The Role of Dice in Election Audits – Extended Abstract

Arel Cordero* arel@cs.berkeley.edu
David Wagner* daw@cs.berkeley.edu
David Dill† dill@cs.stanford.edu

June 16, 2006

Abstract

Random audits are a powerful technique for statistically verifying that an election was tabulated correctly. Audits are especially useful for checking the correctness of electronic voting machines when used in conjunction with a voter-verified paper audit trail (VVPAT). While laws in many states already require election audits, they generally do not address the procedure for generating the random sample [12]. The sample generation procedure, however, is critical to the security of the audit, and current practices expose a security flaw. This paper examines the problem of sample selection in the context of election audits, identifies necessary requirements for such a procedure, and proposes practical solutions that satisfy those requirements.

1 Introduction

Many things can go wrong in an election, whether intentionally or unintentionally. The ability to detect (and correct) errors is critical. Recent decades have seen widespread use of computers and automation in elections, including use of optical scan machines, DRE (direct recording electronic) machines, and computerized election management systems to record, tabulate, and report votes. Unfortunately, this trend has come at some cost to transparency, as the automation of these processes reduces opportunities for observers and interested members of the public to monitor the operation of the election. Random audits, performed after the election but before certification, remain as one of few defenses for ensuring fairness and for building public confidence in the result. Consequently, the details of these audits are of increasing importance to election integrity.

When done right, and in a transparent and publicly observable way, random audits can establish *objective* and *quantifiable* measures of election accuracy[1] However, without a transparent process, there is no reason to believe an audit will correctly represent the election, which would defeat objectivity, rendering it meaningless as an assurance of fairness. Even if correct, a non-transparent audit would have trouble quelling skepticism and could thereby fail to provide confidence in an election.

All parts of an audit must be performed correctly—and transparently—for the the audit to mean anything. In particular, getting the random sample selection process right is critical. To start with, the selection process must ensure every vote has an equal (or minimal) probability of being selected. Moreover, like any fair lottery, a second requirement is that no party be able to bias or predict the selection in any way. An important implication is: *for an audit to give every party confidence in an election, every party must also have confidence in the fairness of the sample selection.* In other words, sample selection can easily become a weak link in the security of an election.

A case in point is the 2004 U.S. presidential contest in Cleveland, Ohio [8]. Cleveland election workers, in an effort to prevent a complete recount, surreptitiously preselected an audit sample (3% of the total precincts) to ensure the ballots recounted by hand would match the initial machine counts. Then, with public observers present, the workers faked a random selection and the preselected sample was used, defeating the purpose of the audit. Because of this deliberate or negligent action, the true result of the race in Ohio, the deciding state in the presidential race, may never be known. To that extent, the

*University of California, Berkeley, CA, 94720
†Stanford University, Stanford, CA, 94305

[1] The theory of statistical polling and statistical quality control provides a quantitative measure of confidence in the accuracy of election results, which can be calculated as a function of parameters such as the sample size [9].

integrity—or security—of the election was compromised.

2 Background

In the United States, elections are conducted at a local level [3]. In California, for instance, specifications for equipment are set by Secretary of State and are implemented by counties according to California election code [1]. As residents of Alameda County, California, we have observed the election and audit procedures used there, and we will use California and Alameda County as a running example of current practices.

2.1 California

California election code has included, for the past four decades, a requirement for a 1% manual tally (or audit) after every election:

> §336.5 "One percent manual tally" is the public process of manually tallying votes in 1 percent of the precincts, selected at random by the elections official, and in one precinct for each race not included in the randomly selected precincts. This procedure is conducted during the official canvass to verify the accuracy of the automated count [2].

The process is *public*, meaning any citizen is invited to observe every step of the audit. This requires a transparent process. The 1% manual audit is performed as part of the official canvass, as one of the last steps before the final election results are certified. Typically, after the sample is selected, election officials print out a report containing the electronic tallies for just the selected precincts, recount by hand (using three- or four-person recount boards) all the paper ballots cast in each selected polling place, and check to make sure that the manual count in each precinct matches the electronic count in that precinct.

The law requires every county to randomly select a minimum of 1% of precincts for the manual recount, but does not stipulate *how* it should be done. As it happens, Alameda County (Figure 1), and presumably many other counties, have turned to computers to perform the random selection.

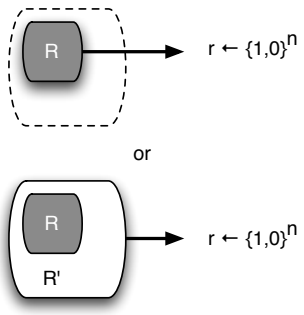

Figure 2: Transparency can be a problem regardless of the quality of the random number generator.

3 Sample Selection

Computers are generally inappropriate for generating random samples in an election audit. Excellent software-based pseudorandom number generators (PRNGs) exist, so it might not be obvious why this use of computers in *election audits* is inappropriate. The inherent problem is transparency: running software is not observable. In fact, this is the exact problem faced by DREs, and the exact problem election audits, together with voter-verified paper audit trail (VVPAT) systems, intend to address. The use of a computer as a random source jeopardizes the integrity of the audit and election.

One devastating threat is that an insider might be able to tamper with the software used for random selection in a way that allows him or her to know in advance the outcome of the selection. An insider could then use this to cheat without—or with a lesser probability of—getting caught. For instance, an insider with advance knowledge of the selected precincts will be free to defraud all other precincts without fear of detection. Or, if the insider has already tampered with the votes in some precincts, the insider could modify the computer's PRNG to exclude the possibility of choosing those precincts. These attacks could be mounted in a way that is very difficult or impossible for observers to detect. Without a verifiably random sample, it is not clear that such an accusation could be defended against.

Dedicated hardware random number generators (perhaps based on physical phenomena such as radioactive decay, or radio static) are subject to the same transparency problem, despite typically being excellent random number generators.

Largest Cities	Oakland, Fremont, Hayward, Berkeley
Population	1,507,500
Registered Voters	714,490
Total Voting Precincts	1,140
Absentee Precincts	240

Figure 1: Some basic facts about Alameda County, CA.

Without transparency, even with a perfect random source, it is hard or impossible to trust the authenticity of its output (Figure 2).

3.1 Requirements

As we saw above, "black-box" random sources are a vulnerability for election audits. What requirements must a viable solution meet? Many choices exist for generating random numbers; proving them unpredictable—in this case to *all* parties—is difficult, if even possible [11]. What matters in the election setting, however, is choosing a procedure that can be used, understood, and trusted by an average member of the voting public, for instance, by an average high school graduate.

A transparent procedure for sample selection involves two problems:

1. *Transparently generating random bits,* and
2. *Transparently turning those bits into a sample.*

Technically, this may seem like a trivial distinction. However, the requirements that follow apply equally to both, and especially because the needs of a general audience must be considered, satisfying the requirements for both is not as easy as it might seem. Many natural schemes have non-obvious problems or pitfalls. In the remainder of the section we lay out several requirements for a transparent random source and selection procedure, and subsequently evaluate possible solutions against these requirements.

Simplicity. *The procedure must be simple to follow and execute.* The public must, without extensive education, understand the procedure, why it is fair, and why they should trust it. Otherwise, confidence in the audit and election may be limited. Additionally, complexity introduces opportunities for error or exploitation.

Verifiability. *The procedure must be verifiable either by inspection (i.e., it is physically observable) or by some other property (e.g., cryptography).* Verifiability is imperative. Every observing party must be assured that the selection is genuinely random with a satisfactory distribution. Otherwise, as in Ohio, the audit could be subverted, or observer's confidence undermined.

Robustness. *The distribution of the procedure must be hard to bias or manipulate.* It should be difficult for any party, particularly someone working from the inside, to affect or gain information about what set of ballots is included or excluded from the audit.

Efficiency. *The procedure must not take an incommensurate amount of resources to prepare or execute.* Election worker's and observer's time, energy, money, patience cannot be taken for granted. As the incident in Ohio demonstrated, the cost of time might tempt election workers to fudge the procedure. Or, perhaps, impatience could reduce an observer's vigilance.

4 Possible random sources

We first examine the applicability of several random sources to election audits, in light of the requirements above.

Cryptography. From the perspective of a cryptographer, the problem of verifiably-random numbers is (at least in principle) solved. One solution [7], for instance, is to have every observer pick a random number of their own and commit to it, perhaps by writing it down and dropping it in a box. When everyone has committed their numbers, they are revealed and summed modulo some number N. The result will be a random number if at least one observer was honest. This could then be built into a very reasonable scheme to perform sample selection.

To an extent, this solution meets all our criteria. However, if we consider trying to explain this to the general public and to election's officials, it might take some work. The concepts of modular arithmetic, independence, and uniformity of distributions must be understood. Also, cryptographic protocols often rely on participants to protect their own interests, which can be a problem when dealing with a non-cryptographer public. For instance, suppose instead of dropping numbers into a box, numbers are written on a board so that the attacker has the "last say," and thus deterministically chooses the outcome. The protocol assumes a savvy enough public to know that commitments are supposed to be hidden.

Drawings. A familiar method of random selection is a drawing in which tickets are mixed then drawn from a hat or box. For instance, we could write all the precincts down on pieces of paper, place them in a box, mix the pieces of paper well, and draw our sample.

Verifiability is difficult when many objects are involved. For instance, if there are 1,000 tickets for 1,000 precincts, every observer must be assured that all precincts are included in the drawing because omissions represent those precincts an attacker could subvert freely. Thus, robustness is an issue as well.

The use of many objects also makes it hard to tell whether they have been adequately mixed. For instance, the Vietnam draft lottery of 1970 used such a scheme, but was later discovered to have suffered from bias: birth-dates later in the year were added last, and due to insufficient mixing, were more likely to be chosen earlier [13].

Lottery-style drawings. Lotteries put a great deal of resources and creativity into maintaining secure random number generators. Lottery machines are culturally familiar and trusted to the extent that the lottery is played by millions of people. It would be prohibitively expensive, however, to replicate and maintain a lottery machine in each Registrar of Voters office, so if we were to create a sampling scheme from lottery numbers, we would want to use the same machinery—and perhaps drawings—used for the actual lottery.

Although simplicity, verifiability, and robustness are excellent, efficiency can become a problem and must be carefully considered in any scheme using lottery drawings. For instance, if a state lottery is chosen, this might impose a heavy travel burden on vigilant observers determined to witness the selection. Another issue might arise because U.S. elections are performed locally: if a single dedicated state-wide lottery were used for selection, coordination between county and state might be a problem (e.g., the lottery could only be done when the last county is ready).

Alternatively, the use of an agreed-upon future drawing as input to a deterministic algorithm that creates the sample, might make an excellent scheme. However, the algorithm must also satisfy the requirements in Section 3.1. Particularly, it must be verifiable and understandable to the general public.

Random number charts. Another idea might be to use a book filled with random numbers [10] as a basis for selection. Of course, printed books or charts are static documents and must be assumed to be known in advance to any attacker. Consequently, any methods that use books of random numbers would have to find another random method of selecting digits within the documents.

Cards. Cards are a time-tested source of randomness. Indeed, they are used in many high-stakes games. However, there are fifty-two cards in a deck, making it hard to verify—especially when subject to sleights of hand, such as those found in card tricks. As with drawings, it might be difficult to ensure the deck is sufficiently shuffled.

Another issue with cards is that not everyone is familiar with their properties (the suits, the ranks, the number of cards), nor how to handle or shuffle cards well. We would prefer a sample selection scheme that as many people as possible understand and could use.

Depending on how the cards are used, efficiency could also come into play, since every action needs to be observed and verified, and preceded by a shuffle.

Coins and dice. Coins are the quintessential random number generator and arguably the most ubiquitous. Methods even exist to mitigate any bias in a coin[2]. The primary drawback of coins is

[2]Assuming independence of coin flips, *pairs* of flips from the same coin can be used to eliminate any bias the coin may have. If we throw out any HEADS-HEADS or TAILS-TAILS combination, we are left with two outcomes that bare equal probabilities. This idea is attributed to John von Neumann.

Figure 3: A ten-sided die, tumbler and tray.

Number	*Precinct ID*
0	P_1
1	P_2
$\vdots$	$\vdots$
$N-1$	P_N

Figure 4: A numbered list of precincts.

that their bandwidth is limited: generating many bits of randomness requires many coin tosses, which takes time.

Dice produce a higher bandwidth of random numbers[3], and are available with different numbers of faces. Ten-sided dice are especially appealing because they map directly onto the decimal numbers. Efficiency can be quite reasonable with dice, especially when multiple dice are rolled at once.

Dice, as with coins, can be biased—intentionally perhaps—so care must be taken to ensure fair dice [5]. Techniques for mitigating potential attacks exist and include: using only new, translucent dice; using a ribbed tumbler to roll the dice; and rolling the dice onto a flat ridged surface such as a dice tray (Figure 3). A dice setup involves the use of few objects, which is good for verification, and dice are simple to understand. They have stood the test of time[4] and are commonly used in games, including high-stakes games. In our proposed solutions that follow we choose dice as the source of randomness primarily because of simplicity, and the verifiability benefits of using few objects.

5 Creating the sample

Once a source of randomness is chosen, a practical procedure must be built around it to perform the sample selection. The procedure, like the random source, must satisfy our above requirements: it must be simple to use and understand, easy to verify, robust against tampering and reasonably efficient.

For the schemes we present, we opted for simplicity because that makes the other requirements easier to reason about. While designing a procedure for Alameda County, our initial proposals involved the use of math (not necessarily complex math). However, the feedback we got indicated the less math, the better. We then considered using computers to perform the math during the selection, but our feedback indicated that no reliance on computers was preferable. We also considered creating worksheets, such as tax worksheets, to guide users through the procedure. However, people generally do not enjoy working on taxes, and if people were to find the worksheets inefficient or frustrating, they might elect to fudge or abandon the procedure.

Keeping that in mind, we tried to separate the "work" from the "sheet" by creating pre-computed lookup tables. The procedure we arrived at requires virtually no math to use.

The basic idea. We begin by preparing a list (numbered $0, 1, 2, \ldots, N-1$) of the population we are sampling from. In California, this would involve preparing a list of the names (or IDs) of every precinct to be included in the audit. Numbering each entry sequentially from 0 to $N-1$ establishes a one-to-one mapping of the integers $[0, N-1]$ to the list of precincts (Figure 4). Election officials commit to this ordering before the audit by publishing the list and providing a copy to each political party and each observer. We assume that the audit is not performed until a final electronic tally is available; the goal is to verify whether these alleged election results match the paper records. Election officials commit to the electronic results before the audit by printing the electronic vote totals broken down by precinct, or by writing them onto write-once media (such as CD-ROM), and providing a copy to every interested observer.

Once these preparation steps have been completed, the random selection process begins. In our scheme, an election official rolls an appropriate number of dice to get a random number between 0 and $N-1$, and then uses the one-to-one mapping established earlier to interpret this number as identifying a single precinct. This precinct is added to the random sample, and the process is repeated until the random sample is of the desired size.

[3] Dice can be considered multi-faced coins.

[4] In fact, dice have existed for thousands of years.

If we used standard six-sided dice, rolling k dice would allow us to randomly choose a number between 0 and $6^k - 1$ by reading off the k outcomes and treating them as a number in base-6. However, requiring election officials and observers to perform base-6 arithmetic may be unreasonable. Therefore, instead of using standard dice, we propose using commonly available ten-sided dice[5], letting each die signify a decimal digit. For example, if we want to select one precinct from a list of 1000, we could use three dice to get a number between 0 and 999.

The above procedure handles the case where N is a power of ten. If N is not a power of ten, one simple method is to roll $\lceil \log_{10} N \rceil$ ten-sided dice, and then re-roll if the resulting number is too large. For instance, suppose we want to choose at random from a list of $N = 750$ precincts. Throwing three dice gives us a random number from 0 to 999. Because only the numbers 0 through 749 correspond to precincts on our list, we ignore and re-roll any time we get a number greater than 749. Because every three-digit number has an equal probability of appearing[6], this method produces a uniform distribution over the 750 precincts.

As we have presented it so far, this scheme works reasonably well for a small number of selections (say $N \leq 1000$). However, re-rolling can become a problem. There is a significant difference between using three dice to select 1% of 1,000 precincts, and using four dice to select 1% of 1,001. Below, we address this issue by using more general range-lookup tables.

5.1 General dice scheme

As discussed above, re-rolling can become inefficient quite quickly. To address this, a natural optimization is to divide the range of the dice into N equal intervals, letting each interval correspond to a precinct on the list[7].

For example, suppose we want to use four ten-sided dice to select a precinct from a list of $N = 1001$ precincts. First, we divide the range $[0, 9999]$ into 1,001 equal sized intervals $[0, 8]$, $[9, 17]$, ..., $[8991, 8999]$, $[9000, 9008]$. We then let each *interval* correspond to a precinct on our list, allowing the remainder, $[9009, 9999]$, to correspond to a re-roll. As a result, we throw out roughly 1 out of 10 rolls instead of 9 out of 10, improving efficiency greatly. See Algorithm 1 for a specification of this scheme.

Algorithm 1 General dice scheme.

INPUTS:
N — number of precincts
n — size of sample
OUTPUTS:
S — set of precincts selected for the sample
ALGORITHM:
$S \leftarrow \emptyset$
$k \leftarrow \lceil \log_{10} N \rceil$
(k = the number of tosses per selection)
$d \leftarrow \lfloor 10^k / N \rfloor$
while $|S| < n$, do:
    Roll k dice to get a random number $x \in \{0, 1, ..., 10^k - 1\}$.
    Set $y \leftarrow \lfloor x/d \rfloor$.
    If $y < N$ and $y \notin S$, then:
        Set $S \leftarrow S \cup \{y\}$.
return S

Numbers	*Precinct ID*
0 . . . 8	P_1
9 . . . 17	P_2
⋮	⋮
9000 . . . 9008	P_{1001}
9009 . . . 9999	RE-ROLL

Figure 5: A list of 1001 precincts labeled with equal-sized ranges.

We can simplify the presentation of this scheme by changing how we prepare the numbered list of precincts. Instead of labeling each precinct with an integer, we could label each with its precomputed range (Figure 5). Now, no arithmetic is necessary to do the selection: one can simply roll k dice and then look up the outcome on the list to obtain a precinct.

5.2 Analysis.

It should first be noted that this simple idea of using range-lookup tables applies to other random sources, and can enable other nice things like allowing for a controlled bias (say to account for precinct sizes).

[5] 10-sided dice can be found at game stores, are often used in board and role-playing games, and typically are numbered 0 through 9.

[6] Subject to any bias present in the random source. If a *verifiable* and *robust* source is used, according to our requirements in Section 3.1, this bias is negligible.

[7] We use division, rather than modulo, because division is a more familiar concept to a general audience.

Advantages. The foremost advantage of this scheme over computer-based PRNGs is the use of an observable random source. Dice are simple, robust and familiar enough to be widely accepted[8]. Dividing up the range of the dice improves efficiency while maintaining simplicity. The efficiency and usability of this scheme is excellent. As shown in Figure 6, election officials can easily select a random sample of 1% of the precincts, and even very large counties will not need too many tosses of the dice. For instance, even with $N = 2000$, only about 25 tosses of the dice (on average) are needed to select a random 1% sample. The fairness of the scheme is also verifiable (to the extent that dice are) because the lookup table is published; it can be inspected after the fact to verify that, for example, every precinct was included and had an equal chance of being selected.

Disadvantages. When the size of the sample is a large ratio of the population, many more re-rolls would be expected, as "collisions" occur in the sample selection. For small ratios this loss is negligible.

Unfortunately, the cost of this scheme scales linearly (keeping the sample to population ratio constant) with the size of the sample. While this might be suitable for selecting 1% of 500 or 1,000 precincts, it might be too time consuming to select 1% of 1,000,000.

A second concern is the threat of biased dice. If a malicious party were able to affect the distributions of the dice (say by weighting a die or by swapping in a die with duplicate faces), that party could affect which precincts get selected, or worse, which precincts *do not* get selected.

Efficiency. In an election setting, dice not only have to be rolled, but inspected by observers. This incurs a cost per roll, so we measure the efficiency of our methods by the number of rolls they require.

We can partially reduce this cost by rolling multiple dice at once. If we do this, we must ensure the order of dice is well-specified in advance. One idea is to use dice with different magnitudes on their faces, such as 1's, 10's, and 100's. Another possibility is to use dice of different colors and establish a clear ordering of the colors before rolling the dice. A good choice in the United States would be red, white and blue for their culturally meaningful order. This would enable us to roll three dice (one red, one white, and one blue) at a time; for instance, if the red one comes up 5, the white one 7, and the blue one 2, that would be interpreted as the three-digit number 572.

[8]The proposal to use dice was met with enthusiasm from election officials in Alameda County.

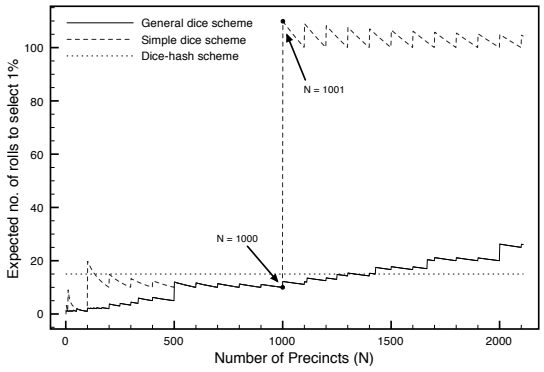

Figure 6: The expected number of tosses of 10-sided dice to select a 1% sample using range-lookup tables (General dice scheme) versus simply re-rolling (Simple dice scheme) as a function of population size (N). The Dice-hash scheme is a deterministic function, similar to [4], to calculate a sample selection using a constant number of dice rolls as input.

6 Related Work

The idea of using auditing to gain statistical confidence in an election is not new. Many states, such as California, already require random audits or manual tallies in their election codes.

Prior work has looked at the issue of the statistical validity of random audits and what size sample is necessary to achieve a given degree of confidence [9]. In these papers, however, the sample selection process is not considered and is assumed to be perfect. Some authors have examined the possibility of auditing individual ballots, stating that if adopted (in a privacy-preserving way), audits of individual ballots could give much higher degrees of confidence with less overall counting [6, 9].

Other people have looked at verifiable methods of sample selection. The IETF uses an algorithm based on a cryptographically strong hash-function, to expand input from a random source into a sample [4]. In the full version of this pa-

per we present a similar scheme, and address its viability in an election setting. Some initial reactions, however, indicated resistance and skepticism from people unfamiliar with hash functions and cryptography, leading us to prefer a dice-only approach.

7 Contributions

Sample selection is a critical part of an audit, but has been overlooked in the election setting, as evidenced by the wordings of election code [12], and existing current practices such as the use of software-based PRNGs. We identify non-transparency as a vulnerability for every part of an audit, particularly sample selection, and lay out requirements that a transparent sample selection procedure must meet.

In addition, we present an algorithm and implementation for transparent sample selection, using dice and a printed lookup table. This scheme is simple and relatively efficient, no calculations are needed to perform the selection, and, if adopted, would make elections more trustworthy.

8 Conclusion

Transparency plays an important role in the security of election audits. Because auditing is a public process capable of establishing confidence in an election's outcome, it is critical that every part of the process, particularly the random sample selection, be transparent, observable, and understandable to all interested parties. While the current use of computers to generate samples is neither transparent nor observable, we have presented a possible remedy that is simple, verifiable, low-cost, and robust.

Ultimately, it is essential that we choose random audit procedures that are capable of convincing the whole population of the genuineness of the audit, and, in turn, of the election.

9 Acknowledgments

We would like to express our deep gratitude for the Alameda County Registrar of Voters office for the generous time and effort they have provided to help us better understand the process of voting in California. In addition we would like to thank the many people who contributed to this discussion of random sample selection in person and by e-mail.

References

[1] California Elections Code, 1965. Section 19201.

[2] California Elections Code, 1965. Section 336.5.

[3] United states elections 2004, 2004. `http://usinfo.state.gov/products/pubs/election04/procedure.htm`.

[4] D. Eastlake. Publicly Verifiable Nominations Committee (NomCom) Random Selection, June 2004. RFC 3797.

[5] Encyclopaedia Britannica. dice, April 2006. Encyclopaedia Britannica Premium Service.

[6] K. C. Johnson. Election certification by statistical audit of voter-verified paper ballots, October 2004. Available at SSRN: `http://ssrn.com/abstract=640943`.

[7] J. Kilian. *Uses of Randomness in Algorithms and Protocols*. MIT Press, 1990.

[8] J. Mazzolini. Workers accused of fudging '04 recount. The Cleveland Plain Dealer, 4 2006.

[9] C. A. Neff. Election confidence. Manuscript, 2003.

[10] Rand Corporation, editor. *A Million Random Digits with 100,000 Normal Deviates*. Rand Corporation, 1955.

[11] B. Schneier. *Applied Cryptography*. John Wiley & Sons, Inc., 2nd edition, 1996.

[12] P. Smith and B. Kibrick. Manual Audit Requirements. Verified Voting Foundation, September 2005.

[13] N. Starr. Nonrandom risk: The 1970 draft lottery. *Journal of Statistics Education*, 5(2), 1997.