

Discussion

1. Chris Caton

I enjoyed reading this paper. It is not often that a private-sector economist gets to spend more than 30 minutes thinking about one topic. I have to say that I can't give equal justice to all parts of the paper, since my level of technical competence has been in decline for some time. Indeed, when I got to Section 3, which outlines the methodology behind the construction of the indices, I was struck by two thoughts. I am an old dog. And this is a very new trick.

The paper begins so well. The two indices constructed from the quarterly data first behave well, and second, behave similarly, which gives one confidence that they must be doing the right thing. As Figure 2 shows, the SW index clearly follows a similar path to GDP, but with a lot less noise, and the requisite tests show that it makes very little difference if one uses more factors to construct the indices, or if one uses a broader panel of data (Figures 3 and 4). So far so good.

It is when we get to the construction of the monthly indices, for the shorter period 1980–2004, that it gets a little more perplexing. First, the authors' methodology leads them to identify the one-factor SW index and the two-factor FHLR index as respective bests of breed. Suddenly, the indices no longer appear similar, as Figure 5 illustrates. It is clear what is causing this: it is not the different methods of construction but the different number of factors. The authors make the point that the behaviour of the two series is very different around 1990. Let me make the same point another way. Suppose you were to ask SW and FHLR for an assessment of the relative strength of the economy in 1986 and 1994. SW would reply that the two years were about equal in terms of growth, while the FHLR index would suggest that 1986 was a very weak year and 1994 a very strong year. All of the comfortable feeling generated by the fact that the two indices were sending a broadly similar message has gone.

It needs somebody smarter than I am to articulate why a small difference in the number of factors makes such a difference to the monthly indices, when it made so little difference to the quarterly indices. This is important, because if we are going to use such indices for current, rather than just historical, analysis, then a monthly index would obviously be far preferable. Indeed, it would need to be the SW variant, because of the lags involved in the FHLR.

I have a small quibble with the construction of the monthly FHLR index. There are many important economic time series (GDP and the CPI are the most obvious examples) that are available only quarterly. They can't be ignored so, in constructing the index, one needs to come up with an equivalent monthly series. This is done by constructing a series that in each month grows at one-third of the quarterly rate of growth for that quarter. A moment's thought will show that such a monthly series doesn't sum to the quarterly series but, more importantly, this methodology in fact imposes an average lag of 1 month on the relevant series. This doesn't seem a desirable property, and it is very easily fixed! You just need to take the quarterly

growth rate, divide it by three, and then apply that rate to the three months beginning the month *before* the quarter begins.

The authors then look at the clear decline in the volatility of GDP growth over the past 45 years, a topic also visited in several of the papers presented earlier. Their conclusion is that, since the coincident indices have experienced no such similar decline over the whole interval, it is quite likely that most of the decline in GDP volatility is a result of improved measurement, rather than a true decline in the volatility of the economy (see Figure 7). But I see two other things when I look at Figure 7. First, the indices appear to show a step up in volatility during recessions (hardly surprising), and a consequent step down 10 years later, when the recession years roll out. Second, I see a trend decline in the volatility of the indices since 1990, which suggests to me that the overall volatility of the economy *has* declined (or have we just had a freakishly long expansion?).

I found the section on the dating of recessions also to be interesting. In brief, the quarterly indices find only three recessions since 1970, while the GDP data find six. I want to return to this count later, but a question that occurred to me, which the authors don't address, is: do the indices have any tendency to act as an early warning for 'GDP recessions'? Or, if they don't, do we at least get the data earlier? At first glance, they both appear to lead both peaks and troughs in two of the three common recessions (Table 6). But there are two problems. First, the indices say that the mid-1970s recession was finished one quarter before the GDP data suggest that it began. Second, of course, we need a real-time analogue of Table 6. We could accept that the indices are less susceptible to revision over time, so what we have now may not be very different from what we would have seen in real time, but we know from Appendix C and the Morse code chart (Figure C1) that GDP has certainly been revised significantly. Indeed, the mid-1970s recession suggested by the indices seems to coincide with one recently revised out of existence.

There are two more broad points that I want to make about recessions. First, the SW index has a transatlantic cousin in the Chicago Fed National Activity Index (CFNAI), which was first constructed some five years ago. This is a monthly series, formed from 85 indicators, and it comes with its own rules of thumb about recessions and recoveries; when a three-month moving average of the series dips below -0.7 , this is taken as a signal of a recession. Such a rule in fact sends a couple of false signals either side of the 1990–1991 recession; a better threshold would be -1 . Now if you go all the way back to Figure 1, such a threshold for the Australian SW index would identify a large number of recessions, including in 1986 and 2001. The indices appear to be standardised in the same way, so why is the threshold apparently so much bigger in Australia? Australian data would naturally be noisier than US data (we're a smaller economy), but the standardisation should take care of this. The SW quarterly series uses 25 components, while the Chicago Fed uses 85, but we've been told that using more series makes very little difference.

Is it something to do with the type of series used? In particular, the Chicago Fed index uses no series relating to overseas transactions or to finance, while the Australian index takes account of six such series. In addition, all the US data are

adjusted for inflation. Intuitively, these differences don't seem to be enough, leaving the question: why do we need a far greater negative reading in the Australian index to signal recession?

The CFNAI also commits itself to two other thresholds. A sustained move above 0.2 in the three-month moving average signals recovery, and a move above 1 well into an expansion signals a pick-up in inflation. Do the authors have any thoughts on similar thresholds for the Australian index, or on the use of simple thresholds in the first place?

And now for something completely different. I can't get away from the feeling that recessions have something to do with unemployment. In this respect, it is perhaps noteworthy that the word unemployment appears just twice in the main body of the paper. Suppose that one in fact defined a recession as having taken place if and only if the unemployment rate has risen by a percentage point. Certainly our American visitors would have little trouble with such a definition. Because unemployment lags, one may have to do further work to determine the timing of the recessions. Recoveries would be defined as periods in which the unemployment rate was stable or falling. Again this would satisfy the Americans.

In the 1970s, there were three occasions in which Australian unemployment rose by more than a percentage point. It rose from 1.6 per cent to 2.9 per cent between the December quarter of 1970 and the September quarter of 1972 (we had only quarterly data on unemployment until 1978). It rose again, from 2.1 per cent to 5.4 per cent between June 1974 and December 1975, and then again, from 4.5 per cent to 5.9 per cent between June 1976 and December 1977. These episodes are consistent with the timing of recessions according to the GDP data, but not according to the coincident indicators. Of course, during the 1970s, real wages increased massively. Presumably some are prepared to say that the huge upward drift in unemployment in that decade was secular rather than cyclical, but to me the behaviour of unemployment adds weight to the GDP timing of recessions rather than that of the coincident indices.

Which brings us to 2000–2001, when the unemployment rate rose from 6 per cent in October 2000 to 7.2 per cent in October 2001. Nobody has ever labelled this episode as a recession to my knowledge, but how close does it come? It used to be said, at that time, that no country had ever introduced a GST (or a VAT) without quickly experiencing a recession. Perhaps none still has.

I am sure there are many more questions raised by this paper. This is not a criticism, in fact it suggests that the authors have done some very useful work – I just want to know more about the applications from here.

2. General Discussion

A number of participants asked about potential uses for the index constructed by the authors, particularly with regards to forecasting. In discussing this issue, one participant questioned whether the predictive capabilities of the index had been tested, while another suggested that it is important to have a better understanding

of the relative weights placed on various series in the index, and which series the index most closely matches, if it is going to be more than a 'black box'. In response, the authors noted that the index was constructed to be a coincident index, rather than a leading index, and that its predictive content had not been a criterion in its construction. Furthermore, Anthony Richards commented that the index approach is still considered experimental, and it remains to be seen whether the monthly index will perform as well as equivalent indices in other countries that are based on many more series.

Related to this, there was much discussion of the likely performance of the index in real time. One participant questioned what effect data revisions would have on the index in real time. In the context of forecasting of current-quarter GDP, another participant commented on some US experience comparing the performance of an index with judgmental forecasts from a team of forecasters, which had shown a number of challenges to be overcome in implementing an index approach. While judgemental forecasts still tend to be more accurate, the performance of these types of indices is likely to improve with further work.

The benefit of using a coincident index instead of GDP to gauge the state of the business cycle was questioned by some participants. One highlighted that GDP is itself a system of around 100 series, but with several theoretical advantages; in particular, the weights are determined by the relative importance of the sector in question in the economy, rather than estimated; and the system for its construction is time-tested. Furthermore, they commented that there are alternative methods for reducing the noise in GDP, such as the use of (Henderson) trend estimates. Anthony Richards responded by saying that it was not intended that these indices would replace GDP, but the analysis so far suggested strongly that they are better means of assessing the state of the business cycle in real time than applying trend filters to GDP, given the end-point problems involved in the latter approach.

A few participants also questioned the logic behind the inclusion of price variables when the index is aimed at representing real activity. In response, the authors noted that the index includes several nominal variables, and that the price series are therefore included to possibly remove trends in these series induced by inflation. This resulted in some discussion about whether the persistence of the index was due to the inclusion of inflation, which is itself very persistent; one participant asserted that at least one series included in the index needed to be persistent if the index was to be persistent itself, a point which was challenged by the authors.