Diagnosing Recessions

BY ÒSCAR JORDÀ

The beginnings and ends of recessions are officially dated about 12 months after the fact. A common rule of thumb declares recessions as two quarters of consecutive negative GDP growth, but this is very inaccurate. A better option is to apply medical diagnostic evaluation methods to the business conditions indexes of the Chicago and Philadelphia Federal Reserve Banks, which suggests the recent recession ended sometime between June and August 2009.

What is a recession? The United States is fairly unique in that an independent group of academic economists, the National Bureau of Economic Research's (NBER) Business Cycle Dating Committee, has been entrusted to maintain a historical chronology of business cycles. NBER defines a recession as "a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in production, employment, real income, and other indicators" (NBER 2008).

In other words, this desire to keep a chronology of economic turning points—peaks and troughs of economic activity, and therefore implicitly expansions and recessions—reflects the notion that there are fundamental differences between these two phases of the economic cycle. And since GDP growth as high as 2% is sometimes observed during recessions, the dating of business cycles is not simply a mindless, mechanical accounting exercise based on when GDP growth is negative.

The NBER's Business Cycle Dating Committee was formed in 1978 to establish a historical chronology of business cycle turning points. The NBER itself was founded in 1920 and published its first business cycle dates in 1929, although records are now available retrospectively starting with the trough of 1854. The committee commonly releases its determinations of cyclical turning points with more than a year's delay, although sometimes committee members make public statements about recession dates in advance of formal statements. The committee's mission is not to serve as an early warning system but to classify economic activity for the historical record. Other countries now have similar committees, including the Euro Area Business Cycle Dating Committee of the Center for Economic Policy Research, founded in 2002. But it's fair to say that the NBER committee's length of historical coverage and experience (its current president, Robert Hall of Stanford University, is one of the original founding members) have no equal.

The NBER committee's definition of recession does not naturally lend itself to simple mathematical formalization. Moreover, since recessions last on average 11 months, the committee often won't make public its determination of when a recession started until after it has already ended. Recent work by Berge and Jordà (2009) found that the NBER committee's business cycle dating method is superior to all the alternatives they investigated and close to an ideal measure of the underlying state of the economy. But its inherent lag can cause problems: policymakers must consider whether to stimulate the economy;

managers must decide whether to open new plants or make other investments; consumers may hold off on buying houses or durable goods until the end of a recession makes their job prospects sunnier. A timelier recession signal is clearly needed. (A way to foretell turning points in advance would be even better.)

Two-quarter rule not very good

Faced with this lack of a timely official business cycle barometer, it is not surprising that the press frequently uses a rule of thumb to define a recession as two consecutive quarters of negative GDP growth. But such a definition does not appear to be very good. From February 1947 to November 2007, the month before the current recession started, this two-quarter rule would have missed the 2001 recession entirely and would suggest that recessions have lasted on average seven months rather than

the 11 months the NBER committee records. A one-quarter rule would have been even worse in that it would have detected 24 recessions over the same period, each lasting an average of five months.

Is there a better alternative? Yes. The problem of having a real-time business cycle indicator is actually very similar to that of developing a diagnostic test to determine whether an individual has a particular disease. To develop a timely recession test and identify periods of economic expansion and contraction, we can apply medical diagnostic evaluation procedures to the current business conditions indexes produced by the Federal Reserve Banks of Chicago and Philadelphia. Known respectively as the Chicago Fed National Activity Index and the Aruoba-Diebold-Scotti index, they are publicly available on the banks' websites and are plotted in Figures 1 and 2.

The figures show shaded areas that correspond to official NBER recessions. High values of the Chicago and Philadelphia indexes are always associated with expansions and low values with recessions, but there is an intermediate region where things are trickier. Identifying a recession is like determining whether to perform a

Figure 1
The Chicago Fed National Activity Index

Optimal threshold = −0.72
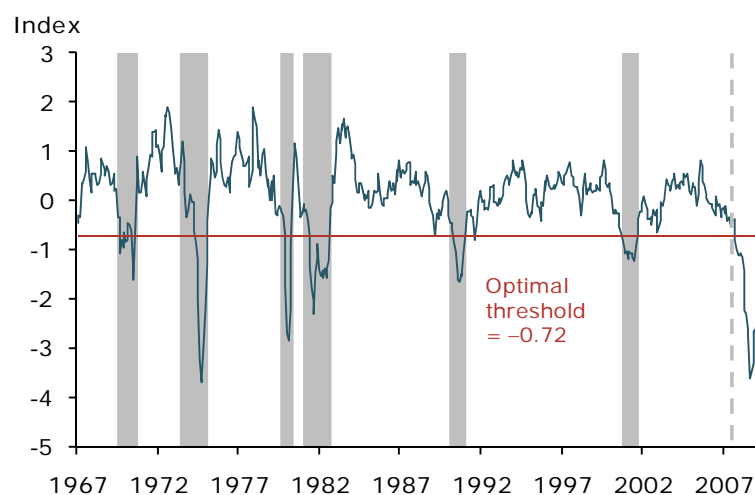
Note: Shaded bars indicate NBER recessions. Dotted line indicates the last peak of economic activity.

Figure 2
The Aruoba, Diebold, and Scotti (ADS) Index

Optimal threshold = −0.80
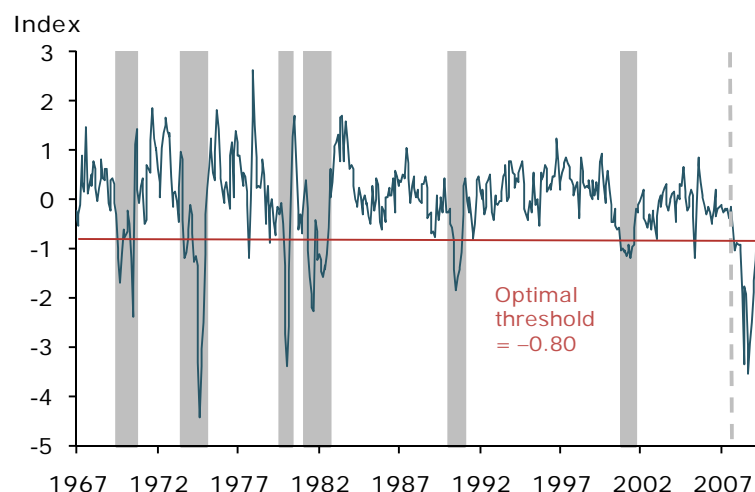
Note: See Figure 1.

biopsy for prostate cancer on a patient who has had a prostate-specific antigen (PSA) blood test. A reading of 4 ng/ml (nanograms/milliliter) or lower is considered "normal," levels above 10 ng/ml are considered "high," but anything in between is considered "intermediate." Now suppose your blood work comes back with a reading of 7ng/ml. What should your doctor do?

Threshold level for action

These two problems share a key element. In both cases, we would like to calculate a trigger level, or threshold, beyond which we would take action. In the case of a business conditions index, what reading would prompt us to call a recession? For the PSA test, what level would cause us to schedule a biopsy rather than send the patient home?

Four possible outcomes are associated with these two decisions, depending on the underlying but unobserved true state: *True positives* would occur when we call a recession or have the patient undergo a biopsy and the economy actually is in recession or the patient does have cancer. *False positives* would result when the underlying states are that the economy is expanding or our patient is cancer free. We can define *true negatives* and *false negatives* symmetrically. Each of these four possible outcomes has a particular set of costs and benefits. Determining the optimum threshold depends on balancing a number of factors, including the rate of true and false readings, the underlying incidences of recessions or cancer, and the comparative costs and benefits of the four outcomes.

In the cases of both recession and cancer tests, the threshold level affects the true/false and positive/negative rates, but not the underlying incidence of recessions or cancer, nor the costs and benefits of each outcome. In other words, the higher the threshold, the higher the rates of true positives and false negatives, and the lower the rates of false positives and true negatives. But obviously the actual incidence of recession or cancer is unaffected by any diagnostic test. Therefore, the net benefits of a particular choice of threshold consist of the following components: First are the benefits of correctly predicting a positive, such as scheduling a biopsy and finding that the patient does have cancer, or calling a recession and implementing a stimulus program. These must be adjusted according to the true positive rate as well as the relative incidence of positives. A similar calculation has to be made for true negatives, for example, sending the patient home when the patient is cancer-free, or foregoing stimulus when the economy is expanding. Adjustments must also be made for the costs of incorrectly predicting positives—conducting a biopsy on a cancer-free patient or applying stimulus to an expanding economy—as well as false negatives.

Recession months represent about 16% of all months since World War II. Now suppose that I give you one dollar for each correct prediction (a true positive or a true negative) and subtract one dollar for every mistake (a false positive or a false negative). Then a recession threshold set so high that all months are found to be in "expansion" would pay off 68 cents per month, since you would gain one dollar 84% of the time and lose one dollar 16% of the time.

You could do better by performing the same kind of calculation medical practitioners use to determine whether to do a biopsy. In medicine, the net benefits of a threshold for action can be calculated by multiplying the expected costs and benefits of each of the four possible outcomes by the frequency of those outcomes and the incidence of the underlying condition. By trying different threshold values, it's possible by trial and error to arrive at an optimal level.

The same can be done with recession tests using the Chicago and Philadelphia indicators. The optimal recession thresholds appear to be −0.72 for the Chicago index and −0.80 for the Philadelphia index

(Berge and Jordà 2009). At these values, you could get pretty close to earning a full dollar on your monthly recession prediction. The two levels correspond to the horizontal lines depicted in Figures 1 and 2. In fact, the Chicago Fed in its website suggests, "When the CFNAI-MA3 value moves below –0.70 following a period of economic expansion, there is an increasing likelihood that a recession has begun."

Using these threshold values and starting with the last official NBER date available (the peak of December 2007), the Chicago index indicates that the recent recession began at the NBER peak and lasted until August 2009, while the Philadelphia index points to the same start date and June 2009 as the trough. These dates correspond well with recent public statements by NBER committee member Robert Gordon who declared that the recession ended in June 2009, committee chair Robert Hall who said the trough probably occurred sometime in the summer, and former Fed Chairman Alan Greenspan who said the recession ended in July 2009 and possibly as early as June 2009. An official NBER statement is not likely to come until the summer of 2010 at the earliest.

Let me end by returning to the hypothetical patient with a PSA reading of 7 ng/ml to note that the costs and benefits of each of the four possible outcomes will vary from person to person. A mechanical rule based on biopsying any patient with a reading of 7ng/ml or above is clearly not the right answer for everyone. For example, an older patient may not tolerate biopsies as well as a younger patient, so the costs of a false positive would be higher. And the costs of an undiagnosed tumor, a false negative, may be lower for him than for his younger counterpart. Hence, for the older patient, the optimal trigger may be higher than for the younger man. So it is with our business index recession tests. If the costs of pursuing an activist policy during a misdiagnosed expansion are not symmetric with the costs of policy inactivity during a misdiagnosed recession, the trigger points for a policymaker will differ from those we have offered. In addition, the cost–benefit calculations of our manager trying to decide whether to open a new factory or our consumer waiting to purchase a new car would be different from those of the policymaker. Application of recession diagnostic tests then is art as well as science.

Òscar Jordà is associate professor of economics, University of California, Davis, and visiting scholar, FRBSF.

References

Berge, Travis, and Òscar Jordà. 2009. "The Classification of Economic Activity." UC Davis working paper 09-18. http://www.econ.ucdavis.edu/faculty/jorda/papers/Berge_Jorda_v4_final.pdf

FRB Chicago. 2010. "Chicago Fed National Activity Index." Accessed February 9, 2010. http://www.chicagofed.org/webpages/publications/cfnai/index.cfm

FRB Philadelphia. 2010. "Aruoba-Diebold-Scotti Business Conditions Index." Accessed February 9, 2010. http://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index/

National Bureau of Economic Research, Business Cycle Dating Committee. 2008. "Determination of the December 2007 Peak in Economic Activity," December 11. www.nber.org/cycles