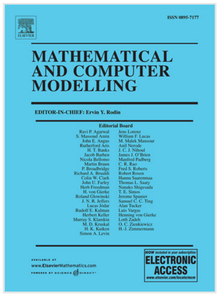

Accepted Manuscript

Median statistics in polling

J. Richard Gott III, Wesley N. Colley

PII: S0895-7177(08)00195-7
DOI: 10.1016/j.mcm.2008.05.038
Reference: MCM 3402

To appear in: *Mathematical and Computer Modelling*

Received date: 19 May 2008
Accepted date: 22 May 2008

Mathematical and Computer Modelling **000**, 1–**??** (2008) Printed 10 June 2008

Median Statistics in Polling

J. Richard Gott, III[1*], and Wesley N. Colley[2]
[1]*Dept. of Astrophysical Sciences, Princeton University, Peyton Hall, Ivy Lane, Princeton, NJ 08544*
[2]*Center for Modeling, Simulation and Analysis, D-15 Von Braun Research Hall, Univ. of Alabama in Huntsville, Huntsville, AL 35899*

10 June 2008

ABSTRACT
In 2004, we used a very simple, but surprisingly effective method to successfully predict the outcome of the U.S. Presidential election. Using the median poll in the last month for each state, we correctly predicted the results in all states but one (Hawaii). Just as we had originally hoped, the method made it possible to predict successfully the results in the large close states (Ohio, Pennsylvania, and Florida) where there were a great many polls taken. States with only a few polls were generally not close, and so the median poll also predicted these states successfully. The method appears particularly well adapted to U.S. Presidential elections where the candidates are chosen well in advance, and where outcomes in individual states determine the winner.

Key words: elections, polling

1 INTRODUCTION

Perhaps no one has ever stated the benefits of the median over the mean better than the great Russian physicist Ya. B. Zeldovich. At an International Astronomical Union meeting in Tallinn, Estonia in 1977, Ya. B. Zeldovich commented approvingly of Gott and Turner's [1] use of median statistics for mass-to-light ratios in groups of galaxies: Zeldovich noted that in Russia the watches are not made very well, so when three friends meet they compare the times on their watches—one says "it's 1 o'clock," the second says, "it's 5 minutes after 1," the third says "it's 5 o'clock." Take the median!

In this paper we will examine the power of median statistics applied to Presidential Election polls.

The usual hypotheses made in evaluating data using a standard χ^2 analysis are: (1) Individual data points are statistically independent (2) There are no systematic effects (3) The errors are normally distributed, and (4) The error bars are well-quantified. These are four extraordinarily potent hypotheses-and lead to powerful results-if the four conditions are indeed true. Even the first two conditions alone can lead to powerful results by themselves. We now consider the implications of relaxing the third and fourth conditions, following the discussion in Gott et al. [2].

As an example of the traditional χ^2 approach, we shall discuss the important analysis of supernova distances by Riess et al. [3], who used all four hypotheses. Distances for each of 16 distant supernovae are combined with data on nearby supernovae to produce a least squares fit for Ω_m and Ω_Λ (the fractions of the cosmological critical density in matter and dark energy respectively) then χ^2 analyses were done to set likelihood ratios for different cosmological models (defined by their values of $\{\Omega_\Lambda, \Omega_m\}$). Now $\Omega_m = 8\pi G\rho_m/3H_0^2$ where G is Newton's gravitational constant, ρ_m is the average density of matter in the universe, and H_0 is Hubble's constant which measures the current expansion rate of the universe, and $\Omega_\Lambda = 8\pi G\rho_\Lambda/3H_0^2$ where ρ_Λ is the density of dark energy. Dark energy is a quantum vacuum state with positive density and and equally negative pressure. The negative gravitational effects of the negative pressure cause a repulsive effect. If $\Omega_\Lambda + \Omega_m = 1$ the universe is infinite and has a flat Euclidean geometry according to Einstein's field equations of general relativity. If $\Omega_\Lambda + \Omega_m > 1$ the universe is positively curved and finite, and if $\Omega_\Lambda + \Omega_m < 1$ the universe is negatively curved and infinite. A cosmogony of $\{\Omega_\Lambda = 0, \Omega_m = 1\}$ would indicate a flat universe with matter only and no dark energy that would expand forever but at a slower and slower rate. Usually it was assumed that $\Omega_\Lambda = 0$. A value of $\Omega_\Lambda > 0$ would indicate a universe containing some dark energy. With enough dark energy, an accelerated expansion would be observed. From their observations of the distances of 16 supernovae, and using some additional Bayesian assumptions, Riess et al. [3] concluded that the probability that $\Omega_\Lambda > 0$ is 97.1%. This was the first direct evidence that the expansion of the universe was accelerating and that most of the density in the universe was in the form of dark energy (as measured by

* E-mail: jrg@astro.princeton.edu

Ω_Λ). These results relied on the important assumption that the errors are normally distributed-as is apparent in their χ^2 analysis. This was a somewhat troubling assumption since the errors in corrected supernovae luminosities are not likely normally distributed. In fact, while there seems to be a rather strong upper limit on supernovae luminosities, there remains a longer low luminosity tail, which implies asymmetric error bars. Also, the corrected supernovae luminosities are calculated using a rather small training set, and so one can never be sure that one is not encountering at least some supernovae that are odd and do not fit the training set. Our results [2] showed that median statistics gave virtually the same results as Riess et al. [3] but without the additional assumptions (3) and (4). Thus the data showed evidence for a universe with accelerating expansion dominated by dark energy: not because of a few outlying data points but because most of the supernovae brightness's were fit better by the accelerating model.

The limits of assuming a normal distribution are illustrated by an example from Hill [4]. Suppose one measured the weights of a million adult penguins and found them to have a mean weight of 100 lbs. with a standard deviation of 10 lbs., and further suppose that the observed data's distribution fits a Gaussian distribution perfectly. Of the million penguins measured, consistent with a Gaussian distribution, suppose the heaviest one weighs 147.5 lbs. Then what is the probability of encountering, on measuring the next adult penguin, a penguin weighing more than 200 lbs.? Well, one might be tempted to say that it was $P = 10^{-23}$, by simply fitting the normal distribution and calculating the probability of obtaining an upward 10-sigma fluctuation. But this would be wrong. There could be a second species of penguin, all of whose adults weighed over 200 lbs. which simply had a population a million times smaller, so that one had not encountered one yet! Then, the probability of encountering a penguin weighing over 200 lbs. is 1/1 million. Even data that fit a normal distribution perfectly cannot be used to extend the range beyond that of the dataset. The correct answer, which does not depend on any assumption (3) is that, the probability that the 1,000,001st penguin weighs more than any of the million penguins measured so far is, quite simply just $P = 1/1,000,001$. This is according to hypothesis (1) that all the data points are independent. After weighing 1,000,001 penguins, since all the penguins are independent, each of the penguins must have an equal chance (1/1,000,001) of being the heaviest one. Thus, the last one must have a probability of 1/1,000,001 of weighing over 147.5 lbs.

People often make similar mistakes by extrapolating distributions beyond their range. The Guinness Book of World Records once found that the oldest living human being was a man in Japan who was at the time 114 years old. After a detailed statistical study showing that the mean life expectancy of people over 100 years old was less than 1 year, and noting the number of centenarians, they calculated that the odds of encountering an individual 115 years of age or older were less than one in 5 billion and that therefore, people claiming to be significantly older that their 114 year old record holder were bound to be fakes. Their 114 year old man had extremely well documented birth records. But then an interesting thing happened: he lived on an additional six years, dying at age 120! Had he defied the odds? If the life expectancy of 114 year olds were really 1 year, then the probability that he would live six more years would be $\exp(-6)$ or 1/400. But the truth was, although we had good life expectancy data on 100–110 year olds, we had essentially no data on the life expectancy of 114 year olds. It could have been 6 years even. Indeed within a few years there was a new record holder, a French woman, who lived to an age of 122 years before her death. This illustrates the difficulty of extrapolating distributions beyond the range over which the data for them has been measured.

χ^2 distributions are particularly prone to such errors. The normal distribution is one which falls off so rapidly at large values (like $\exp[-x^2/2\sigma^2]$) that it is easy to get very dramatic results, results which are not justified if the underlying distribution turns out in the end not to have Gaussian errors, but to have longer tails.

When Gaussian errors are assumed, one of the great benefits is that the errors go down as one over the square root of N, where N is the number of measurements. Thus, with 16 supernovae one can get estimates of $\{\Omega_\Lambda, \Omega_m\}$ that are 4 times as accurate as with a single measurement. Median statistics can take advantage of a similar one-over-square-root-of-N factor to produce surprisingly accurate results, even while not relying on hypothesis (3) and (4) that the errors are Gaussian and that we know their magnitude. Indeed, hypotheses (1) and (2) are sufficiently powerful by themselves to produce dramatic results, results that are only slightly more conservative than by adopting hypothesis (3) and (4) as well, but ones in which we may have more confidence because two significant, and perhaps questionable, assumptions have been relaxed.

In section 2, we shall outline how median statistics can be used with N estimates (adapting a discussion in our paper on median statistics in astronomy [2]). Then we will tell how we applied median statistics to predict the 2004 Presidential Election results.

2 MEDIAN STATISTICS

We will assume the truth of hypotheses (1) and (2): that our measurements of a given quantity are independent and that there are no systematic effects. Suppose we were to take a large finite number of measurements. We will then assume-call this hypothesis (2a)-that the true median value thus obtained as the number N of measurements tends to infinity will be the true value. This is almost equivalent to the hypothesis that there are no systematic effects, for if we make a trillion measurements and there are no systematic effects, and each is a fair measurement, we might naturally expect half to be above the true value and half to be below the true value. Indeed this might be one's definition of "no systematic effects." Note also that there are some distributions like the Cauchy distribution which have no defined mean value in the limit of large N, but still have perfectly well defined median values. We will suppose that after some very large number of measurements, as N tends to infinity, there would be a true median which we will call TM. Now by hypothesis (1) each individual measurement will be statistically independent; thus, each has a 50% chance to be above or below the true median (TM). Suppose we make N independent measurements M_i where $i = 1...N$. Where is TM likely to be? Well, the probability that exactly n of the N measurements are higher than TM is given by the well-known binomial distribution: $P = 2^{-N} \cdot N!/[n!(N-n)!]$, because there is a 50% chance that each measurement is higher than TM; thus P is equal to the probability of getting exactly n heads out of N coin flips. Given the fact that we have taken N measurements, M_i, and these are later ranked by value such that $M_j > M_i$ if $j > i$, then the probability that the true median TM lies between M_i and M_{i+1} is

$$P = \frac{2^{-N} N!}{i!(N-i)!}, \tag{1}$$

where for convenience we can set M_0 equal to minus infinity and M_{N+1} equal to plus infinity. For example, if $N = 2$, the probability is 50% that the true median lies between M_1 and M_2, which is to say that one measurement is too high and the other is too low. There is a 25% chance both measurements are too high, and a 25% chance that both are too low. In the coin flip analog, there is a 50% chance of getting one heads and one tails, a 25% chance of getting two heads, and a 25% chance of getting two tails.

Importantly, the distribution of the true median, TM, is much narrower that the distribution of the measurements themselves. For comparison the probability that the next individual measurement you take will lie between M_i and M_{i+1} is just

$$P = \frac{1}{N+1}. \tag{2}$$

If we set $r = i/N$, then we can define $M(r) = M_i$. Then the distribution width can be defined by the variable r. Then any measurement m may be associated with a value r by the inverse function $r(m)$ such that $M(r) = m$ with suitable interpolation applied. In the limit of large N we find that the mean expectation value of r for the next measurement m is $\langle r \rangle = 0.5$ and its standard deviation is $\left\langle r^2 - \langle r \rangle^{1/2} \right\rangle = 1/\sqrt{12}$ [since the distribution is uniform in r over the interval 0 to 1]. On the other hand, in the limit of large N, the expectation value of r for the true median TM is $\langle r \rangle = 0.5$ and its standard deviation is $\left\langle r^2 - \langle r \rangle^2 \right\rangle^{1/2} = 1/\sqrt{4N}$. (In fact, in the limit of large N the distribution in r approaches a Gaussian distribution with the above mean and standard deviation). Thus, as we take more measurements we see that the standard deviation in r of the TM is proportional to $1/\sqrt{N}$. So if we use median statistics, we find that our accuracy in determining the true median (as measured by the percentile r in the distribution of measurements) improves as $1/\sqrt{N}$ as N grows larger. So we do get our factor of one over square root of N: improvement just like people using mean statistics would be getting.

Let us apply this to the previously mentioned penguin problem. Suppose we have measured the mass of 1,000,000 penguins and find that they follow well a normal distribution with a mean of 100 lbs. and a standard deviation of 10 lbs. Now applying the standard formula, we would compute the mean to be 100 lbs, and the standard deviation of the mean to be $\sigma_{\text{mean}} = 10$ lbs.$/(999,999)^{1/2}$ = 0.01 lbs. Thus, we would deduce that there is a 95% probability that the true mean for the population of penguins lies between 99.98 and 100.02 lbs. But this result will be true only if the distribution in penguin masses beyond the limits seen in the first million penguins is well behaved, in particular falling off more rapidly than 1/mass. Suppose one penguin in a million weighs 100,000,000 lbs. Since we have examined only 1 million penguins, there would be an appreciable chance ($P = e^{-1} = 0.38$) that we would have missed one of the super massive ones). Yet, these super massive penguins make the true mean = 200 lbs. (Lest this example seem extreme, consider that in calculating the mean mass of all atomic nuclei in the universe, one would have to consider neutron stars as well as those nuclei in the periodic table! Neutron stars are, after all, just big atomic nuclei stabilized by gravity; but instead of having atomic weights less than about 250, their atomic weights are of order 10^{57}. They would dramatically influence the mean but would hardly influence the median.) So, even if the already measured data are well behaved, it is easy to be fooled by extreme cases falling beyond the observed distribution. One is less likely to be fooled about the median mass.

In the above example we would deduce that with $N = 1,000,000$ that the expected r value of the true median (TM) and its standard deviation would be 0.5 and 0.0005 respectively. Thus, we would expect that there would be a 95% chance (2σ) that the true median has an r value between 0.499 and 0.501. In other words, we expect the true median weight of penguins to lie between the weight of the 499,000th and the 501,000th most massive of the million measured penguins. These are distributed approximately normally so the 499,000th most massive weighs 99.975 lbs. and the 501,000th weighs 100.025 lbs. Thus, with 95% confidence we would say that the true median lies between 99.975 lbs. and 100.025 lbs. Note that these limits are only slightly more conservative that the 95% confidence limits we made on the mean earlier (which were 95% that the true mean was between 99.98 and 100.02). Furthermore, these limits are not invalidated by the super massive one-in-million penguins. Their existence only changes the true median to 100.000025 lbs. If one's data points are independent and there are no systematic effects then, the median value is not going to be greatly perturbed by rare data points lying beyond the range of observed values—whereas the mean can always be significantly perturbed.

Thus, interestingly, the 95% confidence limits on the true median are not much wider than those we would have made for the mean assuming Gaussian distributions and they are more secure since the hypothesis of Gaussian distributions with known error bars is dropped.

We can translate these results into a Bayesian Analysis. Bayesian statistics say that the posterior probability of a particular model after analyzing the data at hand is proportional to the prior probability of that model multiplied by the likelihood of obtaining the observational data given that model. So effectively we are assuming as a prior a distribution which is equal to the probability that the next observation will be between M_i and M_{i+1} (which is $1/N+1$). Then we multiply this by the likelihood of having n observations higher that the true median value as given by the binomial distribution. For example, suppose that we have only two independent measurements of a given quantity like the Hubble constant (H_0) (as was essentially the situation for decades). This quantity measures the expansion rate of the universe. The two values measured were $M_1 = 50$ km/s/Mpc and $M_2 = 100$ km/sec/Mpc. We would say that if the data points were independent and there were no systematic effects that according to the binomial distribution in Eq. (1), there was a 25% chance both measurements were too high, a 50% chance one was too high and the other was too low, and a 25% chance that both were too low. So we would estimate $P = 25\%$ that

$H_0 < 50$ km/s/Mpc, $P = 50\%$ that 50 km/s/Mpc $< H_0 < 100$ km/s/Mpc, and $P = 25\%$ that $H_0 > 100$ km/s/Mpc. This result using Eq. (1) is reasonable. By using the likelihood (of high and low values) as the probability estimate, we are effectively assuming a prior of 1:1:1 odds that the value is below 50, between 50 and 100, and above 100. Multiplying by the binomial likelihoods of 25%, 50%, and 25% gives us posterior probabilities of 25%, 50%, and 25%.

Note for comparison that if we had used a vague Bayesian prior for a positive quantity (uniform in the logarithm), we would have prior probabilities of 50%, 0% and 50%, because there is only a finite logarithmic interval between 50 and 100, while there are infinite logarithmic intervals both above 100 and below 50. Thus, even after multiplying by the binomial likelihoods of 25%, 50%, and 25%, we would still deduce zero chance that the true value was between 50 and 100. That would be unreasonable, so we prefer Eq. (1), which gives reasonable results for this situation.

3 MEDIAN STATISTICS IN OTHER APPLICATIONS

Median statistics can be useful in a variety of applications. For example, Gott and Turner [1] used median estimates for mass-to-light ratios of groups of galaxies. This minimized the effects of occasional groups containing contaminating effects of background and foreground galaxies which gave discrepant values. This was the application Zeldovich referred to in his comments dicussed previously. In his 1995 book *How Many People Can the Earth Support?*, Joel Cohen [5] catalogs 66 different estimates of the carrying capacity of the Earth (how many people the Earth could possibly support) made in the previous 316 years; he takes the median. The first estimate was 13.4 billion by Leeuwenhoek, the inventor of the microscope, in 1679. The second estimate was 12.5 billion by King in 1695. Cohen notes that the values depend on the assumptions made about how people will choose to live, and a variety of assumptions lead to a variety of answers, but he notes that the median estimate of this quantity has hardly changed in over 300 years. The median value of all the 66 expert estimates (spanning more than 300 years) is 12 billion. This value of 12 billion is remarkably close to the very first estimate made by Leeuwenhoek in 1679 (at a time when the world population was only 640 million)! The estimates have a large scatter. One high estimate assumes that the land area of the earth is *entirely* covered by crops, and their leaves intercept *all* the available light from the sun (letting none be wasted by hitting the ground), that furthermore these plants are all converted for human consumption (no other land animals) and that all humans are vegetarians (that led to an estimate of a carrying capacity of 1 trillion people). The most extreme estimate was from a physicist, who considered only the limits imposed by heat dissipation. He imagined the Earth covered by air conditioned structures, with a limiting outer skin temperature of 1000°C (that led to an estimate of 6×10^{16} people). Most estimates made more realistic assessments, including in their calculations the fraction of arable land and what crops might actually be grown, and making plausible assumptions about the space allocated to human living areas, etc. Some estimates for the maximum number of people supportable on the earth, on the other hand, are even below the current world population. How can this be? Because the authors of those studies believed that the current population has already overshot the maximum carrying capacity of the Earth, is not sustainable, and will surely crash below that value shortly. Carrying capacity estimates how many people could be supported in a steady-state, or sustainable condition. Ehrlich and Ehrlich [6] believe that a maximum of only 1.2 billion people can be sustained over long periods of time, and that today we are using up nonrenewable resources—for example, by burning the rainforest and turning it into not very good farmland—and that as pollution soars and waste accumulates, we are actually lowering the capability of the earth for long-term sustenance of people to a level below 1.2 billion. But the Ehrlichs [6] have one of the low estimates . How should the estimates be combined? Using the average of the 66 estimates (that would give 9×10^{14} people!) would certainly be wrong, as it would be entirely dominated by the one physicist's extreme estimate based on waste heat alone. No, following Zeldovich, one should use the median; that's 12 billion.

A rather different type of median estimate can be made using the Copernican Principle (cf. Gott [7,8]) to address the question "What is the maximum that human population will likely ever achieve?" If population growth follows a simple model of exponential growth and decay, 50% of all observers will be born when the population is greater than 1/2 the maximum population. And so, if a baby born today is, by the Copernican Principle, not special, then there is a 50% chance of that baby finding the current population within a factor of two of its maximum value. Thus (with these simplifying assumptions about exponential growth and decay—which may be a reasonable approximation in the brief population spike where most people are born) the "Copernican" prediction is: given that the population is now 6 billion, if a baby born today is not special, there is a 50% chance that the maximum population will not exceed 12 billion. This gives a median estimate of 12 billion for the maximum population—we expect this estimate to underestimate the true value half the time, and overestimate the true value half the time.

Let us consider a third independent method using the median. Lutz, Sanderson, and Scherbov [9] estimated the future population history of the earth based on experts' opinions of future trends in fertility, mortality and migration. When experts made different estimates, a probability distribution for that quantity was used. This allowed the authors to make 4,000 different simulations covering in each case the probable ranges of birth rates, death rates and migration rates according to the experts. They derived probability distributions for population sizes and age distributions for both sexes in 13 different regions of the world as well as considering migration between regions, all the way to the year 2100. Starting with a population of 5.8 billion in 1997 their different simulations fanned out in the future. Their median-estimate population curve (a curve such that half of their simulations lie below it and half lie above it) is quite interesting, showing the world population increasing to 7.9 billion by 2020 and 10.0 billion by 2050. Their median curve reaches a peak value of about 10.9 billion in approximately 2070-2080 and then begins to decline gently. Therefore, their median estimate for the maximum population of the Earth is 10.9 billion. Thus, three dramatically different methods (carrying capacity, Copernican, and demographic) with different assumptions give median estimates (12 billion, 12 billion, and 10.9 billion) which are in surprisingly close agreement.

Median statistics have also been applied in politics. Duncan Black (1958) in his book *The Theory of Committees and Elections* [10] quotes Francis Galton, who notes that the median is optimum when, for example, a jury has to assess damages, or a council has to allot a sum of money for a particular purpose: "... *not* the *average* of all the estimates, which would give a voting power to 'cranks' in proportion to their crankiness... I wish to point out that the estimate to which least objection can be raised is the middlemost estimate, the number of votes that it is too high being exactly balanced by the number of votes that it is too low. Every other estimate is condemned by a majority of voters as being either too high or too low." It is in just this context that Cohen's [5] use of the median is sensible. Any other estimate would be condemned by a majority of his experts as too high or too low.

In many sports ratings systems it is recognized that outlying ratings should be excluded. In Olympic diving, the high and low scores among the judges are eliminated and the remainder are averaged. A similar system is used for the computer rankings used to pick the top contenders for the National Championship game of the Bowl Championship Series in college football. The high and low ranks are eliminated out of 6 computer ranking methods and the rest averaged. But of course, there might always be more than one outlier, and the median could also serve well in scoring as outlined in Galton's argument.

4 ELECTION POLLING RESULTS

We decided to apply median statistics to predict the U.S. Presidential Election of 2004. Please note that all of our data on the polls was gathered from the website Electoral-Vote.com [11]. Our idea was that median statistics applied to polls would be interesting and informative. In the Presidential race it is the electoral vote that is important. That means that the important polls are those done state by state. Such polls usually have smaller sample sizes than national polls and larger expected errors (typically $\pm 4\%$). If a state is close, an individual poll will not declare a winner, since the separation of the candidates is typically less than the margin of error of the poll. However, there are many polls. If we can use all the polls we can obtain a more accurate estimate. Second, in state-by-state polls, often done by newspapers, there is a danger of an outlying poll skewing the data. Some polls may have a partisan ax to grind or have bad polling techniques. Median statistics would protect as well as possible against the occasional bad poll. If all the polls are bad, of course one will not achieve success, but the median statistics protect against bad polls as much as possible. Finally, state polls usually give data only to even percentiles: i.e. 48% to 47% and so ties are possible (i.e. 49% to 49%). One could not even define the average more accurately than 1% given that reporting is only given to 1%, but we can easily tell whether a given candidate has won or lost or tied an individual poll.

4.1 The One-Month Window

Our first thought was simply to use the median of all polls in a given state to predict the ultimate winner. This would use the most polls, achieving the highest precision, and we hoped that there would be many more polls taken as the election approached, so that we would be somewhat protected against trends. However, trends were still the major worry with this technique. Suppose a candidate won 14 polls in a row in a given state, but then lost the next 12 polls in a row as the election approached. On the day before election, one would tend to not pick him as a winner, because the most reasonable hypothesis would be that he had simply lost popular support with time. Trends can sink the median method. We thought of various possible ways around this problem. Perhaps one could apply a rank sum test to the data ordered by time, to see if there was a statistically significant correlation between the order of the wins and the time. In other words, was the median time of the losses statistically significantly earlier than the median time of wins? If that were true, then one could discard the first half of the data. The trouble with this was that we anticipated that many states would have only a handful of polls and so no statistically significant trends could be detected. Also any of the systems we considered, seemed ad hoc, needlessly complex, and hard to explain to the public. In the end we decided to protect against trends by simply adopting a short period over which to count the polls (one month). This was long enough to give us many polls in the close states and still short enough, we hoped, to protect us against trends. It was simple and easy to understand, and the period of one month did not seem ad hoc. We decided to report both results. We reported the median of all polls in each state from July 7 onward (after both conventions when both tickets were known) and for just the last month. We put the results online in early October, and kept them up to date until election eve. We regarded the election as a test of which method would perform better. When we started, some states had not taken any polls yet, so for these we used as a first poll the way the state went in the last election. But these were quickly replaced when actual polling results from 2004 appeared. We used all polls and gleaned the data from the internet. We made an election map where the area of each state was proportional to the electoral vote and colored them in red or blue depending whether the Republican (Bush) or Democrat (Kerry) was leading at the time in the median statistics-i.e. had won more polls in that state (overall, or within the last month). See Figure 1 for our election eve prediction using the last month's polls median.

4.2 Vagaries of the Electoral System

We were careful to follow the rules in each state in predicting the electoral results. We thought that was important. Most states have a winner-take-all system, but Colorado had a ballot initiative that would, if passed, split Colorado's votes according to the proportion of the

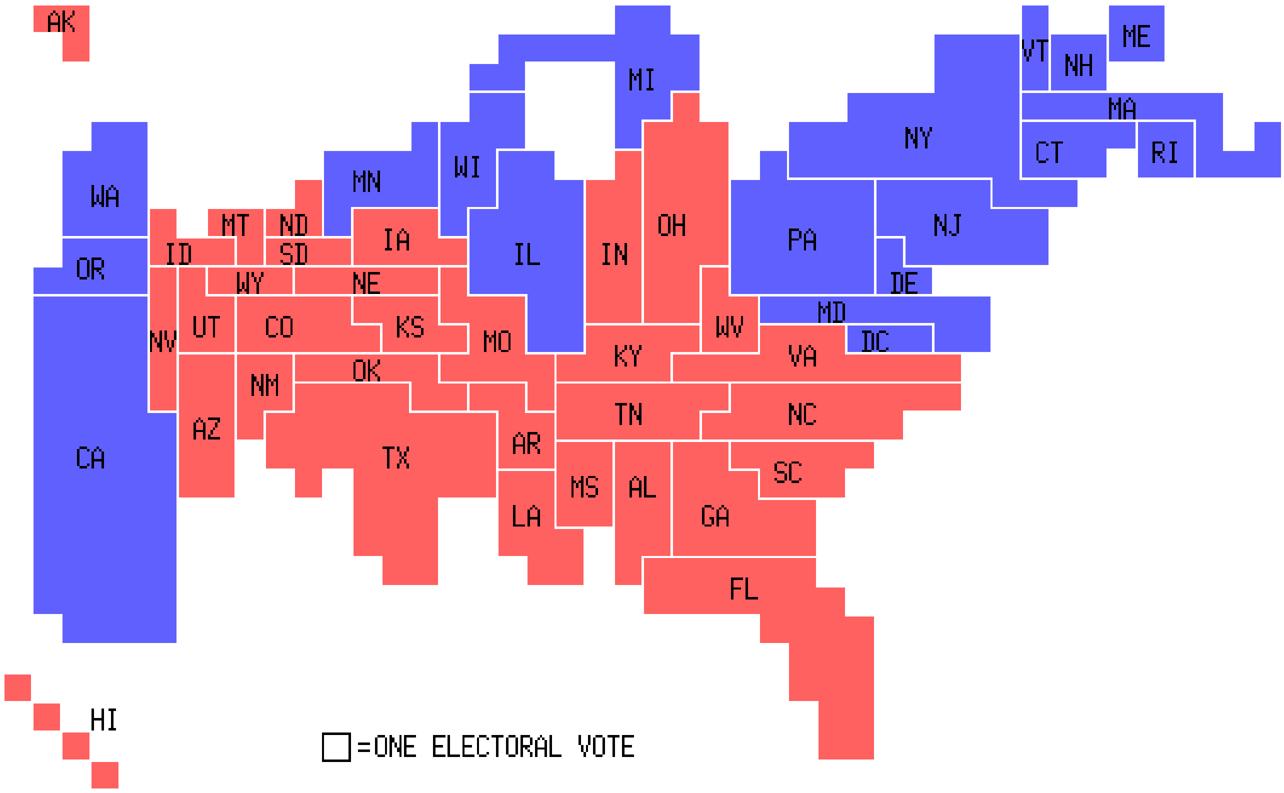

Figure 1. Election results predicted by the last month median method, with Bush states in red and Kerry states in blue. Note that each block represents one electoral vote.

vote gained by each candidate in that state. So in addition to tracking the median percentage of votes garnered by Kerry, Bush, and Nader, we also tracked whether that ballot initiative was passing or not in the median according to the several polls that were taken. If it was ahead, we allocated the votes according to that rule. As it turned out, most polls in Colorado by election eve, both overall, and in the last month showed the ballot initiative losing, so we noted that and used the winner-take-all system for Colorado. (This could be an important point in predicting future national elections as a similar bill splitting California's electoral votes has been mentioned as a possibility. Again, it would have to be followed in the median to see if it was passing.)

Maine and Nebraska award two electoral votes to the popular vote winner in the state, and the other electoral votes to the winner in each congressional district. Ideally we would have had polls to tell us who was ahead in each congressional district, but we did not. We only had statewide polls. So what we did was award two electoral votes to the candidate who was leading in the median. We had median state percentiles for each candidate. We also had the percentiles won by the Democrat and Republican in the last election (2000) by congressional district. So we took differences in the winning margins between each congressional district and the state total from 2000, and used those differences to correct the state median values from the 2004 polling data, to predict the median winning margin in each congressional district. It was the best we could do. In the end, both methods by election eve, predicted a sweep of electoral votes for each candidate in each state. But we would have split the electoral votes in these three states if it had looked like that would occur. It is important to follow the actual rules in each state.

In states where there were no polls in the last month, we used the most recent previous poll. In this case the idea was to get as close to last month's polls as possible. These few states were typically ones in which few polls were taken because the outcome was well decided, such as Arkansas where there were no polls in the last month and the most recent previous poll was a win for Bush with a margin of 27%.

5 RESULTS

Both the overall median and the median of the previous month's polls worked well. But the median of the past month's polls worked slightly better. We present the results from this (last month's median) method in both a map form (Figure 1), and in a table form (Table 7) as they appeared and as we published on WNC's website on Election Eve [12]. The table gives the number of wins for the Republican, Democrat, and the ties in each state. (A tie occurred when the poll said, for example, Bush 47%, Kerry 47%, and Nader 3%). Nader won none of the states; so it was either a Republican or a Democratic win or a tie for each of the polls. We predicted the state for the candidate who was ahead

Table 1. Subset of Table 7, showing only PA, OH and FL.

	Kerry Wins	Bush Wins	Ties	Predicted Winner
Pennsylvania	52	3	5 ties	Kerry
Ohio	14	39	3 ties	Bush
Florida	16	40	12 ties	Bush

in the median-that is had won most polls in the state. It quickly became apparent that the election hinged on the three big close states where the results were as shown in Table 1 (posted the morning of Election Day, 8:17 am, Nov. 2, 2004).

Kerry had a lead in the number of polls won in Pennsylvania straight through our entire run. Each day we would add the new polls taken that day, and drop the ones now over a month old. Kerry was always ahead in Pennsylvania. Similarly, Bush had an uninterrupted lead in both Ohio and Florida. These leads were substantial in terms of wins and losses each case. If you tossed a coin 56 times there was very little chance, according to the binomial distribution to get 40 or more heads and 16 or less tails, for example. Bush must be leading in Florida, or else why would he be winning the lion's share of the polls?

5.1 Media Rerpots

But what were the media saying all during this time? The news reports went something like this: "New Polls out today. In Ohio, Bush 47%, Kerry 45%, but the margin of error is 4% so Ohio is too close to call. In Florida, Bush 46%, Kerry 45% but the margin of error is 4% so it is too close to call. In Pennsylvania Kerry leads 48% to 45% but the margin of error is 4% so Pennsylvania is too close to call as well. Which ever candidate takes two or more of these three toss-up states is likely to win the election but all three states are too close to call. Bush campaigns in Pennsylvania today while Kerry will visit Ohio again." Every newscast was basically the same. They all said these three states were too close to call. Our instructions for people listening to the news were simple: ignore the margins of victory and the polls' margins of error; just note who was winning. To us that newscast sounded like: "Bush wins again in Ohio and Florida, while Kerry wins again in Pennsylvania. Those states remain locked up." We knew for over a month that Kerry would likely win Pennsylvania and Bush would likely win Ohio and Florida-and the election. Basically everyone in the country thought the election was too close to call (according to the polls) but at least the 300,000 or so people visiting our website had access to better information. If there were no systematic effects, Bush would win and Kerry would lose. On election eve (8:17am Nov 2), our final prediction (illustrated in Figure 1) showed Bush was leading with 290 electoral votes to Kerry's 248. Every day from early October when we had over a month's data, and started publishing our results on the web, Bush had on each and every day more than the required 270 electoral votes needed to win (see Table 2). On some days, a particular state might show as "too close to call" according to our data. Some states, such as New Hampshire, oscillated from the Democratic win column to "tied" and back, Hawaii from the Democratic column to the Republican column, and Wisconsin from Republican to tied, to Republican, to tied, to Democratic, but these were too small to turn the election. (Using all polls [back to July 7] the median results were similar. During the same period using all polls Kerry ranged from 247-257, while Bush ranged from 274-284. Bush was always above the 270 needed to win, with toss-ups ranging from 0-17) Although most people no doubt thought the election "too close to call" at least the users of our website knew how things really stood.

Curiously, the major media outlets never seemed to use averages of several polls or medians to get at the answers in the three big states. They were so concerned to get the up-to-date data that they were calling the states ties or seeing them switch from side to side as the latest polls came in. One particularly good website (Electoral Vote) had a colored map which called each state for the candidate who was leading in the latest poll. They even made a movie of this with time. Some states were constant, but big important states like Pennsylvania, Ohio, and Florida oscillated wildly. It made it look like America could not make up its mind, day to day. This website ended on election eve calling the entire election for Kerry because he had a last minute win in Florida (although in the total number of wins there he was still way behind). But these oscillations were just the margin of error of 4% in each state, randomly moving the polls back and forth. Integrate over a month to get enough precision to call the big states, and it was clear who was winning in each; the leads did not change at all.

5.2 How Could Kerry Win?

In fact, most states voted exactly the way they had 4 years earlier. Kerry's trouble was that if Bush simply carried each of the states he did 4 years earlier he would win the election, because the census of 2000 had actually added a number of electoral votes to the states he already carried. Kerry was unable to change this. To win, Kerry would have to steal away more than 20 electoral votes from Bush according to our predictions on election eve. His prospects for doing this were dim. First Kerry had to win all the states in which he was leading-he was likely to do this, but some were close enough to cause concern.

Now, there were a number of states that took only a few polls, such as Kentucky which had only 4 polls; but because Bush won all 4, and his median margin of victory was 20% no further polls were thought necessary. His victory there was assured. Fortunately, for our method, states that took only a few polls usually did so because the margins of victory were so dramatic that it was not worth the effort taking more

Table 2. Predictions as a function of time for total electoral votes based on the last month's polls, plus the actual electoral votes for the election at the end.

Date	Kerry E.V.	Bush E.V.	Toss-Up
Oct. 12	246	288	4
Oct. 13	246	288	4
Oct. 15	246	288	4
Oct. 20	250	288	0
Oct. 21	250	288	0
Oct. 22	253	285	0
Oct. 24	242	292	4
Oct. 25	243	295	0
Oct. 26	243	295	0
Oct. 27	243	295	0
Oct. 28	238	295	10
Oct. 29	238	290	10
Oct. 30	238	290	10
Oct. 31	248	290	0
Nov. 1	248	290	0
Nov. 2	248	290	0
Actual	**252**	**286**	**0**

Table 3. Best prospects for a Kerry "steal" vs. our election eve last month median predictions.

	Electoral Votes	Bush Wins	Kerry Wins	Ties	P(**Kerry Lead**)
Colorado	9	14	6	2	5%
Iowa	7	17	11	6	15%
Nevada	4	16	1	1	0.01%
New Mexico	5	12	6	1	11%

polls. In the extreme case where there was only one poll in a state our method would work no better (or worse) than just taking the last poll. Of course the median method still works when there are just a few polls, it is just not as powerful in that case. But since the states with only a few polls were usually states that were not close the median predicted the outcome correctly in the great majority of those states as well. Close states engendered a lot of interest, and many polls were taken, allowing one to call them even if they were close. An individual poll in Pennsylvania might have a margin of error of 4%, but with 60 polls there we improved the accuracy by a factor of $\sqrt{60}$. If Kerry was leading by over a half a percent we could detect it. If we ruled out Kerry steals of Bush-leading states where Kerry had never won a poll, that left only the prospects for him listed in Table 3. With each we give the probability that the true median actually put Kerry ahead given that Kerry had no more wins in the polls than he did, counting just wins and losses using the binomial distribution. Or equivalently, given our method, the probability of obtaining that big or bigger margin of number of polls wins by Bush, if Kerry were actually ever-so-slightly in the lead.

To win Kerry would have to win 21 or more of these electoral votes. He would have to win Colorado, Iowa, and New Mexico. The odds were stacked against him. The chance he was actually ahead in these three states was small. If he did not have these three states, and lost Nevada, even winning Hawaii (4 Electoral votes-where Bush had 1 win, no losses and 1 tie) would not help him.

As we observed the polls continuing in the median to strongly predict a Kerry loss, we wondered if Kerry and his team of advisors knew the election was lost. Near the end of the campaign, Kerry did act as if he knew about some of our results. He suddenly went to campaign in Iowa and Wisconsin which we found were both actually close (in Wisconsin, Kerry had 16 wins, 13 losses, and 2 ties-that was a close state where he was leading but had to defend). Kerry canceled a scheduled campaign trip to Missouri as Election Day approached. It had 11 electoral votes, but we showed that he had had 12 losses and no wins and no ties there in the last month, and Kerry also apparently did give up on it as hopeless.

5.3 The Electoral System and Concentrated Campaigning

Our observations highlight a perhaps unfortunate aspect of the electoral system. As campaigns become more and more sophisticated, they concentrate more and more effort on those few states that are actually close. Bush and Kerry spent an inordinate amount of time in Ohio, Pennsylvania, and Florida. For example, Bush and Kerry visited Ohio 52 times! Their efforts to sway voters in those three states by trips there either had no effect or canceled each other out, as there was no systematic change in the median results in those states with time. Except for fund-raising purposes, large states like Texas, California, Illinois, and New York, with their large populations and large number of electoral votes were almost completely neglected by the campaigns because the outcome in those states was basically already decided. Our results indicated that the campaigns could have been even more restricted in their trips. If the election were decided by popular vote, each

Table 4. Discrepancies between predictions by the last month median and (the median of all polls from July 7 on).

	Electoral Votes	Bush Wins	Kerry Wins	Ties	Predictions
Wisconsin	10	13 (25)	16 (23)	2 (4)	Kerry (tie)
Iowa	7	17 (26)	11 (22)	6 (8)	Bush (tie)
Hawaii	4	1 (1)	0 (2)	1 (1)	Bush (Kerry)
New Mexico	5	12 (13)	6 (15)	1 (2)	Bush (Kerry)

voter would get equal attention by the campaigns. With the electoral system and accurate and extensive polling, voters in close states get an inordinate amount of attention during the election campaign. Of course, a remaining advantage of the Electoral College is that if the popular vote were nearly a dead heat, there would have to be a recount of the entire country, rather than just recounting one state.

5.4 Last Month Median vs. Median of All Polls

The last month's data method which included by election eve only polls from Oct. 3 to Nov.2, and the overall median method which included all polls from July 7 to Nov 2 gave equal predictions in all states but four, and in all of those cases where they were in agreement both methods predicted the correct winner. What about the four states where the predictions were different?

In Wisconsin for all polls for example, Bush had 25 wins, 4 were tied, and Kerry had 23 wins. That's 52 polls. Rank them from the biggest Bush victory to the biggest Bush loss. The median is between the 26th and 27th highest poll for Bush. That is between two polls that tied. So the prediction was a tie, in other words, we said that state was too close to call. Iowa in the last month's polls, on the other hand had 17 Bush victories, 6 ties, and 11 Kerry wins. The median poll there was between a Bush win and a poll that was tied-that gave an advantage to Bush and so we called it for him. As you can tell, the binomial distribution told us that these states were all close, too close to call with 95% confidence for example. But we gave the best prediction we could make. We realized that since percentages in polls are given just to even percentiles there would be some ties (say 47% to 47%). There was often a real winner in these polls-they typically polled of order several hundred to a thousand people, but because their statistical accuracy was only 3% to 4% they rounded the numbers to the nearest percent-causing a number of ties, and preventing us from finding out which candidate actually won that poll (by however slight a margin). This caused us to call some ties-states too close to call. And we felt that was appropriate. Thus with all polls, Wisconsin and Iowa were both called ties even though Bush had slightly more wins in each. A number of the ties could have been slight Kerry victories that were rounded to ties, so if the median poll was actually in the tied region we felt justified in saying that state was too close to call.

New Mexico showed a slight edge for Kerry (15 wins, 13 losses, and 2 ties) in all polls, while for polls in the last month Kerry was behind with (6 wins, 12 losses, and 1 tie). Clearly a trend developed in the last month-the data previous to the last month had Kerry leading with 9 wins, 1 loss, and 1 tie. That was a significant shift. Here we felt the last month's data would prove better because a shift was clearly visible. Now Hawaii was an interesting case. This was a traditionally Democratic state and all polls showed a Kerry win there-he had two victories to Bush's 1, with one tie. So the median was between a Kerry win and a tie and we gave it to Kerry. Kerry's median margin of victory was 3.5%. Now the two Kerry wins were not in the last month. In the last month there was 1 Bush win and 1 tie. So the median was between a Bush win and a tie and we gave it to Bush. The median Bush margin of victory was 0.5% (in other words Bush won one poll by 1% and tied in the other to make the median equal to 0.5%—in such cases we took the average of the two polls adjacent to the median). Now obviously, this was close. We were quite nervous about the last month Hawaii prediction because of the unfortunately small number of polls (just 2). And Bush's margin of victory could not have been smaller. Plus, it is a traditionally Democratic state, and all polls showed a larger median Kerry win. However, these were the results of applying the median test over the last month in Hawaii.

How did it turn out? Wisconsin and Hawaii went for Kerry, Iowa and New Mexico went for Bush. In the case of Wisconsin, Iowa, and New Mexico, the last month's polls got all three right, whereas all polls would have made the first two too close to call and missed the third. (Although to be fair, New Mexico was decided by a razor thin margin.) In the case of Wisconsin and New Mexico, limiting the data to the last month appears to have successfully caught trends. Kerry did perform more strongly in Wisconsin during the last month, and Bush performed more strongly in New Mexico. Wisconsin is one case where the Kerry campaign (recall Kerry's last minute trip to Wisconsin) seems to have changed peoples minds and turned an early Bush lead into a Bush loss in that state. In Iowa, Bush already had a slight lead, and in the last month he lengthened his lead enough to overcome the few ties. In the end, we think it no accident that the last month's polls got these three states right. If two had really been toss-ups, and the other a Kerry win the chance of the last month's polls winning all three by chance would have been less than 12.5%. The last month's polls failed in Hawaii however, because of the small number of polls and their closeness. Using all polls would have been a better predictor for Hawaii.

Which method did better? How do we score success? Picking the Democrat or the Republican should be scored just like a true-false test. Score the number of correct answers minus the number of incorrect answers (in this case about individual electoral votes). That is the correct penalty for guessing. If one guessed about all electoral votes, one should get a score of zero. If one fails to call a state, you should get zero points for that state. You should be rewarded for predicting a state for the Democrat instead of the Republican only if your probability for getting it right was greater than 50%. If you think a state is a toss-up you should not be encouraged to guess. We are interested in predicting

Table 5. Comparison of median methods: median of all polls taken, and median of polls in the last month.

System	E.V. Correct	E.V. Missed	Score
All polls from July 7 – Nov. 2	516	5	511
Polls from Oct. 3 – Nov. 2	534	4	530

electoral votes-so the score should be number of electoral votes predicted correctly minus the number of electoral votes predicted incorrectly. If you fail to call a state, you do not get any points for that state-but you accumulate no minus points either. A perfect score is 538 electoral votes: all 538 electoral votes predicted correctly, no states not called, and no misses.

So although both methods performed very well, the last month's poll method performed better. It got Wisconsin, Iowa and New Mexico correct at the expense of missing Hawaii which had only 4 electoral votes. On the true-false test, the last months' method got a better score because it didn't leave two high point questions (10 points and 7 points) blank, while missing a 5 point question; rather it answered those correctly, while missing only a 4 point question. So we would regard the last month's polls method as better.

Would trend detecting algorithms have worked better? Significant trends might have been picked up in New Mexico making us go with only later (last month's) data, and there would not have been enough of a significant trend picked up in Hawaii to drop any data, so it would have been correct also. But, Wisconsin and Iowa would have remained ties, because there were no significant late trends at the 95% confidence level. So the score with trend spotting would have been 521. That's still less than simply limiting the data to the last month.

6 COMPARISON WITH OTHER FORECASTS

We found 36 other websites making electoral vote predictions. Plus, one could always simply predict that each state would just go as it had in the 2000 election. We have compiled all of these final predictions and present the results in Table 6. They are ranked by score in the table below, together with our two methods, making 39 in total. Of the 39 methods our last month's median came in 2nd out of 39. Only one website beat us: Presidential Elect. This was just one person's opinion—it got a perfect score, predicting all 50 states correctly. The Wall Street Journal did well, coming in 8th, one place ahead of our all polls median estimate. Following that was O'Reilly who missed New Hampshire, and failed to call Ohio. A number of media outlets, The Washington Post, CNN, CBS, USA Today, and Time got lower scores by being conservative and leaving a number of states (including at least one of the big three: Ohio, Pennsylvania, and Florida) as too close to call. Twenty-six of the 39 methods failed to call Ohio, Pennsylvania, and Florida correctly. Here the median method scored well in calling the big close states, as we had hoped. Electoral Vote, which had an excellent website that included a movie of polls with time, used only the most recent poll in each state (the ultimate website in trying to go with the most recent trend). It failed badly, coming in 35th out of 39. It predicted Kerry would win the election by predicting a win for him in Florida. His last minute poll win there did not overbalance the many poll losses he had had in that state. Here taking the median was superior to taking just the latest poll. Princeton (Wang) used the average of several recent polls, and factored in a hypothesized Democratic advantage among undecided voters—this failed badly by predicting Kerry would win the election; this prediction came in 36th out of 39. Predicting that each state would go the same way it did in 2000, got a score of 506, and came in 14th out of 39, by missing Iowa, New Hampshire, and New Mexico.

7 ELECTION NIGHT

Early on election day, exit polls had shown a Kerry lead in both Florida and Ohio, and many reports actually had him winning the election. This caused a brief drop in stock prices on Wall Street. By contrast, our median statistics had predicted a Bush victory for weeks.

These exit poll results were not presented by the networks. Instead, NBC and others just began to call states as the polls closed in each state (the final exit polls, by the way, eventually did match the outcome of the election by the time they all came in). Early in the evening, the states were falling exactly as predicted by the median statistics; Hawaii, our lone error, wouldn't be reported until early the next morning.

As election night wore on, NBC was slow to call Florida even with 97% of the vote counted and Bush leading 3,600,969 to 3,277,411, but eventually did Florida for Bush, at which point Bush led 246 to 206. Then at 12:59 pm NBC gave Ohio to Bush. Just after 1 am, they gave Bush Alaska, bringing Bush to 269, just one electoral vote short of victory. NBC to its credit had called no states incorrectly.

CNN, meanwhile, was changing its New Hampshire prediction for Bush to a win for Kerry (which was correct). That was the only state CNN had miscalled. In Ohio, Bush was leading 51% to 49% with 91% of the vote counted, but neither CNN, ABC, nor CBS had called Ohio. Then CNN did something remarkable—it called Ohio "too close to call"—and rather than coloring it in blue (Democratic) or red (Republican) it colored Ohio green. They had never done this before. They explained that the green meant they were projecting it would require recounts and perhaps court cases like Florida in the 2000 election. Our median statistics had projected a Bush victory in Ohio all along.

Fox, like NBC, had called Ohio for Bush, but the Kerry campaign would not concede Ohio, saying it would wait at least a day. NBC and Fox stood by their predictions on Ohio. Peter Jennings said goodnight for ABC news. Fox had called the election for Bush. The president issued an optimistic statement.

By the next morning, the Kerry campaign could see that it would not win, and shortly after 11 am Kerry conceded the election in a phone call to President Bush. According to CNN, during the last week of the campaign the Kerry campaign had thought it would win, and Kerry himself had thought he would win until about 6 or 7 or 8 o'clock election night. Even though the Kerry campaign had access to all the information we were seeing and more, our median statistics had told us that the election was lost to Bush for almost a month, with no changes significant enough to change the outcome occurring during that time. Did the Bush campaign feel it had the election sewn up all along? We don't know. But, on election afternoon, the Bush campaign argued that the exit polls did not reflect heavy Republican turnout, and by 9:30 pm on election night Bush had a solid lead in Florida and seemed confident.

8 CONCLUSIONS

Median statistics worked well to predict the outcome of the 2004 U.S. Presidential Election, and we look forward to applying them again in 2008. Using the median poll in the last month proved very effective at predicting the outcome in individual states. Just as we had originally hoped, this method made it possible to predict successfully the results in the large close states (Ohio, Pennsylvania, and Florida) where there were many polls. States with only a few polls were generally not close, and so the median poll predicted these states successfully as well, missing only Hawaii. The method appears particularly well adapted to U.S. Presidential elections where the candidates are chosen well in advance, and where outcomes in individual states determine the winner. Our advice to voters looking at polling results in the general Presidential election is simple. Ignore the margins of victory and the statistical uncertainty in the polls, and just notice who is racking up more victories in each state day by day. This would have told you the winners in Ohio, Pennsylvania, and Florida, and the winner of the election.

Our final prediction of Bush 290, Kerry 248, was only off by 4 electoral votes (Hawaii). Using the last month's polls seemed to provide a long enough time to allow correct predictions in the large close states, and yet a short enough time to correctly capture some trends that showed up in a few small states.

The 2004 election showed that there are surprisingly few states that are really up for grabs. Our results may be of interest to campaign staffs, telling them which states may actually be uncertain and therefore where they may best spend their campaign efforts. Most observers thought that Ohio, Pennsylvania, and Florida were toss-up states, but actually there were well-established, if small, leads in each of those states all along. The remaining few states that were uncertain did not have enough electoral votes to swing the election. It points up a weakness in the electoral system, that as campaigns become more sophisticated at using polls to predict outcomes (using our method, and others) more and more attention and money during the campaign will be spent on fewer and fewer states.

We used median statistics to protect against undue influence by discrepant polls in the 2004 U.S. Presidential election (as in an astronomical application, where we used median statistics to protect against discrepant mass-to-light ratios of groups of galaxies caused by background/foreground contamination). Wang of Princeton University used averages of recent polls (plus a hypothesized fraction of the undecided voters to vote Democratic) and erroneously forecast a Kerry victory. Using just the latest polls in each state, the Electoral Vote website also failed to call the winner, and simply predicting that each state would vote as it did in the previous election (2000) also did worse than our median method. Using the median instead of the average (as would be used in a standard χ^2 analysis) protects against the possible overinfluence of a single biased or erroneous poll. In a median analysis, one still must worry about any systematic effects or non-independence of different polls, but assumptions about normal distributions and knowing the errorbars can be dropped. We hope this paper will encourage further mathematical investigation into use of the median in similar situations.

REFERENCES

[1] Gott, J. R., and Turner, E. L., 1977, "Groups of Galaxies. III. Mass-To Ratios and Crossing Times", *Astrophysical Journal*, 213, 309–322.

[2] Gott, J. R., Vogeley, M. S., Podariu, S. & Ratra, B. 2001, "Median Statistics, H_0, and the Accelerating Universe," *Astrophysical Journal*, 549, 1–17.

[3] Riess, A.G., et al., 1998, "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant," *Astronomical Journal*, 116, 1009–1038.

[4] Hill, B. M. 1992, in *Bayesian Analysis in Statistics and Econometrics*, eds. P. K. Goel & N. S. Iyengar (New York: Springer-Verlag), p. 43.

[5] Cohen, J. E., 1995, *How Many People Can the Earth Support?*, W. W. Norton, New York.

[6] Ehrlich, P. R. and Ehrlich A. H., 1990, *The Population Explosion*, Simon and Schuster, New York.

[7] Gott, J. R., 1999, "The Copernican Principle and Human Survivability, Human Survivability in the 21st Century," *Transactions of the Royal Society of Canada*, Series IV, Volume IX, ed D.M. Hayne, University of Toronto Press, Toronto, 131–147.

[8] Gott, J. R., 2001, *Time Travel in Einstein's Universe*, Houghton Mifflin, Boston.

[9] Lutz, W., Sanderson, W., and Scherbov, S., 1997, "Doubling of world population unlikely," *Nature*, 387, 803–805.

[10] Black, D., 1958, *The Theory of Committees and Elections*, Cambridge University Press, Cambridge, p. 188.

[11] Electoral-Vote.com, 2004, http://www.electoral-vote.com/.

[12] Colley, W. N., & Gott, J. R., 2004, http://www.colleyrankings.com/election2004/.

Table 6. Comparison of election prediction methods.

Rank	Method	Record	Score	Missed	Uncalled
1	Presidential Elect	538 - 0	538		
2	Colley/Gott Last Month	534 - 4	530	HI	
3	Cold Hard Truth	528 - 10	518	WI	
4	Strategisphere	518 - 0	518		OH
5	Silvey	528 - 10	518	WI	
6	Federal Review	528 - 10	518	WI	
7	Calhoun	526 - 12	514	IA, NM	
8	Wall Street Journal	526 - 12	514	IA, NM	
9	Colley/Gott All	516 - 5	511	NM	IA, WI
10	O'Reilly	514 - 4	510	NH	OH
11	My Election	524 - 14	510	HI, WI	
12	Poll Booth	508 - 0	508		OH, WI
13	Daly	517 - 10	507	WI	HI, IA
14	2000 Election Results	522 - 16	506	IA, NH, NM	
15	Tripias	522 - 16	506	HI, IA, NM	
16	Election Projection	521 - 17	504	IA,WI	
17	Electoral Survey	494 - 0	494	FL,IA,WI	
18	Washington Post	486 - 0	486		IA, MN, NM, OH, WI
19	CNN	475 - 0	475		FL, IA, MN, NH, NM, WI
20	ArrowHead	475 - 0	475		CO,FL,IA,OH
21	Leip	506 - 32	474	IA, NM, OH	
22	Iomada	504 - 34	470	FL, IA	
23	New York Times	469 - 0	469		FL, IA, NM, OH, WI
24	BattleGround	499 - 34	465	FL, IA	NM
25	Sabato	510 - 37	464	FL, WI	
26	Search the Links	464 - 4	460	HI	CO,FL,NH,OH,WI
27	PollKatz	499 - 39	460	FL, IA, NM	
28	Mind the Gap	499 - 39	460	CO, OH, WI	
29	CBS	433 - 0	433		FL, IA, MI, MN, NV, NH, NM, OH, WI
30	Electoral Expectation	432 - 0	432		CO,FL,IA,MI,NM,OH,PA
31	National Journal	484 - 54	430	FL, IA, OH	
32	Cook	429 - 0	429		FL, IA, MN, NV, NH, NM, OH, PA, WI
33	Boston Globe	434 - 7	427	IA	CO,FL,NH,NM,OH,PA,WA
34	DC2	425 - 0	425		AR, CO, FL, MN, NV, NM, OH, PA, WI
35	Electoral Vote	475 - 58	417	FL, HI, IA,	OH
36	Princeton (Wang)	474 - 64	410	FL, IA, NV	NM,OH
37	USA Today	407 - 0	407		CO, FL, HI, IA, MI, MO, NM, OH
38	DC	422 - 20	402	OH	CO, IA, MI, MN, NH, NM, PA, VA, WI
39	Time	349 - 0	349		AZ, AR, CO, FL, IA, ME, MI, MN, MO NV, NH, NM, OH, OR, PA, WA, WV, WI

Table 7. Last month median results for all states as of Nov. 2, 2004.

State	Bush	Kerry	Ties	Prediction
AK	1	0	0	Bush
AL	5	0	0	Bush
AR	8	0	2	Bush
AZ	10	0	0	Bush
CA	0	9	0	Kerry
CO	14	6	2	Bush
CT	0	1	0	Kerry
DC	0	1	0	Kerry
DE	0	1	0	Kerry
FL	40	16	12	Bush
GA	11	0	0	Bush
HI	1	0	1	Bush
IA	17	11	6	Bush
ID	1	0	0	Bush
IL	0	8	0	Kerry
IN	6	0	0	Bush
KS	3	0	0	Bush
KY	4	0	0	Bush
LA	6	0	0	Bush
MA	0	1	0	Kerry
MD	0	7	0	Kerry
ME	0	4	0	Kerry
MI	5	42	5	Kerry
MN	16	30	6	Kerry
MO	12	0	0	Bush
MS	1	0	0	Bush
MT	2	0	0	Bush
NC	10	0	0	Bush
ND	1	0	0	Bush
NE	1	0	0	Bush
NH	2	10	4	Kerry
NJ	1	18	2	Kerry
NM	12	6	1	Bush
NV	16	1	1	Bush
NY	0	7	0	Kerry
OH	39	14	3	Bush
OK	8	0	0	Bush
OR	1	14	0	Kerry
PA	3	52	5	Kerry
RI	0	2	0	Kerry
SC	6	0	0	Bush
SD	3	0	0	Bush
TN	8	0	0	Bush
TX	3	0	0	Bush
UT	2	0	0	Bush
VA	8	0	0	Bush
VT	0	1	0	Kerry
WA	0	10	1	Kerry
WI	13	16	2	Kerry
WV	5	0	0	Bush
WY	1	0	0	Bush