11
Business-Cycle Indicators

Tools: Basic statistics (standard deviation, correlation); cross-correlation function.

Key Words: Volatility; procyclical and countercyclical; leading, lagging, and coincident.

- Business cycle indicators are characterized by several properties: procyclical and countercyclical, leading and lagging.
- Cross-correlation functions identify these properties.

Probably the leading use of macroeconomic data (and macroeconomists) is forecasting: predicting future movements in economic variables so that businesses can decide how much to produce, investors can decide how to allocate their assets, and households can decide how much to spend. The good news is that forecasting is possible; we're not simply throwing darts at a board. The bad news is that it's not easy; even the best forecasters are far from perfect.

This chapter is devoted to short-term business-cycle indicators — variables that indicate changes in near-term economic conditions — and how to use them. In principle, we could be interested in many features of the economy: output, inflation, interest rates, exchange rates, and so on. We'll focus on output, but the methods can easily be applied to other variables. We look at the US, but similar ideas and methods apply to any country with reliable data.

11.1 Terminology

We refer to the properties of economic indicators with two related sets of terms. One set of terms describes whether an indicator's movements tend to come before or after movements in output. We say an indicator *leads* output if its ups and downs typically precede those of output, and *lags* output if they come after. An indicator whose movements are contemporaneous with those of output is referred to as *coincident*. Thus, the adjectives leading, lagging, and coincident describe the timing of an indicator's movements relative to those of output. Looking ahead, you might guess that leading indicators are most useful in forecasting. The stock market, for example, is a common leading indicator; it leads output by six to eight months, as we'll see shortly.

A second set of terms refers to whether an indicator's movements are positively or negatively correlated with output. If the correlation is positive, we say it is *procyclical*; if the correlation is negative, we say it is *countercyclical* . Most indicators are procyclical: employment, stock prices, housing starts, and so on. The most common countercyclical indicators have to do with unemployment: Both the unemployment rate and new claims for unemployment insurance rise during recessions.

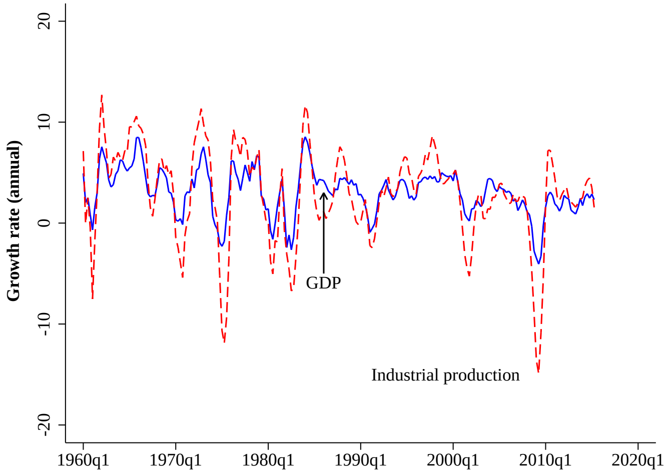
11.2 Forecasting

The classic forecasting problem goes something like this: What do we expect the value of [some economic variable] to be k periods in the future? Here, k is any period of time you like, but we're usually interested in anything from next week to a few years in the future.

If we're forecasting GDP, there's an extra difficulty because we don't know the present or the recent past, much less the future. We've seen, for example, that fourth-quarter GDP is first reported near the end of the following January, and even that number is a preliminary estimate. From the perspective of mid-January, then, we need to "forecast" the previous quarter.

We're going to shortcut this difficulty (somewhat) by using the monthly Industrial Production (IP) index as a substitute for real GDP, but the issue is a general one, in that the time lag in getting data is both an issue in its own right and a constraint on forecasting the future. IP measures output in manufacturing, mining, and utilities. More important, its fluctuations are strongly correlated with those in GDP. You can see that in Figure 11.1, which compares year-on-year growth rates in GDP and IP (aggregated to a quarterly frequency). You will notice that IP is more volatile than GDP but otherwise follows its ups and downs reasonably well. You may also

Figure 11.1: US GDP and industrial production.



notice some differences between them in the recent past, which have been traced to the rising importance of services in the US economy. In the US, IP is reported by the Federal Reserve in the middle of the following month. Data for December, for example, are available in mid-January. Using IP, therefore, gives us a shorter information lag than GDP. In addition, the monthly frequency gives us a finer time interval for near-term forecasting. For both reasons, we will focus our discussion of forecasting on IP rather than, GDP, although the same principles apply to both, as well as to other macroeconomic and financial variables.

11.3 Good indicators

Good forecasts require good inputs. One way to forecast a variable is with its own past. Future growth rates of IP, for example, might be related to current and past growth rates. We can usually do better than that by adding other indicators to our analysis. Speaking generally, a good indicator should have one or more of these properties:

- **Correlation.** A good indicator is correlated with the variable we are forecasting.
- **Lead.** A good indicator leads the variable we are forecasting.
- **Timeliness.** A good indicator is available quickly.

- **Stability.** A good indicator does not undergo major revisions subsequent to its initial release, and its relationship with the variable we are forecasting doesn't change over time.

On the whole, measures of economic activity (employment, for example) tend to be strong on correlation and weak on timeliness (see the discussion of GDP above) and stability (many economic series are revised frequently). The best ones lead the business cycle. In contrast, financial indicators (equity prices, interest rates) are weaker on correlation but stronger on the other three properties: They're typically available immediately, often lead the cycle, and are not revised. Various indexes of leading indicators combine multiple series with the hope of getting the best from each. The Conference Board's quasi-official index of leading indicators is the most common example.

11.4 Identifying good indicators

How do we identify indicators with high potential? We'll use another bit of terminology that leads to an extremely useful graphical representation of the dynamic relation between two variables: the *cross-correlation function* (ccf).

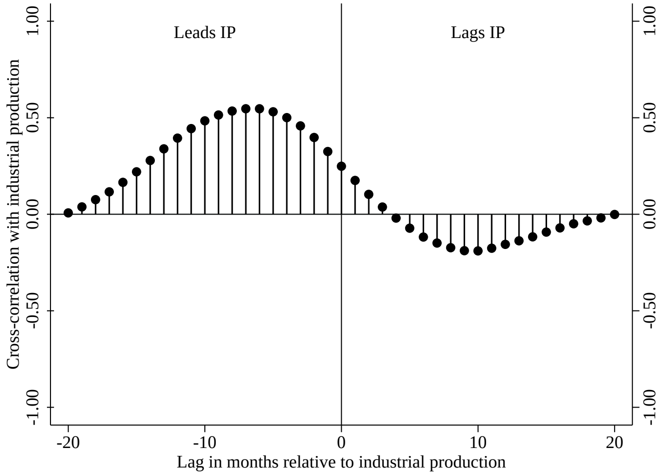
You may recall that the correlation between two variables (x and y, say) is a measure of how closely they are related in a statistical sense. If the correlation is (say) 0.8, then observations with large values of x tend also to have large values of y. If the correlation is 0.4, this association is weaker. And if the correlation is –0.8, observations with large values of x tend to have small values of y — and vice versa.

The cross-correlation function extends the concept of correlation to the timing of two indicators. Specifically, consider the correlation between x at date t and y at date $t - k$. If k is negative, then we're talking about the correlation between x now and y k periods in the future. If k is positive, we have the correlation between x now and y k periods in the past. By looking at the pattern of correlations, we can identify indicators x that tend to lead the variable y. We refer to k as the lag of y vs x, but if k is negative it refers to a lead. Mathematically, we write

$$\text{ccf}(k) \;=\; corr(x_t, y_{t-k}).$$

Typically, we would graph this against k, with k starting with a negative number and moving to positive numbers. The pattern of correlations tells us whether an indicator x leads or lags (on average) a variable y.

Figure 11.2: Cross-correlations: the S&P 500 and industrial production.



Both series are year-on-year growth rates for the period 1960-present. The large correlations to the left tell us that the S&P 500 index is a good indicator of future industrial production.

Let's move from the abstract to the concrete to make sure we understand what the ccf represents. [You might want to work your way through this paragraph slowly, it's important.] We calculate the year-on-year growth rates of the S&P 500 index and industrial production and compute their ccf using the S&P 500 for x and industrial production for y. Figure 11.2 is a plot of their correlations against the lag k. There's a lot of information here, so let's go through it one dot at a time. The dot at $k = 0$ (on the vertical line at the center of the figure) shows that the contemporaneous correlation is about 0.2. Contemporaneous means that we're looking at the two variables at the same time: March 2001 industrial production is lined up with March 2001 S&P 500, and so on. Next, consider the dot corresponding to $k = -10$ on the left side of the figure. The correlation of (roughly) 0.5 pictured in the figure shows the growth rate of industrial production with the growth rate of the S&P 500 index dated ten months earlier. Evidently high growth in equity prices now is associated with high growth in IP 10 months later. Finally, consider a dot on the right side of the figure. The dot at $k = +10$ suggests that the correlation of industrial production growth with equity price growth ten months later is about –0.2.

This pattern of correlations tells us a lot about the timing of movements in the two variables. In general, negative values of k (the left side of the figure) indicate correlations of the S&P 500 with future industrial production; we would say that they reflect the tendency of stock prices to lead output. Positive values of k (the right side of the figure) indicate correlations of the S&P 500 with past industrial production; they reflect the tendency of stock prices to lag output. What we see in the figure is a strong correlation of the S&P 500 index with industrial production seven to eight months later. Evidently, the stock-price index is a leading indicator of industrial production.

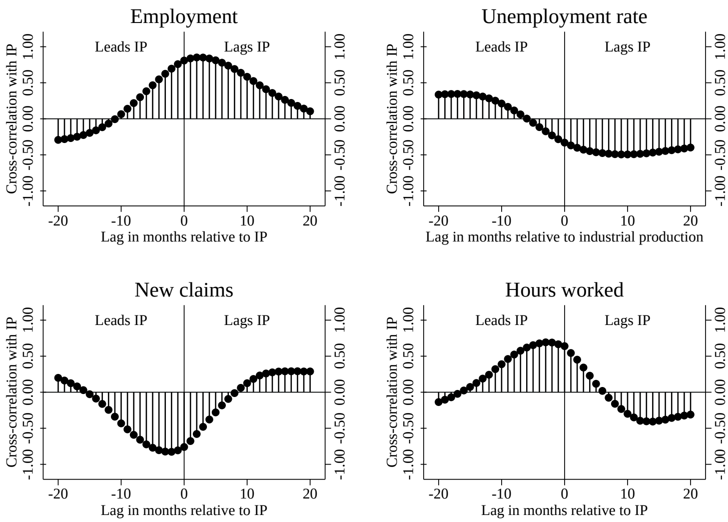
We'll use the cross-correlation function to identify whether an indicator is leading or lagging, procyclical or countercyclical.

To do this, we find the largest correlation in absolute value. If it occurs to the left of the figure, we say it's a leading indicator; if on the right, lagging. Similarly, if the (largest) correlation is positive, we say the indicator is procyclical; if negative, countercyclical. In principle an indicator could be both leading and lagging, or both pro- and counter-cyclical, but we'll deal with that if and when it happens.

Digression. We snuck something in here that we should mention again, although it's not particularly important for our purposes. We used year-on-year growth rates instead of monthly growth rates. We could use either, but the year-on-year pictures are smoother and, in our view, more attractive. We'd see a similar pattern with monthly growth rates, but the correlations would be both smaller and choppier.

Let's look at some other indicators and see which ones lead IP. Some of the most common indicators are labor-market variables, constructed by the Bureau of Labor Statistics. Cross-correlation functions for four of them are pictured in Figure 11.3. Nonfarm payroll employment (a measure of employment constructed from a survey of firm payrolls) is a slightly lagging indicator since the ccf peaks with a lag of one to two months. It is, nevertheless, useful because the correlation (over 0.8) is unusually strong. And even a two-month lag is more timely than the GDP numbers. The unemployment rate is countercyclical (note the negative correlations) and lags IP in the sense that the largest correlation comes at a lag of three to four months. It seems that a rise (fall) in output is associated with a fall (rise) in the unemployment rate three to four months later. New applications ("claims") for unemployment insurance are also countercyclical, but the correlation is stronger than for the overall unemployment rate, and it leads industrial production by two to three months. Another popular indicator is average hours worked per week in manufacturing. This indicator is strongly procyclical and leads industrial production by two to four months. The labor market, in short, provides a good overall picture of the economy and, in some cases, supplies indications of future movements in industrial production. The leading variables ("new claims" and "average weekly hours") are more highly

Figure 11.3: Cross-correlation functions: labor market indicators.

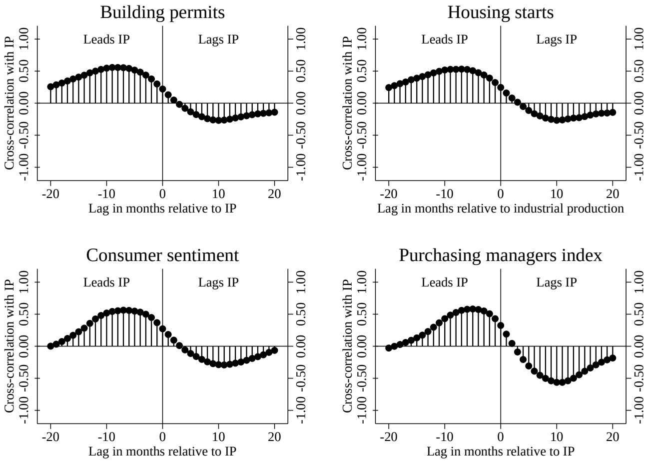


correlated with industrial production than the S&P 500 index, but the leads are shorter.

Other sources of useful information are various measures and surveys of economic activity conducted by the Bureau of the Census and private organizations. Cross-correlation functions for four common ones are pictured in Figure 11.4. The first two are building permits and housing starts, two indicators of new home construction reported by the Census. Two ideas lie behind their use: that construction of new capital is more volatile than other sectors of the economy and that decisions to build new homes reflect optimism about the future. The cross-correlation functions suggest that they work; while the correlations are smaller than with (say) employment, the leads are substantial (ten months or so). The next two are popular private surveys. Consumer sentiment, based on a survey of consumers collected by the University of Michigan, reflects consumers' optimism about current and future economic conditions. The purchasing managers index is what we call a "diffusion index." It's based on a survey of purchasing managers who report whether they see economic activity increasing or decreasing. Each is used as is. We see in the figure that both are procyclical leading indicators.

We could go on. There are hundreds of indicators, more all the time. The most common one we've skipped is the slope of the yield curve: Flat or downward-sloping yield curves are associated with slower-than-usual future

Figure 11.4: Cross-correlation functions: surveys of economic activity.



growth in output. More on this in the Appendix.

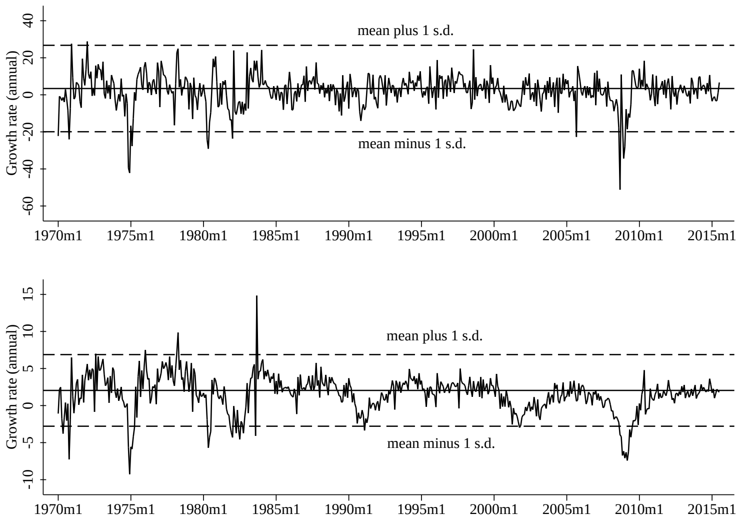
11.5 The business-cycle scorecard

Now that we understand how to identify good indicators, how do we put them to work? The central question here is how to combine the inputs of multiple indicators. One way to do that is to summarize them informally, which is what we do here. Another is to use multivariate regression, which is the next topic, but not one we'll spend much time on in this course.

The business-cycle scorecard is a summary of what selected indicators tell us about near-term economic conditions. We'll use the four monthly indicators pictured in Figures 11.5 and 11.6. In the first figure, we see the monthly growth rate of IP (top panel) and the change in (nonfarm) employment for the period 1960 to the present. They show similar patterns, with the major postwar downturns evident in each. Evidently, employment is procyclical, rising in good times and falling in bad times. Industrial production is a "noisier" series, which is one reason that many analysts prefer employment as a measure of current economic conditions. The lines show us the mean value (the solid line) and plus and minus one standard deviation (the dashed lines). The lines are useful benchmarks for telling how strong the current value of an indicator is relative to past experience.

In the second figure (Figure 11.6), we see similar data for new claims for unemployment insurance and housing starts. New claims are reported weekly;

Figure 11.5: Industrial production and employment.

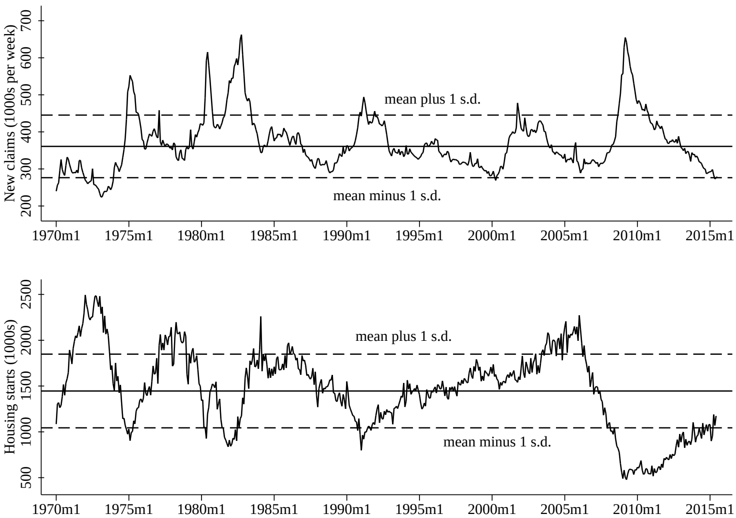


The two panels show, respectively, the annual growth rate of industrial production and the year-over-year change in the number of people employed.

the figure is based on the four-week moving average. Remember, they are countercyclical: they rise when the economy weakens. The second panel is housing starts. You can see in the figure that housing starts don't always go up and down with the economy. In the 2001 recession, housing starts fell only slightly. In 2008, we made up for that, with housing starts falling to their lowest point since (at least) 1960. None of that will come as a surprise to you. These four indicators come from the Federal Reserve (industrial production), the Bureau of Labor Statistics (employment, new claims), and the Bureau of the Census (housing starts). These government agencies are the primary sources of economic indicators in the US. There are private indicators also, but the government indicators are widely used and publicly available.

In the business-cycle scorecard, we rate each pro-cyclical indicator as strong positive if the current value of the indicator is above the "mean plus one standard deviation" line, weak positive if it's between the mean line and the one above it, weak negative if it's below the mean line but above the "mean minus one standard deviation" line, and strong negative if it's below the bottom line. For countercyclical indicators we reverse the direction: for example, strong positive means below the "mean minus one standard deviation" line.

Figure 11.6: New claims and housing starts.



The two panels show, respectively, new claims for unemployment insurance and housing starts, two popular indicators of economic conditions.

This is a rough cut, to be sure, but a useful one. It leads to this summary of economic conditions as of August 2014 based on the four indicators we have seen so far:

- Industrial production: Growth has recently been close to the mean. Assessment: weak positive.
- Employment growth: The most recent numbers show steady moderate growth. Assessment: weak positive.
- New claims: They are the only countercyclical indicator in our list. They have fallen dramatically over the last two years, and are now well below the long-run average. Assessment: strong positive.
- Housing starts: They remain low by historical standards, although there has been improvement since 2009. Assessment: weak negative.

These assessments are collected in Table 11.1. Overall, we see one strong positive, two weak positives, and one weak negative, a mixed set of signals that's not unusual. A more extensive analysis would use more indicators, decide how much weight to give each one, assess how far into the future they point, and so on.

Table 11.1: Business-cycle scorecard in action

Indicator	Strong Negative	Weak Negative	Weak Positive	Strong Positive
Industrial production			x	
Employment			x	
New Claims				x
Housing starts		x		
Summary	0	1	2	1

11.6 Regression-based forecasting

A more formal statistical approach is to include as many indicators as we like in a multivariate regression. We estimate the regression by some appropriate method and use it to forecast the future. Here are the steps we might follow in constructing a forecast of (say) industrial production k months in the future.

The first step is to construct the variable we're forecasting. Let us say that we're interested in the growth rate of industrial production between now and k months in the future. You can do what you want, but we compute the (annualized) growth rate this way:

$$\gamma_{t,t+k} \;=\; \ln(IP_{t+k}/IP_t) \times (12/k).$$

We refer to k (here measured in months) as the *forecast horizon*. The adjustment factor "$12/k$" converts the growth rate to annual units. For a one-year forecast, then, we would set $k = 12$ and compute the year-on-year growth rate.

The second step is to find some variables you think would be useful in forecasting. The previous section might give you some ideas. There's a half-step that sneaks in about here, too: what form of the indicator to use. In most cases, we use growth rates of the indicators, too, either over one period or a year, whichever you think works best. But some variables are used as is. In Figure 11.3, for example, the cross-correlation for the unemployment rate is for the rate, period — not its growth rate, change, or other transformation.

Third, you put all the ingredients into a statistical package and run a regression. For example, to forecast IP growth, we would estimate the regression

$$\gamma_{t,t+k} \;=\; a + bx_t + \text{residual},$$

where x_t is the value of the indicator we have chosen. We use a sample of data to estimate the parameters a and b. Note well: The growth rate is between now (date t) and a future date $(t+k)$, but the indicator is observed now (at t). This is central to the exercise: We use what we know now to predict the future. It's not kosher to use future variables to predict the future because we don't know the future when we make the forecast (duh!).

Fourth and last: Once we have estimates of the regression parameters ($\widehat{a}$ and $\widehat{b}$, say), we use them and the current value(s) of the indicator(s) (x, say) to compute the forecast:

$$\widehat{\gamma}_{t,t+k} \;=\; \widehat{a} + \widehat{b}\, x_t.$$

The "hats" remind us that we are using estimates; $\widehat{\gamma}_{t,t+k}$ is our forecast of future growth. There are lots of variants of this approach — you can add multiple indicators, lags of the indicators $(x_{t-1}, x_{t-2}, ...)$, and even past values of the growth rate of industrial production. We recommend all of the above.

The result of such an exercise is generally a useful forecast — useful in the sense that it tells us something about the future. Something, but not everything! Over periods of a year or two, forecast accuracy is usually modest. Even in-sample, the regressions rarely have R^2s above 0.25, which tells us that most of the variation (at least 75 percent) in our forecast variable is unexplained. Some people see a lesson in this: It might be more important to know how to respond when the unexpected occurs than to have better forecasts. In practice, both are useful: knowing something about the future, and having backup plans to deal with the inevitable forecasting errors. It pays to carry an umbrella when the forecast calls for rain.

11.7 Aggregation and prediction markets

There's another appealing approach to forecasting: Let markets do the work. Most of the best forecasts aggregate information from multiple indicators and sources. Indexes of leading indicators do this one way by combining multiple indicators to produce an index, which is then used to forecast the future. Or we could use multiple indicators in regression-based forecasts, as we suggested above.

Another approach is to aggregate the forecasts themselves — that is take several forecasts, perhaps based on different indicators, and average them. The business-cycle scorecard is a simple version of this. The so-called "Blue Chip" forecast is an average of forecasts generated by experts, and it performs better than any single forecaster. Some statistical forecasters do the same sort of thing on their own. They generate multiple forecasts with methods like our forecasting regression, and then average them to generate a final aggregate forecast. Again, the aggregate tends to do better than the individual forecasts.

A related idea is to rely on markets, which aggregate information from the people using them. Presidential futures markets, for example, have predicted the popular vote in the last four elections more accurately than any of the major polls. In the economic arena, there are a growing number of markets in which you can trade futures contracts whose payoffs are tied to the value of specific economic numbers: the consumer price index, the fed funds rate, and so on. These markets are increasingly used as forecasts themselves, with one wrinkle. The simplest interpretation is that the futures price is a market forecast of the relevant economic number. For example, if we are interested in the value of an economic number y to be released in 6 months (y_{t+6}, say), we might use its current futures price (f_t, say):

$$f_t \;=\; \text{Market's Current Forecast of } y_{t+6}.$$

Experience (and possibly some insight) tells us that we may want to make a correction for the risk of the contract:

$$f_t \;=\; \text{Market's Current Forecast of } y_{t+6} + \text{Risk Premium}.$$

There's no limit to the amount of sophistication we can bring to bear on the last term, but for now, you can simply note that we probably want to address it in some way. Once you do, markets are an extremely useful source of information about the future.

Executive summary

1. Fluctuations in economic activity can be (partially) predicted by a number of indicators.

2. The cross-correlation function is a tool for describing the timing of the relation between two indicators: for example, whether one indicator leads another.

3. Markets are useful aggregators of information — and increasingly popular sources of economic forecasts.

Review questions

1. Terminology. Consider economic indicators in general.

(a) What is a procyclical indicator? A countercyclical indicator?

(b) Give an example of each.

(c) What is a leading indicator? A lagging indicator?

(d) Give an example of each.

Answer.

(a) A procyclical indicator moves up and down with GDP. A countercyclical indicators goes up when GDP moves down. We typically identify this feature with the sign of the correlation.

(b) Most indicators are procyclical : employment, the S&P 500, and so on. The unemployment rate is the classic countercyclical indicator.

(c) A leading indicator is correlated with future GDP growth, a lagging indicator is correlated with past GDP growth.

(d) The stock market is a leading indicator, the unemployment is a lagging indicator. We typically identify this feature with the cross-correlation function.

2. Housing starts. We mentioned housing starts as an indicator of future economic activity. In what ways do you think it's a good indicator? A bad one? (For further information, see the US Census Bureau's web site.)

   Answer. Good: connected to housing, which, as a durable good, should be cyclically sensitive and volatile; available quickly; it leads the cycle (as you can see from its ccf). Bad: based on a sample, which leads to short-term noise; revised periodically; strong seasonality; possibly misleading now that we have a glut of housing to work off.

3. Unemployment. The unemployment rate is widely reported in the press, but professionals rarely use it. Why do you think that is?

   Answer. One reason is that the unemployment rate understates the change in employment in a downturn. Some people who lose jobs leave the labor force, so they're not included in the unemployment rate. Another reason is that the unemployment rate is a lagging indicator. It falls slowly, well after the economy turns around. Employment (the number of people actually working) is the preferred indicator for both reasons.

4. Terrorism futures. In 2002, a government agency recommended that we establish a futures market in terrorist attacks, on the grounds that it would give us a useful public indicator of their likelihood. The idea was widely criticized. Do you think it was a good idea or a bad one? What would you need to do to implement it?

   Answer. Another case of a good idea thrown out because it sounded bad to politicians. It's not clear that such attacks are predictable, but if they are, we'd expect futures markets to do as well as any other method.

To implement the idea, you'd need to define (and possibly quantify) a terrorist event.

If you're looking for more

There are many sources of leading indicators around the world and almost as many guides to them. Among them:

- The best book we've seen on the subject is Bernard Baumohl, *The Secrets of Economic Indicators*. If you use economic indicators in your job, you should buy this book.
- The Bloomberg Economic Calendar gives release dates and short summaries of a wide range of indicators. Ditto the WSJ, Yahoo, etc.
- The CME has a nice report, "Impact of economic indicators," on the information content of common indicators for futures prices.

Most statistical software packages have one-line commands to compute cross-correlation functions. You can also do it in a spreadsheet, but it's a lot more cumbersome.

To do more with this topic, you need some knowledge of time series statistics. If you'd like to learn more about forecasting economic and financial variables specifically, we recommend "Forecasting Times Series Data," course STAT-GB.2302, taught in alternate years by Professors Deo and Hurvich, two of our best statisticians.

Symbols and data used in this chapter

Table 11.2: Symbol table.

Symbol	Definition
ccf	Cross-correlation function
ccf(k)	Correlation of (x_t, y_{t-k}) at lag k
$\gamma_{t,t+k}$	Continuously compounded growth rate from t to $t+k$
$\widehat{x}$	Estimate of x
f_t	Futures price at time t

Table 11.3: Data table.

Variable	Source
Industrial production	INDPRO
Real GDP	GDPC1
S&P 500	SP500
Employment	PAYEMS
Unemployment rate	UNRATE
New claims	IC4WSA
Hours worked	AWHMAN
Building permits	PERMIT
Housing starts	HOUST
consumer sentiment	UMCSENT
Purchasing managers' index	NAPM
10-year Treasury yield	GS10
2-year Treasury yield	GS2
Federal funds rate	FEDFUNDS

To retrieve the data online, add the identifier from the source column to http://research.stlouisfed.org/fred2/series/. For example, to retrieve nonfarm employment, point your browser to http://research.stlouisfed.org/fred2/series/PAYEMS

Index