1 It's the Economy, Stupid

Election night at midnight:
Boy Bryan's defeat.
Defeat of western silver.
Defeat of the wheat.
Victory of letterfiles
And plutocrats in miles
With dollar signs upon their coats,
Diamond watchchains on their vests and spats on their feet.

Vachel Lindsay, from Bryan, Bryan, Bryan, Bryan

A common pastime in the United States every four years is predicting presidential elections. Polls are taken almost daily in the year of an election, and there are Web sites that allow betting on elections.

Some of the more interesting footnotes to presidential elections concern large errors that were made in predicting who would win. In 1936, the *Literary Digest* predicted a victory for Republican Alfred Landon over Democrat Franklin Roosevelt by a fairly large margin, when in fact Roosevelt won election to a second term by a landslide. The *Literary Digest* polled more than 2 million people, so the sample size was huge, but the sample was selected from telephone directories and automobile registrations, which overrepresented wealthy and urban voters, more of whom supported Landon. In addition, the response rate was higher for voters who supported Landon. The *Literary Digest* never really recovered from this error, and it ceased publication in 1938.

Another famous error was made by the *Chicago Tribune* in 1948, when it ran the headline "Dewey Wins." After it became clear that Thomas Dewey had lost, a smiling Harry Truman was photographed holding up the headline.

A more recent large error was made in June 1988, when most polls were predicting Michael Dukakis beating George Bush by about

17 percentage points. A few weeks later, the polls began predicting a Bush victory, which turned out to be correct.

While interesting in their own right, polls are limited in helping us understand what motivates people to vote the way they do. Most polls simply ask people their voting plans, not how or why they arrived at these plans. We must go beyond simple polling results to learn about the factors that influence voting behavior. This is where tools of the social sciences and statistics can be of help.

A Theory of Voting Behavior

To examine the question of why people vote the way they do, we begin with a theory. Consider a person entering a voting booth and deciding which lever to pull for president. Some people are dyed-in-the-wool Republicans and always vote for the Republican candidate. Conversely, others are dyed-in-the-wool Democrats and always vote Democratic. For some, one issue, such as abortion or gun control, dominates all others, and they always vote for the candidate on their side of the issue. For these people, there is not much to explain. One could try to explain why someone became a staunch Republican or Democrat or focused on only one issue, but this is not the main concern here. Of concern here are all the other voters, whom we will call *swing voters*. Swing voters are those without strong ideological ties and whose views about which party to vote for may change from election to election. For example, Missouri is considered a swing state (that is, a state with many swing voters). It sometimes votes Democratic and sometimes Republican. Massachusetts, on the other hand, almost always votes Democratic, regardless of the state of the economy or anything else, and Idaho almost always votes Republican. The percentage of swing voters in these two states is much smaller than the percentage in Missouri.

What do swing voters think about when they enter the booth? One theory is that swing voters think about how well off financially they expect to be in the future under each candidate and vote for the candidate under whom they expect to be better off. If they expect that their financial situation will be better off under the Democratic candidate, they vote for him or her; otherwise, they vote for the Republican candidate. This is the theory that people "vote their pocketbooks." The theory need not pertain to all voters, but to be of quantitative interest, it must pertain to a fairly large number.

The Clinton presidential campaign in 1992 seemed aware that there may be something to this theory. In campaign headquarters in Little Rock, Arkansas, James Carville hung up a sign that said, "It's the Economy, Stupid"—hence the title of this chapter.

The theory as presented so far is hard to test because we do not generally observe voters' expectations of their future well-being. We must add to the theory an assumption about what influences these expectations. We will assume that the recent performance of the economy at the time of the election influences voters' expectations of their own future well-being. If the economy has been doing well, voters take this as a positive sign about their future well-being under the incumbent party. Conversely, if the economy has been doing poorly, voters take this as a negative sign. The theory and this assumption then imply that when the economy is doing well, voters tend to vote for the incumbent party, and when the economy is doing poorly, they tend to vote for the opposition party.

We now have something that can be tested. Does the incumbent party tend to do well when the economy is good and poorly when the economy is bad? Let's begin by taking as the measure of how the economy is doing the growth rate of output per person (real per capita gross domestic product [GDP]) in the year of the election. Let's also use as the measure of how well the incumbent party does in an election the percentage of the two-party vote that it receives. For example, in 1996, the growth rate was 3.3 percent, and the incumbent party candidate (Bill Clinton) got 54.7 percent of the combined Democratic and Republican vote (and won—over Bob Dole). In 2008, the growth rate was –2.3 percent, and the incumbent party candidate (John McCain) got 46.3 percent of the combined two-party vote (and lost—to Barack Obama). (The reasons for using the two-party vote will be explained later.)

Figure 1-1 is a graph of the incumbent party vote share plotted against the growth rate for the 24 elections between 1916 and 2008. The incumbent party vote share is on the vertical axis, and the growth rate is on the horizontal axis. According to the above theory, there should be a positive relationship between the two: when the growth rate is high, the vote share should be high, and vice versa.

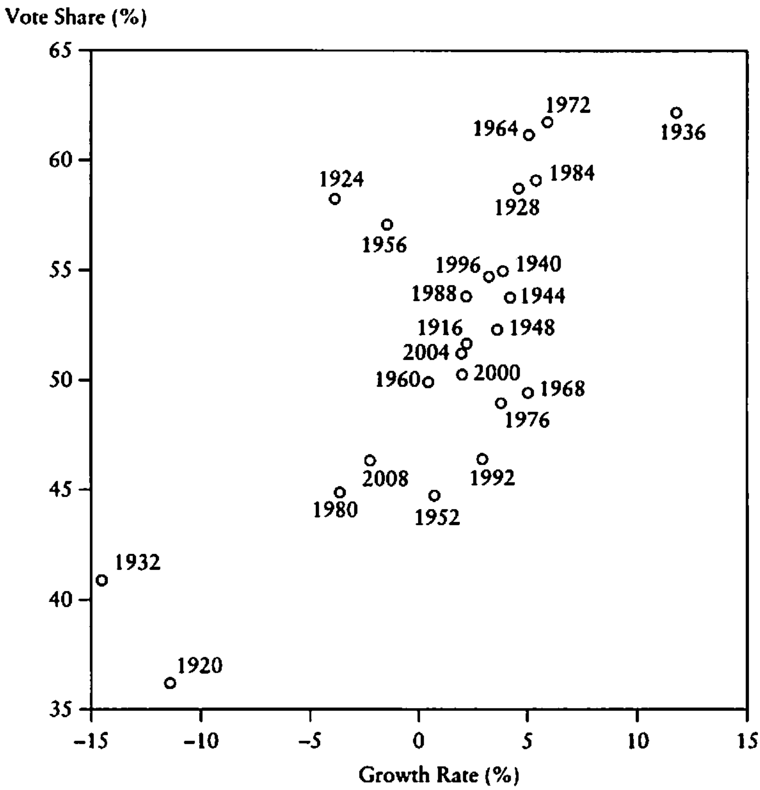

FIGURE 1-1 Incumbent party vote share graphed against the growth rate of the economy

Two observations that stand out in Figure 1-1 are the elections of 1932 and 1936. In 1932, the incumbent party's candidate (Herbert Hoover) got 40.9 percent of the two-party vote, a huge defeat in a year when the growth rate of the economy was –14.6 percent. (Yes, that's a minus sign!) In 1936, the incumbent party's candidate (Franklin Roosevelt) got 62.2 percent of the two-party vote, a huge victory in a year when the growth rate of the economy was an exceptionally strong 11.8 percent. (Although 1936 was in the decade of the Great Depression, the economy actually grew quite rapidly in that year.)

Figure 1-1 shows that there does appear to be a positive relationship between the growth rate in the year of the election and the incumbent vote share: the scatter of points has an upward pattern. Voters may thus take into account the state of the economy when deciding for whom to vote, as the theory says.

The growth rate is not, however, the only measure of how well the economy is doing. For example, inflation may also be of concern to voters. When inflation has been high under the incumbent party, a voter may fear that his or her income will not rise as fast as will prices in the future if the incumbent party's candidate is elected and thus that he or she will be worse off under the incumbent party. The voter may thus vote against the incumbent party. Many people consider inflation bad for their financial well-being, so high inflation may turn voters away from the incumbent party.

Deflation, which is falling prices, is also considered by many to be bad. People tend to like stable prices (that is, prices that on average don't change much from year to year). There have been some periods in U.S. history in which there was deflation. For example, prices on average declined during the four-year periods prior to the elections of 1924 and 1932. If voters dislike deflation as much as they dislike inflation, then inflation of −5 percent (which is deflation) is the same in the minds of the voters as inflation of +5 percent. Therefore, in dealing with the data on inflation, we will drop the minus sign when there is deflation. We are thus assuming that in terms of its impact on voters, a deflation of 5 percent is the same as an inflation of 5 percent.

To see how the incumbent party vote share and inflation (or deflation) are related, Figure 1-2 graphs the vote share against inflation for the 24 elections between 1916 and 2008. According to the theory just discussed, there should be a negative relationship between the two: when inflation is high, the vote share should be low, and vice versa. You can see from Figure 1-2 that there does seem to be at least a slight negative relationship between the incumbent party vote share and inflation: the scatter of points has a slight downward pattern.

We are not, however, limited to a choice between one or the other—the growth rate or inflation. It may be that *both* the growth rate and inflation affect voting behavior. In other words, we need not assume that voters look at only one aspect of the economy when they are considering their future financial well-being. For example, if both the growth rate and inflation are high, a voter may be less inclined to vote for

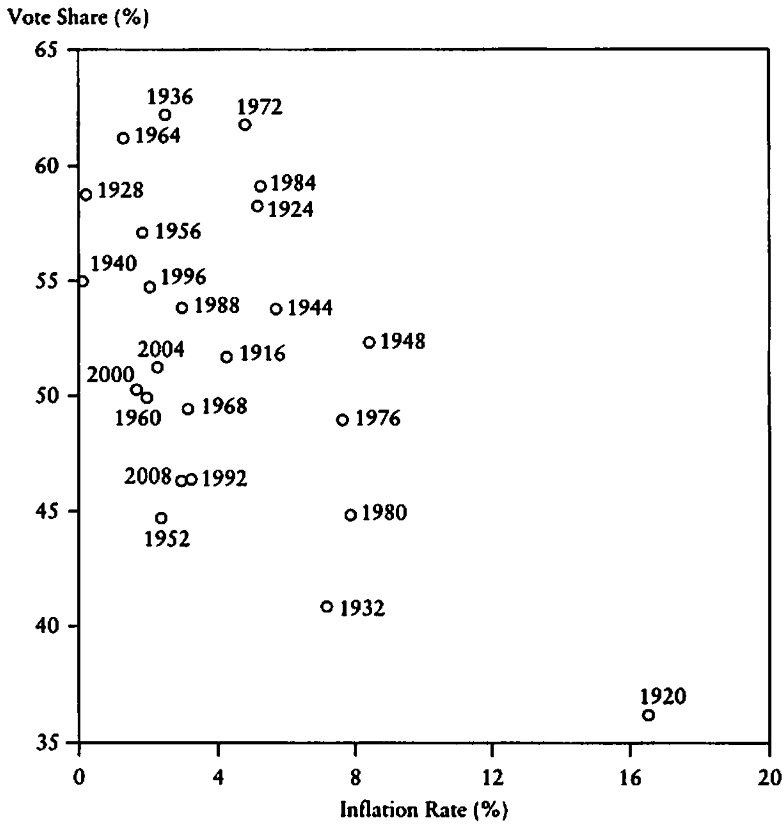

FIGURE 1-2 Incumbent party vote share graphed against the inflation rate

the incumbent party than if the growth rate is high and inflation is low. Similarly, if both the growth rate and inflation are low, a voter may be more inclined to vote for the incumbent party than if the growth rate is low and inflation is high.

An example of a high growth rate and low inflation is 1964, where the growth rate was about 5 percent and inflation was about 1 percent. In this case, the incumbent, President Johnson, won by a landslide over Barry Goldwater—receiving 61.2 percent of the two-party vote. An example of a low growth rate and high inflation is 1980, where the growth rate was about –4 percent and inflation about 8 percent. In this case, the incumbent, President Carter, lost by a large amount to Ronald Reagan—receiving only 44.8 percent of the two-party vote.

To carry on, we are also not limited to only two measures of the economy. In addition to observing how the economy has grown in the year of the election, voters may look at growth rates over the entire four years of the administration. In other words, past growth rates *along with* the growth rate in the election year may affect voting behavior. One measure of how good or bad past growth rates have been is the number of quarters during the four-year period of the administration in which the growth rate has exceeded some large number. These are quarters in which the output news was particularly good, and voters may be inclined to remember these kinds of events. There is some evidence from psychological experiments that people tend to remember peak stimuli more than they remember average stimuli, a finding that is consistent with voters remembering very strong growth quarters more than the others. Therefore, voters may be more inclined to vote for the incumbent party if there were many of these "good news quarters" during the administration than if there were few.

Noneconomic factors may also affect voting behavior. If the president is running for reelection, he (or maybe she in the future) may have a head start. He can perhaps use the power of the presidency to gain media attention, control events, and so forth. He is also presumably well known to voters, and so there may be less uncertainty in voters' minds regarding the future if he is reelected rather than if someone new is elected. Voters who do not like uncertainty may thus be more inclined to vote for the incumbent party if the president is a candidate than otherwise.

It is also possible that voters get tired, or bored, with a party if it has been in power for a long time, feeling that time has come for a change. Therefore, the longer a party has been in power, the less inclined many voters may be to keep it in power. The vote share may thus depend on a measure of the duration of the incumbent party.

The theory of voting behavior that has just been presented can be summarized in Box 1-1. The items in the box are called *variables*. A variable is something that changes, or varies. For example, the vote share is different for different elections; it varies across elections. Likewise, the growth rate is different for different elections; it also varies. Both the vote share and the growth rate are thus variables.

BOX 1-1

vote share depends on:	
	growth rate
	inflation
	good news quarters
	president running
	duration

The variable to be explained—in this case, the vote share—is called the *dependent* variable. According to the theory, it "depends" on the other variables in the box. The other variables in the box—growth rate, inflation, good news quarters, president running, and duration—are called *independent* or *explanatory* variables. They help "explain" the dependent variable. The independent or explanatory variables are not themselves explained by the theory; they simply do the explaining.

The theory we have proposed may not, of course, capture well the way that voters actually behave. Maybe the economy plays no role. Maybe the vote share depends on a completely different set of factors—personality factors, foreign policy issues, social welfare issues, and so forth. We must collect data and test the theory.

The Data

To test the theory, we need data on past election outcomes and on what the economy was like for each election. We have already seen data on the vote share and the growth rate in Figure 1-1 and data on the vote share and inflation in Figure 1-2. Table 1-1 presents a more detailed picture of the data. The two-party vote share is presented for each election from 1916 to 2008, along with the growth rate, the inflation rate, the number of good news quarters, and a measure of duration.

(Regarding the poem at the beginning of this chapter, William Jennings Bryan was the Democratic Party candidate in 1896, 1900, and 1908. He was defeated in the first two campaigns by William McKinley and in the 1908 campaign by William H. Taft. These elections are not in the sample in Table 1-1, but economic topics played a key role in Bryan's campaigns. In the election of 1896, for example, he made his famous cross of gold speech: "You shall not press down upon the brow of

TABLE 1-1 Data used for explaining the outcome of presidential elections, 1916–2008

Year	Party in Power 1 = D −1 = R	President Running 1 = yes 0 = no	Election Outcome	Incumbent Vote Share (%)	Growth Rate (%)	Inflation Rate (%)	Good News Quarters	Duration Value	Democratic Vote Share (%)
1916	1	1	President Wilson (D) beat Hughes (R)	51.7	2.2	4.3	3	0.00	51.7
1920	1	0	Cox (D) lost to Harding (R)	36.1	−11.5	16.5	5	1.00	36.1
1924	−1	1	President Coolidge (R) beat Davis (D) and LaFollette	58.3	−3.9	5.2	10	0.00	41.7
1928	−1	0	Hoover (R) beat Smith (D)	58.8	4.6	0.2	7	1.00	41.2
1932	−1	1	President Hoover (R) lost to F. Roosevelt (D)	40.9	−14.6	7.2	4	1.25	59.1
1936	1	1	President F. Roosevelt (D) beat Landon (R)	62.2	11.8	2.5	9	0.00	62.2
1940	1	1	President F. Roosevelt (D) beat Willkie (R)	55.0	3.9	0.1	8	1.00	55.0
1944	1	1	President F. Roosevelt (D) beat Dewey (R)	53.8	4.2	5.7	14	1.25	53.8
1948	1	1	President Truman (D) beat Dewey (R)	52.3	3.6	8.4	5	1.50	52.3
1952	1	0	Stevenson (D) lost to Eisenhower (R)	44.7	0.7	2.3	7	1.75	44.7
1956	−1	1	President Eisenhower (R) beat Stevenson (D)	57.1	−1.5	1.9	5	0.00	42.9
1960	−1	0	Nixon (R) lost to Kennedy (D)	49.9	0.5	1.9	5	1.00	50.1
1964	1	1	President Johnson (D) beat Goldwater (R)	61.2	5.1	1.3	10	0.00	61.2
1968	1	0	Humphrey (D) lost to Nixon (R)	49.4	5.0	3.1	7	1.00	49.4
1972	−1	1	President Nixon (R) beat McGovern (D)	61.8	5.9	4.8	4	0.00	38.2

1976	−1	0	Ford (R) lost to Carter (D)	49.0	3.8	7.6	5	1.00	51.0
1980	1	1	President Carter (D) lost to Reagan (R)	44.8	−3.7	7.9	5	0.00	44.8
1984	−1	1	President Reagan (R) beat Mondale (D)	59.1	5.4	5.2	8	0.00	40.9
1988	−1	0	G. Bush (R) beat Dukakis (D)	53.8	2.2	3.0	4	1.00	46.2
1992	−1	1	President G. Bush (R) lost to Clinton (D)	46.4	2.9	3.3	2	1.25	53.6
1996	1	1	President Clinton (D) beat Dole (R)	54.7	3.3	2.0	4	0.00	54.7
2000	1	0	Gore (D) lost to G. W. Bush (R)	50.3	2.0	1.6	7	0.00	50.3
2004	−1	1	President G. W. Bush (R) beat Kerry (D)	51.2	2.0	2.2	1	0.00	48.8
2008	−1	0	McCain (R) lost to Obama (D)	46.3	−2.3	3.1	1	0.00	53.7

NOTE: See Chapter 3 for description of the data.

labor this crown of thorns, you shall not crucify mankind upon a cross of gold." The 1896 campaign in part pitted northeastern creditors against southern and western debtors. Bryan wanted the dollar to be backed by silver as well as gold, which would have put more money in circulation and led to lower interest rates, thus benefiting the South and West over the Northeast.)

The data in Table 1-1 are discussed in more detail in Chapter 3, and so only a few points about the data will be made now. As mentioned earlier, in the periods before the elections of 1924 and 1932, inflation was negative (that is, there was deflation). Positive values are, however, listed in Table 1-1, reflecting the assumption that voters dislike deflation as much as they dislike inflation. If the president is running for reelection, the word *President* precedes the name. A candidate who was elected vice president and became president during the administration was counted as the president running for reelection. (Gerald Ford was not counted because he was not elected vice president; he was appointed vice president after Spiro Agnew resigned.)

The duration variable in Table 1-1 has a value of 0.0 if the incumbent party has been in power only one term before the election. It has a value of 1.0 if the incumbent party has been in power two terms in a row, a value of 1.25 for three terms in a row, a value of 1.50 for four terms in a row, and a value of 1.75 for five terms in a row. We will use this variable in Chapter 3; it can be ignored for now.

Because we are taking the vote share to be the share of the *two-party* vote, we are ignoring possible third-party influences. We are in effect assuming that third-party votes are taken roughly equally from the two major parties. Again, we will come back to this in Chapter 3.

Table 1-1 gives a good picture of what is to be explained and the variables that the theory says may be important in the explanation. In the election of 1916, the incumbent party was Democratic, and President Wilson beat the Republican Hughes with 51.7 percent of the two-party vote. The growth rate was 2.2 percent, and inflation was 4.3 percent. In 1920, the incumbent party was Democratic and the Democrat Cox lost to the Republican Harding by a landslide. Cox got only 36.1 percent of the two-party vote. For this election, the growth rate was –11.5 percent and inflation was 16.5 percent—hardly a good economy! This outcome is, of course, consistent with the theory: the economy was very poor, and the incumbent party got trounced.

You may want to go through the rest of the elections and see what the story is for each. Only three more will be mentioned here. As we noted above, in 1980 President Carter lost to Reagan with a vote share of 44.8 percent. The economy was not good for Carter—the growth rate was –3.7 percent and inflation was 7.9 percent. This outcome is again consistent with the theory. On the other hand, the theory has trouble with the 1992 election, when President G. Bush lost to Clinton with a vote share of 46.4 percent. Unlike 1980, the economy was not that bad in 1992—the growth rate was 2.9 percent and inflation was 3.3 percent. There is thus a puzzle as to why President G. Bush lost, or at least why he lost by as much as he did. We will return to this question in Chapter 3. In 2008, the incumbent party was Republican and the Republican John McCain lost to Democrat Barack Obama with a vote share of 46.3 percent. The economy was not good in 2008 regarding the growth rate, which was –2.3 percent.

On to the Tools

We have presented a theory of voting behavior in this chapter, and we have presented data that can be used to test the theory. This is, however, as far as we can go without the tools, and so it is on to Chapter 2. Once the tools have been explained in Chapter 2, they will be used in Chapter 3 to test the theory of voting behavior and to predict the 2012 election.

Plato, despair!
We prove by norms
How numbers bear
Empiric forms,

How random wrong
Will average right
If time be long
And error slight,

But in our hearts
Hyperbole
Curves and departs
To infinity.

Error is boundless.
Nor hope nor doubt,
Though both be groundless,
Will average out.

J. V. Cunningham, Meditation on Statistical Method

(Monday) Lesson 1: Begin with a Theory

We begin with a theory of what we are trying to explain. In the previous chapter, a theory of voting behavior was presented—a theory of what motivates people to vote the way they do. Although it does not matter where a theory comes from, it should have a ring of plausibility. A theory that seems completely at odds with how something works is not of much interest to test and is not likely to get us very far.

We are also interested in causation, not correlation. Two variables are positively correlated if large values of one are associated with large values of the other and small values of one are associated with small values of the other. (Two variables are negatively correlated if large values of one are associated with small values of the other and small values of one are associated with large values of the other.) If you look across cities, you will likely find that those with a relatively high percentage of air conditioners also have a relatively high percentage of people with skin cancer. Air conditioners and skin cancer are positively correlated across cities. But this does not mean that air conditioners cause skin cancer or that skin cancer causes people to buy air conditioners. In this case, the climate of a city is causing both. A theory is a statement that one or more variables cause another variable—the variable we are interested in explaining. A statement that air conditioners cause skin cancer is not a theory, or at least not a theory that is interesting to test, because it confuses correlation with causation.

It is easiest to explain the following tools using a particular theory as an example, and we will use the theory of voting behavior. To begin, we simplify matters and assume that the vote share depends only on the economy's growth rate in the year of the election. The theory we begin with is that the growth rate has a positive effect on the incumbent party vote share.

(Tuesday) Lesson 2: Collect Data

Empirical evidence or information is needed to test a theory. This requires collecting observations on the variable to be explained (the dependent variable) and the variables that, according to the theory, do the explaining (the independent or explanatory variables). Observations on variables are called *data*. Table 1-1 provides a good example of data collection. Observations are available for the period 1916–2008 on the vote share, the growth rate, and a number of other variables. This table contains the data that we will use to test the theory of voting behavior.

Data collection is perhaps the most important step in testing a theory. A good test requires that the variables for which observations are collected match closely the variables that the theory is talking about. It is not always easy to find good matches, and much of social science research consists of the nitty-gritty job of finding appropriate data. Tuesday is thus a critical day. Always keep in mind when deciding how much to trust the results of a test that a test is no better than the data behind it.

(Wednesday) Lesson 3: Fit the Data

This is the hardest lesson except for Lesson 4. If you have not worked this kind of material before—essentially fitting points to a line—you may need to go through it more than once. Keep in mind that the main aim of this lesson is to fit as closely as possible the theory from Lesson 1 to the data from Lesson 2.

We begin with Table 1-1, where we have 24 observations on the vote share and on the growth rate. In Figure 2-1, the vote share is graphed

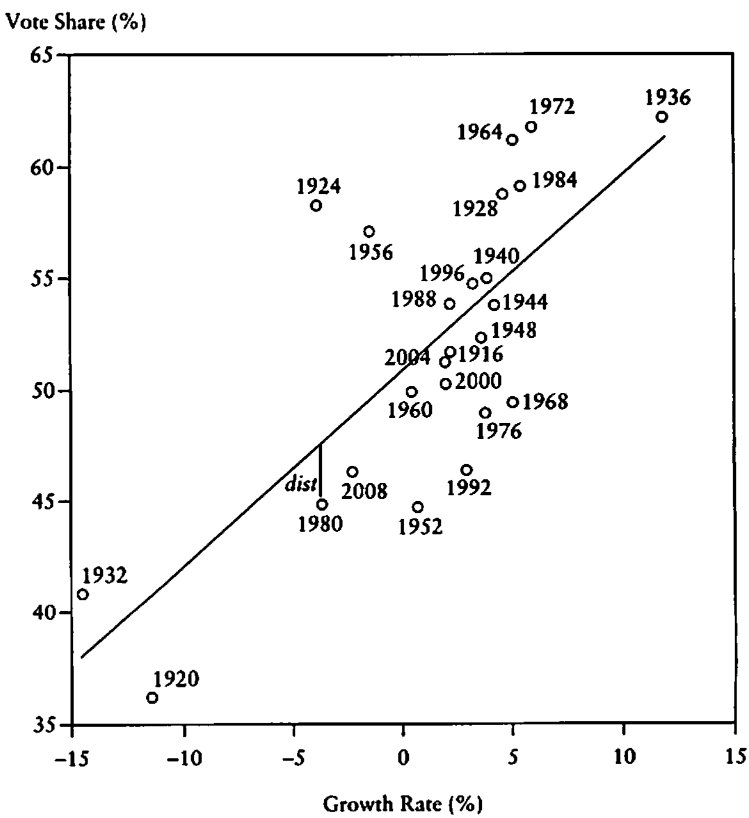

FIGURE 2-1 Incumbent party vote share graphed against the growth rate of the economy

against the growth rate. This is the same graph as Figure 1-1, although now we are going to do more with it. The vote share is on the vertical axis, and the growth rate is on the horizontal axis.

Note that the points in Figure 2-1 have an upward pattern (rising from left to right). In other words, the vote share tends to be larger when the growth rate is larger. This is, of course, exactly what our theory says should be the case. In some cases, however, one vote share is larger than another, even though it has a smaller growth rate associated with it. For example, the vote share for 1956 is larger than the vote share for 1996, even though the growth rate for 1956 is smaller than that for 1996. It is clear that the points in the figure do not all lie on a straight line, and in this sense, the relationship between the vote share and the growth rate is not exact.

Even though all the points in Figure 2-1 do not lie on a straight line, we can ask which straight line the points are closest to. Put another way, we can ask which line best "fits" the points. To see what is involved in finding the best fitting line, consider drawing some line through the points, such as the line drawn in the figure. You should think for now of this line as being any arbitrary line, not necessarily the best fitting line. Some points on the particular line in the figure, such as the one for 1940, are almost exactly on the line. Others, such as the one for 1920, are farther away from the line. The vertical distance between a point and the line, such as that labeled *dist* for 1980, is a measure of how far the point is from the line. This distance is sometimes called an *error*. If all points were exactly on the line, all errors would be zero, and the relationship between the vote share and the growth rate would be exact. Otherwise, the larger the distances are on average, the less precise is the relationship.

If we define an error as the difference between the point on the line and the actual point, then errors corresponding to the points below the line are positive, and errors corresponding to the points above the line are negative. Now, if we change the sign of the negative errors to positive and add up all the errors, the answer (the sum) is a measure of how closely the line fits the points. We are just adding up all the distances of the points from the line. One line can be said to fit better than another line if it has a smaller sum. In practice, a more popular way of measuring how well a line fits the points is first to take each error and square it (that is, multiply each error by itself) and then to add up all the squared errors. Either way, the general idea is the same. In terms of deciding which line fits best, a line in which the points are far from the line (that is, when the distances from the points to the line are large) is not as good as a line in which the points are close to the line.

We will follow the convention of squaring the errors before summing them. Now, imagine drawing thousands of straight lines in Figure 2-1, each with a different position in the figure, and for each line taking the 24 errors, squaring them, and then adding up the squared errors. We can thus associate one number (the sum of the errors squared) with each line, and this number is a measure of how well that particular line fits the points. If the number for a given line is large relative to numbers for other lines, this means that the given line is not positioned well in the figure. For this line, the distances from the points to the line are on average large, so the line does not fit the points well. The best fitting line out of the thousands of lines is simply the line with the smallest number (that is, the smallest sum of squared errors). Although it would take hours to draw thousands of lines by hand and compute the number for each line, computers can find the best fitting line very quickly. It takes almost no time for a computer to find the line with the smallest sum of squared errors.

Assume that we have found the best fitting line with the help of a computer. (This is in fact the line drawn in Figure 2-1.) Associated with this line is, of course, its sum of squared errors. Although the sum of squared errors is a measure of how well the line fits the points, it does not give a sense of what a typical error is. A better sense of the size of a typical error can be obtained by dividing the sum of squared errors by the number of observations, which gives the average squared error, and then taking the square root of the average squared error, which gives the average error. It turns out, for example, that the sum of squared errors for the line in Figure 2-1 is 486.9. If we divide this number by 24, the number of points in the figure, we get 20.3, which is the average squared error. If we then take the square root of this number, we get 4.5, which is the average error.

Although we might use the average error of 4.5 as the measure of a typical error, in practice a slightly different measure is used. In the present example, instead of dividing the sum of squared errors of 486.9 by 24, we divide it by 22. In other words, we subtract 2 from the number of points before dividing. We use 2 because the line is determined by two points.

If there were only 2 points in Figure 2-1, the line would fit perfectly—2 zero errors. So we start off in this sense with 2 zero errors, leaving 22 to play with. If we divide 486.9 by 22, we get 22.1. The square root of this number is 4.7, slightly larger than 4.5. We then take 4.7 as the measure of a typical error. This error measure is sometimes called a *standard error*, and we will use this terminology. A standard error is just a measure of the average size of a typical error for a line like that in Figure 2-1.

Is the standard error of 4.7 for the line in Figure 2-1 large or small? We can see from Table 1-1 in the previous chapter that the vote share ranges from a low of 36.1 percent in 1920 to a high of 62.2 percent in 1936. The difference between the high and the low is thus 26.1, and so 4.7 is fairly small compared to this range. On the other hand, many elections have been decided with a margin less than 4.7 percentage points, and so on this score, 4.7 is fairly large. We will see in the next chapter that when the other explanatory variables are taken into account, the standard error is in fact much smaller than 4.7. For now, however, we will stay with using only the growth rate and thus the standard error of 4.7.

We have so far found the best fitting line and calculated the standard error. Another way to get a sense of how well the line fits the points is simply to examine individual points. In particular, it is of interest to see which points are far from the line (that is, which points have large errors associated with them). For example, the error for the election of 1924 is quite large, as we can see in Figure 2-1. For this election, the vote share is high, and the growth rate is low (in fact negative), resulting in a point far above the line. The election of 1956 is a similar case, where the vote share is high, and the growth rate is low. A point that is fairly far below the line is 1952, where the vote share is low, and the growth rate is modest. On the other hand, a number of points—such as 1936, 1940, 1944, and 1988—are very close to the line.

Another useful number we can get from Figure 2-1 is the *slope* of the line. The slope of a line is the measure of how steep it is. The slope of the line in Figure 2-1 is positive: it rises upward and to the right. A negative slope is one in which the line moves downward and to the right. The slope of the line in Figure 2-1 is 0.9. This means that if you move along the horizontal axis by 1 unit, the vote share on the line will increase by 0.9 units. For example, an increase in the growth rate of 1.0 percentage point will increase the vote share by 0.9 percentage points on the line.

You can probably see already why the size of the slope is important. If the slope is large (that is, the line is steep), the growth rate has a large effect on the vote share according to the line. If, on the other hand, the slope is zero (that is, the line is horizontal), the growth rate has no effect on the vote share according to the line. Clearly, if the slope in Figure 2-1 were zero or close to zero, there would be no support for the theory that the growth rate affects the vote share.

The slope of the best fitting line is sometimes called the *estimated slope*, and we will use this terminology. The slope is estimated in the sense that it is computed by finding the best fitting line given whatever observations are at our disposal.

This is it for Lesson 3. The basic idea is to see how well a theory fits the data. In the simple case considered so far of a dependent variable and one explanatory variable, the theory is used to choose the two variables (in our example, the vote share and the growth rate). The data are then used to find the best fitting line. The main point of this lesson is to show how best fitting lines are found.

(Thursday) Lesson 4: Test

Just because we have found the best fitting line does not mean we have discovered anything interesting or useful. One of the main issues social scientists worry about when analyzing relationships such as the one between the vote share and the growth rate is that a relationship may have been uncovered by coincidence and that it is in fact not truly valid. In the present example, we worry that the vote share is not really affected by the growth rate even though for the 24 observations it looks like it is.

One way to express our anxiety is by stating our concern that the slope of the true line in a figure like Figure 2-1 is zero. If the true slope is zero, there is no relationship between the growth rate and the vote share. Although the slope of the line in Figure 2-1 is positive, as the theory of voting behavior says it should be, perhaps the positive slope is just a fluke. It may be just by chance that the 24 available observations (points) show a positive effect. If we had 24 other observations (say by waiting for 24 more elections to take place), they might show a much smaller positive slope or even a negative slope. We need to test whether the positive slope is or is not a fluke.

The main point of this lesson is to explain how we can test whether the true slope is zero. As noted at the end of Lesson 3, the slope in Figure 2-1 is 0.9, and we want to see how likely it is that the true slope is zero, even though we have estimated it to be 0.9. We will show that we can compute for an estimated slope its *t-statistic.* We will then see that if the value of the *t*-statistic is greater than about 2.0, it is very unlikely that the true slope is zero. A *t*-statistic greater than 2.0 is thus good for the theory. It says that the slope we have estimated by finding the best fitting line is unlikely to be truly zero.

The rest of this lesson is difficult, and if you are willing to take the result about the *t*-statistic on faith, you may skip to Lesson 5. Or you may skip this material for now and come back later, once you have seen, in Chapters 3 and beyond, the use of *t*-statistics in action.

Thursday Morning

We begin with the errors in Figure 2-1. Remember that an error for a point is the distance from the point to the line. Also remember that an error is negative if the point is below the line and positive if the point is above the line. Now, say that instead of 24 errors, we had hundreds of them (from hundreds of elections), and we recorded how many of them were between 0.0 and 0.1, how many were between 0.1 and 0.2, how many were between –0.1 and 0.0, and so on. In other words, suppose we have intervals of size 0.1 and we record how many errors are in each interval: we record how many very small errors there are, how many fairly small errors, how many medium errors, and so on.

In most cases, we will find that there are more very small errors than small ones, more small errors than medium errors, more medium errors than large ones, and so on. In fact, in many cases, if we graph the number of errors in each interval against the position of the interval on the horizontal axis, we will get points that lie on a curve that is approximately like that in Figure 2-2. The curve in Figure 2-2 is a bell-shaped curve, which is a curve you may have seen in other contexts. If, for example, you divided scores on intelligence tests (IQ scores) into small intervals and graphed the number of scores in each interval against the position of the interval on the horizontal axis, you would get points that lie approximately on a bell-shaped curve.

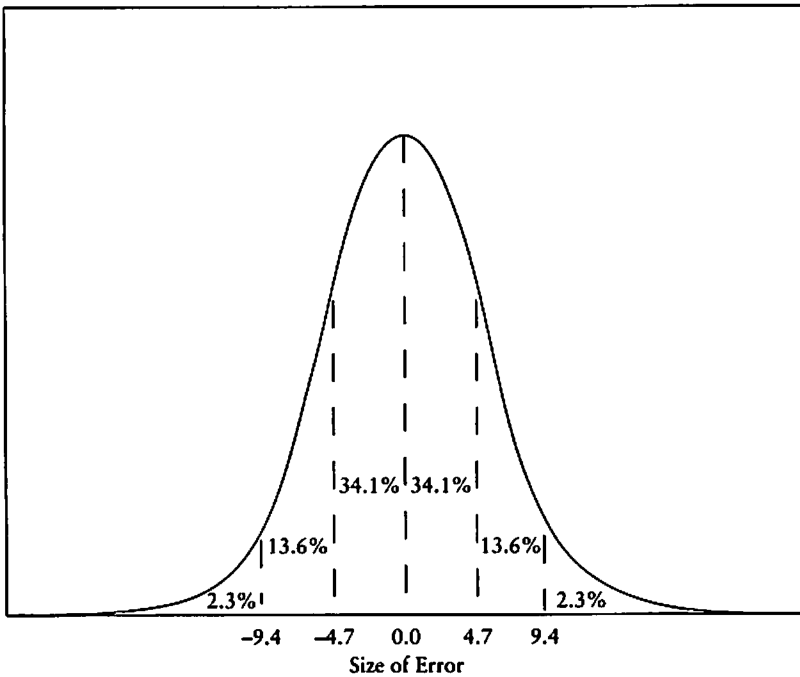

FIGURE 2-2 Bell-shaped curve for errors in Figure 2.1

We will assume that the curve in Figure 2-2 is an exact bell-shaped curve. We will also assume that the peak of the bell-shaped curve corresponds to an error of zero and that the standard error is, as computed in Lesson 3, 4.7. A bell-shaped curve has some useful characteristics concerning the size of the area under the curve that we need to know about. The total area under the curve is the space under the curve and above the horizontal axis. Consider starting from zero and moving a distance of 4.7 to the right and 4.7 to the left, where 4.7 is the standard error. Doing this sweeps out 34.1 percent of the area to the right of zero and 34.1 percent of the area to the left of zero, for a total of 68.2 percent of the area under the curve. This area is shown in Figure 2-2. In other words, 68.2 percent of the errors are between –4.7 and 4.7, where again 4.7 is the standard error.

If we sweep out the area between 4.7 and 9.4, we get another 13.6 percent, and if we sweep out the area between –4.7 and –9.4, we get another 13.6 percent. The total area between –9.4 and 9.4 is thus 95.4 percent. Looked at another way, the area to the right of 9.4 is 2.3 percent and the area to the left of –9.4 is 2.3 percent. That is, for a bell-shaped curve with a standard error of 4.7, only 2.3 percent of the errors are larger than 9.4 and only 2.3 percent are smaller than –9.4.

So, you might ask, why are you telling me this? In particular, what does this have to do with the question of whether the positive slope in Figure 2-1 is a fluke? Well, we are making progress, but patience is still needed. The next step is to consider the following thought experiment. Imagine there is another universe in which the true relationship between the vote share and the growth rate is the line in Figure 2-1 (slope equal to 0.9). Imagine also that in this universe there are 24 presidential elections with the same 24 growth rates as in Figure 2-1. For our thought experiment, we want this universe to have different errors than the errors depicted in Figure 2-1.

Consider, for example, the 1992 election, when President George H. W. Bush lost to candidate Bill Clinton. As in any election, there were many things that affected voting behavior aside from the growth rate, and for this election, these other things were a net negative for Bush. This can be seen by noting that the point for 1992 in Figure 2-1 is below the line. Remember that the distance from a point to the line is the error for that point. Because the point for 1992 is below the line, the error is positive (the difference between the point on the line and the actual point). Although not directly recorded in Figure 2-1, the error for 1992 is 7.1 percent, which is fairly large. In 1992, President Bush got 46.4 percent of the two-party vote, which is the 1992 point in Figure 2-1. The position on the line corresponding to the 1992 growth rate is 53.5 percent. The 7.1 error is the difference between 46.4 and 53.5. In the present context, we can look upon an error in any given election as reflecting all the other things that affect voting behavior aside from the growth rate. The more these other things matter, the larger on average will the error be. We have, fortunately, a measure of how large the errors are on average, which is the standard error of 4.7.

Now comes the key step. We want to draw for the other universe a different error for the 1992 election from the one that actually occurred (which was 7.1). We are going to draw this error from a bell-shaped curve with a standard error of 4.7. This error will in general be different from the actual error of 7.1. We are imagining an election in 1992 in the other universe with the same growth rate but a different set of the other things that affect voting behavior. Maybe in the other universe Clinton did very poorly in the debates, leading fewer people to vote for him. Or maybe President Bush successfully toppled Saddam Hussein in the Gulf War, leading more people to vote for him. We are thus imagining a different set of other things, which implies a different error, and we are drawing this error at random from a bell-shaped curve with a standard error of 4.7. By using a standard error of 4.7, we are assuming that the same bell-shaped curve pertains to the other universe as it pertains to ours. In other words, the average size of the effect of the "other things" is assumed to be the same.

Say we drew an error of 3.0 for 1992. What would be the actual vote share in this case? Remember we are assuming the line in Figure 2-1 is correct, and given the growth rate for 1992, the point of the line for 1992 is 53.5, as noted above. For the actual error of 7.1, the actual vote share is 46.4, which is simply 53.5 minus 7.1. If instead we have an error of 3.0, the actual vote share is 53.5 minus 3.0, or 50.5. In this other universe, the vote share for President Bush is larger than it is in the actual universe because the error is smaller. To get the vote share in the other universe, we thus take the point on the line in Figure 2-1, which we are assuming is correct, and subtract the error that we draw.

We draw an error not only for 1992, but for the other 23 elections as well, each time using the bell-shaped curve with a standard error of 4.7. From these 23 errors, we can get 23 vote-share values as just discussed. This means that we have 24 different points (vote shares) in Figure 2-1. Using these new points, we can find the line that best fits the points. Because the points are different, the best fitting line will generally be in a different position from the line in Figure 2-1. So generally, it will have a different slope. The slope in the other universe will generally not be 0.9.

We now go to another universe and draw a new set of 24 errors. We get a new set of vote-share values, again assuming that the line in Figure 2-1 is correct, and we find the best fitting line for the new vote-share values. This gives us another slope. We keep doing this until we have run through many universes, say 1,000. We have thus computed 1,000 different estimated slopes.

The next step is to examine the 1,000 slopes. As we did above for the errors, we divide these slopes into small intervals and then graph the

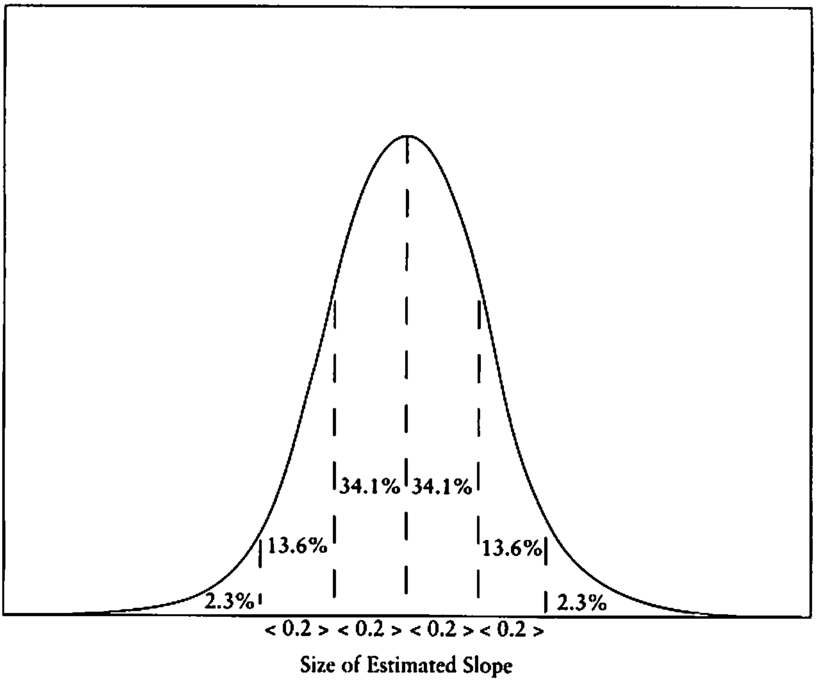

FIGURE 2-3 Bell-shaped curve for estimated slopes

number of slopes in each interval against the position of the interval on the horizontal axis. When we do this, we get points that lie approximately on a bell-shaped curve. This curve is drawn in Figure 2-3. We will assume that this curve is exactly bell-shaped. In practice, it is not quite bell-shaped, but for most applications, the bell-shaped curve is a very close approximation.

It is important to keep in mind that we have moved from errors to slopes. The bell-shaped curve in Figure 2-2 is for the errors from Figure 2-1 (that is, the distances from the points to the line), whereas the bell-shaped curve in Figure 2-3 is for the slopes we have estimated from our different universes.

We use the bell-shaped curve in Figure 2-3 in the following way. We start from the estimated slope that is exactly under the peak of the curve, and we sweep out 34.1 percent of the area to the right. We know from the characteristic of a bell-shaped curve that the distance we travel along the horizontal axis to sweep out 34.1 percent of the area is one standard error. In the present example, this distance turns out to be 0.2. This means that the standard error of the estimated slope is 0.2. As shown in the figure, if we move another 0.2 to the right, we sweep out 13.6 percent of the area. If we move left from the center by 0.2, we sweep out 34.1 percent of the area, and if we move left another 0.2, we sweep out another 13.6 percent of the area.

This standard error of 0.2 for the estimated slope in Figure 2-3 is, of course, different from the standard error of 4.7 in Figure 2-2, which is for the line in Figure 2-1. Each bell-shaped curve has its own standard error. We always need to be clear in what follows as to which standard error we are talking about. When there is possible confusion, we will use the phrases *standard error of the line* and *standard error of the estimated slope.*

Let us recap what we have done so far. We have assumed that the line in Figure 2-1 is correct, and using it and the bell-shaped curve in Figure 2-2, we have generated many sets of 24 vote shares by drawing errors. For each set of 24 vote shares, we have found the best fitting line, which gives us an estimated slope. Once we have computed many slope values, we use the bell-shaped curve in Figure 2-3 to find the standard error of the estimated slope. All the work so far has simply been to compute the standard error of the estimated slope. In practice, this standard error can be computed in a way that does not require drawing errors from hypothetical universes, but the answer is the same in either case.

We still have one more key step to take, but you may want to take a break for lunch before finishing. You should have a good idea of what the standard error of the estimated slope is before reading further.

Thursday Afternoon

The main issue we are worried about from the point of view of the theory of voting behavior is whether the true slope in Figure 2-1 is zero. If the slope is truly zero, it means the growth rate has no effect on the vote share: changes in the growth rate have no effect on the vote share if the line is completely flat. If the slope is zero, then the theory is not supported by the data.

Fortunately, we can now test whether the slope is zero. This is done in Figure 2-4. The figure is based on the assumption that the true slope is zero. We know from Figure 2-1 that the slope we have computed is 0.9. To test whether 0.9 is a fluke, we need to know the chance that we would get a value of 0.9 if the true slope were zero. This is where we need our standard error of the estimated slope of 0.2. Figure 2-4 shows it is very unlikely with a standard error of 0.2 that we would get an estimated slope of 0.9 if the true slope were zero. The probability that we would get a slope of 0.9 if the true slope were zero is the area under the curve to the right of 0.9 in the figure, which is very close to zero. (The area is so small that it can't even be seen in the figure!) The data thus support the theory that the growth rate affects the vote share. To repeat, if the growth rate did not affect the vote share, it is unlikely we would get a slope of 0.9 in Figure 2-1. It is thus likely that the theory is true.

Now comes the punch line. If we divide the slope (0.9) by the standard error of the estimated slope (0.2), we get what is called a *t-statistic* (4.5 in this example). The *t*-statistic is a highly useful concept. Assume

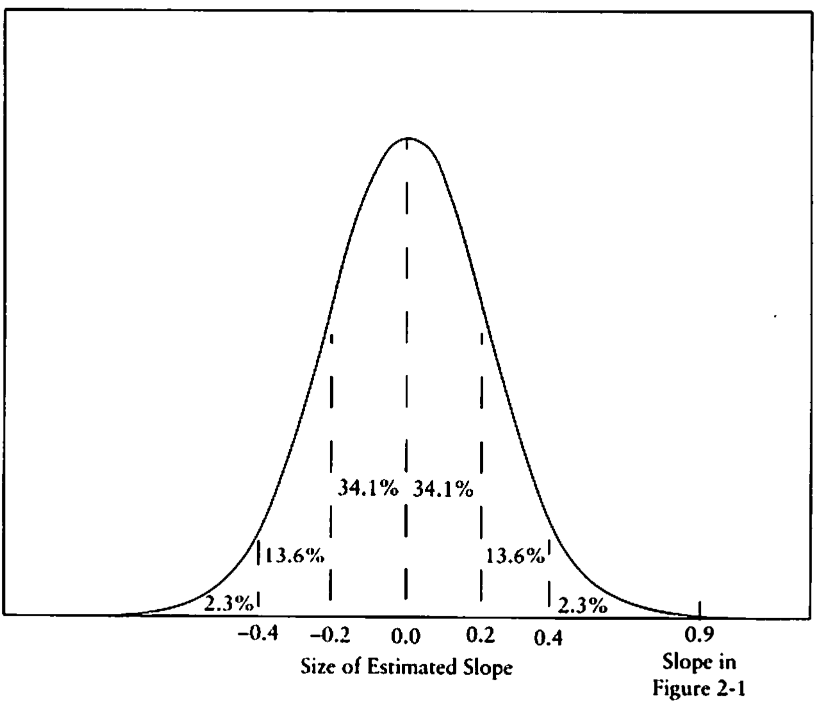

FIGURE **2-4** Bell-shaped curve for estimated slopes if true slope were zero for the sake of argument that we had an estimated slope of 0.4 instead of 0.9 with the same standard error of 0.2. This would mean that the *t*-statistic is 2.0. We can see from Figure 2-4 that for an estimated slope of 0.4, 2.3 percent of the area under the curve lies to the right of 0.4. Therefore, the probability that we would get a slope of 0.4 if the true slope were zero (and the standard error of the estimated slope were 0.2) is just 2.3 percent. For a *t*-statistic of 2.0, it is thus unlikely that the true slope is zero and thus likely that the theory is true. In practice, a *t*-statistic of 2.0, or around 2.0, is used as a cutoff regarding whether a theory is supported or not supported. A theory is not supported if the *t*-statistic is less than 2.0 or thereabouts, and a theory is supported if the *t*-statistic is greater than 2.0 or thereabouts. This convention is, of course, somewhat arbitrary, and other cutoffs could be used. If a *t*-statistic less than 2.0 were used as a cutoff, more theories would be supported, and if a *t*-statistic greater than 2.0 were used as a cutoff, fewer theories would be supported. We will use a cutoff of 2.0 in this book.

If a slope is negative, the *t*-statistic is negative. In this case, the cutoff is a *t*-statistic of –2.0. A *t*-statistic less than –2.0 (such as –3.0) means it is unlikely that the true slope is zero; thus, a theory stating that the slope were negative would be supported when the *t*-statistic is less than –2.0.

Using a cutoff of 2.0 (or –2.0), an estimated slope is said to be *significant* if it has a *t*-statistic greater than 2.0 (or less than –2.0). Significant means it is unlikely that the true slope is zero. Significant is also used sometimes to refer to a variable that has a significant estimated slope. In our present example, the growth rate is significant because the *t*-statistic of the estimated slope is 4.5, which is greater than the cutoff of 2.0. The results are summarized in Box 2-1. The estimated slope of the line is 0.9, and it has a *t*-statistic of 4.5. Since the *t*-statistic is greater than 2.0, we can say that the growth rate is significant. The standard error of the line, a measure of a typical error, is 4.7, and there are 24 points, or observations.

BOX **2-1**

vote share depends on:		*t*-statistic
0.9	growth rate	4.5
50.9	intercept	51.4
standard error: 4.7		
number of observations: 24		

The term in the box that we have not yet discussed is the *intercept*. The line in Figure 2-1 has both a slope and an intercept. The intercept is the point on the line where the growth rate is zero. When we find the best fitting line, we find both its slope and its intercept. The intercept has a standard error associated with it just like the slope does, and its standard error can be computed in the same way as was done for the slope's standard error. Once the intercept's standard error is computed, its *t*-statistic can be computed. As can be seen in the box, the *t*-statistic of the intercept is 51.4, which is huge. Not surprisingly, the data strongly reject the case that the true intercept is zero. The intercept is 50.9, which means that if the growth rate were zero, the vote share would be 50.9 percent according to the line. The line thus says that the incumbent party wins by a small margin if the growth rate is zero.

As a final point, it is important to see how the size of a *t*-statistic is related to the size of the standard error of the line, both of which are presented in Box 2-1. Consider, for example, a case in which the points in Figure 2-1 lie on average much closer to the line than is actually the case in the figure. In particular, say that the size of a typical error (that is, the standard error of the line) is 1.7 rather than 4.7. This means that the bell-shaped curve in Figure 2-2 is much less spread out. Now, when we are computing the standard error of the estimated slope by drawing many sets of 24 errors (from many universes), the errors are on average smaller. With smaller errors, the computed vote-share values are closer to the line in Figure 2-1. The best fitting line using the new values will thus be closer to the line in Figure 2-1, and so the estimated slope will be closer to 0.9. In other words, the estimated slopes will vary less when the errors that are drawn are smaller, and so the standard error of the estimated slope will be smaller (Figure 2-3 will be less spread out). Therefore, the *t*-statistic of the slope will be larger. This, of course, makes sense. If the errors in Figure 2-1 are on average small, we have more confidence that the slope is not a fluke than if the errors are on average large.

If a slope has a *t*-statistic that is very large in absolute value—much larger than 2.0—it is sometimes said to be *highly significant*. It is also sometimes said to be *precisely estimated*.

We are finally finished with Lesson 4. We have shown how to test whether the true slope of a line is zero given the slope that was computed by finding the best fitting line. The test is to compute the *t*-statistic of the estimated slope and see if it is greater than 2.0 or less than –2.0. If it is, then it is unlikely that the true slope is zero, which is support for the theory. One should not, however, get carried away with this result. The test is not necessarily the final answer, which is the subject matter of Lesson 5.

(Friday) Lesson 5: Think About Pitfalls

Since the *t*-statistic for the growth rate in our example is much greater than 2.0, are we almost positive that the growth rate affects the vote share? Figure 2-4 would suggest yes because there is a very small probability that we would get a slope of 0.9 if the true slope were zero. Alas, life is not quite this simple. Mistakes are possible that would not be caught by the test using the *t*-statistic. The theory may be wrong even with a *t*-statistic of 4.5. The following are possible pitfalls that you need to be aware of even with large *t*-statistics.

Perhaps the truth is that what affects the vote share is not the growth rate but the size of the armed forces. The size of the armed forces is large during wars, and if people rally around the flag during wars, they may be more inclined to vote for the incumbent party during wars than otherwise. Wars also tend to stimulate the economy, and so the growth rate is on average higher during wars than at other times. The growth rate and the size of the armed forces thus tend to move together: when one is high, the other tends to be high, and when one is low, the other tends to be low. (The growth rate and the size of the armed forces are positively correlated.) If this is true, then high growth rates will be associated with large vote shares and low growth rates will be associated with small vote shares. We may thus get a pattern of points like that in Figure 2-1 (namely, an upward pattern), but what is really affecting the vote share is not the growth rate but the size of the armed forces. We would be fooling ourselves in thinking that the growth rate is the cause. We thus have to be careful that the variable we think is affecting the vote share is not in fact acting as a proxy for something else.

It is also possible that *both* the growth rate and the size of the armed forces affect the vote share. In other words, there may be both a rally-around-the-flag effect and a separate growth-rate effect. We discuss how to treat two explanatory variables instead of just one later in this chapter. What we can say now is that if both variables matter but only the growth rate is included, the growth rate will get too much credit regarding its effect on the vote share. The estimated slope will be too large because by omitting the size of the armed forces, the growth rate is picking up part of the effect of the size of the armed forces on the vote share. This problem is called *omitted variable bias*. The estimated slope for an explanatory variable is wrong (biased) because some other explanatory variable has been omitted that truly affects the dependent variable and that is also correlated with the included explanatory variable.

The use of polling results provides a good example of another possible pitfall. Say that for each election one had polling data one week before the election. If in a figure like Figure 2-1 one graphed the vote share against the polling results, there is likely to be an upward pattern and fairly small errors on average around the best fitting line. Polls are usually fairly accurate one week before the election. The slope of the best fitting line is likely to be positive and have a large *t*-statistic. It is not the case, however, that the polling results are *causing* the voters to vote the way they do. The polls are just asking voters one week ahead how they are going to vote. Therefore, good fits and large *t*-statistics do not guarantee that one has explained anything. There can be correlation without causation.

Another possible pitfall is that voting behavior may have changed sometime during the period of the data. Possibly, the true slope was large during the 1916–1944 period and small (perhaps zero) after that. The best fitting line in Figure 2-1 would be based on the incorrect assumption that the true slope is the same for both periods. Although this mistake will make the average error larger than it would be if there were no shift of behavior, the *t*-statistic for the estimated slope may still turn out to be greater than 2.0. Possible shifting behavior is a nightmare for social scientists trying to explain behavior over time because stability is needed to learn very much.

It is also possible that the true relationship is not a straight line in Figure 2-1 but a curved line. Maybe the line begins to curve up at a growth rate of about 5 percent. Incorrectly using a straight line may still result in a positive estimated slope and a *t*-statistic greater than 2.0.

Another interesting way we might be fooled goes under the name of *data mining*. Say that we have observations on 100 possible variables that we think may affect the vote share. We then try each variable, one by one, and see what its estimated slope and *t*-statistic are. Let's say we then pick the variable with the largest t-statistic and conclude that this variable affects the vote share. The potential problem with this procedure should be obvious. By trying so many variables, it is likely that just by chance we have found one that results in a good fit and an estimated slope with a t-statistic greater than 2.0. In other words, even if none of the 100 variables truly affect the vote share, we may find some that by chance fit well.

There are two ways to mitigate the data mining problem. One is to use theory to reduce the number of variables to try. Stupid variables, such as a candidate's eye color, can be eliminated. We can use theory to narrow the list of possible variables to those that have some plausibility. The other way concerns prediction, which is the topic of Lesson 7.

To conclude, one must take any t-statistic with a grain of salt. High t-statistics are not a guarantee of success, and any result must be examined carefully for possible pitfalls.

(Saturday) Lesson 6: Examine the Results

If the tests of a theory reveal that the data do not support it, then this lesson and the next are of no interest. It is of no interest to examine results like the size of the slope or to use the slope to make predictions if the theory is not supported. Similarly, if the possible pitfalls seem really serious, we may not want to continue even if the theory were supported using the tests in Lesson 4. We need to have some confidence in the theory before we can use it.

If the results support a theory and the possible pitfalls do not seem serious, the next step is to examine the implications of the results. In Figure 2-1, for example, it is of interest to see what the line says about the size of the effect of the growth rate on the vote share. If, say, the growth rate increases from 2.0 percent to 3.0 percent, how much does the line say the vote share should increase? We have already answered this question in our discussion of the slope at the end of Lesson 3. The size of the effect is measured by the slope of the line. A steep slope implies a much larger effect than does a mild slope. The slope of the line in Figure 2-1 is 0.9. The line thus tells us that if the growth rate increases by 1.0, the vote share should increase by 0.9. If the slope were instead steeper, say 1.5, then an increase in the growth rate of 1.0 would mean that the vote share should increase by 1.5.

A slope of 0.9 is fairly large in the context of this example. For instance, it can be seen in Table 1-1 in Chapter 1 that the growth rate was –3.7 in 1980 and 5.4 in 1984, which is a difference of 9.1. A change in the growth rate of 9.1 implies, according to the line, that the vote share should change by 0.9 times 9.1, or 8.2, which is a large change in the vote share. It is interesting to note that the actual vote share was 44.8 percent in 1980 (President Carter lost to Reagan) and 59.1 percent in 1984 (President Reagan beat Mondale), a difference of 14.3. The actual change in the vote share was thus larger than 8.2 implied by the line, but the line got quite a bit of the increase.

In our simple example here, all we really have to examine is the size of the slope, which we have done. In most applications, there is more to be done (that is, more implications of the results to consider). We will see examples in the following chapters.

(Sunday) Lesson 7: Predict

Once we have done all the above work, prediction is easy and fun. We must not get too excited, however. The possible pitfalls we discussed in Lesson 5 are always lurking in the shadows, and we must not become overconfident.

Say that we want to use the line in Figure 2-1 to predict the 2012 election, which was not one of the 24 observations (the last election used was 2008). To make a prediction, we must first choose a value for the growth rate. Suppose we are making a prediction in January 2012, and we think the growth rate in 2012 will be 3.0 percent. We know, from Box 2-1, that the intercept is 50.9. This means that the value of the vote share on the line at a growth rate of 0.0 is 50.9. How much will the vote share increase if the growth rate is 3.0 rather than 0.0? Since the slope is 0.9, the vote share increases by 0.9 times 3.0, or 2.7. The vote share on the line at a growth rate of 3.0 is thus 50.9 plus 2.7, or 53.6. The incumbent party (Democratic) is thus predicted to win with 53.6 percent of the vote for a growth rate of 3.0 percent.

Say instead you thought in January 2012 that there was going to be a recession in 2012, and the growth rate was going to be –3.0. In this case, the predicted vote share is 50.9 minus 2.7, or 48.2, which would mean a Republican victory.

It is thus easy to make a prediction. Pick a growth rate, and find out where on the line in Figure 2-1 you are for this growth rate. Note that any prediction assumes that the error for 2012 will be zero (that is, the vote share will be exactly on the line). This is where one needs to be cautious. A typical error is 4.7 (the standard error of the line), so the actual outcome could differ from the predicted value by quite a bit. The standard error of 4.7 incorporates all the factors that affect voting behavior other than the growth rate, and these other factors have on average an effect of 4.7 on the vote share. You can see from Figure 2-2 that 68.2 percent of the time the error will be between –4.7 and 4.7. So with a predicted vote share of 53.6 percent, we can say that 68.2 percent of the time the actual vote share will be between 48.9 and 58.3 percent. This is, of course, a fairly wide range. (We will see in the next chapter that the range narrows considerably when we depart from assuming that the vote share depends *only* on the growth rate.)

In addition to taking into account the size of the standard error of the line when thinking about a prediction, we must also be cautious about the possible pitfalls from Lesson 5. If any of the pitfalls are relevant, we will at a minimum be using a standard error that is too small.

Prediction can help us see if the data mining problem is serious, at least if we are willing to wait for another observation. Say that by trying many variables we have found a variable that fits the 24 elections very well. If the truth is that this variable has no effect on the vote share, then the line that we have chosen is not likely to predict the next election well. We have searched using the 24 elections to find a line that looks good, but it is in fact spurious. Because the line is spurious, there is no reason to think it will do well for the next election, since the next election's outcome has not been used in the search. Observing how well a line predicts outside the period of the fit is thus a way of checking for possible data mining problems.

We have one more very important point to make about prediction. Remember that all we need to make a prediction of the vote share is a value for the growth rate. We used values of 3.0 and –3.0 earlier, but we could easily use other values. Once the actual growth rate for 2012 is known, we can also use it. The important point is that in terms of testing the theory, the actual growth rate should be used, not just some prediction of the growth rate. We want to compare the actual vote share for

2012 with the predicted vote share using the actual growth rate for 2012. Using any other growth rate would not be a test of the theory because the test would be based on an incorrect growth rate.

Finally, note that a prediction can be made for any observation in the sample period. That is, a prediction can be made for any of the 24 elections in Table 1-1 once the slope and intercept are estimated. This can be done because we know the actual value of the growth rate for each election. We just take the intercept, which is 50.9 from Box 2-1, and add to it 0.9 times the actual growth rate, where 0.9 is the estimated slope in Box 2-1. We in fact did this earlier when discussing the 1992 election. The predicted value for President Bush in 1992 is 50.9 plus 0.9 times 2.9, which is 53.5, where 2.9 is the actual value of the growth rate. The actual value of the vote share is 46.4 percent, and so the error is 7.1 percent. We can thus say that an error for any observation is simply the difference between the predicted value and the actual value.

Adding More Variables

The lessons in this chapter have been presented under the assumption that there is just one explanatory variable—the growth rate. We now must relax this assumption and consider more than one explanatory variable. In practice, there is almost always more than one explanatory variable. The theory of voting behavior outlined in Chapter 1, for example, is not that the vote share depends *only* on the growth rate. Other variables that were put forward are inflation, the number of good news quarters, whether the president is running for reelection, and a measure of duration. Fortunately, it is fairly easy to extend the analysis in this chapter to more than one explanatory variable. As an example, let's assume that the vote share depends on both the growth rate and inflation, as shown in Box 2-2.

BOX **2-2**

vote share depends on:	
	growth rate
	inflation
	intercept

When there is more than one variable, we can no longer use a graph like Figure 2-1 to help see what is going on. The line in Figure 2-1 is determined by two numbers, the intercept and the slope, but now we have three numbers to determine. From now on, we will use the word *coefficient* instead of slope to refer to the size of the effect associated with a variable. Using this terminology, the three numbers we have to determine in the present example are (1) the intercept, (2) the coefficient for the growth rate, and (3) the coefficient for inflation. Keep in mind that a coefficient is just a number, like 0.9.

The line that fits the points best in Figure 2-1 is the one that has the smallest sum of squared errors. We imagined a computer trying thousands of lines, computing the sum of squared errors for each line, and choosing the line with the smallest sum. Each line is characterized by a value for the intercept and a value for the slope (that is, a value for the coefficient for the growth rate).

For the present example, imagine the computer choosing three numbers: the intercept, the coefficient for the growth rate, and the coefficient for inflation. Given these three numbers, the computer can compute the error for each of the 24 elections. Consider, for example, the 1916 election, and assume that the three numbers are 51.0, 1.0, and –1.0. From Table 1-1 in Chapter 1, we see that the actual vote share for 1916 is 51.7, the growth rate is 2.2, and inflation is 4.3. Using the three numbers we have chosen, the predicted vote share for 1916 is 51.0 plus 1.0 times the growth rate of 2.2 and –1.0 times inflation of 4.3. The predicted vote share is thus 51.0 plus 2.2 minus 4.3 equal 48.9. The error for 1916 is then the predicted value of 48.9 minus the actual value of 51.7, or –2.8. Using the same three numbers, the computer can compute the errors for the other 23 elections in the same manner. Each error is the difference between the predicted value for that election and the actual value. These predictions are based on using the actual values of the growth rate and inflation. Once the 24 errors are computed, they can be squared and then summed.

We have so far seen that we can go from three numbers—the intercept, the coefficient for the growth rate, and the coefficient for inflation—to a value for the sum of squared errors. Now consider the computer doing this thousands of times, each time for a different set of three numbers, and in the end choosing the set of three numbers that has the smallest sum of squared errors associated with it. This best fitting set of three numbers is the analog of the best fitting line when there is only one explanatory variable.

Once the best fitting set of coefficients is found, we can compute the standard error. We first divide the sum of squared errors, which for our example turns out to be 486.1, by the number of observations (24) less the number of coefficients (3), or 21. This gives 23.1. We then take the square root to get 4.8. The standard error (a measure of the size of a typical error) is thus 4.8.

The thought experiment in Lesson 4 to derive the standard error of the estimated slope can be modified to incorporate more than one explanatory variable. For each set of 24 drawn errors (that is, for each universe), a best fitting set of coefficients is computed. After, say, 1,000 sets have been computed, a figure like Figure 2-3 can be drawn for each coefficient in the set. The standard error for a coefficient can then be determined by sweeping out 34.1 percent of the area as shown in Figure 2-3. The *t*-statistic for a coefficient is the coefficient divided by its standard error.

The results for the current example are presented in Box 2-3. The coefficient for the growth rate is 0.9, with a *t*-statistic of 4.8. Since the *t*-statistic is greater than 2.0, it is unlikely that the true coefficient for the growth rate is zero. The growth rate is significant. The coefficient for inflation is 0.1, with a *t*-statistic of 0.2. The *t*-statistic is very small, and so inflation is not significant. Also, the coefficient is positive, contrary to what theory says. There is thus no support for the theory that inflation has a negative effect on vote share. We will return to this result in the next chapter, where we will see that inflation is in fact significant with a negative coefficient when other explanatory variables are added. The main point here is that it is possible to compute the coefficient and its associated *t*-statistic for each explanatory variable. There is nothing new in principle here from the case of just one explanatory variable.

BOX **2-3**

vote share depends on:		*t*-statistic
0.9	growth rate	4.8
0.1	inflation	0.2
50.6	intercept	30.1
standard error: 4.8		
number of observations: 24		

Adding a third variable is also straightforward. There are then four coefficients to be computed instead of three, but all else is the same. In practice, it is quite easy for a computer to find the set of coefficients that leads to the smallest sum of squared errors. There are faster ways than the search procedure just discussed, but the answer is the same either way.

As a final point, we will sometimes use the phrase *other things being equal* in this book. This means we are engaging in the thought process of changing one explanatory variable without changing any of the others. For example, we might want to think about what happens to the vote share if the growth rate changes but inflation does not. We might then say that if the growth rate changes by such and such, the vote share will change by such and such, other things being equal. Other things being equal would mean that inflation does not change, nor does anything else that might affect the vote share.

Testing for Other Variables

We saw in the last section that each explanatory variable has associated with it a t-statistic. If the t-statistic for a particular variable is greater than 2.0 or less than –2.0, it is unlikely that the coefficient for the variable is zero. If the coefficient for the variable is not zero, then the variable has an effect on whatever is being explained, such as the vote share. It is thus easy to test whether a variable is supported by the data by including it in the fitting process and seeing if its t-statistic is greater than 2.0 or less than –2.0. This is a way of testing whether variables belong in the explanation (that is, whether variables are significant).

To give an example, say we are interested in whether the size of the armed forces affects the vote share. It may be that both the growth rate and the size of the armed forces affect the vote share. We can test for this by simply including the size of the armed forces along with the growth rate in the fitting process and seeing what the size of the t-statistic for the armed forces variable is.

Let's return now to possible pitfalls. Since it is so easy to try different variables or different sets of variables to see if they have large t-statistics, you can see why possible data mining is such a concern. Trying 100 combinations of variables and choosing the combination that leads to the best fit increases the chance that the chosen combination is spurious (that is, it is just a fluke). Again, theory needs to narrow the list of possible combinations.

We can now also be more precise about the problem of omitting the size of the armed forces from the explanation if it really belongs. If the growth rate tends to be high when the size of the armed forces is high and low when the size of the armed forces is low, and if both variables affect the vote share, then omitting the size of the armed forces will lead to a coefficient for the growth rate that is too large. For example, the 0.9 coefficient for the growth rate in Box 2-3 would be picking up the true effect of the growth rate on the vote share, which might be 0.6, and part of the effect of the size of the armed forces on the vote share, which might be 0.3. This would be an example of omitted variable bias, as mentioned in Lesson 5.

It is important to be clear on when there might be omitted variable bias. If an explanatory variable is omitted that truly affects the dependent variable, there will be no bias for the included explanatory variables if the omitted variable is uncorrelated with the included variables. The fit will not be as good because an explanatory variable has been omitted, but this will not bias the estimated effects of the other variables if they are not correlated with the omitted variable. The included variables will not be picking up any of the effect of the omitted variable because there is no correlation. The omitted variable needs to be correlated with the included explanatory variables for there to be omitted variable bias.

Horse Races

Many times in social science research we have two or more competing explanatory variables. For example, perhaps it is not the growth rate that affects the vote share but instead the change in the unemployment rate. These two variables are highly correlated because when the growth rate is high, the unemployment rate tends to decrease and vice versa. The two variables are not, however, perfectly correlated, and either variable is a plausible candidate for affecting votes for president.

Fortunately, it is easy to test between the growth rate and the change in the unemployment rate. We simply include both variables in the fitting process and compute the *t*-statistic for each. If one is significant (*t*-statistic larger than 2.0 or smaller than –2.0) and the other one is not, the one variable has dominated the other. We run, in other words, a *horse race* to see which variable dominates. We will perform a number of horse races in this book. A horse race is an effective way of allowing the data to choose which variable to keep. It is possible, of course, that both variables are significant or that neither variable is significant. In these two cases, the horse race is a dead heat (that is, the test is inconclusive).

Sensitivity or Robustness Checks

Related to horse races is the procedure of adding variables that you think from theory should not matter and thus should not be significant. If such variables are in fact significant, there is a possible problem with your theory. It is even a more serious problem if such variables are significant and adding them results in large changes in the other coefficients. In these cases, the results are said to be *sensitive* or to lack *robustness*. Adding variables that you think are unimportant is thus a way of looking for pitfalls. If the results are sensitive to such additions, this is a pitfall to worry about.

Another sensitivity check on possible data mining problems is the following. Say that we have found that the growth rate in the three-quarter period before an election is a significant explanatory variable in explaining the vote share in that election. This could be a fluke in the sense that we may have tried many variables before arriving at this one, and just by chance this one turned out to be significant. If the result is truly a fluke, then replacing the three-quarter growth rate with, say, the two-quarter growth rate or four-quarter growth rate should give very different results. On the other hand, if the growth rate does in fact affect the vote share, we would expect similar results using slightly different measures of the growth rate. For example, the two-quarter growth rate should give results similar to those using the three-quarter growth rate. Some sensitivity work is thus to replace certain explanatory variables with others that are similar to see how the results change. If the changes are large, this may be a cause for concern. We would like our results to be robust to small changes in the explanatory variables. More generally, we would like our results to be robust to small changes in the assumptions behind our theory and to small changes in the choice of data used.

Conclusion

We have covered a lot of material in this chapter, and if you have not seen any of it before, everything may not have completely sunk in. A good way to get an understanding of the tools is to see them in use, which is what the rest of this book is about. In the next chapter, we test the theory of voting behavior outlined in Chapter 1. In the present chapter, we simplified the theory to help in understanding the tools, but from now on, we will cut no corners.

As we go through each chapter, keep in mind the seven lessons: proposing a theory, collecting data, fitting the data, using the data to test the theory, thinking about pitfalls, examining the results, and making a prediction.

3 Presidential Elections

How can I, that girl standing there,
My attention fix
On Roman or on Russian
Or on Spanish politics?
Yet here's a travelled man that knows
What he talks about,
And there's a politician
That has read and thought,
And maybe what they say is true
Of war and war's alarms,
But O that I were young again
And held her in my arms!

William Butler Yeats, Politics

A Theory of Voting Behavior

The theory of voting behavior that we are going to test is discussed in Chapter 1. We summarized the theory in Box 1-1, which is repeated here as Box 3-1. The dependent variable is the vote share, and the explanatory variables are the other five variables in the box.

BOX 3-1

vote share depends on:	
	growth rate
	inflation
	good news quarters
	president running
	duration

In Chapters 1 and 2, we have defined the vote share to be the incumbent party's share of the two-party vote (Democratic plus Republican). In this chapter, we will instead define the vote share to be the *Democratic* share of the two-party vote. This requires changing the sign of the economic variables when the incumbent party is Republican. If, for example, there is a large growth rate and the Republicans are in power, this has a positive effect on the Republican vote share and so a negative effect on the Democratic vote share. Therefore, if the Democratic share of the two-party vote is the dependent variable, the growth rate variable when the incumbent party is Republican should be the negative of the actual growth rate. A 4 percent growth rate would be –4 percent. We will continue to call the explanatory variables the growth rate, inflation, and good news quarters, but remember that the variables are the negative of the actual values when the incumbent party is Republican. This change makes no difference to the theory and results. For example, the coefficients of the economic variables corresponding to the best fit are the same, as are their *t*-statistics. The standard error is also the same. It is just sometimes easier to talk about the Democratic share across all elections than the incumbent-party share, so we will use the Democratic share.

We have discussed the theory in Chapter 1, and so we can move immediately to the data.

The Data

We are going to use the presidential elections from 1916 to 2008 to test the theory. The relevant data are listed in Table 1-1. Since any test of a theory is no better than the data behind it, we need to be clear regarding the exact variables that were used. We will discuss each of the variables in turn. Refer back to Table 1-1 as the various variables are discussed.

As mentioned in the previous section, the vote share that is used is the Democratic share of the two-party vote. For reference purposes, Table 1-1 lists both the incumbent-party share of the two-party vote and the Democratic share of the two-party vote. When the Democratic Party is the incumbent, the two shares are the same; otherwise the incumbent-party share is 100 minus the Democratic-Party share. The vote share is the two-party share, not the share of the total vote. Which vote share should be used depends on who the third-party voters are. To take an example, assume that before a third-party candidate came along, 30 percent of the voters were staunch Republicans, 30 percent were staunch

Democrats, and 40 percent were swing voters who were influenced by the economy. Assume also that the economy was neutral in the sense that half the swing voters were for the Democratic Party and half were for the Republican Party. In this case, the vote share for the Democratic Party would be 50 percent. This is, of course, both the total vote share and the two-party vote share, since there is no third party.

Now assume that a third-party candidate comes along and takes half of the swing voters who would have voted Republican and half who would have voted Democratic. In this case, both major parties get 40 percent of the total vote, with the third party getting 20 percent. The two-party vote share for both major parties is still 50 percent. The third-party candidate has thus lowered the total vote share for the Democratic Party from 50 percent to 40 percent, but the two-party vote share remains the same at 50 percent. It is thus clear that a third-party candidate can affect the total vote share in a big way, but the effect on the two-party vote share may be modest, depending on how much the candidate takes from one party versus the other. By using the two-party vote share, we are in effect assuming that third-party candidates take about the same amount from each major party.

The one exception to the use of the two-party vote share is the election of 1924. There is some evidence that LaFollette (the third-party candidate, running for the Progressive Party) took more votes from Davis (the Democrat) than from Coolidge (the Republican). It has been estimated (see the notes to this chapter) that 76.5 percent of the votes for LaFollette would otherwise have gone for Davis, with the remaining 23.5 percent going for Coolidge. The vote share for the Democratic Party was thus taken to be the number of votes that Davis got plus 76.5 percent of the votes that LaFollette got divided by the total number of votes for all three.

The growth rate that is used is the growth rate in the first three quarters (nine months) of the election year. Also, it is the *per capita* growth rate of real GDP, the growth rate of output per person. If the economy is growing only at the rate that population is growing, this is obviously less beneficial to the average person than if output per person is growing.

The measure of inflation is the average inflation rate over the 15 quarters prior to the election (that is, all the quarters of the administration except the last one). The variable used for inflation is the percentage change in the GDP deflator. As noted in Chapter 1, inflation and deflation are treated symmetrically: deflation is assumed to be just as bad in voters' minds as inflation.

The good news quarters variable is the number of quarters out of the 15 quarters before the election in which the growth rate exceeded 3.2 percent. (Remember, this is the per capita growth rate.) These are quarters in which the economy did exceptionally well, and the theory is that voters remember these kinds of events.

If the president is running for reelection, the president running variable is given a value of 1; otherwise, the value is 0. Vice presidents who became president during the administration were given a 1 if they ran again, except for Ford. Ford was given a 0 because he was not on the elected ticket but was appointed vice president when Agnew resigned. As with the economic variables, the president running variable is taken to be the negative of the value if the incumbent party is Republican.

The duration variable, which is listed in Table 1-1, is given a value of 0.0 if the incumbent party has only been in office for one consecutive term, 1.0 for two consecutive terms, 1.25 for three consecutive terms, 1.5 for four consecutive terms, and 1.75 for five consecutive terms. The duration variable is taken to be the negative of the value if the incumbent party is Republican.

This completes the discussion of the variables in Box 3-1, but there are two other variables that need to be mentioned. First, a party variable was included that has a value of 1 if the incumbent party is Democratic and –1 if the incumbent party is Republican. This variable is listed in Table 1-1. It tests whether there is a pure party effect. Second, the elections of 1920, 1944, and 1948 were treated somewhat differently because of the wars. The period prior to the 1920 election was dominated by World War I, and the periods prior to the 1944 and 1948 elections were dominated by World War II. The inflation variable and the good news quarters variable were assumed not to be relevant for these three elections. In other words, voters were assumed not to take into account past inflation and good news quarters when deciding how to vote during the three war-dominated periods. This treatment requires that a war variable be included that has a value of 1 for the 1920, 1944, and 1948 elections and a value of 0 otherwise. It also means that the values used for inflation and good news quarters for these three elections are 0, not the values listed in Table 1-1.

Fit and Test and Examine

We are now ready to see how well the explanatory variables do in explaining the vote share. Their coefficients are determined in the manner discussed in the last chapter. The set of coefficients is found that results in the smallest sum of squared errors, and the t-statistics are computed. The data used are from Table 1-1, with negative values used for the three economic variables, the president running variable, and the duration variable when the incumbent party is Republican. The results are shown in Box 3-2. In this case, there were eight coefficients to determine, one for each of the eight variables in the box (counting the intercept). There are thus eight corresponding t-statistics.

BOX **3-2**

vote share depends on:		t-statistic
0.67	growth rate	6.22
−0.65	inflation	−2.31
0.99	good news quarters	4.31
2.9	president running	2.18
−3.4	duration	−2.87
−1.9	party variable	−0.84
5.1	war variable	1.99
47.4	intercept	77.56
standard error: 2.5		
number of observations: 24		

It will be convenient after considering how well we have fit the data to discuss the size of the coefficients at the same time as we discuss their t-statistics. If we were exactly following the lessons in Chapter 2, we would not discuss the size of the coefficients until Saturday, after we had discussed the t-statistics and thought about possible pitfalls. Only if the t-statistics look good and the possible pitfalls seem minor should we care about the coefficients. We will be jumping ahead only for convenience of discussion.

Let's begin with the standard error in the box, which is 2.5. A typical error is thus 2.5, which is fairly small. As discussed in the last chapter,

TABLE 3-1 Actual and predicted Democratic vote share

Year	Party in Power	Election Outcome	Actual Vote Share (%)	Predicted Vote Share (%)	Predicted Minus Actual (%)
1916	D	President Wilson beat Hughes	51.7	50.1	−1.6
1920	D	Cox lost to Harding	36.1	39.4	3.3
1924	R	President Coolidge beat Davis and LaFollette	41.7	42.4	0.7
1928	R	Hoover beat Smith	41.2	42.8	1.5
1932	R	President Hoover lost to F. Roosevelt	59.1	61.2	2.0
1936	D	President F. Roosevelt beat Landon	62.2	63.6	1.4
1940	D	President F. Roosevelt beat Willkie	55.0	55.5	0.5
1944	D	President F. Roosevelt beat Dewey	53.8	52.0	−1.7
1948	D	President Truman beat Dewey	52.3	50.8	−1.5
1952	D	Stevenson lost to Eisenhower	44.7	45.4	0.7
1956	R	President Eisenhower beat Stevenson	42.9	43.6	0.7
1960	R	Nixon lost to Kennedy	50.1	48.7	−1.4
1964	D	President Johnson beat Goldwater	61.2	60.9	−0.3
1968	D	Humphrey lost to Nixon	49.4	50.3	0.9
1972	R	President Nixon beat McGovern	38.2	41.5	3.3
1976	R	Ford lost to Carter	51.0	50.2	−0.9
1980	D	President Carter lost to Reagan	44.8	45.7	0.9
1984	R	President Reagan beat Mondale	40.9	38.2	−2.6
1988	R	G. Bush beat Dukakis	46.2	49.2	3.0
1992	R	President G. Bush lost to Clinton	53.6	48.8	−4.8
1996	D	President Clinton beat Dole	54.7	53.2	−1.5
2000	D	Gore lost to G. W. Bush	50.3	49.3	−1.0
2004	R	President G. W. Bush beat Kerry	48.8	45.5	−3.3
2008	R	McCain lost to Obama	53.7	55.2	1.5

we know that 68.2 percent of the time the error in predicting the vote share will be between –2.5 and 2.5.

We can get a more detailed picture of the errors by looking at Table 3-1. The table shows, for each of the 24 elections, the actual vote share, the predicted vote share, and the error (predicted minus actual). These predictions are based on the actual values of the explanatory variables. For each election, each coefficient in Box 3-2 was multiplied by the actual value of the respective variable for that coefficient, with the answers then added to get the predicted value for that election. If this is not completely clear, we will go over this procedure in more detail later in this chapter when we make a prediction for the 2012 election. You may also want to review the discussion of predictions at the end of Lesson 7 in Chapter 2.

The largest error in absolute value in Table 3-1 is for the 1992 election, when President George Bush got 48.8 percent of the vote and was predicted to get 53.6 percent, an error of –4.8 percentage points. The next three largest errors in absolute value are for the elections of 1920, 1972, and 2004. In 1920, Cox was predicted to get 39.4 percent and got 36.1 percent, an error of 3.3 percentage points. In 1972, McGovern was predicted to get 41.5 percent and got 38.2 percent, also an error of 3.3 percentage points. And in 2004, Kerry was predicted to get 45.5 percent and got 48.8 percent, an error of –3.3 percentage points. All but 3 of the remaining 20 errors are less than 2 percentage points in absolute value.

Not counting 2000, three elections were predicted incorrectly as to the winner: 1960, 1968, and 1992. In 1960, Kennedy got 50.1 percent of the vote, a win, but was predicted to get 48.7 percent, an error of –1.4. In 1968, Humphrey got 49.4 percent of the vote, a loss, but was predicted to get 50.3 percent, an error of 0.9. Even though the winner was predicted incorrectly in these two elections, the errors are small, so in this sense the elections were predicted well. The 1992 error, on the other hand, was large, and the winner was also predicted incorrectly. For the 2000 election, Gore got 50.3 percent of the vote and was predicted to get 49.3 percent, an error of –1.0. In this case, the vote-share prediction was on the wrong side of 50 percent, predicting a Gore loss regarding the two-party vote share. In fact, Gore won the two-party vote share but lost the election in the electoral college.

Figure 3-1 plots the last column of Table 3-1. The errors are plotted for each election starting from the earliest. The large positive error for 1992 stands out in the figure.

Having examined the errors, let's now turn to the coefficients and their associated *t*-statistics. The coefficient for the growth rate is 0.67, which says that if the growth rate increases by 1.0, the vote share is predicted to increase by 0.67. The coefficient for inflation is –0.65, which says that if inflation increases by 1.0, the vote share is predicted to decrease by 0.65. The growth rate and inflation thus have similar effects on

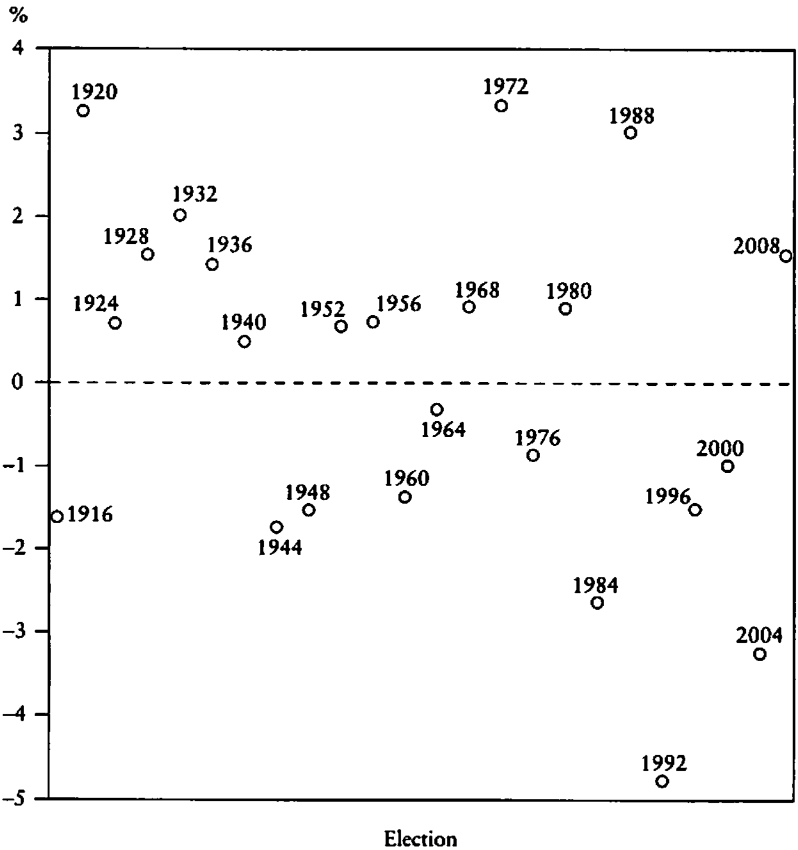

FIGURE 3-1 Predicted minus actual vote share

the vote share except that one effect is positive and the other is negative. The t-statistic for the growth rate is 6.22, and since this value is considerably greater than 2.0, it is highly unlikely that the true coefficient for the growth rate is zero. In other words, it is quite likely that the growth rate affects the vote share. The t-statistic for inflation is –2.31, and since this value is less than –2.0, it is unlikely that the coefficient for inflation is zero. In other words, it is likely that inflation affects the vote share. The growth rate and inflation are thus both significant.

The significance of inflation in Box 3-2 is contrary to the results in Box 2-3, where inflation is not significant. This is an example of omitted variable bias. Box 2-3 has omitted a number of significant explanatory variables, namely, the additional variables in Box 3-2, and these omitted variables are correlated with inflation in such a way as to lead the coefficient on inflation in Box 2-3 to be too small in absolute value. Once the variables are added, the inflation coefficient increases in size in absolute value and becomes significant.

The coefficient for good news quarters is 0.99, which means that each additional good news quarter is predicted to increase the vote share by 0.99. This is a fairly large effect. The *t*-statistic for good news quarters is 4.31, so it is also unlikely that the true coefficient for good news quarters is zero.

The coefficient for the president running variable is 2.9, which means that a president running for another term is predicted to have a head start of 2.9 percentage points. The *t*-statistic is 2.18, so it is unlikely that the true coefficient is zero. If the true coefficient were zero, the president running would have no head start.

The duration variable has a coefficient of –3.4 with a *t*-statistic of –2.87. If a party has been in power for two consecutive terms, it is predicted to start out behind by 3.4 percentage points because the duration variable has a value of 1.0 in this case. If a party has been in power for three consecutive terms, it is predicted to start out behind by 3.4 times 1.25, or 4.25 percentage points, because the duration variable has a value of 1.25 in this case. For four consecutive terms, the value is 3.4 times 1.50, or 5.1 percentage points, and so on. The duration variable is significant because its *t*-statistic is less than –2.0.

The coefficient for the party variable is –1.9 with a *t*-statistic of –0.84. Remember that the party variable has a value of 1 if the Democratic Party is the incumbent party and –1 if the Republican Party is the incumbent party. The variable is not significant because the *t*-statistic is less than 2.0 in absolute value. It could be deleted without changing the results much. If it is retained, the coefficient of –1.9 means the following. Say that all the other variables have a value of zero. Then if the incumbent party is Democratic, the predicted Democratic vote share is the intercept, 47.4, minus 1.9, or 45.5. If the incumbent party is Republican, the predicted Democratic vote share is the intercept, 47.4, plus 1.9, or 49.3. The predicted Republican vote share is 100 minus the predicted Democratic vote share, or 50.7 when the Republican Party is the incumbent party.

Therefore, other things being equal, the Republican Party starts off when it is the incumbent party with a predicted vote share that is larger than the predicted vote share for the Democratic Party when it is the incumbent party (50.7 versus 45.5). If the party variable is dropped and the coefficients are reestimated, the intercept is 47.5. This implies that if all the other variables have a value of zero, the Democratic Party starts off when it is the incumbent party with a predicted vote share of 47.5 and the Republican Party starts off when it is the incumbent party with a predicted vote share of 100 minus this, or 52.5. Either with or without the party variable included, the results show a bias in favor of the Republicans, other things being equal.

The coefficient for the war variable is only relevant for three elections. Its t-statistic is 1.99. The war variable is thus significant if we use a cutoff of 2.0 and round 1.99 up to 2.0.

Overall, the results seem very good. The errors are small except for 1992, and all the variables are significant except for the party variable. In particular, the t-statistics for the economic variables (the growth rate, inflation, and good news quarters) suggest that the economy does have an effect on the vote share; it is very unlikely that we would get these t-statistics if the economy did not affect the vote share. We cannot, however, relax because of the possible pitfalls lurking in the background. To these we now turn.

Possible Pitfalls

The main pitfall that we need to worry about is the possibility of data mining. Many variables were tried in arriving at the final results, and we have only 24 observations. It may be that by chance we have fit the data well, but in fact, the vote share is determined by other things. The following are the main things that were tried that may be subject to the data mining problem.

- Increments other than 0.25 were tried for the duration variable, and 0.25 was chosen because it gave the best results in terms of fit.
- Values other than 3.2 percent were tried for the cutoff for good news quarters, and 3.2 was chosen because it gave the best results in terms of fit.

- The particular treatment for the wars for the three elections was done because this led to an improved fit.
- Different periods for the growth rate were tried, and the particular one chosen, the first three quarters of the election year, gave the best results in terms of fit.
- Different periods for inflation were also tried, and the particular one chosen, the entire four-year period except for the last quarter, gave the best results in terms of fit.
- After the large error was made in 1992, an attempt was made to find reasons for it. This effort led to the choice of the good news quarters variable, which prior to 1992 had not been thought of. The good news quarters variable helps make the error for 1992 smaller because, as you can see from Table 1-1, there were only two good news quarters for the George Bush administration. President Bush is still predicted to win in 1992 in Table 3-1, but by less than he would be predicted to if it were not for the good news quarters variable.

With only 24 elections and all this searching, it is certainly possible that the results in Box 3-2 are a fluke and are not really right. As discussed in Chapter 2, one way of examining the seriousness of the data mining problem is to see how well future observations are predicted. If the results are a fluke, future predictions should not in general be very accurate. In particular, if the results are a fluke, the prediction for the 2012 election is not likely to be accurate, since no information about this election was used in getting the results. The prediction of the 2012 election is discussed later in this chapter.

An alternative approach to examining the data mining problem is to use only part of the observations to get the coefficients and then see how well these coefficients do in predicting the other observations. This is not as good a check as waiting because we have used information in the whole sample (both parts) to decide which variables to include, but at least the coefficients are obtained using only the information in the first part of the observations.

To perform this check, the best fitting set of coefficients was obtained using only the elections through 1960. In other words, the best fit was obtained for the 1916–1960 period (12 elections), and no data from

1964 or later were used. The best fitting coefficients for this set of observations are shown in Box 3-3. These coefficients obviously differ from the earlier ones because they are based on only 12 observations rather than 24. The coefficient for the growth rate still has a large *t*-statistic, and its value has changed only slightly—from 0.67 to 0.80. The *t*-statistic for the inflation coefficient, however, is now only –1.15. The inflation coefficient has changed from –0.65 to –0.42. The results are thus weak for inflation affecting the vote share if the observations used stop in 1960.

BOX 3-3

Estimation period 1916–1960		
vote share depends on:		*t*-statistic
0.80	growth rate	7.78
–0.42	inflation	–1.15
0.72	good news quarters	2.99
5.0	president running	3.99
–2.2	duration	–2.42
–2.9	party variable	–1.35
4.3	war variable	1.85
46.4	intercept	79.58
standard error: 1.5		
number of observations: 12		

Perhaps the most important result is that the *t*-statistic for good news quarters is still fairly large (2.99). Even though the good news quarters variable was not formulated until after the 1992 election, its significance does not depend on the 1992 observation. Even stopping in 1960, the results say that a zero coefficient for goods news quarters is unlikely. The size of the coefficient has fallen from 0.99 to 0.72, although it is still significant.

We can use this new set of coefficients to predict the outcomes of the elections after 1960. These predictions are presented in Table 3-2. They are based on using the actual values of the explanatory variables, as are those in Table 3-1. The predictions in the two tables differ because they are based on the use of different coefficients. If data mining is a

TABLE 3-2 Actual and predicted Democratic vote share: 1960 coefficients

Year	*Party in Power*	*Election Outcome*	*Actual Vote Share (%)*	*Predicted Vote Share (%)*	*Predicted Minus Actual (%)*
1964	D	President Johnson beat Goldwater	61.2	59.2	−2.0
1968	D	Humphrey lost to Nixon	49.4	49.1	−0.3
1972	R	President Nixon beat McGovern	38.2	38.7	0.5
1976	R	Ford lost to Carter	51.0	48.0	−3.0
1980	D	President Carter lost to Reagan	44.8	45.9	1.0
1984	R	President Reagan beat Mondale	40.9	36.4	−4.5
1988	R	G. Bush beat Dukakis	46.2	48.1	1.9
1992	R	President G. Bush lost to Clinton	53.6	44.6	−9.0
1996	D	President Clinton beat Dole	54.7	53.1	−1.6
2000	D	Gore lost to G. W. Bush	50.3	47.3	−3.0
2004	R	President G. W. Bush beat Kerry	48.8	43.0	−5.8
2008	R	McCain lost to Obama	53.7	53.9	0.2

serious problem, the predictions in Table 3-2 should not in general be very good because the coefficients have been estimated using only data through 1960.

The predictions in Table 3-2 are in fact fairly good. The largest error is for 1992, where it is –9.0. The next largest error is for 2004, when Kerry got 48.8 percent of the vote and was predicted to get only 43.0 percent, an error of –5.8. The error for 1984 is –4.5 percent, and otherwise, the errors are less than or equal to 3 percentage points in absolute value. Remember that by the 2008 election, the prediction was based on a set of coefficients that was chosen using data ending 48 years earlier!

Sensitivity checks to small changes in some of the assumptions have also been made. The coefficients are fairly robust to (1) the use of 2.7 or 3.7 percent instead of 3.2 percent as the cutoff for good news quarters, (2) the use of 0.00 or 0.50 instead of 0.25 as the increment for the duration variable, and (3) counting Ford as an incumbent running again for the president running variable. The results are more sensitive to the treatment of the two world wars. If the adjustment for the wars is not made, the t-statistic for inflation falls in absolute value to –1.67, although both the growth rate and good news quarters remain significant with only slightly smaller coefficients. The fits are worse if the growth rate is only for the second and third quarters of the election year or for the four quarters before the election, but the growth rate always remains highly significant. The inflation variable loses its significance if only 11 quarters or only 7 quarters before the election are used instead of 15, although its coefficient is always negative. The results are thus fairly robust to small changes in some of the assumptions.

The analysis in this section thus suggests that data mining may not be a serious problem, but one can never be sure. The results need to be interpreted with some caution. This is particularly true in light of the large error for 1992. Perhaps 1992 is just an unusual draw—something unlikely to happen very often—but it could also signal something more serious. It is encouraging, however, that for the four elections since 1992, the errors are modest in Table 3-1.

What About the Electoral College?

Nothing has been said so far about the electoral college. The dependent variable (that is, the variable to be explained) has been taken to be the Democratic share of the two-party vote, not the Democratic share in the electoral college. A candidate can get more than 50 percent of the two-party vote and yet lose the election in the electoral college, as happened to Gore in 2000. This also happened in 1876 (Hayes versus Tilden) and in 1888 (Harrison versus Cleveland).

The aim of our analysis is to explain voting behavior in the sense of explaining how many votes one party gets relative to the other in the whole country. The aim is not to explain how many states go for one party over the other.

Nor is the aim of the analysis to explain who wins the election. The theory is judged by how close the predicted values of the two-party vote share are to the actual values (that is, by the size of the errors). Consider two hypothetical elections in which the incumbent party is the Democratic Party and in which it wins both times. Say the party's vote share is 61 percent in the first election and 51 percent in the second. Say also that the predicted vote share is 54 percent for the first election and 49 percent for the second. The winner was thus correctly predicted for the first election, but the error was fairly large at –7 percentage points; much of the large margin for the Democrats was not predicted. On the other hand, the winner was incorrectly predicted for the second election, but the error was small at 2 percentage points. The election was predicted to be close and it was. The second election is thus predicted better than the first even though the winner was predicted incorrectly in the second. As a social scientist trying to explain the vote share, I care about the size of my errors, not about who wins. This point is not always easy to get across to reporters, and I sometimes sound uncaring.

Do Campaigns Matter?

Nothing has been said about campaigns. Does this mean that campaigns don't matter, and all that matters are the incumbency information and the economic variables? What if one party campaigned and the other did not? Wouldn't this make a difference? This is a commonly asked question, especially by those considering working on a campaign. A related question that is sometimes asked is whether it matters who is nominated. What if some extreme left- or right-wing candidate were nominated? Wouldn't this have a big effect on the outcome?

The answer is that campaigns are likely to matter and that the nomination of an extreme candidate is likely to make a difference. This does not, however, call into question the above analysis, and the reason is somewhat subtle to grasp. Remember what is being explained—the voting that actually takes place on Election Day. It seems safe to say that in all 24 elections, both parties campaigned hard. Each party did its best, given the issues and so forth. Each of the 24 vote shares is thus based on strong campaigns by both parties. In none of the 24 elections did one party not campaign, so our analysis has nothing to say about what would happen if one party did not campaign. We are asking the question of what determines the vote share *given* that both parties campaign hard (which they always do).

Similarly, extremists are not nominated by the two parties (yes, I know, some will disagree with this statement), so our analysis has nothing to say about what would happen if a party nominated one. Again, we are asking the question of what determines the vote share *given* that both parties nominate nonextremists.

To take one more example, say that two months before an election, some new campaign tactic of one of the parties, such as negative ads, appears to be working well. The other party is likely to counter with its own negative ads, and in the end, the new tactics of the two parties would probably roughly cancel each other out. Again, each party has done its best by election time, and we are looking at voting behavior after all the hard campaigning has been done.

Manipulating the Economy to Increase Votes

As previously discussed, different periods for the growth rate were tried, and the period that led to the best results in terms of fit was the first three quarters of the election year. The growth rates in the other quarters matter in that they can contribute to the number of good news quarters, but they do not get extra weight beyond this. Voters appear to give the more recent experience more weight—"recent" in this case being the election year. This weighting is consistent with a number of psychological experiments, where what happens at the end of an experiment is remembered more than what happens otherwise (see the Notes to Chapter 1).

Because a strong growth rate in the year of the election is good for the incumbent party, there is an incentive for the incumbent party to try to stimulate the economy in the election year to help get reelected. One strategy would be to slow the economy in the beginning of the four-year term and then stimulate hard in the final year and a half or so. Slowing the economy in the beginning would allow more room for rapid growth rate at the end. If this were done, there would be a four-year business cycle in which the trough would be reached near the middle of the period and the peak would be reached near the end. This type of business cycle is called a *political business cycle*. The cycle is political in that it is induced by the incumbent party manipulating the economy for its own political purpose.

Whether this strategy has been pursued at some points in the past (that is, whether there are in fact political business cycles) is difficult to test. There are four-year periods in which a trough was reached in the middle and a peak at the end, but it is hard do know if this was done deliberately by the party in power. There is also the question of whether an administration has the ability to manipulate the economy in such a precise way, especially if the other party has control of Congress. Possibly, some have tried and failed. This question of how parties behave regarding the economy once in power is not the same as the question of how voters behave, and we have only been concerned in this chapter with the behavior of voters.

Does Congress Matter?

An implicit assumption behind our theory of voting behavior is that voters praise or blame the White House, not the Congress, for the state of the economy. If one party controls the White House and the other controls the Congress, who should be judged? The above theory is obviously wrong if the answer is Congress, and this is another possible pitfall that should be kept in mind. Casual observation suggests that the buck stops at the White House, but we have simply assumed this to be true and have not tested it.

Do Voters Really Know the Growth Rate and Inflation Numbers?

I have often been asked if I really believe that many voters know the actual values of the growth rate and inflation when they enter the voting booth on Election Day. How many people really look up the numbers in their newspapers before voting? (This is usually a hostile question.) It is surely unlikely that many voters know the exact numbers. They form their opinions about the economy by looking at the conditions around them—how their friends and neighbors and employers are doing—not by looking at the numbers themselves. They may also be influenced by the media, especially radio and television commentators. The numbers, however, are likely to be related to the conditions that people are observing. If the growth rate is negative, for example, it is likely that the conditions that people see around them are not so good and that commentators are saying bad things about the economy. It is thus not necessary that voters know the exact numbers as long as the numbers accurately reflect what the voters are actually paying attention to.

Real-Time Predictions of the Eight Elections Between 1980 and 2008

A prediction of a dependent variable, like the vote share, is based on a particular choice of explanatory variables—a particular box in a book if you will, a particular set of coefficients, and a particular set of values of the explanatory variables. A special case of a prediction, which we will call a *real-time prediction*, is one that was actually made before the event in question, like an election, took place. For example, I made a real-time prediction of the 2008 election on October 29, 2008, a few days before the actual election occurred. This prediction used the explanatory variables in Box 3-2, coefficients that were based on observations through the 2004 election, and values of the economic variables that were available at the time. This predicted value is not the same as the predicted value for 2008 in Table 3-1. The predicted value for 2008 in Table 3-1 is based on the same choice of explanatory variables, but the coefficients differ slightly because they are based on using one more observation—the 2008 observation that was not available until after the election. Also, the values of the economic variables used for the predictions in Table 3-1 are the latest revised values. They are not exactly the same as those that were available on October 29, 2008, because economic data get revised over time.

The voting analysis in this chapter was first developed in 1978. This allowed a real-time prediction to be made for 1980. Since that time, I have made real-time predictions for seven other elections, 1984 through 2008. In each case, I used the coefficients and economic data that were available before the election. As mentioned earlier in this chapter, the choice of explanatory variables was changed slightly following the 1992 election. In particular, the good news quarters variable was added. No changes in explanatory variables have been made since then. This means that the explanatory variables in Box 3-2 were slightly different for the elections before 1996. The eight real-time predictions are presented in Table 3-3 along with the actual values and the corresponding errors.

Also presented in Table 3-3 are the predictions in Table 3-1. These predictions use all the information through the 2008 election. The economic data are the latest revised data, and the coefficients are based on using observations through 2008. The explanatory variables are those in Box 3-2. These predictions obviously use more information than the real-time predictions, which only use the information available before the particular election predicted. Because of this, it seems likely that on average the real-time predictions will not be as accurate as the predictions in

TABLE 3-3 Real-time and Table 3-1 predictions: Democratic share of the two-party vote

		REAL-TIME		TABLE 3-1	
Year	*Actual*	*Predicted*	*Error*	*Predicted*	*Error*
1980	44.8	46.4	1.6	45.7	0.9
1984	40.9	43.2	2.3	38.2	−2.6
1988	46.2	48.1	1.9	49.2	3.0
1992	53.6	43.1	−10.5	48.8	−4.8
1996	54.7	51.0	−3.7	53.2	−1.5
2000	50.3	50.8	0.5	49.3	−1.0
2004	48.8	42.3	−6.5	45.5	−3.3
2008	53.7	51.9	−1.8	55.2	1.5
Average absolute error			3.60		2.33

Table 3-1. This is in fact the case. The average absolute error in Table 3-3 for the real-time predictions across the eight elections is 3.60 percentage points, which compares to 2.33 percentage points for the Table 3-1 predictions.

The largest errors in Table 3-3 are for the elections of 1992 and 2004. In 1992, Clinton, running against President George Bush, got 53.6 percent of the vote and was predicted to get much less—the real-time error is −10.5 percentage points. In 2004, Kerry, running against President George W. Bush, got 48.8 percent of the vote and was predicted to get much less—the real-time error is −6.5 percentage points. The other six elections are predicted fairly well. Excluding 2000, which is hard to know how to count regarding the winner because of the difference between the popular vote outcome and the electoral college outcome, only 1992 was forecast incorrectly as to the winner.

Although the choice of explanatory variables was changed slightly following the 1992 election, in part to try to improve the prediction of the 1992 election, this election remains a poorly predicted one. The error in Table 3-1 for 1992 is the largest of the 24 errors (ignoring sign) at −4.8 percentage points.

The main point of this section is to realize that the accuracy of real-time predictions is likely to be worse that the accuracy of predictions like those in Table 3-1, which are based on more information.

Predictions of the 2012 Election

If President Obama runs for reelection, we know the following before the election of 2012: the president running variable has a value of 1; the duration variable has a value of 0.0; the party variable has a value of 1; and the war variable has a value of 0. However, we do not know the values of the economic variables until a few days before the election. At the time of this writing (November 2010), economic data are available through the third quarter of 2010. There are eight more quarters before the election in November 2012. So far, we know that there has been one good news quarter since the first quarter of 2009, and inflation has been low at a rate of about 1 percent.

To predict the 2012 vote share, we need to choose values of the growth rate in 2012, of inflation over the 15 quarters, and of the number of good news quarters. So we need three numbers before we can predict the vote share. Table 3-4 shows three predictions. The first is based on the assumption of a robust recovery in 2011 and 2012, the second is based on the assumption of a very modest recovery, and the third is based on the assumption of a recession in 2012—a so-called double dip recession.

In the case of a strong recovery, the predicted vote share for President Obama is 55.9 percent. In this case, he is predicted to get a larger vote share than he did in 2008, which was 53.7 percent. In the case of a modest recovery, Obama is predicted to get 49.1 percent, which is a narrow loss. In the case of a double dip recession, he is predicted to get 46.4 percent, which is a fairly large loss.

Whether President Obama is predicted to win or lose thus comes down to what the economy will be in the next two years. This is, of

TABLE 3-4 Predictions for 2012 Democratic share of the two-party vote

	ECONOMIC VALUES			
Predicted share	*Growth Rate*	*Inflation*	*Good News Quarters*	
55.9	3.7	1.4	6	Robust economic recovery
49.1	1.0	1.4	1	Modest economic recovery
46.4	−3.0	1.4	1	Double dip recession

course, what the analysis in this chapter is all about—the economy matters. By the time you are reading this, you will know more about the economy in 2011 and 2012 than I know now. You can use the coefficients in Box 3-2 and your own assumptions about the growth rate, inflation, and the number of good new quarters to make your own prediction of the vote share. To make this clear, the following discussion spells out exactly what you need to do to make a prediction for 2012.

We begin with the coefficients in Box 3-2. To make a prediction for 2012, we multiply each coefficient by a chosen value of its corresponding variable, with the answers then being added to get the predicted value. Box 3-4 lists the calculations using the economic values corresponding to a strong recovery. These economic values are 3.7 percent for the growth rate, 1.4 percent for inflation, and 6 for the number of good news quarters. You can see that the sum of the individual multiplications is 55.9 percent—the predicted percentage for Obama.

BOX 3-4

Real-time prediction for 2012

coef.	value	coef. × value	
0.67	3.7	2.5	growth rate
−0.65	1.4	−0.9	inflation
0.99	6	5.9	good news quarters
2.9	1	2.9	president running
−3.4	0.0	0.0	duration
−1.9	1.0	−1.9	party variable
5.1	0.0	0.0	war variable
47.4	1.0	47.4	intercept
		55.9	TOTAL (vote share)

A calculator on my Web site does this work for you. If you give it your own values of the three economic variables, it will calculate the predicted vote share. It is simply doing what is done in Box 3-4 for your particular values of the growth rate, inflation, and the number of good news quarters.

You might ask how much confidence should you put on any particular prediction for 2012. The standard error in Box 3-2 is 2.5 percentage points; thus, from the analysis in Chapter 2, we should expect the prediction error to be between –2.5 and 2.5 about 68.2 percent of the time and to be between –5.0 and 5.0 a little over 95 percent of the time. The actual degree of uncertainty is, however, greater than this. First, there are the possible pitfalls, which could be important. Second, we know from Table 3-3 that real-time predictions are in general less accurate than predictions like those in Table 3-1. This was discussed earlier. (Any prediction before the 2012 election is obviously a real-time prediction.) The standard error of 2.5 is estimated using the errors in Table 3-1, and so it is too low when applied to real-time predictions. Third, any prediction that you make before all the economic data are in (which is not until right before the election) will be using economic values that are themselves subject to error.

If we look at the real-time predictions in Table 3-3 for the last four elections, the average error is 3.1 percentage points. These are the elections that were predicted after the changes that were made following the 1992 election. This suggests that you might think of the standard error as being about 3 percentage points rather than 2.5 percentage points for your real-time prediction for 2012. And then, added on to this degree of uncertainty should be the uncertainty that you attach to your choice of economic values.

Conclusion

This chapter is an example of the use of the tools in Chapter 2. A theory of voting behavior is proposed, data are collected, and the theory is tested. This example is interesting because there are many possible pitfalls to think about. It is also interesting in that once we get to prediction (on Sunday), an important event can be predicted—namely, the outcome of the 2012 presidential election. We now turn to Congress.