

The New York Times

FiveThirtyEight

Nate Silver's Political Calculus

Measuring the Effect of the Economy on Elections

By Nate Silver July 5, 2012 2:36 pm

The most significant change to our presidential forecast model this year is that it contains an economic index, which is used to guide forecasts along with the polls.

In fact, as you may have seen since we began our short daily summaries of the model's output, new economic data often has just as much influence over the forecast as the latest poll from Ohio or Florida.

I have some fairly strong views about the right way to use economic data in a forecasting model like this one. This is fundamentally a very challenging problem because there have been only 16 presidential elections since World War II, and yet there are dozens and dozens of plausible economic variables to pick from. (The Federal Reserve's Web site, in fact, now publishes about 45,000 economic statistics.)

The historical evidence is robust enough to say that economic performance almost certainly matters at least somewhat, and that poorer economic performance tends to hurt the incumbent party's presidential candidate. Likewise, it seems clear that the trend in performance matters more than the absolute level — otherwise, Franklin D. Roosevelt would not have been re-elected easily with an unemployment rate well into the double digits (although rapidly declining) in 1936.

But we just do not have anywhere near enough data to make confident claims about exactly *which* economic variables are important. For that matter, most of the more obvious choices for economic variables have performed about as well as one

another on the historical data anyway. Each one gets some elections right and some wrong.

Let me explain some of the choices I made about the model in light of this problem.

First, I wanted a composite economic index rather than picking just one or two variables.

If it is hard to tell exactly which economic variables are most important to elections, it seems far better to take some kind of aggregate or average of them rather than arbitrarily picking one. As you will see later, different economic variables this year would give you a radically different outlook on how likely President Obama is to be re-elected.

One of the reasons that some of the presidential forecasting models published by academics have not performed all that well is because they have not adopted this consensus approach. Instead, they will use just one or two variables. Yet rarely do they use exactly the same variable, and sometimes the choice is something fairly exotic rather than commonly-used measures like jobs growth, inflation, or gross domestic product.

In some of these models, it seems evident that the modeler has searched through hundreds of different model specifications to get the best 'fit' on past data. When dealing with noisy data and a limited number of observations, however, that can be a problem, because it will introduce significant selection bias or overfitting. Over a small sample size, it is inevitable that some variables will have performed much better than others because of luck alone. If you calibrate the model based on the lucky variables, their luck will eventually run out and your model will not be as accurate as claimed when used to make real predictions.

In addition, the American public experiences many different facets of the economy in many different ways.

It's a basic rule of sound forecasting that data subject to measurement or modeling error becomes more robust if you average or aggregate it together.

Typically, this averaging or aggregation process reduces forecast error by about 20 percent.

Second, I wanted to use relatively broad-based measures of economic performance rather than narrowly tailored ones..

I wanted to look at the sorts of measures that investors and economists might weigh most heavily in gauging the performance of the economy, and I wanted an index that makes good economic sense, rather than being cherry-picked to fit elections data in particular.

I also wanted to pick variables that reflected different aspects of economic activity without double counting them, although there is certainly some overlap in the ground the variables that I chose cover.

Third, I wanted data that is updated regularly — monthly or more often.

If you did want to use just one variable, then gross domestic product might be a reasonable choice, being the broadest-based measure of economic activity in the United States. (Although G.D.P. still contains plenty of noise, like the inventories adjustment.) The problem is that G.D.P. is updated only once per quarter, and then with a significant lag. If it becomes clear in July that the nation is experiencing an economic collapse, it doesn't make much sense to have to wait until late October (when third-quarter G.D.P. is finally reported) to have that reflected in an economic model.

Fourth, and related to this goal of building a model that could make realistic forecasts in real time, I wanted to use data as it was *initially reported* during election years to calibrate the model, and the data as it was revised after the fact.

Most economic data series are subject to revisions; the process can persist for months or even years after the fact. Sometimes, these can be very severe — turning a quarter that was originally thought to provide average growth in to a recession, or vice versa.

But the magnitude of the revisions — and therefore, the reliability of a data point based on its initial print — can vary a lot from indicator to indicator. Some series are revised much more than others, and a few even have a history of biased revisions (meaning that the revisions usually tend to go in one direction).

This problem gets too little attention, in my view. I will sometimes visit macroeconomic forecasting Web sites where economists and investors get into detailed debates about which variables are more lagging and leading. But usually they are arguing about revised data, *all* of which is lagging in the sense that it is not available to forecasters seeking to make predictions about the economy in real time. The lack of sensitivity to these data-quality issues may hinder economic planning and is one reason that economic forecasts are often much less accurate than advertised.

The initially reported data, meanwhile, is also what is available to candidates and voters at the time of the election. Some econometric models score the 1992 election as a “miss,” because revised economic data shows roughly average growth during that year, when the incumbent, the elder President Bush, was defeated. However, the data was still quite poor as it was reported during 1992 itself. The initially reported data represents a closer approximation of what voters would have been weighing at the time.

Fortunately, the Federal Reserve is doing a better and better job of making archived economic data available through its Alfred Web site. Not all 45,000 variables are archived, but most of the major variables have reasonably good coverage, especially from the mid-1960s onward.

What I settled upon is a series of seven variables. The variables are weighed equally in the model, with one slight exception. All are relatively broad based, and are available in archived form going back to at least the election of 1968.

The first four variables are among the monthly indicators that economists most commonly use to help date recessions.

Nonfarm payrolls. This is the jobs figure that is commonly reported in news accounts, as in “100,000 jobs were added last month.” I prefer this figure to a

calculation based on the unemployment rate, which comes from a separate survey but which is subject to larger measurement error.

Personal income. Many academic election models use this variable or close cousins of it, like disposable personal income. I use the personal-income version because the archived record for it is more complete.

In theory, this variable has a lot of merit, since it reflects the different income streams coming to voters. A pay increase at work will be reflected. So would things like stock dividends or rental income.

In practice, however, measuring all these different income streams is challenging for the government. So this variable can fluctuate wildly from month to month and is subject to severe revisions. (These revisions, moreover, have been upwardly biased in the past, meaning that the government initially tended to underestimate income as measured in this way.) This variable can also be sensitive to changes in government policy, like stimulus payments, which voters may not react to in the same way as other types of income.

Still, this variable is certainly useful as long as it is not treated as some sort of magic bullet, and it is included in our index.

Industrial production. This is the granddaddy of economic variables — the government has kept track of it since 1919. Industrial production is the government's broadest measure of activity in manufacturing and related fields like mining. It is generally timed well to the business cycle, or can sometimes slightly lead it as it can reflect businesses' estimates of consumer demand for durable goods in the near future. It is also subject to fewer revisions than many other data series.

Personal consumption expenditures This measures household consumption of all kinds and goods services, which represents about 70 percent of gross domestic product. This variable is often strongly correlated to consumer confidence. But it arguably provides for a more tangible measure of the consumer, as it reflects how they are actually behaving with their dollars.

Inflation. Inflation, as measured through the Consumer Price Index, is the fifth economic variable in the model. Some of the academic models include measures of inflation and some do not. I think the case for doing so is reasonably clear. Inflation is among the most visible economic measures, and is among the most central in setting policy. It has also had a strong correlation with presidential approval ratings in the United States and has had a strong correlation with election outcomes in other countries.

The relationship between the inflation rate during the election year and election outcomes in the United States has been somewhat weaker, although when inflation is high — like in 1980 or in many elections before World War II — it usually has meant trouble for the incumbent president. Having one measure of inflation, as compared to six variables that measure growth, seems like a reasonable compromise.

Unlike for the other variables, higher inflation is worse for the incumbent, while lower inflation is better; the model, of course, considers this. However, our version of the variable gives a president no additional credit if the inflation rate is *below* 2 percent, since having a small amount of inflation is considered ideal by the Federal Reserve. A president gets no “extra credit” for deflation or near-deflation, in other words.

I get a lot of e-mail and Twitter questions about whether gas prices are included in the model. This is where they fit in, since gas prices will be reflected in the Consumer Price Index. Gas prices, of course, can also have indirect effects on variables like consumption.

The last two variables are forward looking.

Forecasted G.D.P. The model uses the forecast of gross domestic product growth over the leading two economic quarters (that is, *not* counting the current quarter) as taken from the median of The Wall Street Journal’s monthly forecasting panel. Right now, these forecasts continue to point toward fairly sluggish growth.

As I mentioned, I do not think it is wise to tweak your model endlessly based on fitting the past data in cases like this where the sample size is limited and the past

data is very noisy. But I did look to see the relative value of current economic measures against forward looking ones.

We found that the current economic measures seem to have more value — perhaps, in part, because economic forecasting is a very rough science. But the forecasts do seem to provide some value if used in moderation.

Stock Market. Likewise, I found that the stock market, as measured by the S&P 500, probably provides some value if used carefully.

The stock market has some unique virtues as a forecasting variable. It is available almost literally instantly, so something like relatively favorable resolution to the European debt summit — which might reduce the economic downside case for the United States in the second half of the year — can be reflected in the forecast almost literally overnight. And the stock market is not subject to revisions of any kind.

The downside is that sometimes the stock market shifts for reasons that have less to do with macroeconomic performance. Shifts in sentiment about Federal Reserve policy, for instance, can influence the market. And sometimes the movement in stock market prices may simply be irrational.

Still, all economic variables have their problems, and our finding is that a small dose of stock market data probably provides some useful information to elections forecasters. Moreover, although the market certainly gets things wrong some of the time, changes in the stock market tend to anticipate changes in several of the other six variables. The idea, in other words, is not so much that the stock market is tremendously important to American voters unto itself (although a fair number of Americans have investments, and the stock market is quite visible and widely-reported upon), but that it provides some signal about how economic conditions are evolving.

As I mentioned, the seven variables are weighted equally, with the exception of the stock market which is weighted slightly less. Specifically, the stock market represents 10 percent of the total index, whereas the other six variables each

represent 15 percent of it. Still, since the stock market is updated daily and can change quickly, it may have a relatively noticeable effect on the forecast.

Why use these largely equal weights? There is support for doing so in the empirical literature on forecasting in cases where the sample size is small and the data is very noisy.

The alternative would be to set the weights by regression analysis, but the data is just not robust enough to do this. You wind up with a big mess when you try to test the relative importance of seven economic variables on 10 or 15 past elections.

Arbitrarily dropping variables may produce a cleaner-looking result, but they do not really prevent overfitting and are the equivalent of putting lipstick on a pig. We think accounting for a larger number of variables, but weighing them about equally, is the better choice.

There are a few other details that need to be resolved. The most important one is what time frame to use in evaluating the data series.

One of the points that the historical record is fairly clear on is that voters have a fairly near-term focus when it comes to the economy. Almost no matter what variable you use or how you specify your model, there just is not much evidence that the economic performance of a president in his first two years in office matters much by the time you get to Election Day. (Some statistical models, in fact — although I do not personally find them all that plausible — even claim that poor economic performance during a president's first two years in office may *help* him.)

With that said, one cannot get too cute about this. Economic data, as we have mentioned, is quite noisy — certainly from month to month and even from quarter to quarter. And if voters are forward looking in their outlook, they can still take some time to parse the changes in the data (as can the news media and economists).

I decided to focus on economic data that looks at the performance of the economy over roughly the past half-year — although the implementation of this gets a bit more complicated to work around another thorny complication in economic modeling.

Below is a chart showing the progression of real personal income over 2004 and 2005. Note that there is a big spike in one month, December 2004, which coincided with a large one-time dividend payment made by Microsoft.

In addition to speaking to how personal income can be a problematic data series to begin with, this spike would also create problems for you if you were trying to measure the *change* in personal income at some later date. For instance, if your model was based at looking at the change in personal income over six months, it would look like something very bad had happened in June 2005, since this would be exactly six-months after the one-time spike. Then in July 2005, the index would suddenly appear to be rising at a healthy clip again.

The solution is to make these comparisons on a rolling basis. Our model calculates the change in each variable at intervals ranging from one month to one year. In other words, it calculates the one-month change in the variable, the two-month change, the three-month change, and so forth, then averages these results together after normalizing them to have the same standard deviation. The result is generally similar to the six-month change, but more robust to short-term blips in the data.

A slight variation is used in the procedure for the stock market. Because the stock market is not subject to measurement error, the model just uses the close of the S&P 500 from the most recent trading day. It then calculates the growth rate in the stock market by comparing this value to that on each of the past 252 trading days on a rolling basis as described above. (Why 252 trading days? That's the average number of days that the stock market is open during a calendar year, excluding holidays, weekends, and so forth.)

In addition, since all the economic measures are on different scales, they are normalized such that they have the same mean and standard deviation. For purposes of legibility, the index is then scaled such that it has the same mean and standard deviation as quarterly changes in gross domestic product. So a value in the low-to-mid 3's represents an economy that is growing at an average rate, while readings at 0 or below are recessionary.

Right now, the different components of the index read as follows.

(Keep in mind that these are *normalized* rather than raw values. Moreover, as I mentioned, the value for inflation is inverted before being normalized, since lower inflation tends to help incumbents.)

Right now, the index reads at 2.5 percent, which means that the economy looked at as a whole is clearly below-average, but not recessionary. There are some bright spots in the data, like the relatively strong rate of industrial activity, very low inflation (in part because of declining gas prices), and a stock market that signals the possibility for a more favorable flow of data in the second half of the year.

On the other hand, growth in income has been very poor — consistent with recessionary conditions, in fact — and consumption has been sluggish. And if the stock market has been bullish lately, the G.D.P. forecasts put out by economists are not so much. The Wall Street Journal panel expects G.D.P. growth of 2.3 or 2.4 percent in the second half of the year, which our model translates to a figure of 1.7 percent after normalizing it.

The normalized value for jobs growth (2.6 percent), interestingly, is right in the middle of these bearish and bullish indicators and almost exactly matches our economic index as a whole. This is another reason this Friday's jobs report is especially interesting; the jobs figures have often seemed to be the "swing vote" in determining whether the economy is getting back to about average growth or slowing down yet again.

With that said, part of the value of building an economic index is that it allows you to avoid getting overly fixated on any one data series, or any one data point. We think we have developed a reasonably well-balanced measure of the economy, which is fairly resistant to noisy data, and which reflects the different types of economic activity that voters will encounter.

In addition, our focus on using real time rather than revised data provides for a more apples-to-apples comparison to previous years. Here is how the economic index would have looked in past elections since 1968, at a range of intervals up to 250 days before the election.

In general, values of the index below 2.0 point to cases where the incumbent president will actually have become the underdog. The economy was quite terrible by the point in 2008 and getting worse, and it was a complete disaster in 1980. Meanwhile, in 1992 when Mr. Bush lost, the economy would have looked quite poor to voters based on real-time data. With the economic index now at 2.5 percent, Mr. Obama is just above this break-even point, but not by much.

© 2017 The New York Times Company