CHANCE

ISSN: 0933-2480 (Print) 1867-2280 (Online) Journal homepage: http://www.tandfonline.com/loi/ucha20

Memories of Election Night Predictions Past: Psephologists and Statisticians at Work

Stephen E. Fienberg

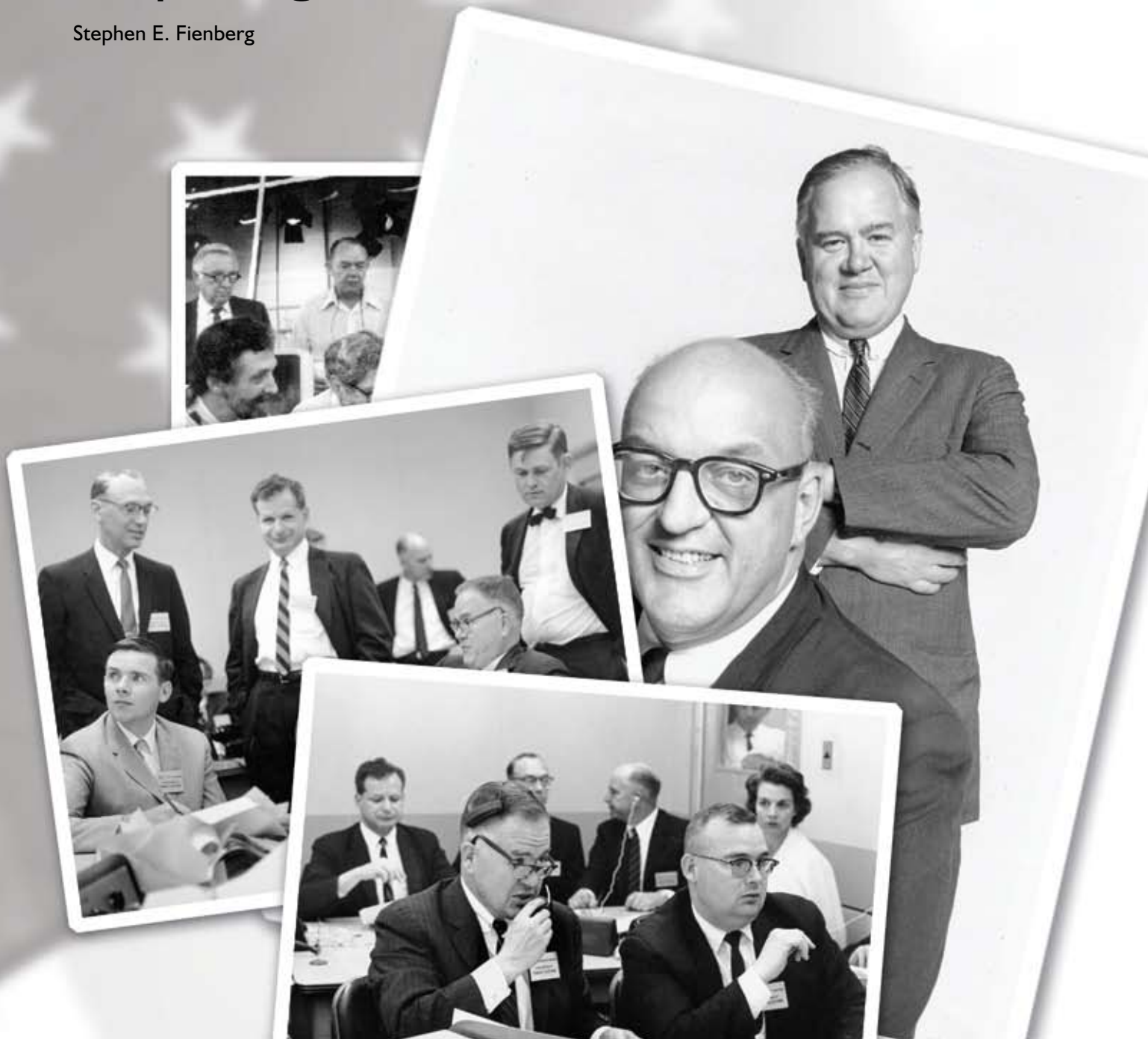
To cite this article: Stephen E. Fienberg (2007) Memories of Election Night Predictions Past: Psephologists and Statisticians at Work, CHANCE, 20:4, 8-17, DOI: 10.1080/09332480.2007.10722868

To link to this article: https://doi.org/10.1080/09332480.2007.10722868

Memories of Election Night Predictions Past:

Psephologists and Statisticians at Work

Stephen E. Fienberg



Psephologist (sē fäl´ə jist) ***n.*** A student of elections or, in today's parlance, a watcher of election polls. [From the Greek *psephos*, meaning "pebbles." The Greeks often used pebbles as ballots to record their votes. Sometimes used to describe statisticians forecasting election results.]

Serious statistical Election Day forecasting is approaching its 50th anniversary. At the same time, we may have come to the end of an era with the passing of two of Election Day's most important figures: John Tukey, who died in 2000, and Warren Mitofsky, who passed away in 2006. Each worked with one of the three major U.S. television networks for an extended period beginning in the 1960s, and they were responsible for the development of two very different statistical tools—one based on hierarchical Bayesian statistical models for early county-level vote returns and the other on exit polls of actual voters. Only pieces of each of their stories have been told, in part because they viewed the technical details as proprietary—especially Tukey. Still, interesting events took place when they went head-to-head in the 1976 presidential election.

Before getting into 1976, however, let's talk about November 2000, when the TV networks called the key state of Florida for Al Gore and then had to retract the call two hours later. All were looking at the same set of data from Voter News Service (VNS)—a single exit-polling operation. Most of us know what ultimately happened in the 2000 election, although many believe Al Gore did, indeed, win Florida and the election. As Philip Meyer of *USA Today* remarked in a November 9, 2000, article:

> … Something happened Tuesday night that made everyone pull back in Florida. One of exit polling's strengths is that the results can be checked against the official count in the precincts covered by the interviewers. In Duval County, where Jacksonville is located, and in the Tampa area, numbers looked wrong.
>
> … Both the components of VNS' projection model and the networks' competitive haste are suspect. … The primary reason for the creation of VNS was to save money and use some of that savings for a higher quality product. It was founded in 1990 by ABC, CBS, NBC, and CNN as Voter Research and Surveys (VRS). Three years later, VRS merged with the News Election Service, whose specialty was high-speed collection of official election returns. The result was VNS, and it is the pooled effort of the original four networks plus Fox and The Associated Press.
>
> It is the combination of fast relays of official returns and exit-poll data that make the broadcasters' election-night projections so spookily fast and accurate. But these advantages come at the cost of competition and the safeguards that it brings. Consider what Tuesday night might have been like if the six news operations each had its own exit-poll operation.
>
> With competition, we might not have needed to wait for the Duval County numbers to turn bad on Tuesday night and put Florida back in the undecided column. If the state were really too close to call, the networks using the other exit-poll services would have been saying so. The one with the questionable numbers might have been motivated to start searching for them earlier.
>
> But competition at that level is meaningless if all of the news organizations are working from the same bad data.

Meyer's observations, while interesting, miss the point. The real culprit was uncertainty and the failure of all the networks to take error and uncertainty into account as they attempted to make predictions. In retrospect, the Florida result was clearly too close to call, perhaps even after the election, but statistical caution lost out. It wasn't always like this.

A Not-so-Brief History

The Arrival of UNIVAC in the 1950s

Election night forecasting arose like a phoenix from the ashes of the 1948 election polls debacle, where the polls mistakenly predicted Thomas E. Dewey's victory over Harry S. Truman. In November 1952, the first commercial computer—the UNIVAC—was used to predict the results of the presidential election. UNIVAC I was unveiled, demonstrated, and

Walter Cronkite with the fake computer before the 1952 election

dedicated in Philadelphia on June 14, 1951. It was designed by J. Presper Eckert and John Mauchly, co-creators of the ENIAC computer.

The following account is drawn almost verbatim from articles by Kevin Maney of *USA Today* and Stephen Wright, who worked in a technical position for UNIVAC.

Prodded by Remington Rand, which was now looking for commercial uses of its new machine, CBS involved Eckert and Mauchly to develop election evening forecasts based on historical data and early returns. They enlisted the help of Max Woodbury, a statistician from the University of Pennsylvania, and he and Mauchly wrote one of the first algorithms for computing. It was essentially a regression analysis program.

Wright recalls, "The last program was finished the day before the election. Quality assurance consisted of a single test run using previous election results. We elected Roosevelt over Wilkie, so we felt pretty good."

On Election Day in 1952, Charles Collingwood set up shop in Philadelphia, working with Woodbury, Mauchly, and the UNIVAC. Meanwhile, CBS Anchor Walter Cronkite sat in the TV studio next to a fake computer—a panel embedded with blinking Christmas lights and a teletype machine. As polls began to close, clerks typed the data into the UNIVAC using three UNITYPER machines, which punched holes in a paper tape to be fed into the computer.

At 8:30 p.m. EST—long before news organizations of the era knew national election outcomes—the team ran its first projection cycle. Woodbury's algorithm produced a startling prediction: a landslide presidential victory for Dwight D. Eisenhower with odds of 100 to 1. Actually, according to Wright, the computer said "00 to 1" for Eisenhower because they had allowed only two digits for the odds. At 9 p.m., they ran a second cycle with additional data and got the same results. Because every poll predicted the race would be tight, CBS refused to air the results. According to Maney:

> Under pressure, Woodbury rejiggered the algorithms. UNIVAC then gave Eisenhower 8-to-7 odds over Stevenson. At 9:15 p.m., Cronkite reported that on the air. But Woodbury kept working and found he'd made a mistake. He ran the numbers again and got the original results—an Eisenhower landslide.
>
> Late that night, as actual results came in, CBS realized UNIVAC had been right. Embarrassed, Collingwood came back on the air and confessed to millions of viewers that UNIVAC had predicted the results hours earlier.

In fact, the official count ended up being 442 electoral votes for Eisenhower and 89 for Stevenson. UNIVAC had been off by less than 1%. It had missed the popular vote results by only 3%. Considering that the UNIVAC had 5,000 vacuum tubes that did 1,000 calculations per second, that's pretty impressive. A musical Hallmark card has more computing power.

The public latched onto the UNIVAC's performance. … "UNIVAC" suddenly became a generic term for those blinking electric brains. Much to IBM's disgust, when it introduced the 701 a few months later, people referred to it as "IBM's UNIVAC." In the public's mind, the UNIVAC was the new leader in computing.

By 1956, the TV networks all used computers and predicted results early, changing the dynamics of Election Day.

It wasn't all smooth sailing, however. In the 1954 congressional elections, according to *TIME*:

> UNIVAC turned Democrat with a vengeance and predicted a Republican disaster. Shortly after 9 p.m., UNIVAC claimed the Democrats would win a majority of 64 seats in the House and 23 in the Senate. But two hours later, the machine completely reversed its field. Commentator Charles Collingwood, who nursemaided the mechanical brain in both 1952 and last week, says "Suddenly UNIVAC said the Republicans were winning the House. We didn't know what to do. Should we change the machine? After all, last time the experts were wrong. I decided to stick with the machine." This particular error turned out to be caused by human frailty: A teletype operator had transposed the Democratic and Republican figures.

Helen Buczek, Max Woodbury, and Stephen Wright with others working for CBS in 1956

As for UNIVAC's mistaken idea that a Democratic sweep was in the making, Collingwood thinks it resulted from the fact that the first two states to report—Delaware and Connecticut—showing a heavier Democratic vote than was true of the national scene. Explains Collingwood defensively: "After all, UNIVAC is only human—that is, it can only make predictions based on the material that humans feed into it."

Tukey and the NBC Election Day Model

By 1960, computers had changed and so had the television networks' Election Day coverage goals. Everything was based on precinct results. At the time, there were well more than 150,000 precincts where votes were cast. Attention was focused on subsets of "key" precincts, chosen in different ways by the three networks according to early precinct results.

In 1960, RCA/NBC hired CEIR, a statistical consulting firm, to develop a rapid-projection procedure. CEIR consultants included Woodbury, Jack Moshman, Mauchly, Richard Scammon, and Tukey. Computers were still large, expensive, and slow, so much of what Woodbury had done for CBS still had to be done by hand. Moshman recalls:

We focused on what John Tukey called "swing-o-metric" precincts, the ones that most often reflected the way the state went. The absolute percentage was unimportant—the precinct could be one that went Republican every year. What mattered was how closely the swings from year to year in that precinct reflected the swing in the state total.

The other critical thing was that the precincts we used had to have machine ballots so we could get their results quickly. We would pick another election that seemed to have similar issues and circumstances, and use that as a starting basis for comparison. Then we would have a big meeting the Sunday before the election and put in the latest poll numbers as our "Time Zero" estimate, then update those numbers and track the differences as actual votes came in. That's essentially still what we do now, except that now we get county and subcounty totals electronically via the media pretty quickly. When we started, we had only statewide totals reported, so we had to rely on a nationwide collection of volunteers—recruited

David Brillinger (left, sitting) and John Tukey (right, sitting) as part of the NBC Election Day team in Rockefeller Plaza Studios with RCA staff Arthur Katz and Holwell Land in the background, 1962. *Photo courtesy of the American Philosophical Society, Tukey Archives*

John Tukey (front, left) and Richard Link (front, right) as part of the NBC Election Day team in Rockefeller Plaza Studios with RCA/NBC staff in the background, 1962. *Photo courtesy of the American Philosophical Society, Tukey Archives*

> mostly by the League of Women Voters—to phone results in from the precincts we used.
>
> In 1960 and 1964, we predicted only nationwide electoral vote totals, not state-by-state. That helped us, because we called both California and Illinois wrong, but the errors more or less canceled each other. Usually, big-city votes come in first and run Democratic, then late returns from rural areas run Republican. We didn't know Cook County's chairman held his vote to make sure Kennedy got enough to win. And in California, we neglected the absentee ballots, which—as they usually do—went heavily Republican. As it turned out, we were the only network that called it for Kennedy early, and we stayed with him all night, so we skunked the competition. But we bit our nails as the returns from the West came in.

In fact, among the insiders, Tukey won renown for preventing NBC from prematurely declaring Nixon the victor in the presidential race. Tukey took a data analysis approach to the problem of election projection, with the added spin that decisions had to be made quickly. He has described the work as "the best education in real-time statistics that anybody could have."

Data of several types were available: history, results of polls preceding the election, political scientists' predictions, partial county returns, and complete results for selected precincts. The data of the analyses were, in many cases, swings from sets of base values derived from past results and from political scientists' opinions. It turned out that the important problem of projecting turnout was more difficult than that of projecting candidate percentage.

Starting with the 1962 congressional election, Tukey assembled a statistical team to develop the required methodology and analyze the results as they flowed in on Election Day. Early members of the team included Bob Abelson, David Brillinger, Dick Link, Mauchly, and David Wallace, who joined for the 1964 primaries. From 1962 through 1966, they were consultants to RCA and interacted with Scammon, who had his own methodology using a collection of key precinct results.

David Brillinger wrote the following in a 2002 *Annals of Statistics* article, titled "John W. Tukey: His Life and Professional Contributions":

> Tukey sought "improved" estimates. His terminology was that the problem was one of "borrowing strength." Nowadays, parts of the work would be described as shrinkage and empirical Bayes. Jargon was developed;

for example, there were barometric and "swing-o-metric" precinct samples. The procedures developed can be described as an early example of empirical Bayes. The uncertainties, developed on a different basis, were just as important as the point estimates. Tukey's attitude to release the techniques developed is worth commenting on. On various occasions, members of his "team" were asked to give talks and write papers describing the work. When Tukey's permission was sought, his remark was that it was "too soon" and that the techniques were "proprietary" to RCA and NBC.

The methods and Election Day forecasting model were indeed novel. They are now recognizable as hierarchical Bayesian methods with the use of empirical Bayesian techniques at the top level. Tukey's students and his collaborators began to use related ideas on "borrowing strength," all of this before the methodology was described in somewhat different form by I. J. Good in his 1965 book and christened "hierarchical Bayes" in the classic 1970 paper by Dennis Lindley and Adrian Smith. The specific version of hierarchical Bayes in the Election Day model remained unpublished, although in an ironic twist, something close to it appeared in a paper several years later written by one of David Wallace's former students and a colleague who were unaware of any of the details of his work for NBC and developed their approach for different purposes. Several other hierarchical Bayesian Election Day forecasting models have been used in other countries.

From 1968 onward, the Tukey team worked directly for NBC and grew and shifted from election to election, although Abelson, Brillinger, and Wallace remained active throughout. The forecasting model was refined. Frank Jordan and Irwin (Bud) Lewis, the top people at the NBC News election unit from 1968 on, were "immortalized" by Tukey in the county-level model through the labels "Burgs" (Bud's urban groupings) and "Jurgs" (Jordan's urban groupings), which are two alternative subdivisions of a state at the top level of the hierarchical model.

I was part of the team in 1976 and 1978, and there were close to 20 PhD statisticians involved in one form or another, working in Cherry Hill, New Jersey, in the RCA Lab—which housed a large mainframe computer (and a back-up) dedicated to the evening's activities—and in New York, interacting with Scammon and the NBC "decision desk." Each analyst had a computer terminal and an assignment of states and races. A summary of each run of the model for a given race could be read easily from the terminal console, but full output went to a nearby line printer and was almost immediately available for detailed examination. Analysts worked with the model, often trying different prior distributions (different past elections chosen as "models" for the ones for which they were creating forecasts) and checking on robustness of conclusions to varying specifications.

John Tukey with Richard Scammon in the foreground, circa 1968. *Photo courtesy of the American Philosophical Society, Tukey Archives*

Sol Dutka, Jack Moshman, and Donald Herzberg with two ABC executives in the foreground, circa 1964. *Photo courtesy of Jack Moshman*

ABC Competes

In 1964, Moshman moved to ABC, where he worked with a group that included Sol Dutka and Irving Roshwalb—from Audits and Surveys—and political scientists Donald Herzberg and Warren Miller. He carried with him many of the basic ideas used by the NBC team, but his regression-style models

Warren Miller in the center with others from the ABC team, circa 1964. *Photo courtesy of Jack Moshman*

took on a somewhat different form than those developed by Tukey, Brillinger, and Wallace.

"We've had the best record of correct calls early," Moshman asserted in 2000, "with maybe five states wrong in presidential, gubernatorial, and Senate elections in the 32 years since we started doing state-by-state projections. And the networks have been moving more and more away from calling states before all their polling places close, so many times we sit for 45 minutes after we have a projection we're ready to release." Moshman remained with ABC through the 2000 election. He describes details for his model in a 1964 paper.

Enter Mitofsky and Exit Polls

The 1960s saw dramatic changes in the effort and extent of Election Day forecasting. CBS began working with pollster Lou Harris and IBM. Harris' version of Election Day forecasting, called Vote Profile Analysis, was based on quota samples of precincts within states. In 1964, he successfully projected the winners in 13 senate and gubernatorial races in seven states. Until 1964, only the wire services were tallying vote counts. The difficulties of that year led to the creation of a jointly sponsored Election News Service to play this role.

In 1967, a small group of survey statisticians led by Mitofsky and Murray Edelman, and with advice from Joe Waksberg with whom they had previously worked at the Census Bureau, carried out the first known exit polls in three off-year elections for CBS. That group soon formed the nucleus of the CBS News Election and Survey Unit in 1968. The sample selection approach they introduced in 1967 was a stratified probability sample with precincts selected in a single stage with probability proportional to the vote total in a recent election. They also used a weighting technique to control the variation in a party's vote, a form of regression estimation to allow for reporting from early return precincts, and a form of ratio estimation to utilize information from a past election in the denominator.

In exit polls, the interviewer stands outside the polling place and hands a questionnaire to a sample of voters throughout the day. The key question asks for whom the individual voted. But the polls quickly became useful for "election analysis" and commentary, not just projections, as other questions probed for the reasons behind voting behavior and attitudes and expectations.

NBC did its first exit poll in 1973, introducing the name. By 1980, all three of the major networks and the Associated Press were conducting and utilizing exit polls, although only Mitofsky and CBS used a true probability sampling procedure to select precincts. Mitofsky regularly emphasized this fact and later described the negative reaction from those

associated with the other networks. As a close colleague of mine remarked, making sure your key precincts are a probability sample is just fine once all the sample precincts have reported, but not until then.

The 1980 pre-election polls predicted a close race between incumbent President Jimmie Carter and then Republican candidate Ronald Reagan. The exit polls caught the real results early in the day, and by the time the networks went on air to broadcast returns, all three knew Reagan was scoring an overwhelming victory. NBC called the election at 8:15 p.m. EST, sounding a death knell to the Tukey statistical group focused on the analysis of early vote totals. It wouldn't be ready to call the election until 9:30 p.m. ABC and CBS waited for almost two hours.

In 1982, all three networks used exit polls in some combination with early vote returns. This was a set of costly enterprises that led to the consolidation in the early 1990s, which adopted the Mitofsky-Edelman approach of probability sampling.

Head-to-Head Competition in 1976

When Mitofsky died in September of 2006, the obituary in *The New York Times*, written by Adam Clymer, told the following story:

> A longtime CBS colleague, Martin Plissner, recalled yesterday how Mr. Mitofsky's insistence on precision caused CBS to be two hours behind ABC and NBC in calling Jimmy Carter's victory in the 1976 election. "About midnight," Mr. Plissner said, "Mr. Carter had secured 265 electoral votes out of the 270 needed for election. The News Election Service, which was counting the hard votes, declared that Mr. Carter had carried Mississippi, which casts seven. That was enough for maestros at the other networks, but not for Warren.
>
> "Warren knew that the electors were elected individually—not as a slate—in Mississippi, and he wanted to make sure Carter had won the five he needed, something NES couldn't tell him. For two hours, Warren and his minions worked the phones until they nailed down for sure the five votes needed to call it a night."

Myths abound in all industries, including the polling industry, and this remembrance is part myth and part fact. First, Mississippi did not cast its electoral votes individually in 1976; that was Maine. And at NBC, Tukey worked until past dawn to get the details correct, even though NBC had called the election for Carter and most NBC election team analysts had called it a night.

Warren Mitofsky (left) and Joe Waksberg working at CBS

Table 1— Forecasts from the NBC Election Day Model for Mississippi in November 1976

Time	%(R-D)/2 Projection	% Vote Reporting	2 s.e.
1:32	-0.6	81	0.7
1:58	-0.6	84	0.7
2:20	-0.3	85	0.6
2:28	-0.6	88	0.5
3:22	-0.8	91	0.4

Second, the News Election Service wasn't the organization that called Mississippi and the election for Carter; it was United Press International (UPI) and then NBC. On the Monday following the election, *TIME* carried the following report:

> The Super Bowl comes but once every four years for network news divisions, and Election Day is it. To call or not to call is the question—first the states, then the presidential winner—and timing with accuracy is everything. It was precisely at 3:30 a.m. when NBC, taking a deep breath and one last look into the oracular recesses of its key precincts, declared that Jimmy Carter would capture an Electoral College majority and be next president of the U.S.
>
> NBC's *pronunciamento* beat ABC to the verdict by seconds, CBS by 15 minutes. [The] three networks trailed United Press International, which declared Carter the winner at exactly 2:57 a.m. "We're hypercautious," admitted Walter Cronkite. "We're always first," said a happy NBC News President Richard Wald as he munched tortilla chips at his Rockefeller Center election command post. To which William Sheehan, Wald's counterpart at ABC, replied, "I'd be satisfied to call it a tie."

I can't speak for UPI, which to my knowledge had no statistical analysts or elaborate prediction models, but I can tell a bit of the story from the NBC perspective, as I was the analyst with primary responsibility from the presidential race in Mississippi. In Table 1, I show the information needed for a Bayesian posterior 95% predictive interval for Ford vs. Carter at various points in the early morning. What was the problem in projecting Mississippi for Carter? Jackson. At 1:32 a.m., only 29% of the vote from the Jackson area was in, and the vote from that area was running counter to the rest of the state. We ultimately called Mississippi for Carter when 51% of the Jackson vote was tallied and when the 95% interval no longer contained zero. But the final decision on the announcement came from New York a few minutes later.

Back to the Exit Polls in 2000

In a November 2000 pre-election tribute to Tukey—several months after his death—David Alan Grier wrote in *The Washington Post*:

> ...we are now living in a post-Tukey age. He departed this Earth just before the nominating conventions last July. As much as his genius and energy will be missed, I cherish the small hope that—absent Tukey—the psephologists will be wrong this year, that their equations will fail and that we will see just a bit more drama when they report the votes.

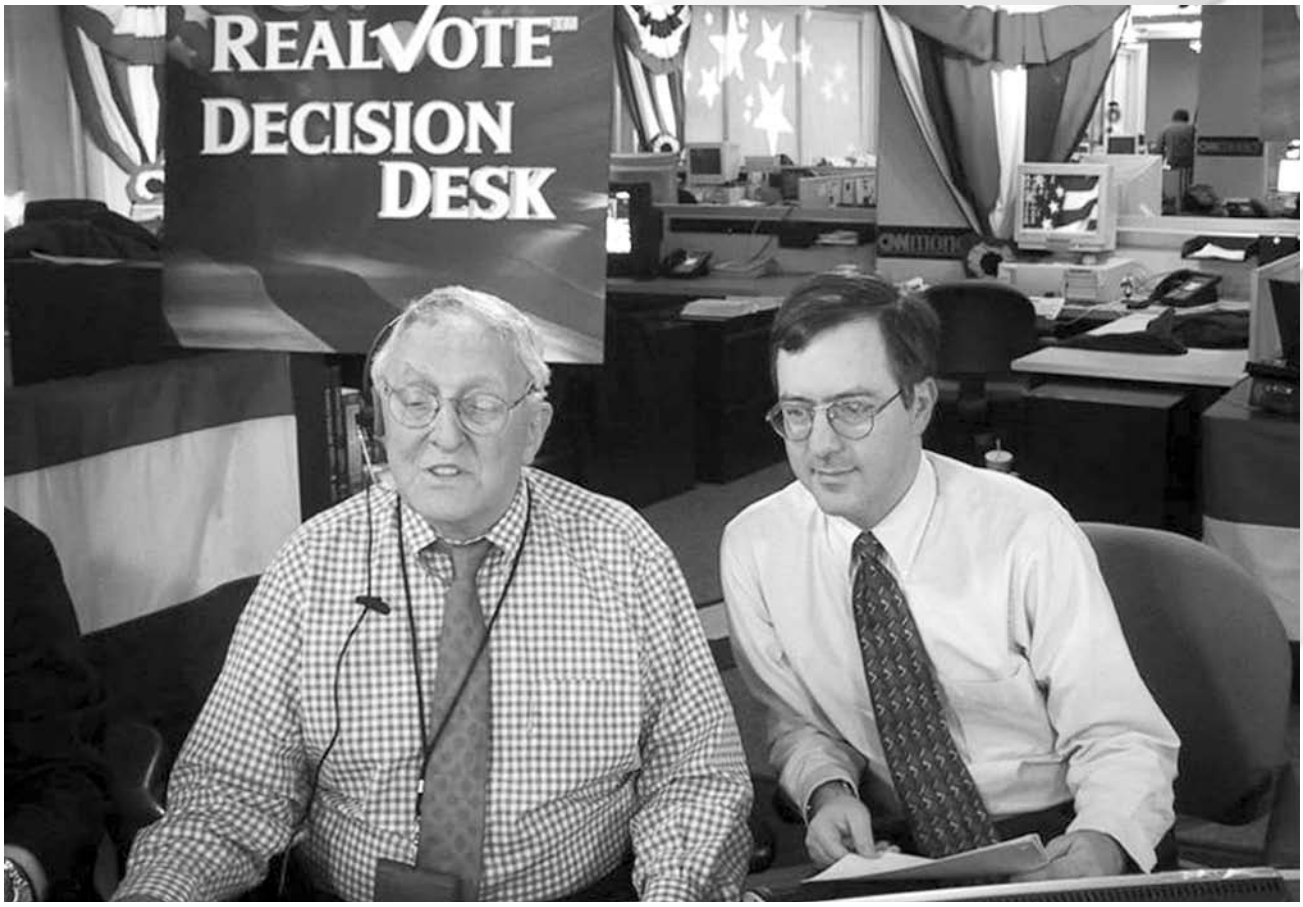
Brillinger observed the following in his 2002 *Annals of Statistics* article:

> NBC stopped involving Tukey after the election of 1980. A stated reason was that exit polls, which interview people directly after they leave a polling station, had become highly refined. Such polls do have the advantage of being based on data from individual people who actually appeared to vote. This measurement device has the further flexibility that, if (because a race is close) there is a need for more data during the day, more voters can be interviewed.

Exit polls remain the principal tool of the television networks, but there is an interesting postscript to this story. In advance of the 2004 presidential election, ABC News posted a statement on its web site about how it would deal with projections:

> To project a race, the Decision Desk analyzes exit polling and actual vote data using a variety of statistical models. The Decision Desk waits until the models indicate that there is at least a 99.5% certainty that the leading candidate is the winner. The team also considers the possible impact of absentee and early voters and a number of other factors, which vary by state.
>
> ABC News will not project any races where the margin between the candidates is less than one percentage point in the tabulated vote, even if 100% of the precincts have reported. This is because many votes may be outstanding, even with 100% of precincts reporting (e.g., absentee and provisional ballots). Also, the county vote data on Election Day are unofficial tallies, so there may be errors that could affect the results in a close race.

We can only speculate on what might have happened had Tukey been involved with one of the networks in making the Florida projections in the fall of 2000. But, we can hope that



Warren Mitofsky and Joe Lenski at CBS on election night, 2002. *Photo used with permission of Joe Lenski.*

his influence will continue to be felt among the psephologists plying their trade on network television. C

Further Reading

ABC News. (2004). "How Projections Are Made: A User's Guide to Understanding How ABC News Projects Election Winners." *http://abcnews.go.com/politics/vote2004/story?id=206326*.

Bernardo, J.M. and Girón, F.J. (1992). "Robust Sequential Prediction from Nonrandom Samples: The Election Day Forecasting." In J.M. Bernardo, J.O. Berger, D.V. Lindley, and A.F.M. Smith (Eds.), *Bayesian Statistics 4*, 61–77. London: Oxford University Press.

Brillinger, David R. (2002). "John W. Tukey: His Life and Professional Contributions." *Annals of Statistics*, 30: 1535–1575.

Brown, Philip J.; Firth, David; and Payne, Clive (1997). "Forecasting on British Election Day 1997." *J. Roy. Statist. Soc., Series A*, 162: 211–226.

Fang, Irving E.; Roschwalb, Irving; Moshman, Jack; and Herzberg, Donald G. (1968). "Election Projection Seminar." *IEEE Spectrum*, September, 40–51.

Levy, Mark R. (1983). "The Methodology and Performance of Election Day Polls." *Public Opinion Quarterly*, 47: 54–67.

Link, Richard, F. (1978). "Election Night on Television." In Judith Tanur, et al. (Eds.) *Statistics: A Guide to the Unknown*, Second Edition. Holden Day: 178–186.

Maney, Kevin (2004). "In '52, Huge Computer Called UNIVAC Changed Election Night." *USA Today*. *www.usatoday.com/money/industries/technology/maney/2004-10-26-univac_x.htm*.

Mitofsky, Warren and Edelman, Murray (2002). "Election Night Estimation." *Journal of Official Statistics*, 18: 165–179.

Moshman, Jack (1964). "The Role of Computers in Election Day Broadcasting." *Advances in Computers*, 5: 1–21.

Samuelson, Douglas A. (2000). "Gore Wins! (At Least That's What the Model Says)." *OR/MS Today*, October 2000. *www.lionhrtpub.com/orms/orms-10-00/samuelson.html*

Scott, Alastair and Smith, T.M.F. (1969). "Estimation for Multi-Stage Surveys." *Journal of the American Statistical Association*, 64: 830–840.

Scott, Alastair and Smith, T.M.F. (1971). "Bayes Estimates for Subclasses in Stratified Sampling." *Journal of the American Statistical Association*, 66: 834–836.

Wright, Stephen. (1988). "Forty Years in the DP Trenches." *http://home.comcast.net/~sew1/CV/40yrs.htm*.