SAMPLING AND TELEVISION ELECTION COVERAGE

Irving Sivin, Louis Harris and Associates

I. Introduction

Good news reporting like good military tactics consists in "getting there fastest with the mostest." The traditional news "scoop" gets to the public first, and contains all the pertinent information - the who, what, when, where, why, and how of the event. The goal of television coverage of election returns is that of good reporting: to give the news with as great speed and accuracy as possible. The news, in this case, is who has won an election, how and why he achieved the victory, and what the victory means for the future.

Until 1948 newspapers and radio dominated national election news reporting. That year the opinion polls were somewhat in error, H.V. Kaltenborn's career as a radio political analyst came to an end, and an early edition of the Chicago Tribune proclaimed that Thomas E. Dewey had been elected president. (Harry Truman kept a copy of that one).

Not because of these events, but certainly soon after them, the television networks began to dominate election news coverage. Television grew up as a contemporary of the electronic digital computer. In 1952 and 1956 television networks consulted their computers during the course of election night broadcasting, but computer abilities were not truly integrated into the broadcast. The computers performed their limited news functions fairly well in those years. In both years Eisenhower won sweeping victories, carrying every state except some southern and border states. By 1960, as another generation of computers was evolving, the networks planned to make fuller use of their capacities, integrating them into the broadcast. The results were not altogether happy. Early in the evening, the computers failed to detect the narrow division that eventuated. When their later analysis proved reasonably accurate, it was too late. The public, like both John Kennedy and Richard Nixon, watched the total vote count and then went to bed -- still waiting for the undetermined outcome.

The networks' evaluation of their 1960 performance led to the introduction of scientific sampling as an important reporting tool for the presidential election in 1964. As a result of this change, all three networks produced remarkably accurate reports throughout election night that year. Some of the credit for this, however, must be given to the Johnson landslide. The fact that it was "LBJ all the way" enabled the presidential election to be reported quickly and accurately. But some of the best reporting in 1964 came about in calling the outcome of contests for Governor and U.S. Senator. Of the 107 contests (inclusive of President) in the old 48 states and the District of Columbia, 106 winners were called correctly by the National Broadcasting Company and the American Broadcasting Company, and all 107 were reported correctly by the Columbia Broadcasting System. ABC and NBC both failed on the same race, the tight contest for Senator in Ohio, which was won by the incumbent, Senator Young, against Robert Taft, Jr.[1]

The accuracy and speed of the 1964 election coverage in contrast with the fuzziness that characterized election night news coverage in previous years was due, in good measure, to the marriage of probability techniques and the statistical analysis of voting groups to computer capacities.

It is my purpose in this paper to formulate in statistical terms the problems of election news coverage and to describe one set of proposed solutions. In so doing, I shall point out some of the limitations of the proposals.

II. The Vote Collection Problem - Sampling, Complete Coverage or Both

Television coverage of the results of a presidential election generally begins between 6 and 7 pm Eastern Standard Time. By then the polls in Kentucky, Connecticut and Maryland have closed, the Eastern voters have returned home and are finished with their evening meal. The California polls close at 11 pm EST. Only Alaska and Hawaii, which have seven electoral votes between them, have polls open after 11 pm EST, a time when the Eastern audience begins to dwindle. A presidential election broadcast is continued on air until the outcome has been made clear, which frequently happens before the West Coast audience goes to bed, or around 3 am EST. The networks broadcast the effect of the presidential election on the senatorial, gubernatorial and congressional contests. In the past the coverage for senatorial and gubernatorial seats has been complete, while the congressional reporting has been selective. These contests are also reported by the networks in non-presidential years.

When the presidency is at stake, national patterns may become clear to the analysts early in the evening, as happened in 1956, when early returns from Connecticut showed that Adlai Stevenson was running behind his own 1952 mark. If the national story is late in breaking, as happened in 1960, the networks may lose large segments of their audience, particularly after 11pm EST, before they are able to announce a winner. In 1960 many people in the East went to bed believing that Kennedy had won a resounding victory, only to discover the next morning, when they turned on their sets, that the outcome was still in doubt.

In order to guarantee that they would be able to tell the news story to the public on the night of the election, the networks took two major steps between 1960 and 1964. First, they allocated a considerable part of their resources to select sample precincts in each state, from which they communicated the results directly to a central network computer input area. Second, the networks engaged in a vote collection effort on an unprecedented scale in order to speed the general reporting of returns.

In 1960 and earlier years the wire services did much of the raw vote collection and the television networks purchased this data from them. The raw vote figures displayed by the various networks could be quite different, since they could reflect different states and different reports within the same state. However, in states which do not have quick-counting automatic voting machines, the wire services' collection was too slow for the networks which might lose their audience before the outcome was known. In some of the slowly reporting states, like California or Massachusetts, a victory of landslide proportions was necessary to determine a winner before the audience went to bed. For these reasons the networks decided in 1964 to invest additional financial resources to speed the reporting of the vote.

The decision as to how much to invest in speeding up the coverage was pending when the California primary of 1964 was held. In that primary, on the basis of sample data, CBS and then ABC declared that Goldwater had won over Rockefeller. Hours later, press service returns, reflecting the vote in northern California, showed Rockefeller with a substantial lead. The California public was confused by the conflicting claims of television and the wire service reports. The sample data proved ultimately to be correct. This situation not only affirmed the networks in their decision to use samples, but also served to spur their efforts to insure that the vote count from all the precincts would be moved to the public as quickly as possible, in order to avoid any similar confusion.

The cost structures of the two types of data collection, sample and complete, are quite different. A very high unit price per precinct is required for the specially designated sample precincts. Election night costs for sample precincts involve the pay of the reporter covering each precinct, the cost of the phone message to the computer input area, and the payment of the operators at the input area. In addition, there may be extra costs for installing a telephone in the vicinity of the polling place. These costs are generally comparable to the complete coverage costs. The major expenses for sample precincts are the pre-election ones. These involve the selection of the sample and research into the voting history of the sample precincts. Such research also involves establishing whether or not the precinct has the same physical boundaries as it had in the recent past, and obtaining a detailed description of the socio-economic, ethnic and religious characteristics of the precinct. The descriptive material is integrated into the analysis on election night. Finally, there is a high cost per precinct associated with the programming of the input quality control and statistical analysis of the sample data. Unless the raw vote data are to be subjected to statistical or detailed political analysis, the cost of handling the raw vote is much smaller on an individual precinct basis.

Despite the differences in unit costs, so few sample precincts are needed to determine the outcome of an election, that for most states the total cost of sampling is far below that of the complete vote collection effort. Since reporting from sample precincts provides accurate and quick results of the outcome of an election, and provides the possibility of detailed analysis, the real question is why have the networks invested in both types of data collection.

There is a threefold answer. I have stated one already: speeding up the general vote collection process means that there will be less possibility that the public will be confused by different reporting sources. Second, and most important, the public knows that an election is really determined only when all the votes are counted, and the public continues to demand that count. Finally, because of sampling error, samples experience difficulties in making determinations of the outcome in close contests. (In such a case a sample can report early in the evening that a contest will be a close one, and a sample will be able to indicate which socio-economic groups and which geographic areas are supporting each contestant, but it still will not be able to name the winner.) In a very close state election, data from all the precincts in states which use automatic (mechanical) voting machines in every precinct will be available before the Eastern audience has gone to bed, and therefore an investment in full collection is worth while. (In states where paper ballots or electronic voting machines are used in part or in whole, the full tabulation of races at the top of the ticket has not yet proved to be really fast enough to determine the outcome of a close election before the audience goes to bed. But the progress of the count has been speeded.)

It is perhaps somewhat ironic that the installation of electronic voting devices in many areas has slowed down the determination of the outcome of elections. With the electronic machines the vote must be counted at a county center. The very act of transporting the cards several miles delays the count. In many counties electronic machine tabulation has meant that several days elapsed before the exact outcome was known. The delay has been due to the fact that the public has not been educated to mark the new ballots properly. The sorting and tabulating equipment have therefore experienced difficulties in reading the ballots.

In sum, between 1960 and 1964 the networks discovered that state by state sampling was an indispensable aid in election coverage. Sampling speeded the determination of the outcome of an election, and this speed, in turn, spurred the effort to obtain even faster raw vote returns.

III Sampling Considerations

A. Initial Considerations - Precision, Risk, and the Underlying Variance

In each race that is being covered the sample must be able to discriminate between a victory and a defeat. It is not feasible for a sample to distinguish between 50.1 and 49.9%, but a sample may be required to distinguish between 51.0% (or at most 52.0%) and 49.9%. By the word "distinguish" I mean that the difference between the sample estimate and 49.9% is statistically significant at the 1 or 2 % level.

Whatever other value an election sample may have, its principal function is to determine the outcome of an election, or the investment in the particular race will not have fully paid off. The risk the networks are willing to take in being publicly in error can be gauged by their 1964 performance, a level of about one or at most two chances out of one hundred of being wrong.

The unit which must be sampled is, of course, the voting precinct. The number of voters in a precinct varies enormously from state to state and almost as much within some states. New Hampshire precincts may have as few as 8 or as many as 5000 voters. The largest precinct in Massachusetts contains fourteen thousand voters, while many towns have fewer than 300 voters. On the average, the 180,000 precincts in the United States in 1964 had about 390 voters each.

A precinct is not a randomly formed cluster of voters. Since precincts represent small geographic areas, they tend to be far more distinct from each other than might be expected from a random formation. In terms of economic characteristics, there are many precincts in which most of the voters have extremely high or low incomes. There are many precincts that are predominantly Negro, or Puerto Rican, Italian, Scandinavian, Polish, or Jewish. As an effect of social, economic, ethnic, and religious clustering, voting precincts show a greater diversity in their support of candidates than could be expected were they randomly formed from the population at large. In a close (50-50) election within a state, some of the precincts give 30% of the vote to the winner while others give 70%. This is such a common phenomenon that we hardly think anything of it. Yet, were precincts randomly formed, few of them (with 400 voters) would differ from the state's average by more than 5%, and almost none of them would differ by more than 8 or 10%. The fact that such vote clustering takes place increases the cost of sampling by requiring a larger number of sample precincts then would be needed were there no clustering. On the other hand, the clustering effect provides an enormously useful tool in the analysis of the voting behavior of the groups which together constitute the American body politic. I will return to the analytical advantages of clustering effects later. For the moment I will concentrate on their effect on sample size.

In a close election it would take a sample of about 10,000 persons selected at random to differentiate an estimate of 51.0% from a value of 49.9% at a 2% (one-sided) significance level. Were precincts randomly formed and of uniform size with 400 voters each, twenty-five precincts per state would be sufficient to determine the outcome of each race. The same precincts could be used for all the contests which one wished to report in a state.

Studies of the outcome of elections indicate that the intra-class correlation coefficient, the measure of voting homogeneity within a precinct, is rather high, attaining values of .05 - .06, for precincts of some 400 voters. This may be translated by saying that the standard deviation in the population of precincts in some states is of about 10 to 12 percentage points. If we suppose a simple random sample of precincts were possible within a state and that we wished to discriminate between 52.0% and 49.9% taking the chance of being in error about 2 percent of the time, the minimum size of sample needed in a state is 100 precincts. A sample size of 100 per state would give rise to a national stratified sample which would estimate the presidential electoral vote with good accuracy. The popular vote for the presidency could be estimated to well within half a percentage point.

An heuristic argument showing the precision to be expected in the national popular vote is shown below:

Let 1) K be a subscript standing for a state

2) $\hat{P}_k$ be the estimated percentage for a presidential candidate within state k

3) σ_k^2 be the variance in a state percentage

4) n_k be the size of sample in a state

5) W_k be the weight (proportion) which a state contributes to the national vote

The estimated national percentage, $\hat{P} = \sum^{51} W_k \hat{P}_k$

The variance of the estimate is $\sigma^2_{\hat{p}} \doteq \sum^{51} W_k^2 \sigma_k^2 / n_k$

If n_k is chosen so that $\sigma_k^2 / n_k = 1$ for all states, then

$$\sigma^2_{\hat{p}} \doteq \sum^{51} W_k^2$$

But in the America of the 1960's only two states, New York and California, have weights that approach one tenth of the total vote. Therefore $\sigma^2_p < .3$.

B. Criteria for Stratification in State Samples

Simple random sampling of precincts is seldom possible in any state. Were it possible, it would seldom be employed, because stratified sample designs are more efficient, particulary when the stratification is accomplished on a correlate of the variable being measured, the percentage vote given to the Democratic (or Republican) candidate. The most efficient stratification is that in which the strata means are widely different. In some states, like New Hampshire, it is possible to accomplish this by allocating precincts individually to strata, because detailed data are published for each precinct, and because the number of precincts is small. For many states, however, the only available published data are those provided by Richard M. Scammon in America Votes or by the Bureau of the Census in its technical studies. The fundamental unit of publication in these sources is the county, although Scammon provides the vote for large units in the principal cities of the country, - the wards or state assembly districts. The large political units, for which past data are available, are the fundamental units used in stratification. Once grouped together in strata, the counties, wards, assembly districts, etc. serve as the primary sampling units for stratified two-stage cluster sampling.

For each primary sampling unit (PSU) the percentage obtained by a Democratic candidate in four recent statewide elections is computed. A time span of six years is used, in order to include the vote in the two most recent presidential election years and in two non-presidential years. The PSU's are then ranked on each election, the rank of 1 being arbitrarily assigned to the PSU with the highest Democratic percentage in a particular election. The PSU ranks are then summed over the four elections used. This sum is used as a basis for the final grouping of the PSU's into strata.

Outside of the South, which is now undergoing what amounts to a political revolution, the rank order correlations of the PSU's have been over .80 for most states and for most pairs of elections. In the states of the Old Confederacy the effect of recent political change has been to produce much lower coefficients, even negatives ones. In Virginia, for example, where the Byrd machine was entrenched until the primary of July 1966, some of the counties which were the strongest supporters of the Democratic party, when Byrd candidates were contending for office, gave relatively low percentages to the Presidential nominees of the Democratic party in 1960 and 1964 when the Byrd machine did not support the Democratic party. Nevertheless, even in the South, the rank order provides a useful stratification device. In some states the rank order goes hand in hand with geography and the degree of urbanization. In New York State and Massachusetts a simple partitioning of the state into a few geographic areas accomplishes a very effective reduction of the electoral variance. The final stratification used in sampling recognizes the three basic stratification variables: 1) average rank order, 2) degree of urbanization, 3) and geographic location.

When the PSU's and strata are formed, an effort is made to keep them equal in size. Counties or wards with very small voting populations are joined together with adjacent counties or wards for this purpose.

Most strata consist of only a few PSU's. There may be 10 to 20 strata in a state. The results of the stratifying process have been to produce PSU's of relatively small size and with small variance between the PSU's. The chief contributor to sampling error has been the variation in the percentage Democratic within the PSU's (the variation among the precincts within the counties, wards and assembly districts). The typical PSU contains somewhere between 25 and 100 voting precincts depending on the state and the number of strata. Stratification of this type has brought the substantial reductions in sampling error. In some states the sampling variance has been cut in half by stratification. This reduction is, of course, equivalent to a reduction of sampling error of about 30% below that of a simple random sample of the same number of precincts.

These estimates of the efficiency of stratification are based on samples used in 1964, in which there was an equal allocation of the number of sample precincts to each stratum. Since all the strata have approximately the same number of voters, and since the within-stratum variances are also about equal, an equal number of sample precincts in each stratum is close both to proportional and to optimum allocation. Although the samples have not departed greatly from the optimum in any state, the gains from stratification vary markedly from state to state.

The effect of the stratification is to reduce the number of precincts required to make a 52.0-49.9% distinction down from 100 to the neighborhood of 50-70 precincts. These gains are translated directly into cost savings.

IV Estimation Procedures

A. Procedures Without Historical Data

The fact of federalism almost insures that there will be no single sampling or estimation procedure that will uniformly produce the smallest possible mean square error for all contests within all states in any year. Even within a single state, a variety of factors make certain estimates more useful for one contest than for another.

Whenever possible both sample design and estimation should take full advantage of historical data. The creation of strata from the PSU's by ranking them is one such use. But below the county level, it may not be possible to find relevant historical voting data. Let me assume for a moment that as a result of 1) the Supreme Court's "one man, one vote" decision, 2) civil rights legislation abolishing the poll tax, 3) the creation of a viable Republican party in the South, 4) a Decennial or local Census, 5) land annexation, or 6) the redrawing of precincts to accommodate new types of voting machines, not a single precinct in the United States has the same land boundaries today as it had in the most recent past election. Revised precinct boundaries do not permit historical data to be used on a precinct level in an unbiased way, and I shall examine the type of estimates which can then be used. (A similar lack of relevant precinct historical data arises when one wishes to determine the outcome of a primary election through sampling, particularly in a state where there has been no primary for the office for that party in the past few years.)

In a simple random sample of the current precincts in a state, or in a stratified self-weighting sample with equal probabilities of selection of precincts at all stages, the mean per precinct estimate would produce an unbiased estimate, if all precincts had the same total number of voters. The term "mean per precinct" refers to the average of the percentages of the vote cast for a candidate in each precinct. Until now my discussion of the size of sample necessary to solve the election night coverage problem has been predicated upon the assumption of equality in the number of voters in each precinct of a state. However, precincts do differ in size, and PSU's within a state differ on three key size variables: the total number of voters in the PSU, the total number of precincts in the PSU, and the average number of voters per PSU. Therefore, the mean per precinct estimate produced by equal probability selections (whether in one or in two stages) can be seriously biased. The bias is a function of the correlation between the percentage Democratic (or Republican) and the number of voters in the precinct. In many states these correlations are sufficiently large, so that, for samples of 40 to 100 precincts, the bias can be the largest contributor to the root mean square error. In New Hampshire, where great variation in the size of the precincts exists, there is a strong association between the number of voters in a precinct and percentage Democratic. This is shown in the table below.

The largest precincts are in the Democratic cities of Manchester, Nashua, and Portsmouth, while the small precincts are identical with the small towns, sturdily Yankee and ever Republican.

If it were possible to select precincts in a single stage process with probabilities proportionate to size, the simple mean per unit estimate (the average of the precinct percentages) from the resulting sample would be unbiased. The considerations which I have adduced with regard to the size of sample necessary to accomplish the election night task would remain substantially valid, and unchanged. But one can never sample precincts with probability exactly proportional to the unknown, future voter turnout. If good measures of the size of the voter turnout in each precinct are available, the mean per unit estimate will be biased, but the bias will be small in comparison to the sampling error for samples of 50 or more precincts. Even under the harsh conditions which I have assumed, the annihilation of virtually all useable past voting data, the current total voter registration, be it for one party in a primary or for all parties in a general election, is available a few months before the actual date of an election and could be used for precinct selection . (The cost of obtaining registration data for all precincts in the sample PSU's is extremely high and the likelihood of 100% coverage is small.)

In the event that no adequate precinct size measures are on hand or that one wishes to produce a consistent estimate, a ratio estimate of the percentage of the vote cast for the Democratic party is always available, regardless of the particulars of the selection procedure. This estimate takes the form

$P' = (X'/Y')100$ where X' is the estimated total Democratic vote
where Y' is the estimated total vote
and P' is the estimated percentage.

X' is an unbiased estimate of the true Democratic total and is derived by inflating the precinct vote to the population or stratum vote. The same is true for the Y' variable.

In stratified samples this estimate is the combined ratio.[2] Since a ratio estimate can be greatly affected by variations in the size of the denominator variable, in states where the precincts differ appreciably in total votes

SIZE AND PERCENTAGE CORRELATION
NEW HAMPSHIRE GUBERNATORIAL ELECTION, 1964

Number of Voters in Precinct	Percentage Democratic	Percent of State's Precincts	Percent of State's Vote
< 50	48.5	2.6	0.05
50- 99	52.7	5.3	0.4
100- 199	54.0	9.9	1.7
200- 399	51.8	16.9	5.0
400- 799	56.5	21.9	13.2
800-1599	67.2	26.2	31.7
1600-3199	70.6	12.9	28.7
3200+	73.0	4.3	19.2

cast, the ratio estimate may best be used in conjunction with sampling with probability proportional to estimated size (PPES). PPES will insure that the coefficient of variation of the sample mean of the denominator is small, less than .1, and therefore that the bias of this estimate will be small in comparison with the sampling error. The detailed form which the ratio estimate assumes in stratified two-stage cluster sampling is shown below:

$$\frac{\displaystyle\sum_{h=1}^{L} \frac{1}{m_h} \sum_{i=1}^{m_h} \frac{\frac{1}{n_{hi}} \sum^{n_{hi}} \frac{X_{hij}}{Z_{hij}}}{Z_{hi}}}{\displaystyle\sum_{h=1}^{L} \frac{1}{m_h} \sum_{i=1}^{m_h} \frac{\frac{1}{n_{hi}} \sum \frac{Y_{hij}}{Z_{hij}}}{Z_{hi}}} = \frac{X'}{Y'} = \hat{P}$$

In the above formulation

X_{hij} is the Democratic vote in precinct j of PSU i in stratum h

Y_{hij} is the total vote in precinct j of PSU i in stratum h

Z_{hij} is the (conditional) probability of selecting precinct j given the fact that PSU i has been selected in stratum h

n_{hi} is the number of precincts selected in PSU hi

Z_{hi} is the probability of selection of PSU hi

m_h is the number of sample PSU's in stratum h

Regardless of the sampling or estimation procedures used, when historical data is not available on a precinct level, stratification is unable to reduce the inherent variance by much more than 50%. The reduction in some states will be appreciably less than that. This fact is due to the high variance within the PSU's.

B. Estimation Using Historical Data

1) The difference or change estimate in single stage sampling.

In any state the relationship between the percentage of the vote cast for a Democratic candidate for office in 1966 and that cast for a Democratic candidate for the same (or similar) office in a contest in 1964 can be expressed as $P_{66} = P_{64} + C$. In this statement P_{66} and P_{64} are the true Democratic percentages in 1966 and 1964 respectively. C is the change in the Democratic percentage in the two years.

Exactly the same relationship holds true on a precinct basis. That is $P_{66j} = P_{64j} + C_j$. Here j is a subscript used to indicate a precinct.

If all precincts had the same number of voters, the arithmetic mean of C_j over all precincts would be the true population value C. A simple random sample of precincts would produce unbiased estimates of the change, C, by using the mean per unit estimate. Since precincts do vary in size, sampling in one stage, with replacement and with probabilites proportional to the turnout (PPES) in the precincts would lead to an unbiased difference estimate of the form:

$$P''_{66} = P'_{66} - P'_{64} + P_{64} = C' + P_{64}$$

P'_{66} is the arithmetic mean of the sample precinct percentages and is an (unbiased) estimate of P_{66}, the true percentage in 1966.

P'_{64} has a similar meaning.

P_{64} is the true Democratic percentage in 1966.

P''_{66}, the difference estimate, will be a better estimate than P'_{66}, the ordinary mean per unit estimate, whenever the variance of the Democratic percentages are about the same in the two years and the correlation between the percentages on a precinct basis is greater than 0.50.

Empirical research indicates that correlations around .85 exist in most elections in the United States. The variance of the difference (change) estimate would be only three-tenths that of the ordinary mean per unit estimate, assuming equality of the variances in the two election years. In general, the variance of the estimate P''_{66} is

$$\sigma^2_{P''_{66}} = \sigma^2_{P'_{66}} + \sigma^2_{P'_{64}} - 2\rho_{66,64}\,\sigma_{P'_{66}}\,\sigma_{P'_{64}}$$

If the variance of the past race is larger than that of the present race, the correlation coefficient must be high for the difference estimate to be useful. For example, if the past race had a standard deviation of 15 percentage points and the current race one of 10, a correlation of .75 would have to be present in order for this difference estimate to be as useful as the mean per unit estimate. Conversely if the past race had a smaller variance than the present one, lower values of the correlation are required.

The following table indicates what values of the variance of the sample estimate can be achieved with different levels of correlation and different relationships between the underlying standard deviations in different years. The sample variances are compared to 100, taken as the variance of the mean per unit estimate based solely on current data. The comparisons are made for the same size of sample.

VARIANCE OF THE DIFFERENCE ESTIMATE
In Comparison to the Mean Per Unit Estimate (Considered as 100) For Different Relations of the Underlying Standard Deviations

RHO	$\sigma_{66} = \sigma_{64}$	$\sigma_{66} = 1.5\,\sigma_{64}$	$\sigma_{66} = 2/3\,\sigma_{64}$
1.00	0.0	11.1	25.0
.95	10.0	17.8	40.0
.90	20.0	24.4	55.0
.85	30.0	31.1	70.0
.80	40.0	37.8	85.0
.75	50.0	44.4	100.0
.70	60.0	51.1	115.0
.65	70.0	57.8	130.0
.60	80.0	64.4	145.0
.55	90.0	71.1	160.0
.50	100.0	77.8	175.0
.45	110.0	84.4	190.0
.40	120.0	91.1	205.0
.35	130.0	97.8	220.0
.30	140.0	104.4	235.0
.25	150.0	111.1	250.0

In Georgia in 1964, the difference estimate had a much higher sampling error than the mean per unit estimate based on PPES sampling. In this state many of the traditionally Democratic precincts voted strongly for the Republican party, while the urban Negro of Atlanta, who had historically given his vote to the party of Abraham Lincoln (Nixon did well among Atlanta's Negro population), turned overwhelmingly against Barry Goldwater. The estimated Johnson-Kennedy correlation was -0.2.

The question of whether or not the correlations between two elections will be sufficiently high to justify the use of a difference estimate need not be answered as late as election night. An opinion poll taken just before the election can determine whether the correlations will be high enough to justify the use of the difference estimate. On election night itself the correlation coefficient may be estimated from the precinct data.

2.) Bias in the Difference Estimate

If simple random sampling of precincts were used, the difference estimate could be biased because of the persistence of the correlations between the number of voters in a precinct and the change in the precinct's percentage. In 1964, outside of the South, Republican voters did not turn out to vote to the same degree that they had in the previous two elections. Some New Hampshire precincts, which were small and heavily Republican in 1960, had an even smaller voter turnout in 1964. The Democratic percentage in these precincts moved upward by a far greater degree than it did in the large, normally Democratic precincts of Manchester and Portsmouth. A precinct which had given Kennedy only 35% of the vote could easily give Johnson 30% higher and remain relatively close to the state and national percentage. But a precinct which had given Kennedy 75% could not possibly give Johnson an increase of more than 25%, and generally did not give him an increase of more than 10 or 15%. Single stage equal probability selections would produce too many large increases and too few small ones, wherever large precincts had been heavily Democratic in the past. In simple random sampling with equal selection probabilities the difference estimate, based on the average change in the precinct percentages, is biased because of the existence of these correlations between the past percentage, the change in the percentage, and the size of the precinct. The mean per unit estimate with PPES sampling reduces this bias to the point where it is small in comparison to the sampling error, for samples of 50 or more precincts.

3.) The Difference Estimate in Stratified Sampling

If no historical data were available for use in estimation, the stratified estimate of the percentage of the vote received by the Democratic candidate might take the form of $P = \sum W_h P'_{66_h}$. With historical data available, the estimate would be:

$$P''_{66} = \sum_{h=1}^{L} W_h (P'_{66_h} - P'_{64_h}) + P_{64} = \sum_{h=1}^{L} W_h C'_h + P_{64}$$

That is, the estimated Democratic percentage is the weighted sum of the estimated changes within the strata. W_h represents a weight and the P'_h is an estimate of the stratum percentage, while C'_h is an estimate of the change in the stratum percentage.

The magnitude of the correlations within most strata tends to be at about the same level as that for the state as a whole. Thus, where the initial introduction of stratification reduces the underlying variance in a state from 30 to 50% of the total variance, the use of a stratified difference estimate within each stratum can accomplish a substantial further reduction. Using a single random sample with no historical data available, 100 sample precincts may be necessary to differentiate an estimate of 52.0% from 49.9%. Using stratified difference estimates, where precincts are selected directly in a single stage by PPES, may mean that only 35-40 precincts are necessary to make that distinction.

In many contests the gains from stratification unfortunately overlap those made by the difference estimate. Samples smaller than 40 precincts may not achieve the needed precision.

I should point out that if stratification is performed on the rank order of the component PSU's, the bias of the ordinary mean per unit or of the difference estimate tends to be small, even when a simple random sample is taken within each stratum. That is, stratification on the basis of the past vote tends to reduce the correlation between the size of the precinct and the change in the percentage. Stratification of precincts by size would also reduce the bias.

C) Estimation Procedures Using More Than One Past Race

In discussing this subject I shall limit my remarks to the simple model of single stage using sampling PPES and the mean per unit estimate.

Estimates of the following linear forms are under discussion:

$$1.)\ P'''_{66} = P'_{66} - 1/2\,(P'_{64} + P'_{62}) + 1/2\,(P_{64} + P_{62})$$

$$2.)\ P''''_{66} = P'_{66} - 1/3\,(P'_{64} + P'_{62} + P'_{60}) + 1/3\,(P_{64} + P_{62} + P_{60})$$

In the above difference formulae, the primes indicate estimates and the P's without primes are true values. The subscripts are chosen to represent different years, but two or more contests from a previous election year may be used instead. The coefficients are pre-assigned values, giving equal weight to each of the past elections used. They may be regarded as regression coefficients. The variance of the above two estimates can be given as follows:

$$1.)\ \sigma^2_{P'''_{66}} = \sigma^2_{P'_{66}} + 1/4\,(\sigma^2_{P'_{64}} + \sigma^2_{P'_{62}}) + 1/2\,\rho_{64,62}\,\sigma_{P'_{64}}\,\sigma_{P'_{62}} - \rho_{66,64}\,\sigma_{P'_{66}}\,\sigma_{P'_{64}} - \rho_{66,62}\,\sigma_{P'_{66}}\,\sigma_{P'_{62}}$$

and

$$2.)\ \sigma^2_{P''''_{66}} = P'^{2}_{66} + 1/9\,(\sigma^2_{P'_{64}} + \sigma^2_{P'_{62}} + \sigma^2_{P'_{60}}) + 2/9\,(\rho_{64,62}\,\sigma_{P'_{64}}\,\sigma_{P'_{62}} + \rho_{64,60}\,\sigma_{P'_{64}}\,\sigma_{P'_{60}} + \rho\,\sigma_{P'_{62}}\,\sigma_{P'_{60}}) - 2/3\,(\rho\,\sigma_{P'_{66}}\,\sigma_{P'_{64}} + \rho\,\sigma_{P'_{66}}\,\sigma_{P'_{62}} + \rho\,\sigma_{P'_{66}}\,\sigma_{P'_{60}})$$

The following tables indicate the efficiencies of these estimates in comparison with the mean per unit estimate, the variance of which is considered as one. The tables assume that the variances of all the races are the same.

VARIANCE ACHIEVED BY MULTIPLE DIFFERENCE ESTIMATES IN COMPARSION TO THE MEAN PER UNIT ESTIMATE (Considered 100) (All Correlations Assumed to be Equal)

RHO	1 Past Race	2 Past Races	3 Past Races
1.00	0.0	0.0	0.0
.95	10.0	7.5	6.7
.90	20.0	15.0	13.3
.85	30.0	22.5	20.0
.80	40.0	30.0	26.7
.70	60.0	45.0	40.0
.60	80.0	60.0	53.3
.50	100.0	75.0	66.7
.40	120.0	90.0	80.0

The next table indicates the effect on the variances if one past race has a correlation with the current race which is smaller by 0.1 than the value(s) of the correlation(s) of the other past race(s).

VARIANCE ACHIEVED WHEN ONE CORRELATION WITH THE CURRENT RACE IS 0.1 BELOW OTHERS

Highest RHO	2 Past RACES	3 Past RACES
1.0	10.0	6.7
.9	25.0	20.0
.8	40.0	33.3
.7	55.0	46.7
.6	70.0	60.0
.5	85.0	73.3
.4	100.0	86.7
.3	115.0	100.0

It is evident from the above tables that where high correlations exist among all races, their use improves the estimates. The most significant reduction in variance comes about from the introduction of the first historical correlate. If a very high correlation is present in one past race, additional historical correlates have only marginal utility. The use of two or three past races, however, lowers the correlation threshold level that is needed to keep the variance of the estimate below that of simple mean per unit estimate.

When one correlate is below that of the others, the advantages of using the multiple correlates are attenuated.

D. Analytic Estimates

One of the most important functions of the sampled precincts is to report how the component groups within the electorate have voted. By classifying the precincts along several analytic dimensions, e.g. degree of urbanization, ethnic and religious make-up, socio-economic status, historical voting patterns and issue orientation, reliable estimates are available for those groups with sufficiently large sample representation.

In addition to providing "hard" news on how these groups have voted, the classification process permits alternative estimates of the state-wide percentages to be made. This is possible through the use of weights associated with each group within each analytic dimension. In effect, this is stratification after sampling.

For those interested in the behavioral aspects of political science, the analytic estimates are among the most interesting and important produced on election night. Far more information is generated on the computer then can be relayed to the public through the television medium on election night. It is to be hoped that the information already culled and reduced to tractable form will be fully explored by political scientists in the future.

V Final Determination of Sample Size and Sampling Procedures

A. A sample must be able to determine the outcome of a contest. Depending upon the type of historical data and on the nature of the analytic estimates that are desired, the required sample size may be as small as 30 precincts or as large as 100 precincts. Generally there is more than a single statewide contest in a state and the same set of sample precincts may be used to determine the outcome of more than one contest. Usually the networks do not cover more than three contests in any state. The sample size necessary is then determined by that contest which looks as if it will be the closest of the three. If none of the contests seem to be close, very few sample precincts are used to determine the victors. In order to maintain the required speed of reporting, however, a minimum of thirty precincts are used in all states for state-wide contests. In most states the information necessary to decide whether contests will be close or not, and whether historical data will prove beneficial for the difference estimates, can be obtained from pre-election polls. Polls determine the optimum solution to the sampling problem in an individual state. Polls also provide the background of issues and personalities as seen by the electorate, and permit interpretation of the "why" of the election returns.

B. Final Sampling Procedures

Within each stratum two PSU's are selected systematically without replacement. Within each sampled PSU two precincts are generally sampled. In states where smaller samples are used, only one precinct may be sampled within each sample PSU. The sampling of PSU's and the sampling of precincts within PSU's are both accomplished by sampling with probability porportional to the size measures assigned. (PPES)

The use of two precincts per PSU is a device to reduce the cost of sampling, which simultaneously permits an unbiased estimate of the within PSU variance to be made. The costs of sampling arise because within PSU precinct data are generally not published. Such data can be obtained only at the county clerk's or town clerk's office, and often only by personal and persuasive visits. Precinct vote history is a matter of public record, but old precinct maps lie, at times, in dusty drawers. County clerks have often asserted that particular precincts have not changed boundaries in a decade, only to have their maps belie them.

In the event of redistricting between the previous and the current elections, precincts are selected within the PSU on the basis of the total current registration. When no current data are available except the number of precincts in the PSU, the sample precincts are chosen with equal probability within the PSU. Except for states like California and Rhode Island, where deliberate care is taken to equalize the number of registered persons per precinct, the variations in the number of voters per precinct is so great that equal selection probabilities for precincts within the PSU's is disadvantageous.

Although the strata are formed in such a way that the total number of voters in previous elections are about the same in each stratum, when only one or two precincts are in sample from each precinct, sampling fractions vary from PSU to PSU, and from stratum to stratum. A five digit weighting factor is therefore associated with each precinct. While some computer storage capacities are therefore occupied by the weights used, this is a small price to pay to avoid the biases that develop by ignoring the different probabilities of selection.

VI Biases Their Origins and Importance

In samples of the size that are necessary to produce reliable estimates of the election outcome, bias may be a factor as large or larger than the sampling error. In samples where the precincts are specified for selection because they have been barometric in the past or because they have proved to move up and down with about the same percentages that the states have moved, the danger of bias is quite large. Only a random sample, employing the known probabilities of selection of the individual precincts can avoid such biases.

The desire to use a good estimate, such as a difference, a ratio or a regression estimate, can induce the sampler to give zero probabilities to precincts whose boundaries have changed since the most recent past election. Since one of the reasons boundaries change is the fact that a precinct may have gained or lost voting population, ignoring such precincts can cause serious biases in the estimate, particularly where the growth (or loss) areas vote differently from the more stable population areas. Given some flexibility in computer programming and in field reporting, it is unnecessary to incur this bias. The field procedures require reporting all the precincts that are now within the boundaries of the old precinct (or to record all the historical vote of the old precincts which now are included in the new precinct). In case the boundaries have severely changed and reconstruction is not possible, the estimate of the past stratum totals may be derived only from those sample precincts with past data. This will increase the variance of the past data and may decrease the correlation between the estimates of past and present derived from the set of sample precincts. Nevertheless, most of the advantages of the use of the past data are retained, without bias.

A different type of bias arises because of the absentee vote. In most states the absentee vote is comparatively small, yet it has been different in terms of the Republican-Democratic vote split from the regular vote. It was the absentee vote which shifted California from the Kennedy to the Nixon column in 1960. In some states the absentee ballot is counted at the same time as the regular vote. Then it can be sampled without bias. In other states the absentee ballot is not counted early on election night, and this causes the problem. In Texas the absentee vote in the larger counties is considerable, amounting to 6% in Harris County and 7% in Ector county in 1964. In a close race in Texas, when this vote is not available one must say that the outcome is undecided. The absentee vote also can prove troublesome when the method of counting it changes. The absentee vote may have been included in the official count of the individual precincts in the previous election, and it may be counted as an absentee box at the town or county level in the current election or vice versa. When this occurs it is difficult to reconstruct the past vote of the individual precinct, although the bias of the resulting procedure should usually be rather small.

VII. The Determination of the Outcome of an Election

When the results of the election have been reported from the sample precincts, a determination must be made of whether or not the sample percentage is significantly above (or below) 50%. This is a question of making an estimate of the sampling error, of judging whether or not the distribution of the normalized variate follows Student's "t" distribution and of judging how much bias, if any, may be in the sample statistics.

For the full model, the estimate of the sampling error is made for all the components of the difference estimate, where the estimated percentage for the two elections in question are themselves ratio estimates. For election night purposes the simple biased method proposed by Des Raj for estimating the variance of stratum totals based on PPES Sampling and systematic selection of PSU's has proved quite useful. [3]

It is often necessary to make an estimate of sampling error before the complete sample has reported on election night. Such an estimate is possible, using the collapsed stratum technique. [4] In this scheme, all the strata in the state have been paired a priori. Errors are calculated whenever one precinct in each of the paired strata has reported. Pairing estimates in this fashion leads to an overestimate of the sampling error, but this has proved to be a useful safeguard against inaccurate declarations of the winner.

Another estimate of sampling error which has proved to be useful, completely disregards the effects of stratification and measures only the unstratified differences in percentages. This estimate is an overstatement of the error, it is not a consistent estimate, but it has proved accurate in calling the outcome of all but the closest races (under 52.0%)

VII. Summary

Election news broadcasting requires quick and accurate determination of the results. Broadcasting is enhanced by the analysis of the voting behavior of different groups within the electorate, and even more so by an analysis of the impact of issues and personalities on the voting groups. Probability sampling is the most dependable way to meet these requirements, and is the least expensive way. No single sampling method or estimation procedure suffices to embrace the variety of situations produced by the volatile American political scene. But sampling techniques are sufficiently flexible to enable one to come close to the optimum solution for each contest in each state.

Notes and References

1) By a correct call I mean a) one in which the winner was correctly named during the broadcast or b) a contest which was stated to be too close to call and which in fact showed a 52-48% split or closer. The counts are taken from reviews of network performance and from published advertisements. Neither ABC nor NBC used sample data alone to make their calls in 1964.

2) Hansen, M.H., Hurwitz, W.N. and Madow, W.G. -- Sample Survey Methods and Theory, pages 189 - 200.

3) Raj, Des, "Variance Estimation in Randomized Systematic Sampling with Probabilities Proportionate to Size", Journal of the American Statistical Association, 60 page 278.

"The Use of Systematic Sampling with Probability Proportionate to Size in a Large Scale Survey", Journal of the American Statistical Association 59. page 251.

4) Sukhatme, P.V., Sampling Theory of Surveys with Applications pages 399-404

Hansen, M. H., Hurwitz, W.N. Annals of Mathematical Statistics. 14, pages 333-362

Appendix

Estimated Johnson-Kennedy Correlations

State	Correlation	State	Correlation
Maine	.91	Missouri	.82
New Hampshire	.90	Arkansas	.22
Vermont	.80	Oklahoma	.83
Massachusetts	.91	Texas	.94
Rhode Island	.85		
Connecticut	.96	Ohio	.91
		Indiana	.86
New York	.85	Michigan	.91
New Jersey	.87	Illinois	.88
Pennsylvania	.90		
		Wisconsin	.88
Maryland	.86	Minnesota	.85
Delaware	.91	Iowa	.84
West Virginia	.90		
		Nebraska	.88
Virginia	.52	Kansas	.86
North Carolina	.87	North Dakota	.71
South Carolina	.74	South Dakota	.88
Georgia	.20		
Florida	.76	Montana	.84
		Wyoming	.90
Kentucky	.87	Colorado	.89
Tennessee	.80		
		Idaho	.67
Alabama	.40	Utah	.94
Mississippi	.00	Nevada	.87
Louisiana	.67		
		Arizona	.95
		New Mexico	.92
		Washington	.85
		Oregon	.84
		California	.90