt in the transparent audits, ensuring outcome integrity (without compromising ballot secrecy). Also, voters can verify that they are picked just as fairly by such outcomes as if they had flipped coins themselves with predefined patterns of heads and tails resulting in being selected to vote.

This work is organized in five sections, each intended to mainly stand alone. In the first section, ten desirable attributes for quality of public-sector elections are introduced and used to compare current elections with the system introduced here. The attributes are then recast slightly so that the system introduced achieves the full desiderata.

The second section summarizes the technical aspects of the system for the non-specialist reader. Emphasis is on the ideas behind some of the more innovative aspects of its security. The various categories of participants in the system are laid out, the new properties achieved are stated informally, and sample-sizes are derived from first principles. (The system's technical feasibility is shown in an appendix detailing it in terms of component parts similar to those of systems already used in governmental elections.)

The third section considers context in terms of the historical and current mechanics of democracy. A variety of situations in which the approach can be applied outside government as well as scenarios for gradual integration into government are sketched in the fourth section. Finally, the fifth section touches on some broader perspectives.

ELECTION QUALITY

Desirable characteristics of public-sector election mechanics defining what is here called their "quality" include: (a) high voter turnout, (b) well informed voters, (c) effectiveness of results in shaping governance, (d) resistance to manipulation through advertising and electioneering, (e) indisputability of tally, (f) protection against voter corruption or coercion, (g) ensured access to vote casting, (h) resistance to voter fraud, (i) decisiveness, and (j) low cost.

Current elections perform egregiously poorly against every single one of these ten positive attributes.

Election reform, however, has been notoriously difficult and slow. One explanation is that improving integrity, and by extension other aspects of elections, implicitly criticizes the soundness of the very system that selected those to the positions of power that enable them to block change. Political parties are also sensitive to and often worried about changes in voter demographics and potential for manipulation, which can result from even small changes in election mechanisms.

Other reasons for the sluggishness of reform may include economic interests. Relatively large amounts of money have been spent on election equipment in some countries, notably the United States, Brazil and India, while recurring expenditures on administration are even higher in aggregate. Elections today entail massive campaign spending, which in large measure benefit media and various experts. (Such advertising and electioneering outlays, however, may be considered undesirable because of the conflict of interest they raise between officeholders/parties and the public with respect to sponsors and lobbies.)

It has been posited that individual voters expending effort to learn about the options presented them in conventional elections is not economically rational, at least not as far as the probability of having an impact. This analysis may account in part for low voter-turnout. It also predicts compounding of the problem of meaningful voter participation as technology continues to increase the complexity of governance and policy options.

Current elections will here be called "mass" elections to contrast with "random-sample" elections, which may broadly be characterized generally as a poll of a sample of the population that is at least as secure against abuse as traditional mass elections.

The significance of each vote is elevated in a random-sample election, compared to a mass election. This allows those voters who are selected to rationally put substantial effort into informing themselves, deliberating and voting, rising to the occasion much like members of juries are known to. Additionally, participants in elections optionally could be paid, but in a way that's verifiably independent of how they vote, detailed later, such as is common with juries and as was practiced in ancient Greece. Even without compensation, a significant boost can accordingly be expected in effective "turnout" and extent to which all contests/questions on ballots are voted. (High voter turnout [a].)

Other current impediments to election efficacy include unclear statement of issues or even complex bundling of issues. Obfuscation for example of beneficiaries of government policy through voluminous regulation, comprised of convoluted and archaic language, is well known. Random-sample voting lets voters peer more deeply into complex ballot questions. To the extent this new type of election is used to frame ballot questions, such as through question ladders mentioned above and discussed further below, it also holds the potential to directly return us to the day when ballot language and legislation were more comprehensible.

Thus, random-sample voting both raise the level of issues that can meaningfully be put before voters and holds potential to bring the statement of issues closer to what voters can understand. (Well informed voters [b].)

The combination of increased participation and meaningful deliberation just described allows random-sample election results to be more fine-grained and useful than those of mass elections in informing and steering governance. Moreover, the episodic election of parties and representatives creates lag, discontinuities and short-term focus of public policy; whereas continual polling facilitated by random-sample voting may foster gradual evolution of longer-term policy while providing lower lag in response to changing circumstances. (Effectiveness of results in shaping governance [c].)

Advertisements and other types of media and campaigning are able to influence mass election outcomes perhaps surprisingly well in this age of unprecedentedly widespread and selective access to information. (In the US, for instance, expenditure for related advertising trumps the mechanics of elections by a significant factor thereby creating decisive leverage for sponsors and lobbies.) This may be due to the shallow diligence of rational voters mentioned above, but it also derives from the episodic nature of election cycles. (Such cycles also in a related aspect disadvantageously focus perhaps much of elected officials' time and attention on campaigning instead of governing.) Influencing an ongoing as opposed to an episodic polling process through media and campaigning is far more costly and difficult, especially when voters are motivated and able to investigate in depth, and random-sample voting may thus reduce such influence. (Resistance to manipulation through advertising and electioneering [d].)

The security features detailed below bring the integrity of a random-sample election up to at least par with those of the best polling-place mass elections, such as so-called cryptographic end-to-end systems. In some such elections that have been conducted, the correctness of the tally is proved mathematically in a way that can be verified easily by voters and also significantly by anyone interested enough to download or even write a modest amount of software. (Indisputable/trustworthy tally [e].)

So-called "improper influence" of voters refers generally to vote buying and coercion of voters. Variants of both types of influence are known in practice in mass elections at polling-places as well as in current vote-by-mail. Vote buying is essentially obviated with random-sample voting by flooding the market with willing sellers of decoy votes that will not be counted but that are enduringly indistinguishable from genuine votes, as detailed later; also, potential coercers are unable to find victims since they are unable to effectively learn who has received a ballot. Influence of groups of voters is also known, where for instance one locality is favored over others because of certain voting patterns; however, integrity of results in random-sample elections obviates mass elections' need to break results down by locality. (Protection against voter corruption or coercion [f].)

Selective denial of voter access to ballot casting is common in mass elections. Physical intimidation of certain demographics along access routes to polling places, outside of polling places, or even within polling places, are known even in countries with developed civil society infrastructure. Non-opening or under-equipping of polling places with particular demographic biases has also impacted election outcomes. Automated calls or letters misdirecting voters to a wrong polling time or place are often reported. The system presented here obviates attacks related to misdirection, by including direct addresses on ballot forms; it also obviates the above denial of access attacks since there are no poling places. Longer voting periods, and even rolling elections, can be expected with random-sample voting, making blocking of online access also considerably more difficult than with online mass elections. (Ensuring access to vote casting [g].)

Abuses sometimes called “voter fraud” involve voting by those who are unregistered or vote more than once. Doubt is sometimes raised as to whether such “retail” threats significantly affect election outcomes; however, there has in some cases been apparent concern sufficient to cause restrictions on participation and even disruption of civil governance. The system introduced here substantially reduces such concerns, as it does not allow voters to select themselves, but instead itself selects voters from the rolls verifiably at random. Of course any voting system’s resistance to voter fraud is only as good as its roster of registered voters, something shared by both types of elections. (Resistance to voter fraud [h].)

Decisiveness, the ability of an election to come to a conclusive result, is limited by what may be called its “precision”: how close a contest it can meaningfully adjudicate. Precision has both a technical and a policy dimension. Best practices with conventional elections include thresholds, though they are often set on unverifiable totals with poor accuracy and thus are of questionable meaningfulness. In case of ties, unverifiable and often flawed random procedures are required in some cases under current law. In theory mass elections could properly adjudicate extraordinarily close contests, even up to tie, yet their precision in practice does not justify this and actually makes it undesirable since this can amplify the efficacy of even small manipulations.

With random-sample voting, it would be preferable to discard and re-run with different parameters when a result falls below a preset threshold of sample size or confidence level. The need for a decisive outcome, however, may also be reduced by the decoupling with election cycles already mentioned. A random-sample election could also be structured so that if the tally is too close, the result is in effect verifiably-randomly determined by reverting to randomness when the underlying precision is exceeded, a better way to structure a vote that must be decisive by a date certain. (Decisiveness [i].)

Cost, compared to a mass election can, as mentioned, be extremely low because the number of voters is far less and cost is primarily per voter. (Low cost [j].)

Thus, random-sample elections come significantly closer to the desiderata above, as more generally stated: (a) higher participation ratios, (b) better informed voters rationally motivated to delve into issues, (c) increased effectiveness of results in shaping governance, (d) improved resistance to manipulation through advertising/campaigning, (e) increased indisputability and trustworthiness of results, (f) anonymity of voters with unsaleable votes, (g) reduced opportunity for selective denial of voter access, (h) voter fraud only through improper voter rolls, (i) equivalent but safer decisiveness, and all with (j) significantly reduced direct and overall cost.

INFORMAL SUMMARY OF TECHNICAL CONCEPT

Indisputability of random-sample elections derives from public verifiability that:
(i) The selection of voters is at random and cannot be manipulated.
(ii) The tally correctly reflects the votes cast by the selected voters.
(iii) Votes remain unlinkable to voters.
(iv) Vote buying is impractical.

How each of these four properties is achieved may be summarized as follows:

(i) A pre-agreed public random event, such as stock-market closing data or weather data, on a day agreed to in advance, determines which voters are to receive the ballots that will be counted. How the public random event data will be interpreted is further randomized in a way that is publicly committed

to before the event but remains encrypted at least until after the voting. Thus, voters are selected essentially independently and uniformly from the list of potential voters in a manner that even the election authority cannot manipulate and remain anonymous at least during the voting.

(ii) Ballots mailed to the randomly-selected voters will be verifiable to those voters as being extremely likely to be counted correctly, as a consequence of published data that can be audited online by anyone. And there is end-to-end verification that voters received correct ballots. Such verifiability is similar to that provided voters in elections conducted using Scantegrity in Takoma Park, Maryland.

(iii) Voters cast votes using "vote codes" printed on ballots. These codes are posted online, but which vote corresponds with which code remains hidden. To allow verification that the selected voters did in fact at least receive the opportunity to vote, though the identity of the selected voters is, at least in the basic system, made public but only once voting is over. (In an optional variant, a number of "verifiers" are selected at random, each provided after close of polls with the identity of a different one of the voters, and later only the identities of these verifiers, but not the voters, are made public. Each verifier is instructed to contact that verifier's voter to check whether the voter has in fact cast a ballot and to raise an alarm if the ballot was not at least received by that voter. This variant thus allows verifiers to be checked on by the public while essentially keeping the identity of voters from becoming public.)

(iv) To keep vote-buying in check, any eligible voter can request a "decoy" ballot. These will not be counted. They can, however, safely be sold because they will remain indistinguishable from actual ballots in principle forever. Requestors of decoys will presumably try to sell the decoys to vote buyers. Those running the election should monitor the market for decoy ballots and endeavor to keep supply such that vote-buyer offers are below a price likely to entice the randomly-selected voters to sell. (Since voters are anonymous before the election and the link from votes to voters is never revealed, manipulating elections by coercion or misinformation would presumably also be difficult.)

Compensation of voters or verifiers was mentioned above as an option. Verifiers, as well as optionally voters who answer verifier queries, may collect rewards conditioned on meeting certain public criteria that obviously must avoid biasing the outcome. Such criteria could, for instance, include consistency of various responses or extra effort by the voter. To allow authentication between verifier and voter and then provide evidence of the successful verification, verifiers can be issued parts of a numeric confirmation code only known fully to the voter. Additionally, if ballots are mailed "signature required," then the authority has some recourse against a voter falsely crying foul.

The participants in a basic random-sample election may be divided into four categories: a single election authority, two disjoint sets of members of the public, and an open-ended collection of auditors. More specifically, the categories and their roles are:

The election authority—commit to data by posting it in encrypted form, certify results of public random events, receive requests for decoy ballots, monitor decoy market, print ballots, mail ballots, and reveal during audit those keys to commitments selected by public random events;

Randomly-selected voters whose votes will be counted—receive a ballot in the mail, vote by providing vote-codes online, and optionally check codes/ballots online and/or respond to audit inquiry;

Self-selected decoy voters whose votes will not be counted—request decoy ballot, receive decoy ballot in mail as if an actual ballot, and try to sell decoy ballot/vote; and

Members of the public who audit the election process—run open-source or self-written software that decrypts (using keys selected for release by a public random event) the published encrypted data and checks its consistency with: election protocol steps, results of public random events, bulletin-

board data, published encrypted data, and keys to published encrypted data once released.
One way to determine how many votes are needed to establish a majority is familiar from coin tossing. Few would dispute that the odds are overwhelming that 20 tosses of a fair coin would include at least one tail, since the chance of all heads is less than one in a million. The same odds hold, for example, for: 22 flips having at least two tails, 25 flips three tails, and 100 flips 28 tails. (In the language of statistics, the relative frequency in the sample space of 100 independent fair coin flips of the event defined by less than 28 tails, calculated using the binomial cumulative distribution with success probability one half, is less than one over one million.)

Suppose organizers of an election believe that about three quarters of voters are in favor. If their belief were reasonable, when they collect 100 randomly-sampled votes it is likely that less than 28 voted against—and such an outcome would itself establish, independent of their belief, and with overwhelming odds, that at least a simple majority are in favor. If they believed support is less than 75%, then larger samples are required. For instance, since 191 out of 300 gives the same overwhelming odds, 2/3 support would likely return a suitable number of affirmations in 300 or so votes. For 60% support, 1000 voters would mean 576 affirmations establish a one-in-a-million chance of error.

Questions that are believed too close may be recast to provide more economical margins. Less compelling odds may also be acceptable or can even result in overwhelming confidence when votes earlier in a series, such as with a ladder or ongoing polling, are all corroborated by subsequent votes. Generalization to super-majorities, multi-way contests, and even correcting for the distribution of voters using different rosters, are all possible though perhaps unnecessary.

CONTEXT

Random-sample voting has context in the historical and current technics of democracy.

Much of what is regarded as Western culture and civilization, including philosophy, constitutional law, science, mathematics, medicine, sculpture, theater, music, literature, sports, and arguably even the phonetic alphabet—akin to the printing press or Internet of its time—can be traced to the 170 years of democracy during the “classical period” of Ancient Greece. All governance, not only adjudication of criminal and civil disputes, was decided by large, randomly-selected juries without judges and using often secret but only yes or no votes. (Such juries could, for instance, be invoked by any citizen to decide constitutionality of legislation and criminal penalties for legislators even only proposing unconstitutional bills.) A then appreciated attribute of their practice of random selection to fill most government posts, called “sortition,” was its resistance to lobbying and corruption—something present in their democracy but now lost in our “representative democracy.”

Juror selection in common law countries, albeit unverifiable and heavily post-culled, is all that remains of Ancient Greek democratic mechanics. Random selection of citizens, however, is widely used by government much more heavily today: conscription for military service, for instance of over fifteen thousand US conscripts who lost their lives in Vietnam; random selection of tax audit subjects; selection of citizens in policing generally, for instance at airport checkpoints; and random selection of poll workers, such as in Brazil. Safety testing and regulation by government in medical, food, transportation, and other sectors is often based on random selection or sampling (as is research underlying much scientific advance). More particularly relevant, random selection is required by law in various jurisdictions around the world in ways that can directly affect election outcome, such as random listing order of candidates, random selection in transferable vote counting, and random choice

in case of apparent tie.

Random selection is also used heavily by political parties, candidates for office, and interest groups—in electioneering—far less so, however, beyond trying to influence election outcomes. Respondents in so-called "public opinion polling" are in principle selected at random and asked to quickly answer a series of questions related to candidates and issues of the day. As currently practiced, however, such surveys are neither transparent, trusted, nor generally believed worthy of public trust.

Public opinion polls are traditionally conducted by banks of questioners calling more or less random phone numbers using a script aimed at identifying a random voter among those present in the household. Mobile phones have rendered this even more difficult and online pools of persons and even persons selected in shopping malls are also used as respondents. Demographic data obtained from respondents, often of dubious quality, is often used presumably in an attempt to correct for the bias in the sampling method, but based on unpublished models that may themselves bias outcome. Turnouts are very low. Respondents have little if any time to explore the meaning or consequences of questions, let alone think about, research or deliberate on, issues in such "opinion" surveys. (The related field called "deliberative democracy" is premised on the utility of group settings and has even been used in binding public-sector elections, but suffers from high cost and is subject to well-studied techniques developed for manipulating juries.)

Surveys have other well-known problems as well. How questions are worded and sequenced is a notorious form of bias. Questioners also bias the results, whether they intend to or not, by how they communicate. The published results of multiple such polls often deviate significantly on the same question, depending on the orientation of the entity conducting or sponsoring the survey, contributing to erosion of public confidence in such polls generally. Opinion polling predicting election results, a significant revenue source for polling organizations, is at best as flawed as the underlying elections. It has a self-fulfilling effect, and consequently some countries ban publication of its results close to elections. It has even been suggested that Asimov's 1955 short story "Franchise," about an artificial intelligence that votes for a whole society based on a few quirky questions asked of an apparently randomly-selected citizen, is a kind of half-hearted straw-man argument by scenario against public opinion surveys.

The techniques presented here could be taken to be merely an incremental improvement on public opinion polling, but are better understood instead as potentially providing so large a qualitative improvement, even surpassing the quality of mass elections themselves, that they offer really new options for democracy.

Some countries, such as Switzerland, routinely have binding referenda on questions at a national level; many countries have them only rarely; and most countries, including the United States, have no provision for national referenda, though the ten westernmost US states do provide for initiatives and many even on constitutional changes. The techniques introduced here go a long way toward mooting traditional first order reservations about referenda: the lack of deliberation, inability of the public to deal with complex issues, and ease with which the public can be manipulated temporarily by media campaigns. Ironically, these same concerns, obviated by random-sample voting, are present in representative democracies, where policies often seem driven by short-term waves of public opinion.

It is known that those who control what candidates or questions are put to vote, how contests are grouped and ordered, and when polls are held, can significantly influence outcomes, even with more direct systems like referenda. Dramatically lower cost might allow referenda to be conducted without

the need to aggregate or sequence contests, fit a sparse schedule, or even let government control the questions put to the electorate. Thus, these second-order concerns may also be obviated.
A related perhaps tertiary problem is what are called election (or more generally aggregation) paradoxes, ways to combine policy choices so that preference for various combinations seem surprisingly at odds with at least some intuition about the original individual policy choices. Arrow's impossibility result from social choice theory is a well known example, but the more indirect the bigger the problem. Such issues are avoided when independent or "separable" yes/no policy decisions are voted individually but in parallel, and non-separable issues voted sequentially. Using random sample elections, these separate contests can readily be conducted in practice and at low cost.

ADOPTION SCENARIOS

Example early adoption scenarios for random-sample voting flowing from the above discussion include interest-group initiated surveys of the public on a particular issue. Such an election offers its sponsors a way to make a statement about the will of the electorate that is irrefutable—far more compelling than petitions and perhaps more effective and less damaging than protests. It could also legitimize a spokesperson or organization, such as to serve as a representative in a consultative process.

Another example use is public opinion surveys related to mass elections. For one thing, this would be more resistant to manipulation and thus potentially of more interest and less likely to be banned. For another, many mass elections around the world are disputed, such as by a mixture of allegations of technical fraud and registration and polling place discrimination. A parallel random-sample election with improved turnout might neither validate nor invalidate a mass election, but it could perhaps more importantly provide indisputable information on the will of the electorate.

An example use, which is generative of ideas or suggestions and competitively filters them, is a ladder of contests to establish issues to be subjected to a general vote. Such a ladder simultaneously and transparently arrives at specific language for questions that emerge from it. The sample size could be small, while the margins required large, at the lower rungs and increase as successful wordings move up the ladder. Electoral rules suitable for allocation, such as the variously re-discovered techniques sometimes referred to as "approval," "range" or "bee" voting, can be more efficient than a tree of binary votes when the number lower-rung questions is large.

One variation for ladder entry uses "CAPTCHAS" or allows each person to pay only a limited amount towards a question or answer such that when a question receives sufficient funding it is put to vote. Another variation uses delegated or liquid proxy voting to fill lower rungs. Providing such avenues for self-selected supporters to push questions to vote advantageously allows motivated or expert groups a way to develop questions without allowing such groups to control outcomes.

The scale of electorates need not be huge, however. Use with smaller populations, such as large volunteer organizations, universities, or corporations, may allow a wide range of less contentious issues to be efficiently decided by what are in effect small, randomly-constituted committees, where near unanimity can keep sample size quite small.

One example kind of transitional scenario towards full use in governance lets candidates for political office or political parties commit in advance or later simply opt to use random-sample voting. They could use it to determine one or more aspects of how they exert their elected power, such as by voting a certain way in a legislative body or implementing a policy. For instance, guaranteeing constituents veto

on any vote for war, a right interestingly granted women in the representative democracy of the Iroquois nation.

Limited initial use by governments is another type of transitional scenario. When, for example, a particular issue must under law be subject to vote, and a mass election is too costly compared to the significance of the issue or the delay in conducting a mass election is problematic, a random-sample election may be used instead. Another example limited use is an initial culling of candidates, perhaps a way to greatly improve political party mechanics, allowing more focused debate and voter focus on a smaller set of candidates chosen without bias. A conventional initiative or referenda usually requires submission of a certain number of signatures, presumably to enhance the chance that the measure is worthy of public attention and winnable. Allowing a random-sample election as a trigger (to say nothing of its ultimate superiority in conducting the final vote) would do a better job and at lower cost.

Yet another kind of transitional scenario involves groups not well aligned with the largely geographical hierarchy of governments today. Some regional groups, for instance, are split between governmental jurisdictions; their elections could be allowed to define their own boundaries. A related use is appointing external representatives of such groups.

There has never been a global referendum. Major issues are increasingly global, but democracy does not currently exist between or above nation states. Moreover, democracies, whose sovereignty derives from that of their citizens, may be hard pressed to justify opposition to the studied and proven will of a majority of the citizens of each major country, at least with respect to global issues. While there are a variety of huge issues that would no doubt engender substantial global interest and support, one kind of issue specially relevant here relates to information policy and rights. Voters being able to obtain information without risk and arrangements that promote transparency and quality of information available, illustrated above, are particularly relevant to effective elections and thus democracy itself.

The cost of even a global election might be as low as ten million dollars for a sample size of a thousand or so, since the lion's share of voters would be in countries with adequate infrastructure. In countries without such infrastructure, voters can still be identified as the person living in a dwelling selected randomly from those in a randomly selected region. (Rules for automatically selecting regions of similar population from satellite imagery and for ordering dwellings within regions would ideally be fixed in advance.) Then an election worker, perhaps with a translator, would go and collect the vote. Example topics include Internet privacy and access, which are additional examples of issues relevant to elections themselves, as well as income distribution, environmental, human rights, drug and food policy as well border and sovereignty issues.

What if someday a government wished to, perhaps because of expression of the will of the electorate after positive experience with non-governmental elections, change the legal framework so that they could start really using random-sample voting. What kinds of elections might be considered? Simply deciding the same questions typical of mass elections would probably rank rather low, as there would be little advantage and less voter involvement. Major policy decisions are an obvious choice for referenda. Switzerland's what is sometimes called "half-direct" democracy lets voters approve laws and this could also take advantage of the increased number of elections possible with random-sample elections. Another example is appointment of the key persons actually running governments, such as cabinet member heads of major departments, key legislative committee leadership roles, and even supreme court judges. The increased number of elections could allow such more fine-grained democracy.

DISCUSSION

Conventional secret-ballot mass elections are a paper-based technology that is 150 years old (concurrent in the U.S. with extension of the franchise to male citizens, roughly, although women's suffrage took another 50 years). Today's sophisticated information technology, however, is used to manipulate political processes by those with more direct access (for example in redistricting, ferreting out and influencing public opinion using random sampling combined with big data, and interest group monitoring of legislator votes). Meanwhile, voters have so far been left with a technology for voting—or a high-tech simulation of it, which makes the point that it is actually only a paradigm for voting—that was developed and promulgated well after the US constitution but before the first telephone call.

Random-sample voting can thus be interpreted more broadly as providing a way forward, from our current paradigm-induced disparity in access to the power of information technology, towards allowing effective voter steering of governance.

In future there may be some who long to vote in mass elections, perhaps romanticizing about the act of casting a secret ballot in person among one's neighbors or at least the chance to submit a vote no matter how futile. But there will likely be few who would oppose the deeper and wider and more continuous monitoring of the will of the electorate provided by random-sample voting—once efficacy is established—at least informing if not being binding on governance.

CONCLUSION

Random-sample elections offer practical low-cost yet unprecedented quality for almost any election, with a range of immediate applications, and each election potentially a step towards more effective and finer-grained democracy at scale.

APPENDIX

DETAILED ELECTION PROCESS

A random-sample election can be conducted as illustrated in the diagram on the last page. It shows an example protocol including all the data and a ballot for an election, using what might be called an Eperio-style variant of Scantegrity I. This concrete example, introduced in the "Informal overview of technical concept" section above, is well suited to practical implementation. Its basic ingredients, concepts, and techniques are first described from a number of perspectives highlighted below, followed by the complete step-by-step protocol recipe.

In terms of the voting experience, each voter receives by mail a paper form like that shown on the left side of the diagram. The voter first chooses freely, and ideally randomly, one of the two ballots printed on the double-ballot form received, in the example either serial number #100a or #100b. The voter then enters this ballot serial number online. To cast his or her vote, the voter next enters online the unique "vote code" printed on the chosen ballot form adjacent the desired "YES" or "NO" vote. For the ballot serial number not voted, at least some digits of both vote codes and their corresponding votes are displayed to voters at some point, so voters can help check that ballot printing does in fact pair votes with codes correctly; if inconsistency were to be detected, voters would have the printed form as convincing evidence of improper printing but that reveals nothing about the vote cast. All other audits can be conducted independently online by any number of interested parties on behalf of all voters.

(The red ovals in the diagram highlight the example choices made by an example voter, with "A" being the choice of the upper ballot of the pair on the double-ballot form and "B" the "NO" vote and its vote code, both of which appear in the table instance shown.)

In terms of who does what to make an election happen, the parties, as introduced earlier, include the Election Authority, called here the "EA" for short, who conducts the election. There are also the voters and decoy voters, who cast their ballots but also optionally check data online and answer queries to help check that their own votes are properly recorded. Anyone can be an "auditor" and check published data as well as check with voters; this role is aimed at ensuring election integrity by trying to find evidence of any deviation from protocol by the EA; the role of the EA in audit steps is that of auditee. The protocol is aimed at ensuring that the EA can neither influence the selection of voters nor change the outcome without an extremely high probability of getting caught by auditors.

In terms of overall election timeline, the centerpiece interval is when voters are allowed to cast their votes online. There are two other prescribed time intervals, one preceding voting and one following voting, during which random values are harvested. The two gaps between these three intervals as well as both the period immediately preceding and following them are when the EA publishes data.

The first of these four publications by the EA defines the election and locks in a hidden mapping from random values to voters. The second publication allows decoy voters to be included, which is followed by printing and mailing of the ballots. The third publication, following the voting interval, translates the vote-codes cast by voters into conformance with the data previously published by the EA.

After the final random interval, as its fourth publication, the EA is required to reveal oly some of the keys used to encrypt previously published commitments. Exactly which keys is determined by the random draw. This release of keys allows auditors to "spot check" with extreme effectiveness that all the encrypted data published by the EA was according to protocol. Combined with cross-checking what voters reported online and report when queried by auditors, this confirms correctness of the random selection of voters and published tally—but does not compromise ballot secrecy.

In terms of the protocol to be followed by the election authority, an election takes place in nine steps. These are enumerated in the frieze across the top of the diagram and are conducted one after another. Before these steps can start, particulars defining the election must be fixed. These comprise the roster of eligible voters, the number of real and decoy voters, the exact language of the question and response options, how the EA will authenticate its postings, when polls will be open, the random draw sources and dates, where voters can vote online, and where all election data will be posted.

The first step consists of the EA using its own private source of random numbers to create tables of data and then independently encrypting and publishing each column of each table. The secrecy of the random numbers used by the EA to create this data protects secrecy of who voted (at least until after close of polls) and which way they voted and indistinguishability of decoys. The second step is a public random event, such as a stock market closing, for which the EA then reformats and publishes the result. In the third step, responsive to requests for decoy ballots, the EA publishes additional encrypted data that assigns the decoy voters to the decoy ballot slots that were dispersed indistinguishably among the real (i.e. non-decoy) ballots in the first step. The third step also includes the EA physically printing and mailing ballots.

The fourth step, the actual voting from opening of polls to closing of polls, accepts online submission of votes on a so-called "electronic bulletin-board." This is followed by the fifth step that translates the resulting posted vote codes and serial numbers into two additional columns that are encrypted and posted to complete the encrypted tables.

The sixth step, like the second, is an independent public random event. It determines exactly which keys the EA must release for each of the three remaining steps. The keys the EA releases in the seventh step let anyone crosscheck consistency between the voter bulletin-board's public data and the published encrypted data; the public yet unpredictable nature of this choice of which parts the EA must make decryptable allows anyone to confirm that the bulletin board is consistent with all the encrypted data. The eighth step is the release of further keys that, in addition to allowing checking internal consistency of the published tables, reveal the tally. The ninth and final step is the release of just enough additional keys to reveal the voters, so anyone can check with them to ensure that they did in fact receive the correct ballots, without revealing which ballots were decoys or who voted which way.

In terms of data posted by the election authority, apart from definition of the election and reformatted public random data, all that the EA posts for an election is a number of identically-formatted encrypted tables and eventually keys unlocking some their encryptions. More specifically, each table consists of eight columns and each column of each table is encrypted with a different secret key chosen randomly by the election authority; the authority only reveals some of these column keys for some tables, and only during audit, with the choice of which keys to reveal determined by the final draw.

To create each table, the EA transforms a copy of a "canonical" table that simply lists in a known order all the serial numbers, vote codes, possible votes, and a pair of numbers. Each of these transformations is randomized by the EA to keep secret the linking between voters and votes, which ballots are decoys,

and (at least until after close of polls) who voted. One transformation is “row-shuffling,” resulting in an unpredictable re-arrangement of the rows, so a particular row almost always appears in a different position in each table. Another transformation further randomizes the pairs of random numbers, here called the first and second “summands,” but keeping their special sum unchanged. (The initial random values of the summands is the same per serial number.) The third and final transformation encrypts each column as a whole, using a secret key that is unique per column and per table. To protect election integrity, the class of encryption used ensures that column content is unalterably committed to, as there is only a single decryption possible per encryption.

The specific example illustrated has 250 tables, the number recommended for an election of any size. This particular example sends double-ballot forms to 1,000 voters, divided somehow between real and decoy. The double-ballots thus result in 2,000 distinct ballot serial numbers and, since there are two choices per ballot, 4,000 rows per table. The 1,000 double ballots are each assigned a voter among the example 10,000 potential voters on the posted list, here called the “roster” (sometimes elsewhere called a voter or registration or electoral roll or file). Instead of posting all columns of the tables at once, four columns are posted first (yellow, 1, 3, 5, and 6), two more to select the decoys (green, 7 and 8) before voting but after the initial draw, and the remaining two (blue, 2 and 4) after voting to lock in votes.

The sum of all three summands corresponding to a serial number yields the position in the roster of the voter who should get that ballot. If the sum is less than the number of voters, it is simply the row number in the roster, numbered from zero to the number of voters minus one. If the sum of the three numbers is larger, then it is reduced to fit, using so-called “modular arithmetic.” This means that each summand can influence the result to be any voter (just as allowing a choice from zero to eleven can shift the hour on the face of a clock to any hour). Ensuring that EA choice of summands does not result in any voter receiving more than one ballot is accomplished by audit.

(The red ovals in the diagram follow the example vote already mentioned along its row in table number one, the table shown on top of the other 249: the ballot serial number and vote code in oval “C”; and in “D” a “NO” vote marked “VOTED.” The second summand pair “E” is 0000 and 5555, with the same sum as the other pairs for this double-ballot, 5555. The entry in the list of third summands, corresponding to serial #100a, “F,” is 2222. Simply summing all three summands, 0000+5555+2222=7777, yields the position number “G” in the “voter roster” address list to which this ballot should be mailed. These example summands were chosen for convenience with repeated digits and not requiring carry; however, as summands are chosen from zero to the number of entries in the roster minus one, 0 to 9,999,999 in the example, carry and modular arithmetic would be needed for typical values.)

In terms of implementation system security, all secrets of the EA can be stored in encrypted form between steps. An election committee then meets for each step and supplies so-called “passphrases” in person to a fresh installation of a computer that they witness. The computer then decrypts the stored EA state, conducts the step, output the new encrypted state for storage, and wipes its own memory.
The online bulletin board need only be provided by the EA with checksums for vote codes and respective serial numbers. This lets it confirm to voters that they have entered the vote code correctly even though there are many vote codes per checksum. The bulletin-board would only be able to learn codes voted and thus be unable to create or change votes. Voters would be instructed to write their voted vote-code on the unvoted individual ballot, destroy the voted individual ballot to protect the secrecy of how they voted, but then keep the unvoted ballot half until they can check it once the choices on the unvoted halves are publicly linked to vote codes in audit.

Random-sample elections lack the time urgency of mass elections and accordingly are less subject to

so-called "denial of service" attacks online, which are typically short lived. Discreditation by voters would involve claims that they did not receive ballots or that the forms were printed incorrectly. These are effectively addressed by certified mailing of ballots and standard document security techniques applied to ballot forms the halves of which voters can keep as evidence without betraying how they voted.

If someone were to wish to bias the selection of voters or change the election outcome, at least with a significant probability of success, compromising all committee members would not help. One potential type of attack, however, would be to somehow substantially change the random draws. Other potential attacks would involve somehow tricking enough voters into not following protocol or into falsely believing that they are in secured communication with the voter bulletin board. Auditors would use so-called "public-key digital signature" authentication specified in the election definition in order to authenticate all information that they receive from the EA, bulletin board, and beacon.

If someone were to wish to learn how particular voters have voted or to distinguish decoy ballots, one attack would be to seek passphrases from sufficient committee members, though this should become impossible once enough members erase their passphrases after the election. Another attack would be to somehow tap or subvert the computer that the committee uses.

If someone were to wish to merely learn the recipients of ballots early, they could do so by observing the envelope printing or mailing. Using this information to try to influence voters, however, might be risky since it could be revealed by voters or even by decoy voters aiming to entrap.

In terms of the random draws, called out as Step (2) and Step (6) in the diagram and recipe below, they are based on publicly-verifiable yet unpredictable data. For instance, closing prices of a set of stocks or temperatures from a number of cities would be obtained and posted by a “beacon” for this purpose, though anyone can also verify the data by tracing to its origin. The beacon uses a pre-defined method that completely determines the public outcome from the raw, definitive, public data of the physical event. The initial draw determines the public list of third summands, used to make the selection of voters unpredictable and unmanipulatable by the election authority; the final draw determines which of the column keys are revealed for audit.

In terms of audit, this process is divided among the last three steps. Step (7) makes sure that the vote-codes cast by voters recorded on the bulletin board were correctly copied into the encrypted tables and that the printed ballots correctly associate votes with vote codes. Step (8) reveals the tally. It also checks that the real ballot rows were not changed when the decoy ballots were interspersed. Step (9) lets anyone check with voters to make sure that they did in fact receive the correct ballot pair. These steps correspond roughly to customary phases in mass elections, where preliminary checking is done before the results are announced and then an extensive canvasing precedes final certification of the result.

The surprising public auditability of election integrity while maintaining ballot-secrecy, decoy-indistinguishability, and, at least temporarily, voter-anonymity, results from use of a so-called “you cut and I choose” protocol. The EA posts the encrypted tables but the random beacon chooses which ones to open. This prevents the EA from posting anything but correct tables—otherwise the discrepancy would, at least with extremely high probability, be detected. The reason is because the final draw determines unpredictably which one of the five “batches” each table ends up in and each batch decrypts a different combination of columns and those combinations cover all columns in a sufficiently interlocked way. An election authority posting table content that deviates enough to likely alter the

outcome or manipulate the choice of voters would be detected with overwhelming odds; very few of the vast number of ways the tables can be divided allow any particular deviation to go undetected. The combinations of columns decrypted per batch, however, reveal neither how any ballot serial number was voted nor which serial numbers are decoys.

Specifically, Step (6) divides the 250 tables into the five batches, each of which contains 50 tables, but in a way that was completely unpredictable to the EA when the tables were locked in by being published in Steps (1), (3), and (5). Each batch corresponds to a particular selection of table columns whose keys are to be revealed by the EA during a corresponding step, in effect "opening" the already committed but hidden content of the selected columns to public inspection. Two batches of tables are opened in Step (7), two in Step (8), and the remaining batch finally in Step (9). The diagram shows lines below the tables grouped in batches, indicating which columns are opened for which batches.

The first audit, Step (7), labeled "audit casting & printing" in the diagram, reveals keys for two batches to allow checking consistency between the tables, the voter bulletin-board, and ballot forms, without revealing the tally or voters. The first batch of columns opened lets anyone check that the vote codes, as posted on the bulletin-board and identified by serial number, are recorded correctly and marked as "voted" (and "not checked") in the decrypted data. The second batch of columns opened lets anyone cross-check, at least for those ballots not voted but checked by voters, that the serial numbers and corresponding vote codes were associated with the correct "YES" and "NO" ballot choice in the encrypted data, ensuring that the ballots were printed with the correct association.

The middle audit, Step (8), "audit voter selection & tally," decrypts the third and fourth batch columns to reveal the votes so that they can be tallied while revealing the decoys to ensure that they are not tallied. Anyone can then add up the revealed votes and check that each of the 100 tables yields the same election result. These two batches also let anyone check that the list of who should get a ballot was correctly copied from the real voters committed to before, and randomized by, the first draw—even though it lets pre-arranged interspersed positions be filled by whatever decoy voters the EA chooses.

The last audit, Step (9), "reveal all voters," opens the fifth and final batch and thereby lets anyone see who ballots were to be sent, so that voters can be contacted and asked whether they received the correct serial number double ballot. It is similar to what in mass elections is referred to as "canvassing." It discloses all roster entries to whom ballots were to be mailed and the respective serial numbers, but nothing about how the ballots were voted nor whether the voters/ballots were real or decoy. It decrypts the serial number column and a pair of summand columns. Also checked (assuming that any duplicate real ballots, those that would result in more than one per voter, should verifiably be deleted as detailed below) is that no voter should get more than one ballot.

In terms of properties achieved, the three audit steps address the first three properties already mentioned in the "Informal Summary of Technical Concept" section above:

Property (i), "The selection of voters is at random and cannot be manipulated," is addressed by Step (7), since the third summands are unpredictable and unmanipulably posted once the first and second summands have, in advance of this step, been publicly committed to. (No voter receives more than one ballot, as Step (8) and Step (9) ensure that ballots are not sent to those voters chosen more than once.)

Property (ii), "The tally correctly reflects the votes cast by the selected voters," is addressed by three aspects of the audits. First, that columns 2 and 3 correctly link "YES" and "NO" to the unique vote codes printed. This is established with high probability by voters checking their unvoted ballot halves against

data revealed in Step (7). Second, that these "YES" and "NO" postings are correctly linked to voted but "non-decoy" column 5 or 6 entries and that all the tables of the third and fourth batches have the same tally, both as checked by audit Step (8). And third, that all voters listed in the roster at the positions determined by the column 5 and 6 summand pairs (which, as verified in Step (8), highly-likely match column 7 and 8 pairs) were provided non-decoy ballots with corresponding serial numbers (which, as verified in audit Step (7), highly-likely match column 2 serial numbers) as verified in audit Step (9).

Property (iii), "Votes can remain unlinkable to voters," still holds even after all the auditing is done. This is addressed by batches 2, 3, and 4 being the only batches including actual votes (column 3) and batch 2 only linking to details of non-voted ballots, and batches 3 and 4 only linking to single random summands and/or decoy indications—and by columns decrypted in these batches being linked neither to voted vote-codes nor voted summand pairs.

The optional variant to the canvasing mentioned earlier, where verifiers provide anonymity of voters, can be substituted by using an elaboration of the same techniques.* That option can also be included in more distributed variants where the election authority is replaced by a multiparty computation that protects privacy from all but collusion or compromise of a majority of potentially multiple election authorities including the option of election authorities per portion of the electorate.

The complete detailed step-by-step recipe, based on the ingredients and diagram already introduced for conducting high-quality random-sample elections, is now detailed in the nine already-mentioned steps:

(1) [EA] Four encrypted columns (yellow, 1, 3, 5 and 6) are posted by the EA in this first step. To generate these, the EA first forms a four-column canonical table. Column 1 of this table is filled by assigning serial numbers for all ballots sequentially down the rows, two for each "a" followed by two for each "b" suffix, and assigning a different random vote-code per row. Column 3 is then filled simply by alternating "YES" and "NO" per row. Next the desired number of decoy ballots determines how many of the adjacent four-row double-ballot "blocks," those with the same numerical serial number, are randomly selected and marked "DECOY BALLOT" in columns 5 and 6. The remaining blocks each get instead a random summand copied across all four of their column 5 entries and another independent random summand copied across all four of their column 6 entries, where summands are chosen from zero up to one less than the number of voters on the roster.

To form each of the actual 250 tables posted, the EA begins with a copy of the canonical table and transforms it, first with two randomizations and then with encryption. One randomization permutes the rows, with the column entries making up a row remaining together in whatever row they are moved to. The other randomization changes the summands of columns 5 and 6, while preserving the sum per row, such as by adding a random value to one summand and subtracting the same value from the other (with summands always represented modulo the roster size). The authority finally then separately encrypts each column 1, 3, 5 and 6 with an independent secret key. Once all 1,000 columns are encrypted using their respective one of the 1,000 keys, the EA posts all 250 encrypted tables.

(2) [beacon; EA] Initial random numbers are drawn in a public and unpredictable manner once both Step (1) is completed and the roster of voters (blue rectangle, left side below ballot) is fixed. These numbers consists of third summands (purple rectangle, lower left corner) each labeled by its respective unique double-ballot serial number. The public rule defining how these third summands are formed from the draw, fixed in advance of this step, preferably ensures that they are independently and uniformly distributed numbers from zero to one minus the number of roster entries. The rule may be as

simple as taking an appropriate number of successive bits of raw random-draw binary output while skipping in their entirety those representing a value equal to or larger than the number of voters on the roster (though a more efficient algorithm could be specified by the election definition).

(3) [EA] Two additional encrypted columns (green, 7 and 8) are posted by the EA, who also then prints and mails out the double-ballot forms. To generate the two columns for a particular table, the EA first copies all the pre-draw summands from columns 5 and 6 to form new respective columns 7 and 8 entries. Where column 5 and 6 are marked "DECOY BALLOT," however, the authority places a random summand in column 7 and then computes the column 8 value of that row so that the sum of all three summands is the position in the roster of a decoy voter assigned that block. (Each voter is ensured of receiving at most one ballot: in the unlikely event that column 5 and 6 entries would cause a voter to receive more than one real ballot, all but one block of rows corresponding to such ballots have their summands marked "DUPLICATE" in columns 7 and 8; an agreed policy limiting the frequency of voter selection across multiple elections can also be verifiably realized by such marking; and decoys are not assigned, though statistically insignificant at scale, to voters who will receive real ballots.) Once all 250 columns 7 and 8 are formed in this way, the election authority encrypts each, using a respective one of the 500 independent secret keys, and then posts the encryptions. After printing the double-ballot forms, the authority finds the respective mailing address for each by looking up the third summand in the summand list by its serial number and then using the sum of the three summands, modulo the roster size, as the row number in the roster.

(4) [voters; bulletin-board] Voters cast votes by posting on the electronic bulletin-board the ballot serial-number along with the corresponding vote code each wishes to vote. Voters are requested to check that the unvoted individual ballot of the double-ballot form is posted with vote codes correctly matching votes. The bulletin-board posts serial number and corresponding vote codes.

(5) [EA] Once the polls close, the EA forms the two final columns (blue, 2 and 4), encrypts them using the final 500 respective independent keys, and posts the result. Column 2 records serial numbers and vote codes for the non-voted ballots of each ballot pair. Column 4 marks, without distinguishing real from decoy ballots, the rows whose vote codes were voted.

(6) [beacon; EA] The final public random draw is performed after the output of Step (5) is posted. This draw is used to unpredictably divide the 250 tables into five batches: 50 tables for each half of Step (7) and (8), with the remaining 50 for Step (9). The rule defining this division can be to place the first 50 table numbers that appear in the successive bytes of the random draw into the first batch, the second into the second, and so on, skipping any repeat or out-of-bound values so that no table appears in more than one batch. So that audit can begin shortly after polls close, this draw is ideally completed quickly (This draw can be shared by multiple concurrent elections and/or multiple such draws can be used for an election in a pre-defined way, such as allowing some tables to be assigned batches early and others only later when more trustworthy random values are available).

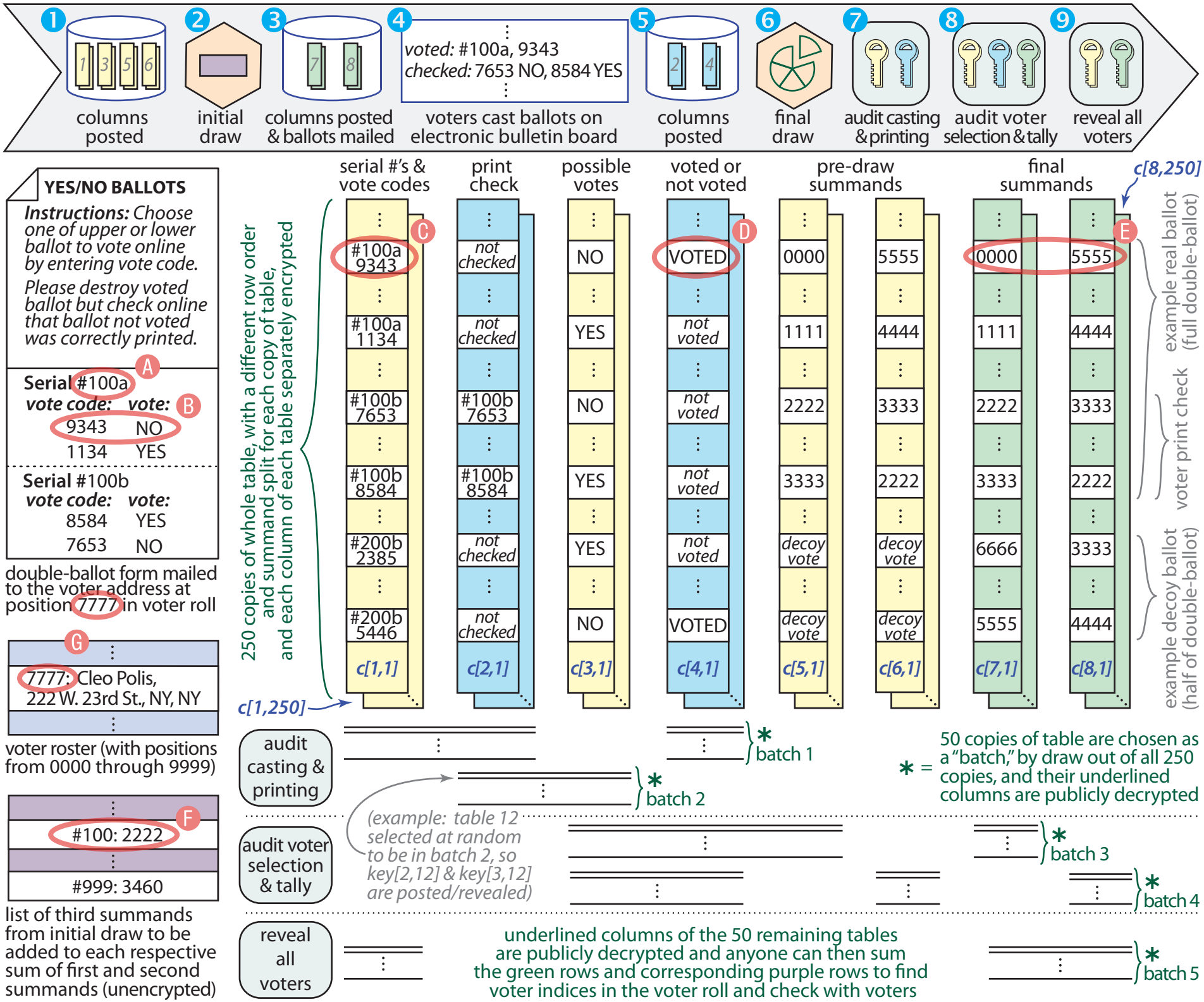
(7) [EA; voters; auditors] This first audit can be conducted as soon as Step (6) is completed. The EA releases keys allowing decryption of columns 1, 2 and 4 for the first batch of 50 tables and columns 2 and 3 for the second 50. The first batch allows voter and auditor crosschecking, bulletin-board against columns, of the voted and not voted serial numbers, as well as that serial numbers of columns 1 and 2 match; the second batch allows crosschecking, bulletin board against columns, of the pairing of vote codes and votes printed on unvoted ballots.

(8) [EA; auditors] This second audit is conducted when the tally is to be revealed. Columns 3, 4, 5 and

7 are publicly decrypted for the third batch of 50 tables and 3, 4, 6 and 8 for the fourth batch. For the third batch, decrypted columns 5 and 7 should be identical, as should columns 6 and 8 for the fourth batch, unless "DECOY VOTE" appears in the first column of the pair. Rows marked "VOTED" in column 4, but not marked "DECOY VOTE" in column 5 or 6, are valid votes and their summation is the tally, which should be the same for all tables in the two batches.

(9) [EA; auditors] The positions in the voter roster containing addresses to which ballots were sent are revealed by the EA in this canvass final audit step, but without revealing whether the ballots were real or decoy or how they were voted. The remaining batch of 50 tables is publicly decrypted in columns 1, 7 and 8. No voter should receive more than one ballot and each such ballot should have the same serial number for its block in each table of the fifth batch. (Summands marked as "DUPLICATE" should also have, in their respective block, the same three-way sum as a mailed ballot.) By using a serial number entry in the first column to look up a third summand in the published list and summing that with the summands in the other columns, any auditor can locate the addresses in the voter roster to which ballots should have been sent. The auditor can then verify, such as by alerting voters, checking mail receipt signatures or online receipt records, or even by directly contacting voters, that these voters did in fact receive corresponding serial-numbered ballots and that the voters did not complain that serial numbers and vote codes were improperly reflected on the bulletin-board.

* Each row entry of column 5, 6, 7 and 8 is appended with a "verifier summand," selected at random just as with the non-decoy summands. The original summands for columns 7 and 8 are encrypted with a special unique key per row. When column 7 or 8 is opened in Step 8, the "master key" for all its original summands is opened and they are revealed. When column 7 and 8 are opened in Step 9, however, the master key is not published but used instead to compute the individual sub keys for the original summands and these keys are encrypted and posted using the corresponding unique key that was mailed at ballot mailing time to the respective verifier.


1 columns posted
2 initial draw
3 columns posted & ballots mailed
4 voted: #100a, 9343
checked: 7653 NO, 8584 YES
voters cast ballots on electronic bulletin board
5 columns posted
6 final draw
7 audit casting & printing
8 audit voter selection & tally
9 reveal all voters
YES/NO BALLOTS
Instructions: Choose one of upper or lower ballot to vote online by entering vote code.
Please destroy voted ballot but check online that ballot not voted was correctly printed.
A
Serial #100a
vote code: vote:
B
9343 NO
1134 YES
Serial #100b
vote code: vote:
8584 YES
7653 NO
double-ballot form mailed to the voter address at position 7777 in voter roll
G
7777: Cleo Polis, 222 W. 23rd St., NY, NY
voter roster (with positions from 0000 through 9999)
F
#100: 2222
#999: 3460
list of third summands from initial draw to be added to each respective sum of first and second summands (unencrypted)
250 copies of whole table, with a different row order and summand split for each copy of table, and each column of each table separately encrypted
serial #'s & vote codes
print check
possible votes
voted or not voted
pre-draw summands
final summands
c[8,250]
C
D
E
#100a 9343
#100a 1134
#100b 7653
#100b 8584
#200b 2385
#200b 5446
not checked
NO
YES
VOTED
not voted
0000
1111
2222
3333
5555
4444
decoy vote
6666
c[1,1]
c[2,1]
c[3,1]
c[4,1]
c[5,1]
c[6,1]
c[7,1]
c[8,1]
c[1,250]
example real ballot (full double-ballot)
voter print check
example decoy ballot (half of double-ballot)
audit casting & printing
batch 1
batch 2
audit voter selection & tally
(example: table 12 selected at random to be in batch 2, so key[2,12] & key[3,12] are posted/revealed)
batch 3
batch 4
reveal all voters
batch 5
* = 50 copies of table are chosen as a "batch," by draw out of all 250 copies, and their underlined columns are publicly decrypted
underlined columns of the 50 remaining tables are publicly decrypted and anyone can then sum the green rows and corresponding purple rows to find voter indices in the voter roll and check with voters