



MEMORANDUM

TO: SCOTT POOL
FROM: BRYON ALLEN
SUBJECT: KENYAN PRESIDENTIAL EXIT POLL
DATE: JULY 17, 2008

This memo summarizes our analysis of the coding of the Kenyan Presidential Exit Poll as well as providing some context about possible sources of uncertainty and error which we cannot test for just by looking at the data set but which you should be aware of as you interpret the data.

The bottom line of our analysis is that there is no one source of error large enough to invalidate the poll. However, there were multiple small errors and issues that we found and legitimate concerns that questionnaire wording, non-response bias, and fielding difficulties could have caused further small biases in the data. While none of these would necessarily be large on its own, taken together they require us to use caution in interpreting the data from the poll.

Coding accuracy

There were some errors in the coding done in Kenya. At least half of the ballots had at least one miscoded question. On the other hand, many of these errors were in not correctly coding refused-to-answer responses. Not counting refused-to-answer issues, only around 20% of the questionnaires had miscodes.

On the presidential ballot, 6.6% of the responses were coded differently on the Kenyan file than in our re-coding. Approximately 1% (58 ballots) were refused-to-answer or “don’t know” responses that had been mis-coded. More of these refusals were coded as Odinga than Kibaki, but it does not seem to have been universally one way or another and because it represents such a small fraction of the ballots it did not change the final numbers.

Question Ordering:

As you and I have discussed, this poll did not follow what would be considered a conventional questionnaire design for an exit poll. Typically an exit poll will ask the ballot questions as either the first questions or the first questions after one or two demographic questions. Issue questions and other items intended to understand why voters made the choices they did are generally asked nearer the end of an exit poll.

My biggest concern with the questionnaire design is that Q3 was asked before the ballots. To my eye, Q3 reads like a message we might test in a pre-election poll. Since messages are obviously intended to change votes, it is possible or even likely that having a message before the ballot recall question could have biased the respondents reports of who they voted for.

Unfortunately since there were no questionnaires without this design, it is impossible to test whether or not it did actual bias the results.



Non-Response Bias

One of the issues that must be addressed in any exit polling situation is how to deal with voters who refuse to take the poll at all. If there are systematic differences in the rates of refusal among sub-groups (for example if men refuse at a higher rate than women) it can bias the results. The typical solution for this issue in exit polls and other in-person interviewing situations is to note some basic demographic characteristics (sex, race, approximate age) of those who refuse and then analyze that data to see if it could be a source of bias. If it is a potential source of bias, the data may be weighted to reflect the actual population of voters including those who refused to participate in the poll.

There is no evidence that the Kenyan data was weighted. This may be because it was not necessary but it could also indicate that no non-response analysis was completed.

Other Difficulties in Administration

Exit polling is difficult under the best circumstances. In recent history in the U.S. the exit poll sponsored by the major news network was completely abandoned in 2002 and had unexplainable divergences from the actual vote in both 2000 and 2004. Given the difficulties involved in executing the logistics of fielding an exit poll, even in a stable environment like the U.S., it is realistic to worry that there could also have been errors and issues in the fielding of the poll in a more volatile and complex environment like Kenya.

Again, this is not something we can test for just by looking at the data. We did examine the issue of some ballots appearing in the original tally that we did not receive and some ballots appearing in our tally which were not in the Kenyan data file. Sixty-two of these differences were due to differences in allocating some ballots to the oversample or the core data for the Nakuru and Mombasa data.

Some of the differences also appear to be the result of miscoding of serial numbers on the Kenyan file—there were multiple duplicate numbers on the Kenyan file which appear to have been miscodings of the serial number. However, even accounting for those causes it does appear there are a handful of ballots we did not receive from Kenya and a handful we received that were not coded in Kenya.

While the small number of ballots that may have been missed in the Kenyan coding was not enough to affect the outcome of the data in a large way, it is indicative of the complications that may have faced the administration of the poll.

