SEP. 12, 2011, AT 4:20 PM

'Keys to the White House' Historian Responds

By Nate Silver

Allan J. Lichtman, the presidential historian at American University whose "Keys to the White House" system we posted a critique of here last week, has kindly prepared a response. (Mr. Lichtman's system forecasts that President Obama is a heavy favorite to be re-elected in 2012.)

Mr. Lichtman's response is reproduced below. My policy is to let authors have their own say in these cases, so any further response I have will be contained in a separate post, with the exception of a brief technical comment at the end.

The Keys Work: A Response to Nate Silver
Allan J. Lichtman

I read with great interest Nate Silver's analysis in the New York Times blog of my system for predicting presidential election results, The Keys to the White House. His work demonstrates the kind of lively debate that the keys can generate. However, his critique of the keys system cannot withstand scrutiny.

Mr. Silver has either failed to read or simply ignored the detailed discussion of the keys in my book "The Keys to the White House" (5th edition forthcoming from Rowman & Littlefield early next year) or in technical articles that I published in peer reviewed journals (for example, International Journal Of Forecasting (April-June 2008) and International Journal Of Information Systems & Social Change (January-March 2010).

Mr. Silver's neglect of the published work on the keys system has resulted in fundamental misconceptions about the development and application of the keys that fatally flaw his critique. Mr. Silver presumes that I took 38 elections from 1860 to 2008 with known outcomes, found 13 key factors out of innumerable combinations of factors and used the impermissible methods of "overfitting and data dredging" to force the calls on these keys to conform to the results of these elections. In fact, the book and articles make clear that I followed a very different and scientifically sound procedure designed precisely to avoid the fallacies of "overfitting and data dredging."

I developed the keys in 1981 in collaboration with Volodia Keilis-Borok, a world renowned authority on prediction methods. The keys are based on a retrospective analysis of elections from 1860 (the beginning of the modern Republican vs. Democratic

era) to 1980 and using a theoretical model, not random data-mining. The theory behind the keys is that presidential elections are determined primarily by the performance of the party holding the White House. This is a very positive message: it suggests that the American electorate makes reasoned, pragmatic decisions in presidential elections and is not manipulated by the pollsters, the admen, and the consultants. It suggests that it is governing not campaigning that counts in presidential elections.

After developing the 13-key model, I then followed proper procedure by testing it repeatedly through advance prediction of seven future elections – 1984 to 2008 — with unknown outcomes. The keys system correctly predicted in advance the popular-vote outcome of each of the seven elections that occurred since its inception in 1981. The probability of attaining this record of seven consecutive correct predictions by chance alone is well under one in 100, a standard measure of statistical significance in social science.

I could not possibly have manipulated and wrenched the calls on the keys to achieve this successful record, since I did not have foreknowledge of the winner of the popular vote, much less the percentages gained by the major party candidates. Moreover, these seven correct predictions were usually made many months or even years prior to the election and often in defiance of the polls and the pundits. In May of 1988, for example, I used to keys to predict the election of Republican George H. W. Bush, when he trailed Democrat Mike Dukakis by double-digits in the polls and the pundits had written off Bush as a hopeless candidate. The table below reports the timing of each advance prediction and the publication cite. Mr. Silver does not even acknowledge in his critique, much less discuss, these seven correct advance predictions.

Keys To The White House: Timing of Predictions

Election Year	Date of Prediction	Source
1984	April 1982	"How to Bet in '84," Washingtonian Magazine, April 1982, 147-49.
1988	May 1988	"How to Bet in November," Washingtonian Magazine, May 1988, 115-24.
1992	Sept. 1992	"The Keys to the White House," Montgomery Journal, Sept. 14, 1992, 12.
1996	Oct. 1996	"The Keys to the White House: Who Will Be the Next American President?" Social Education, Oct.1999, 358-360.
2000	Nov. 1999	"The Keys to Election 2000" Social Education, Nov./Dec. 1999, 422-24.
2004	April 2003	"The Keys to the White House," Montgomery Gazette, April 25, 2003, 4.
2004	March 2006	"The Keys to the White House: Forecast for 2008," Foresight: The International Journal of Applied Forecasting (Feb. 2006), 5-9.

Attention to the material in my book and articles would also have assuaged much of Mr. Silver's worries about the subjectivity of some keys. A degree of judgment is required to answer some of the key questions, because the real world cannot accurately be captured by so-called "objective" questions alone. However, the amount of subjectivity is far less than meets the eye, given the careful definition of each key in the published material, the record of how each key was turned in the 38 elections from 1860 to 2008, and the successful predictions from 1984 to 2008 – elasticity on calling the keys cannot explain the correct prediction of elections with unknown outcomes.

For example, Mr. Silver charges that my failure to count Republican candidate John McCain as charismatic was arbitrary and subjective because "Mr. McCain, with his service in the Navy, might have met Mr. Lichtman's description of a 'national hero.'" In fact, the published definition of the challenger charisma/hero key rules out counting McCain as a national hero since he did not lead the nation through war, however admirable his record in war may have been. In first edition of the Keys to the White House published many years before the 2008 election, I wrote that candidates attain heroic stature only "through vital leadership in war like Ulysses Grant and Dwight Eisenhower." I noted that "many other candidates like William McKinley, George McGovern, and George Bush have impressive military records but have fallen far short of the heroic status obtained by Grant and Eisenhower." Clearly, McCain falls into the former, not the latter category.

Mr. Silver also takes me to task for crediting President Obama with the policy change key, even though his "initiatives, especially the health care bill, are rather unpopular." However, the definition of the policy change key refers only to the magnitude of the change and the departure from prior policy. The definition of the key does not include the "popularity" of policy initiatives, which is extremely difficult to measure accurately. Thus, the call on the policy change key is fully consistent with the keys system. Mr. Silver may not like the definition of the key, but that is quite different from objecting to the call on the key.

Mr. Silver also claims that the keys are flawed because of inaccuracies in predicting the percentage of the vote received by the incumbent party. However, the keys were not designed to predict vote percentages. They were designed to forecast whether the incumbent or challenging party will prevail in the popular vote, regardless of the margin of victory. Mr. Silver's critique of the keys system is akin to critiquing a pregnancy test, not for its failure to detect pregnancies, but for its failure to determine the day of conception.

The keys are a purely binary system and do not include percentages of the vote share. The dependent variable is a vector of 0s and 1s, indicating whether the incumbent party won or lost the popular vote. The independent variables are vectors of 0s and 1s, indicating whether each of the 13 keys favored the incumbent or the challenging party. Thus the keys are not designed to estimate percentages, but only popular vote winners and losers. Obviously, if I had sought to estimate percentages, I would have included such measures in the dependent variable. To reason backwards from percentage estimates to the likelihood of correctly identifying winning candidates is a flawed approach to evaluating the keys system.

However, it is possible purely as a by-product of the system to use the number of keys falling against the party in power to estimate the percentage of the two-party vote achieved by the incumbent party candidate. I developed such a formula in the two recently published articles cited above and applied it to both the retrospective predictions from 1860 to 1980 and the advance predictions from 1984 to 2008.

The formula that I developed and published comes much closer to replicating actual results than the one used in Mr. Silver's analysis. (My formula predicts the two-party percentage for the incumbent party and Mr. Silver's predicts the margin of victory, however, the two can be made comparable by converting the two-party percentage into a victory margin). For example, Mr. Silver takes me to task for vastly underestimating Herbert Hoover's landslide defeat in the presidential election of 1932. In fact, the correct prediction using the published formula of $V = 36.75 + 1.84L$, where V = the percentage of the two-party split going to the incumbent and L = the number of Keys favoring the incumbent party, is 46 percent, just 5 percentage points higher than his actual percentage of 41 percent and a clear loss for the incumbent party. It is true that the predictions on the keys flatten at the extremes and the greatest numerical errors are for landslide victories like 1924, 1932 and 1972. However, such errors are of absolutely no consequence in recognizing the popular vote winner of elections. In every instance of a landslide election, the keys show a very substantial victory by the winning candidate. Indeed, all of Silver's examples of supposedly large errors in numerical predictions come from landslide elections.

The keys are extremely accurate when it matters for relatively close elections, including elections predicted in advance. For the seven elections with advance predictions, the mean absolute error is only 1.9 percent, despite predictions made months and years in advance. In 2008, the system predicted the actual incumbent party percentage with an error of 0.3 percent. (see the Table below) By the logic of Mr. Silver's critique, he

somehow would have had more confidence in the system if it had produced one to two errors (20 percent) in these advance predictions as his flawed table of probabilities would suggest.

The Keys to the White House, Advance Predictions Only, 1984-2008

Year	Predicted Result	Actual Result	Difference
1984	59.2	57.0	-2.2
1988	53.9	55.2	1.3
1992	46.6	49.6	3.0
1996	54.7	51.5	-3.2
2000	50.3	51.5	1.2
2004	51.2	53.3	2.1
2008	46.3	46.0	0.3
Absolute Mean Diff			1.9

Overall, for all 38 elections, including the landslides, the 95 percent confidence level for predicting the two-party percentage for the incumbent party is 6.3 percent (much of it generated by the landslide elections), much lower than Mr. Silver's calculations produced. Only two elections (6 percent of all elections), 1912, when the Republican Party split, and 1924, when the Democratic Party split fall outside this range. Although I reject Mr. Silver's method of assessing the keys through numerical predictions that it was not designed to generate, this much tighter error band (about 12.6 percent, not 16 percent when converted to margins of victory or defeat) would produce much higher probabilities of correct predictions than those in Mr. Silver's table of probabilities – President Obama's chance of victory, for example, based on the current assessment would be close to 95 percent (depending on the call on the election year recession key), not 79 percent as Mr. Silver's invalid table of probabilities would indicate.

Finally, Mr. Silver suggests that including only two economic keys in the 13 key system understates the importance of the economy in determining electoral outcomes. Again, this critique is covered in the book and articles, which note that the economic keys often have trigger effects on turning other keys. For example, in 1932 the sour economy led to losses in the midterm election, social unrest, and the emergence of a charismatic challenging candidate (FDR).

Mr. Silver suggests that his own research shows that economic factors account "for about half of a voters' decisions." I would ask only, based on this research, how many elections

he has correctly forecast in advance of Election Day and how far in advance he published his predictions.

—

Thank you to Mr. Lichtman for his response.

The technical matter: there's some dispute here about the margin of error associated with Mr. Lichtman's formula when applied to predict the incumbent party's margin in the popular vote.

Some of this, as Mr. Lichtman alludes to, is because his formula is based on predicting the vote *share* received by the incumbent party's candidate (disregarding votes for third-party candidates) whereas mine is based on the *margin* separating the Democratic and Republican candidates (but not removing third-party candidates from the denominator). Errors calculated using the margin rather than the vote share will generally appear to be about twice as high: if you predict the election to be a 50-50 tie and the Republican candidate wins by 4 points, 52-48, you will have missed the Republican's vote share by 2 points but his margin of victory by 4 points. But they're basically measuring the same thing.

Even after accounting for that, however, there is still a modest discrepancy in the margins of error that we respectively identify.

I've double-checked my data and can't find any errors, so I've simply re-produced it, per the table below, for others to peruse.

The margins of victory in my figures should match those as recorded in Dave Leip's Atlas of U.S. Presidential Elections. Note that there are a few ambiguous elections in the sample — like 1868, 1872 and 1912 — that scholars, can handle differently, so the assumptions I've made about each one are also described in the table.

Year	Keys Favoring Incumbent Party	Incumbent Party Popular Vote Margin	Year	Keys Favoring Incumbent Party	Incumbent Party Popular Vote Margin
1860	6	-10.1	1936	12	+24.3
1864	10	+10.1	1940	11	+10.0
1868*	11	+5.3	1944	11	+7.5
1872*	10	+11.8	1948	8	+4.5
1876	4	-3.0	1952	5	-10.9
1880	9	+0.1	1956	12	+15.4
1884	6	-0.6	1960	4	-0.2
1888	8	+0.8	1964	10	+22.6
1892	7	-3.0	1968	5	-0.7
1896	5	-4.3	1972	9	+23.2
1900	10	+6.1	1976	5	-2.1
1904	13	+18.8	1980	5	-9.7
1908	10	+8.5	1984	11	+18.2
1912*	7	-18.7	1988	10	+7.7
1916	10	+3.1	1992	7	-5.6
1920	5	-26.2	1996	8	+8.5
1924	9	+25.2	2000	8	+0.5
1928	10	+17.4	2004	9	+2.4
1932	5	-17.8	2008	4	-7.3

Notes:

1868: Republican Party treated as incumbent party although the party affiliation of Andrew Johnson, who succeeded Lincoln, is ambiguous.

1872: Horace Greeley, formally of the Liberal Republican party but endorsed by Democrats, is treated as the Democratic nominee.

1912: Margin separating the major-party candidates, Taft (Republican incumbent) and Wilson (Democrat). Taft finished in third behind both Wilson and Teddy Roosevelt.