BANK FOR INTERNATIONAL SETTLEMENTS

BIS Working Papers

No 148

The 2001 US recession: what did recession prediction models tell us?

by Andrew J Filardo

Monetary and Economic Department

March 2004

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The views expressed in them are those of their authors and not necessarily the views of the BIS.

Copies of publications are available from:

Bank for International Settlements
Press & Communications
CH-4002 Basel, Switzerland

E-mail: publications@bis.org

Fax: +41 61 280 9100 and +41 61 280 8100

This publication is available on the BIS website (www.bis.org).

© *Bank for International Settlements 2004. All rights reserved. Brief excerpts may be reproduced or translated provided the source is cited.*

ISSN 1020-0959 (print)

ISSN 1682-7678 (online)

Abstract

How predictable was the recent US recession? This paper evaluates the accuracy of several recession prediction models. In particular, traditional rule-of-thumb models using the composite index of leading indicators (CLI), Neftçi's sequential probability model, a probit model, and Stock and Watson's experimental recession indexes are compared. Despite the relatively mild depth of the recession, the models using the CLI performed particularly well. The results are robust across different types of models and with respect to the use of real-time data. The strong real-time performance stands at odds with earlier sceptical claims about the marginal usefulness of the CLI in predicting cyclical turning points, and complements the results in the earlier research of Filardo (1999). At a more conceptual level, the paper provides general support to the classical business cycle view that turning points of business cycles from expansion to recession are complex, possibly endogenous and non-linear, phenomena. The results also suggest that the impressive insights of Geoffrey Moore into the theory and construction of the CLI will continue to shape our understanding of business cycles well into the future.

Table of contents

Introduction[1]

Geoffrey Moore (1963) in his essay "What is a recession?" chose to answer this question by focusing on the famous description of the business cycle from Burns and Mitchell (1946):

> Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; this sequence of changes is recurrent but not periodic; in duration business cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar character with amplitude approximately their own.

The National Bureau of Economic Research (NBER) – considered by many to be the official arbiter of business cycle peaks and troughs – continues to use this guidance to define the phases of business cycles in the 21st century.

The choice of this description, then as well as now, was not motivated by its precision and irrefutable quality. As economic definitions go, it is quite inexact and vague. But the description is apt now for the same reason it has been circulating largely unchanged for nearly 80 years – it has resonated with those who study the recurrent ups and downs in US economic activity.

There has been far less success in translating this qualitative description into quantitative models to predict recessions. Various models have been offered over time with varying degrees of econometric sophistication to capture the salient features of the business cycle. Some are simple, maybe too simple. Some are quite sophisticated, reflecting state-of-the-art econometric modelling methods. Simple or sophisticated, at the end of the day, the yardstick with which to measure the performance of any recession prediction model is its ability to provide reliable, advanced warning of a recession.

The recent downturn in the United States offers another, though increasingly rare, opportunity to examine the "out-of-sample" reliability of recession prediction models. According to the NBER business cycle dating committee, the United States entered recession in March 2001. The economy prior to this date showed tangible signs that the risk of recession had risen. After surging in the late 1990s, economic activity began its rapid deceleration in the second half of 2000 as many economic factors weighed heavily on the expansion. Key economic factors included the dramatic decline in the stock market – especially the collapse of high-tech stock prices, the jump in oil and gas prices to very high levels, the increase in interest rates, a significant drop in industrial activity, and a build-up of what proved to be excessive inventories. Despite expansionary monetary and fiscal policies in 2001, the economy continued to decelerate, with consumption moderating and capital spending contracting. This recession was somewhat atypical in several dimensions. It has been one of the mildest on record, with the housing sector remaining remarkably healthy, durable goods consumption faring relatively well when compared to non-durable consumption spending, and the brunt of the recession disproportionately hitting the industrial sector.

How well did recession models predict the end of the longest US expansion on the record books? This paper examines the empirical performance of four popular recession prediction models. The next section describes the alternative models. The third section evaluates their empirical performance in the late 1990s and early 2000s using the revised and "real-time" data. The final section concludes that the superior performance of several "old-fashioned" recession prediction models highlights the continuing need to reconcile modern econometric methods with the insights and valuable intellectual contributions to our understanding of the business cycle by Geoffrey Moore and other early business cycle pioneers associated with the NBER.

[1] The views expressed are those of the author and not necessarily the views of the Bank for International Settlements. The author thanks Bob McGuckin and Ataman Ozyildirim of the Conference Board, and Jim Stock and Mark Watson for providing data on their respective business cycle measures. The author also thanks Palle Andersen, Claudio Borio, Bill English, Gabriele Galati and Steve Landefeld for helpful discussions. This paper was prepared for a book honouring Geoffrey H Moore.

Description of the four recession prediction models

This section briefly describes four recession prediction models considered by Filardo (1999). They include a simple rule-of-thumb model using the Conference Board's composite index of leading indicators (CLI), Neftçi's sequential probability model, a probit model, and Stock and Watson's experimental recession indexes.[2] These models all share a common heritage in the empirical business cycle literature pioneered during the 20th century. See Filardo (2003) for a recent discussion of various intellectual traditions, highlighting some of the main differences between the endogenous and exogenous views of business cycles. The endogenous view emphasises a complex interplay of confidence, real activity and the financial factors, while the exogenous view emphasises the role of independent shocks as the main source of business cycles. The conceptual differences translate into important implications for business cycle prediction because of the different roles business cycle phases play in the joint data generating process for economic activity under the two views. In the endogenous view, expansions and recessions play an intrinsic role in determining economic outcomes. Knowing the state of the business cycle helps to explain the likely direction of the economy. In contrast, the exogenous view puts weight on the extrinsic nature of cycles; that is, business cycle fluctuations produce patterns that exhibit features consistent with a definition of expansions and recession but the denotations are simply labels rather than an intrinsic part of the data generating process. The first three models using the leading indicators can be thought of as falling into the category of the intrinsic models of business cycle phases and the Stock-Watson model into the category of the extrinsic models.[3] To preview the central findings in this paper, the intrinsic business cycle models performed relatively well during the run-up to the recent recession, providing evidence that intrinsic business cycle models using the real-time CLI are useful macroeconomic monitoring tools.

Simple rules of thumb using the composite index of leading indicators

The composite index of leading indicators has played a central role in the long history of business cycle prediction at the NBER (Moore (1961), Zarnowitz (1992)). The composite index methodology was developed in the mid-20th century as a means to provide a summary of economic series that exhibited a leading relationship with the business cycle. Designing an index was no simple task, especially in an era when computers were rare and exotic electronic machines. The question was how to best summarise the information about the state of the business cycle contained in potentially hundreds of time series.[4] Moore, who some consider the father of the leading indicators, argued that a small set leading indicator series could be combined to yield a useful quantitative index that would provide early signals of changing economic activity (Moore (1961), Banerji and Klein (2000)). In addition, it was recognised at the time that a simple average of the various leading indicator series would put too much weight on series with high volatility relative to those with low volatility. Moore, therefore, advocated a volatility adjustment, which still today is the basis for averaging the component series of the composite index. In terms of modern econometric modelling methods, the weighting scheme is crude. Nonetheless, the composite index of leading indicators continues to attract considerable attention, especially at times thought to be turning points in the economy.

By construction, declines in the CLI are supposed to give advanced warning of an economic downturn, and hence should provide useful information as a signal of a future recession. In this sense, the rules of thumb are non-parametric models of recession prediction. The rules of thumb in this chapter are restricted to those using consecutive declines in the CLI as an indicator of imminent recession. For

[2] While these are a core set of empirical business cycle models constructed to predict recessions, the particular choice of models was restricted in order to offer a true out-of-sample forecasting evaluation of those models examined in Filardo (1999). The out-of-sample nature of the empirical exercise is meant to underscore the earlier finding that the composite index of leading indicators provides valuable information (when properly filtered) for business cycle phase prediction.

[3] See Filardo and Gordon (1999) for a comparison of several alternative intrinsic models of business cycles based on multivariate time-varying Markov switching models. The findings bolster the case for intrinsic business cycle models.

[4] Moore (1950) chronicles the early efforts of the NBER to sift through hundreds of economic series available at the time to find the important statistical indicators of business cycle expansions and recessions. Burns and Mitchell (1946) used roughly 500 series and identified 71 as useful. Moore winnowed the list further and found eight series that provided good leading properties. They included business failures, stock prices, new orders for durable goods, residential building contracts, commercial and industrial building contracts, average hours worked per week, new incorporations and wholesale prices.

example, a *k*-month rule of thumb would signal an imminent recession if there were *k* consecutive declines in the CLI:

<u>K-month rule-of-thumb model</u>

If {CLI_t<0, ..., CLI_{t-k} <0)}, then a recession signal is sent.

In general, multi-month rules of thumb that require consecutive declines are considered to be more reliable predictors of imminent recession than the month-to-month changes in the CLI (Hyman (1973), Vaccara and Zarnowitz (1978), Wecker (1979), Zarnowitz and Moore (1982)) because the month-to-month changes in the CLI often produce many false signals. To further filter out false signals, I consider rules of thumb that include a constraint that the consecutive CLI decline must be sufficiently large to send a valid signal of imminent recession.[5]

Neftçi's sequential probability model

Neftçi's sequential probability model is a non-linear method that provides an inference about a regime shift in the data generating process of the CLI data, which then can be used to infer a turning point in economy-wide activity. The theory behind this model comes from the literature on optimal stopping time and provides algorithms to assess the likelihood of a regime shift within a particular time-series data subsample. To use this model to predict recessions, several assumptions are required. First, a downturn in the CLI data can be accurately and reliably characterised as a shift in the distribution of the CLI data from an expansion distribution to a recession distribution. Second, because the method provides information only about a regime shift somewhere in the data subsample but not at an exact date, the lag between a downturn in the CLI data and its detection via the model is short. Third, a turning point in the CLI data provides reliable information about an imminent turning point in general economic activity.

With these assumptions, Neftçi (1982) and Diebold and Rudebusch (1989) provide a method to draw inferences about the likelihood of imminent recession. Technically, the model is a Bayesian recursion that uses CLI data to update the probability at time *t-1* that a turning point in the CLI data had occurred at some point in the subsample of data and can be calculated by the following equation:

$$P_t(t-\tau < Z \le t \mid CLI_{t-\tau},...,CLI_t) = \frac{[P_{t-1} + \pi^r(1-P_{t-1})]F^r}{[P_{t-1} + \pi^r(1-P_{t-1})]F^r + (1-P_{t-1})(1-\pi^r)F^e}, \qquad (1)$$

where P_t is the conditional probability at time *t* of a turning point (represented by the integer-values random variable *Z* that is a time index for the first period after a regime switch from an expansion distribution to a recession distribution) having occurred in the data subsample $\{CLI_{t-\tau},...,CLI_t\}$, P_{t-1} is the analogous probability at time *t-1*, π^r is the unconditional transition probability of the economy entering a recession under the assumption that the economy is in expansion, and F^e and F^r are the density functions of the CLI data under the assumption that they came from an expansion distribution or a recession distribution, respectively.[6] This equation highlights the feature of the model that the exact time of the turning point is not estimated. Rather, the method only provides an inference about whether a turning point occurred at some time between $t-\tau$ and *t*.

To use this model to predict recessions, the estimated probability of a turning point, P_t, is compared to a prespecified threshold level of confidence which is intended to allow a small probability of type I

---

5 Other rules might include more elaborate sequences of the CLI (Zarnowitz and Moore (1982)) to help filter out the false signals. In addition, Klein and Moore (1983) have pointed out that the CLI also provides information about growth cycle rather than business cycle turning points – and hence the fluctuations in the series can send false signals about imminent business cycle recessions.

6 The probability distribution functions of the CLI data, F^e and F^r, are modelled as being normally distributed around mean growth rates of the three-month moving average of the CLI during expansionary and recessionary periods. The three-month moving average of the CLI smooths the wiggles, or noise, in the CLI data. Following Diebold and Rudebusch (1989), the transition probability from expansion to recession, π^r, is assumed to be independent of the time elapsed in the phase and is set to 0.02, which is consistent with results from Hamilton (1989). Alternatively, the transition probabilities could be modelled as being time-varying as in Filardo (1994).

error. This level of confidence reflects two types of inference: confidence of a statistically significant regime shift in the CLI data and confidence that such a shift portends a turning point in general economic activity. Following Diebold and Rudebusch (1989), this threshold is assumed to be 95%, which represents a conventional burden of proof for this type of model, ie a small probability of Type I error. Operationally, soon after the model's probability, P_t, exceeds the threshold level of confidence, the recursion is reinitialised to search for another turning point over a subsequent subsample of CLI data.[7]

Probit model

Consistent with previous research by Estrella and Mishkin (1998) and Lamy (1997), the probit model is a (non-linear) regression model that translates information contained in leading indicators into a probability of recession at a particular time horizon:[8]

$$P(recession \mid X_{t-k}) = F(\beta_0 + \beta_1 TS_{t-k} + \beta_2 CS_{t-k} + \beta_3 SP500_{t-k} + \beta_4 CLI_{t-k}) \qquad (2)$$

The variables that are assumed to help to predict a recession are the change in the term spread (*TS*), change in the corporate spread (*CS*), S&P 500 return (*SP500*), and growth rate of the CLI. To predict a recession *k* months ahead, the model is estimated using lagged information as represented in the vector $X_{t-k} = \{TS_{t-k}, CS_{t-k}, SP500_{t-k}, CLI_{t-k}\}$. The threshold criterion for this model is 50%. If the probability is less than 50%, the model signals an expansion because an expansion is more likely than a recession; if the probability is above 50%, a recession is more likely.[9]

Stock and Watson's experimental recession indexes

Stock and Watson (1989, 1991, 1993) built a sophisticated econometric time series model to infer the probability of recession, which is estimated in two steps. First, business fluctuations are viewed through the lens of a multi-equation unobserved variable model. The unobserved component is assumed to represent the common business cycle factor shared by four cyclically sensitive variables. The cyclically sensitive variables are industrial production, real personal income less transfer payments, real trade sales, and employment hours in non-agricultural establishments. The model has the following structure:

$$\begin{aligned}
\Delta \mathrm{X_t} &= \beta + \gamma(\mathrm{L})\Delta \mathrm{C_t} + \mathrm{u_t} \\
\mathrm{D(L)u_t} &= \varepsilon_\mathrm{t} \\
\Delta C_t &= \mu_c + \lambda_{cc}\Delta C_{t-1} + \lambda_{CY}(L)Y_{t-1} + \nu_{Ct} \\
Y_t &= \mu_Y + \lambda_{YC}(L)\Delta C_{t-1} + \lambda_{YY}(L)Y_{t-1} + \nu_{Yt}
\end{aligned} \qquad (3)$$

The growth rates of the cyclically-sensitive variables are stacked in the vector ΔX_t, the growth rate in the unobserved coincident indicator of economic activity is $\Delta \mathrm{C}_t$, ε is an uncorrelated error term, and

---

7 In practice, the recursion was reinitialised 18 months after the trough and a year after a false signal. The exception to this rule occurred during the 1980 recession because it was followed so closely in time by the 1981-82 recession. In this case, the recursion was reinitialised three months after the July 1980 trough.

8 For an international perspective, see Bernard and Gerlach (1996).

9 The 50% threshold is different than the threshold level in the Neftçi model because of the different type of model inference about the business cycle. In particular, the probit model assesses the probability of being in a recession *k* periods ahead rather than assessing the probability that sufficient information has become available to infer that the CLI distribution had switched and that the switch is accurately signalling an imminent regime switch in economic activity. In the probit model, the state of the economy *k* periods ahead is assumed to be in only one of two states: either recession or expansion. Hence, for example, the inferred probability of recession of less than 50% indicates that the inferred probability of expansion is greater than the probability of recession. With a symmetric loss function, the best assessment of the state of the economy *k* periods ahead is expansion. It should be noted, however, that the 50% threshold may be considered somewhat arbitrary because in an optimal decision-making setting, the optimal threshold would depend on the nature of the loss function. If the loss function is not symmetric, the optimal threshold would generally not be 50%. Other loss functions might justify a neutral range where it might be optimal not to call a turning point.

$\gamma(L)$, $D(L)$, $\phi(L)$ are standard lag polynomials. The lag polynomials, λ_{ij}, are estimated using statistical criteria and the error terms (ν_{Ct}, ν_{Yt}) are assumed to be uncorrelated and independent of ε_t. The last two equations in the model provide a link between the leading indicator series and the coincident index of economic activity, where leading indicators, *Y*, are used to help to predict the growth rate of the unobserved coincident indicator ΔC_t.[10] With this model, Stock and Watson define a leading index, not as the simple volatility-weighted sum of the leading indicator variables, but rather as a weighted average of the indicators that have weights chosen to minimise the mean squared forecast error of the coincident indicator ΔC_t six months ahead.[11]

In the second step, Stock and Watson use the estimated model in the first step to generate forecasts of C_{t+k}. Defining a multi-period pattern for C_{t+k} consistent with past recessionary episodes, a probability index of recession six months ahead, XRI, is constructed. This second step is called the *pattern recognition step*.

How well did the models predict the 2001 turning point?

The NBER Dating Committee announced that the recent recession began in March 2001, but as is usual, dating the initial month of the recession is subject to considerable uncertainty. While there is little controversy that the US economy contracted in 2001, there are some questions about when the contraction began. For example, industrial production peaked in June 2000 as did real manufacturing and wholesale-retail sales, real personal income less transfers peaked in November 2000, and total non-agricultural payroll employment peaked in March 2001.[12] The large benchmark revision to real GDP in July 2002 raised further questions about the starting date of the recession. Prior to the benchmark revisions, the Bureau of Economic Analysis estimated that real GDP contracted only once during the year, in the third quarter. This new snapshot of the data showed that the US economy contracted during three consecutive quarters starting in the second quarter. The revised picture also showed an economy that experienced a deeper and longer contraction than previously estimated.

Given the uncertainty about the starting date of the recession, it should not be surprising that the extent of the advanced warning from different recession prediction models could vary considerably. Nor should it be surprising that, for a given model, differences between the results using real-time data and the recent data vintages (ie subsequently revised data) could be significant. This section reviews the performance of the four recession prediction models in light of these possibilities.[13]

Did the recession prediction models predict the start of the 2001 recession?

The *CLI rules of thumb* performed fairly well in predicting the March 2001 business cycle peak. Table 1 shows that the two-month rule signalled an imminent recession eight months prior to the NBER-denoted starting date. The performance, however, may be somewhat suspect because the two-month rule has had the tendency to send frequent false signals. For example, over the past four decades, the two-month rule produced 19 false signals.

[10] The components of their leading index are building permits, real manufacturers' unfilled orders (smoothed), the trade-weighted index of the nominal exchange rate, part-time work because of slack, the 10-year treasury bond yield, and the yield spread between the 10-year bond and the one-year Treasury bill.

[11] Technically, the parameters are estimated by minimising the mean of the squared errors over the sample period. It is well known that a good in-sample fit may not guarantee a good out-of-sample fit, ie a good predictor of turning points. Conversely, a good model of turning points may not do well at forecasting economic activity during "normal" economic times, ie those periods not subject to turning points in the business cycle. See also Kling (1987) and Wecker (1979).

[12] In addition, the Federal Reserve's aggressive easing actions in early 2001, especially the one at the unscheduled January meeting, suggested a certain urgency to act at the time. These actions also support the view that there was evidence of a significant downdraft in economic activity prior to the NBER-designated turning point date.

[13] The concern about using real-time data has been well established in the literature on the leading indicators. Moore (1950, 1961) and Moore and Shiskin (1967) performed some of the original out-of-sample studies of the CLI. Moore and Zarnowitz (1982) compare the performance of the CLI in real time and propose robust criteria (ie the "band approach") to deal with data uncertainty.

In contrast, the two-month rule with threshold and the three-month rules provided an early warning of imminent recession in November 2000, four months ahead of the official starting date. These rules were also subject to many false signals in the 1960-2002 period. What do the false signals indicate about the performance of these rules? The false signal tally for these rules appears high on first examination, but it is important to note that most of the misses were influenced by the economic slowdown in 1966. This episode accounts for seven, seven and six of the false signals for the two-month rule with threshold, the three-month rule and the three-month rule with threshold, respectively. Excluding the 1966 episode, for example, the three-month rule of thumb produced only two false signals: one at the end of 1991 when the economy was experiencing anaemic economic activity and one in the spring of 1995 when the economy experienced a mild slowdown due in part to the rise in commodity prices and interest rates. The four-month rules sent no advanced warning of an imminent recession. These rules, however, may be too stringent. The post-1960 record shows that that the power of these rules to predict recessions was spotty.

The usual complaint about the rules of thumb is that their predictive power may be overstated when using the latest vintage of data because of the sensitivity of the rules' performance to revisions in the CLI data. Figure 1 shows the number of false signals from the rules using different vintages of the CLI data. In this figure, the number of false signals is calculated for each vintage of data (denoted by the date of the final observation in the series) from March 1992 to July 2002. For each vintage of data, the number of false signals since March 1992 was tabulated. To be sure, the number of false signals varies with the vintages. For example, in early 1994, the then-published CLI data provided three false signals of imminent recession using the two-month rule. The number of false signals rose to seven by early 1999 and to 10 by 2001. Subsequent CLI data revisions led to a reduction in the number of false signals to five. One way to interpret this figure is to note that if false signals arose randomly with a constant probability of occurrence, the bars would generally rise monotonically, except possibly when the source data or methodology underlying the CLI series was fundamentally changed, as in mid-2001.[14]

While the real-time analysis of the two-month rule might suggest the possibility of random errors, the other rules do not. In fact, several of the rules exhibit systematic "errors" related to the slowdown in economic activity during 1995. In the case of the three-month rule, the 1995 slowdown accounts for the majority of the false signals. Moreover, from a policymaker's point of view, the false signal in 1995 shared by most of the rules might have been somewhat useful in the sense that the signal reflected a period of softer economic activity. Having such information even in the form of a false signal may not be undesirable. Overall, the real-time performance of the CLI rules of thumb in the 1990s and early 2000s is quite strong compared to what might have been expected based on some of the research around the time of the 1990-91 recession into the predictive power of the CLI (eg Diebold and Rudebusch (1992), Emery and Koenig (1993), Hamilton and Perez-Quiros (1996)).

The *Neftçi model* predicted the March 2001 peak five months in advance of the NBER-designated peak. Figure 2 shows the general performance of the Neftçi model since late 1959. The model using the CLI has provided a median lead time of seven months using the 95% threshold. However, the model provided false signals in the mid-1960s during the dramatic industrial slowdown. As discussed above, such a false signal should not necessarily be interpreted as a failure of the model. There were also two false signals in the first half of the 1990s. These false signals coincided with slower economic activity but the findings certainly raise questions about the sensitivity of the Neftçi model.

This model also has been subject to complaints that its signals of recession often look much better in retrospect because of revisions to the CLI data (Diebold and Rudebusch (1992), Filardo (1999)). Figure 3 shows that this conclusion is not warranted in the latest recessionary episode. The dotted line is the probability of imminent recession at time t using the time t vintage of the CLI data. While there are some differences between the dotted line and the thick line (the estimated probability using the July 2002 vintage of the CLI data), the discrepancies are small and the lead time of five months is unaffected.

The *probit model* also performed reasonably well in predicting the 2001 turning point. The dark (horizontal) hash marks in Figure 4 represent the probability of a future recession at each specified forecast horizon. For example, the three-month-ahead forecast model indicates that the probability of

[14] In August 2001, the US Department of Commerce's Census Bureau and Bureau of Economic Analysis converted their industrial classification scheme from an SIC classification system to the NAICS classification system. This change significantly affected some component series of the CLI (see Conference Board).

recession initially exceeded 50% in October 2000. In other words, this model predicted that the economy would be in recession in January 2001.[15] Largely consistent with this finding, the six-month-ahead forecast model's recession probability initially exceeded 50% in December 2000, and the nine-month-ahead forecast model's probability exceeded this threshold in October 2000. The 12-month-ahead forecast model never signalled an imminent recession using the 50% rule but came close with a reading of 49.1% in December 2000. Taken together, the probit model sent fairly clear signals of an imminent recession by the end of 2000. By autumn 2001, the recessionary signals began to wane for the shorter horizon versions of the model despite the economic disruptions that followed the terrorist attacks in September 2001.[16]

Figure 4 also provides a summary of the real-time performance of the probit model as indicated by the dispersion of the probability estimates at each date associated with the various data vintages. The probit model exhibits signs of considerable sensitivity to real-time data, even though three of the variables are not subject to revision. Part of the sensitivity is due to CLI data revisions, but part of it reflects sampling error of the parameters arising from longer samples. In each of the four panels in Figure 4, the high and low probability estimates from each model for the relevant vintages are denoted by the top and bottom of the vertical lines at each date. The variation is relatively large near the NBER-designated turning point and generally larger for the longer forecast horizons. The variation is so high in the 12-month-ahead forecast model that it began sending false signals in late 1998 with an earlier vintage of data. Overall, the probit model appears to be sensitive to real-time data but it nonetheless sent advanced warning, albeit noisy, of the 2001 peak.

Stock and Watson experimental recession indexes did not perform well in the recent recessionary episode. Figures 5 and 6 plot the actual and real-time estimates of their experimental recession indexes (XRI and XRI-2). While Stock and Watson (1993) recommend no particular probability threshold to trigger a signal of imminent recession, the historical XRI in pre-1990 recessions generally exceeded 50% prior to the starting date. Using this 50% convention, the XRI missed calling the recent downturn by a wide margin. Stock and Watson also produced an alternative recession index, called XRI-2. The XRI-2 puts less weight on financial variables in its list of leading indicators than does the XRI. Part of the motivation for an alternative was the poor performance of the XRI in the 1990-91 recession, which was partly attributed to the atypical timing of interest rate swings at the time. Despite the heavier weight on quantity-based leading indicators, the XRI-2 does not perform much better than the XRI. In addition, the real-time performances of the XRI and XRI-2 do not differ remarkably from that of the June 2002 vintage.

In early 2001, the *Stock and Watson Indicator Report* began including a new recession index, the XRI-C. The XRI-C measures the contemporaneous probability of recession and is calculated as the probability of recession at time t using information up to time t. In contrast, the XRI and XRI-2 assess the probability of recession six months ahead. Figure 7 shows the superior performance of the XRI-C, relative to that of the XRI and XRI-2. The XRI-C rises abruptly in late 2000 and early 2001 - suggesting that the recession may have started somewhat sooner than the NBER date. The real-time XRI-C (denoted by the dotted line) shows, however, that the XRI-C is subject to large revisions. For example, the XRI-C declined sharply in March 2001 to 13%, but was revised later to 56%.

The results of the Stock and Watson recession indexes represent a significant challenge for business cycle researchers. Without a doubt, the Stock and Watson experimental recession indexes are built on one of the strongest scientific foundations in this literature. To be sure, missing a turning point (or two) by itself does not necessarily reveal a fatal flaw in a recession prediction model. Recessions are complex economic phenomena that are sufficiently different from episode to episode to humble even the best built recession prediction models.

However, dismissing the latest empirical failure of the Stock and Watson recession indexes as a chance miss may be too facile. One early concern about the indexes was that they were too ambitious because they were designed to focus their predictive power at a six-month horizon. Emphasis on a fixed lead might be too stringent a constraint for recession prediction models because lead times between leading indicator series and coincident measures of economic activity may exhibit substantial

15 It is interesting to note that the last signal of recession at the three-month horizon occurred in October 2001, indicating that the last month of recession was likely to be January 2002.

16 In general, the three financial variables in the model are statistically significant. The CLI tends to be statistically significant at the shorter horizons.

variability from recessionary episode to recessionary episode. In the latest episode, however, this constraint does not seem to be particularly binding. The other recession prediction models considered in this chapter provided relatively clear signals of imminent recession at a horizon close to six months. Therefore, fine-tuning the horizon of the Stock and Watson model may not be the solution to the conundrum. Zarnowitz (1992) has also suggested that the Stock and Watson leading indicators may put too much weight on interest rates and financial variables and not enough on the traditional variables in the CLI. Given the performance of XRI-2, which was constructed to deal with such criticisms, further data mining along this dimension may be of limited value.

Several lines of inquiry might deserve further exploration. First, it might be useful to extend the estimation period of the Stock and Watson model back before 1960. While there would be a host of econometric and data issues to deal with, the longer estimation period and the greater number of turning point episodes may increase the power of the model to predict turning points. Second, the apparent increase in the trend growth rate in the late 1990s may have been sufficiently at odds with the fixed parameter assumptions of the Stock and Watson model that allowing for time-varying parameters or a structural break in the parameters could resurrect the basic approach. Third, and possibly most fruitful, it might be important to incorporate non-linear features of business cycles. As Stock and Watson (1993) noted, their model does not include an intrinsic role for phase dependence. However, there is some evidence that the conditional means of economic variables may be phase dependent and there may be unconditional and conditional dependence in the phases (Hamilton (1989), Filardo and Gordon (1998)). If a data generating process underlying business cycle fluctuations is highly non-linear, then Stock and Watson's linear model with pattern recognition may be missing important features that could help increase the predictive power for turning points. Whatever the solution, further investigation into the 2001 recession may shed more light on the appropriate modern econometric approach to modelling business cycle fluctuations.

Conclusion

The various recession prediction models had mixed success in predicting the start of the 2001 recession. The non-parametric CLI rules of thumb provided a four- to eight-month early warning of the peak. The advanced warning from the Neftçi model was also consistent with this range. It is not always true that these two models send similar lead times, but in this recent recessionary episode the fluctuations in the CLI data were strong enough to affect both models in a similar way. Moreover, the results from these two models were relatively robust to the use of real-time data. Even though the probit model does not provide a perfectly comparable lead-time calculation, it too generated probabilities of recession that rose significantly prior to March 2001, but the sensitivity of the results to the use of real-time data might raise some concerns about the model's reliability and deserves further research. The Stock and Watson experimental recession probability indexes performed less well.

In light of recent history, the superior performance of the CLI-based models may help to resurrect their perceived usefulness. At the very least, those who track the economy – and attempt to predict turning points – may renew their interest in the value of the CLI data and the traditional analysis of business cycle fluctuations that has sometimes been belittled as "measurement without theory." For business cycle researchers, the results stand as a challenge to those who have had reservations, if not doubts, about the marginal predictive content of the CLI. While predictions about the future are always subject to considerable risk, it seems reasonable to expect with some confidence that the impressive contributions of Geoffrey Moore to the theory and construction of the CLI and to our understanding of business cycles will help to lead to a better synthesis of traditional business cycle analysis and modern econometric practice of recession prediction.

References

Auerbach, Alan (1982): "The index of leading economic indicators: measurement without theory thirty-five years later", *Review of Economics and Statistics*, November.

Banerji, Anirvan and Philip A Klein (2000): "Father of leading indicators: the legacy of Geoffrey H Moore", *International Journal of Forecasting*, April-June.

Bernard, H and S Gerlach (1996): "Does the term structure predict recessions? the international evidence", *BIS Working Papers*, no 37, September.

Braun, Philip and Victor Zarnowitz (1989): "Comment on Stock and Watson," *NBER Macroeconomics Annual*, Cambridge: MIT Press.

Burns, Arthur F and Wesley C Mitchell (1946): *Measuring business cycles*, New York: NBER.

The Conference Board: *Business Cycle Indicators*. Various issues

Diebold, Francis X and Glenn D Rudebusch (1989): "Scoring the leading indicators", *Journal of Business*, July.

——— (1991): "Forecasting output with the composite index of leading indicators: a real-time analysis", *Journal of American Statistical Association*, September.

——— (1992): "Turning point prediction with the composite leading index: an ex ante analysis", in Kajal Lahiri and Geoffrey H Moore (eds), *Leading economic indicators,* Cambridge: Cambridge University Press.

Dueker, Michael J (2002): "Regime-dependent recession forecasts and the 2001 recession", *Federal Reserve Bank of St Louis Review*, November/December.

Emery, Kenneth M and Evan F Koenig (1993): "Why the composite index of leading indicators doesn't lead", *Federal Reserve Bank of Dallas Research Paper* no 9318, May.

Estrella, Arturo and Frederic S Mishkin (1998): "Predicting US recessions: financial variables as leading indicators", *Review of Economics and Statistics*, February.

Filardo, Andrew J (1994): "Business cycles and their transitions", *Journal of Business and Economic Statistics*, July.

——— (1999): "How reliable are recession prediction models?", *Federal Reserve Bank of Kansas City Economic Review*, second quarter.

——— (2003): "The enduring core of business cycles", unpublished *BIS Working Paper*, June.

Filardo, Andrew J and Stephen F Gordon (1998): "Business cycle durations", *Journal of Econometrics*, July.

——— (1999): "Business cycle turning point: two empirical business cycle model approaches", in Philip Rothman (ed) *Nonlinear time series analysis of economic and financial data*, Boston: Kluwer Academic Publishers.

Fredman, Giela T and Michael Niemira (1991): "An evaluation of the composite index of leading indicators for signalling turning points in business and growth cycles", *Business Economics*, October.

Hamilton, James D (1989): "A new approach to the economic analysis of nonstationary time series and the business cycle", *Econometrica*, March.

Hamilton, James D and Gabriel Perez-Quiros (1996): "What do the leading indicators lead?", *Journal of Business*, January.

Hymans, Saul H (1973): "On the use of leading indicators to predict cyclical turning points", *Brookings Papers on Economic Activity*, no 2.

Klein, Philip A and Geoffrey H Moore (1983): "The leading indicator approach to economic forecasting – retrospect and prospect", *Journal of Forecasting*, 2, April-July.

Klein, Philip A, and Michael P Niemira (1994): *Forecasting financial and economic cycles*, New York: John Wiley and Sons.

Kling, John L (1987): "Predicting the turning points of business and economic time series", *Journal of Business*, April.

Lamy, Robert (1997): "Forecasting US recessions: some further results from probit models", *Finance Canada*, working paper, May.

McGuckin, Robert H, Ataman Ozyildirim and Victor Zarnowitz (2001): "The composite index of leading economic indicators: how to make it more timely", *NBER Working Paper*, 8430, August.

Moore, Geoffrey H (1950): "Statistical indicators of cyclical revivals and recessions", *NBER Occasional Paper*, 31.

——— (1961): Business cycle indicators, Princeton, NJ: Princeton University Press for NBER.

——— (1963): "What is a recession?", *The American Statistician*, October.

——— (1978): "Why do the leading indicators lead?" *Across the Board*, May.

——— (1979): "The forty-second anniversary of the leading indicators", in W Fellner (ed), *Contemporary economic problems,* Washington, DC: American Enterprise Institute.

——— (1983): *Business cycles, inflation, and forecasting*, Cambridge: Ballinger Publishing Company.

Moore, Geoffrey H and Julius Shiskin (1967): "Indicators of business cycle expansions and contractions", *NBER Occasional Paper*, 103.

Moore, Geoffrey H and Victor Zarnowitz (1982): "Sequential signals of recession and recovery", *Journal of Business*, January.

——— (1986): "Major changes in cyclical behaviour", in Robert Gordon (ed), *The American business cycle: continuity and change*, Chicago: University of Chicago Press.

——— (1992): "Forecasting recessions under the Gramm-Hollings law", in Kajal Lahiri and Geoffrey H Moore (eds), *Leading economic indicators*, Cambridge: Cambridge University Press.

Neftçi, Salih (1982): "Optimal prediction of cyclical downturns", *Journal of Economic Dynamics and Control*, August.

Sims, Christopher A (1989): "Comment on Stock and Watson", *NBER Macroeconomics Annual*, Cambridge: MIT Press.

Stock, James H and Mark W Watson (1989): "New indexes of coincident and leading economic indicators", *NBER Macroeconomics Annual*, Cambridge: MIT Press.

——— (1991): "Turning point prediction with the composite leading index: an ex ante analysis", in Kajal Lahiri and Geoffrey H Moore (eds), *Leading economic indicators*, Cambridge: Cambridge University Press.

——— (1993): "A procedure for predicting recessions with leading indicators: econometric issues and recent experience", in James H Stock and Mark W Watson (eds), *Business cycles, indicators, and forecasting*, Chicago: University of Chicago Press.

——— Various issues, "Stock and Watson indicator report".

Vacarra, Beatrice N and Victor Zarnowitz (1978): "Forecasting with the index of leading indicators", *NBER Working Paper*, 244, May.

Wecker, William E (1979): "Predicting the turning point of a time series", *Journal of Business*, January.

Zarnowitz, Victor (1992): *Business cycles: theory, history, indicators, and forecasting*, Chicago: University of Chicago Press.

——— (1998): "Has the business cycle been abolished?", *Business Economics*, October.

——— (1999): "Theory and history behind business cycles: are the 1990s the onset of a golden age?", *NBER Working Paper*, 7010, March.

——— (2000): "The old and the new in US economic expansion of the 1990s", *NBER Working Paper,* 7721, May.

Table 1

CLI rules of thumb and the start of the 2001 recession

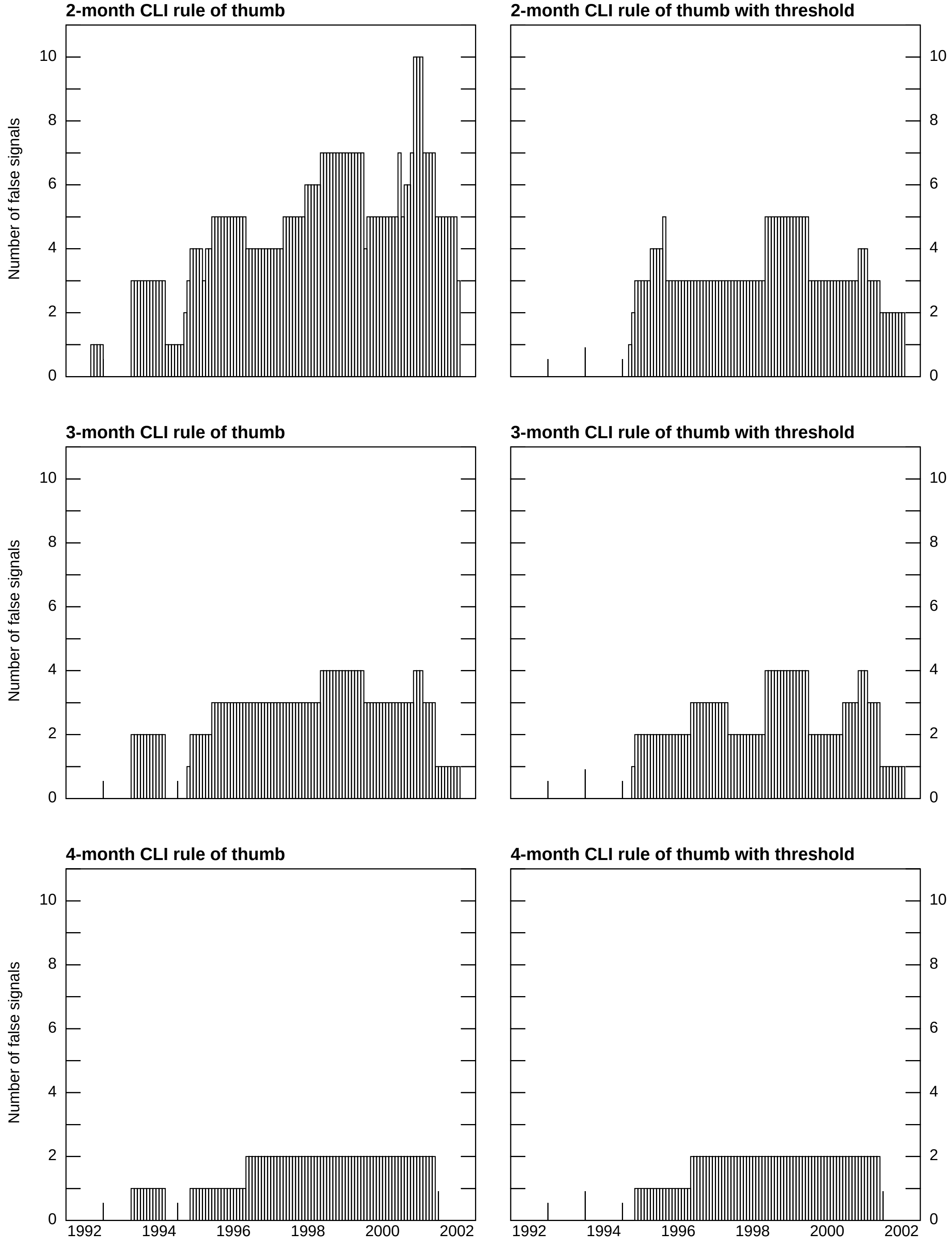
Timeliness and accuracy of various CLI rules of thumb

	Advanced warning of the start of a recession (in months)					
	2-month rule	**2-month rule with threshold**	**3-month rule**	**3-month rule with threshold**	**4-month rule**	**4-month rule with threshold**
Start of recession						
May 1960	10	9	9	9	8	7
January 1970	7	7	6	6	no signal	no signal
December 1973	8	8	7	7	6	6
February 1980	14	14	13	13	no signal	no signal
August 1981	7	7	6	6	no signal	no signal
August 1990	3	3	–3	–3	1	–1
Mean (lead time)	8	8	7	7	5	5
April 2001	8	4	4	4	no signal	no signal
	Number of signals of recession without an imminent onset of recession					
False signals	19	13	9	8	6	5

Note: False signals are defined as those signals that fall outside a one-year period before or after a peak date.

Figure 1

Variation in the number of false signals by data vintage

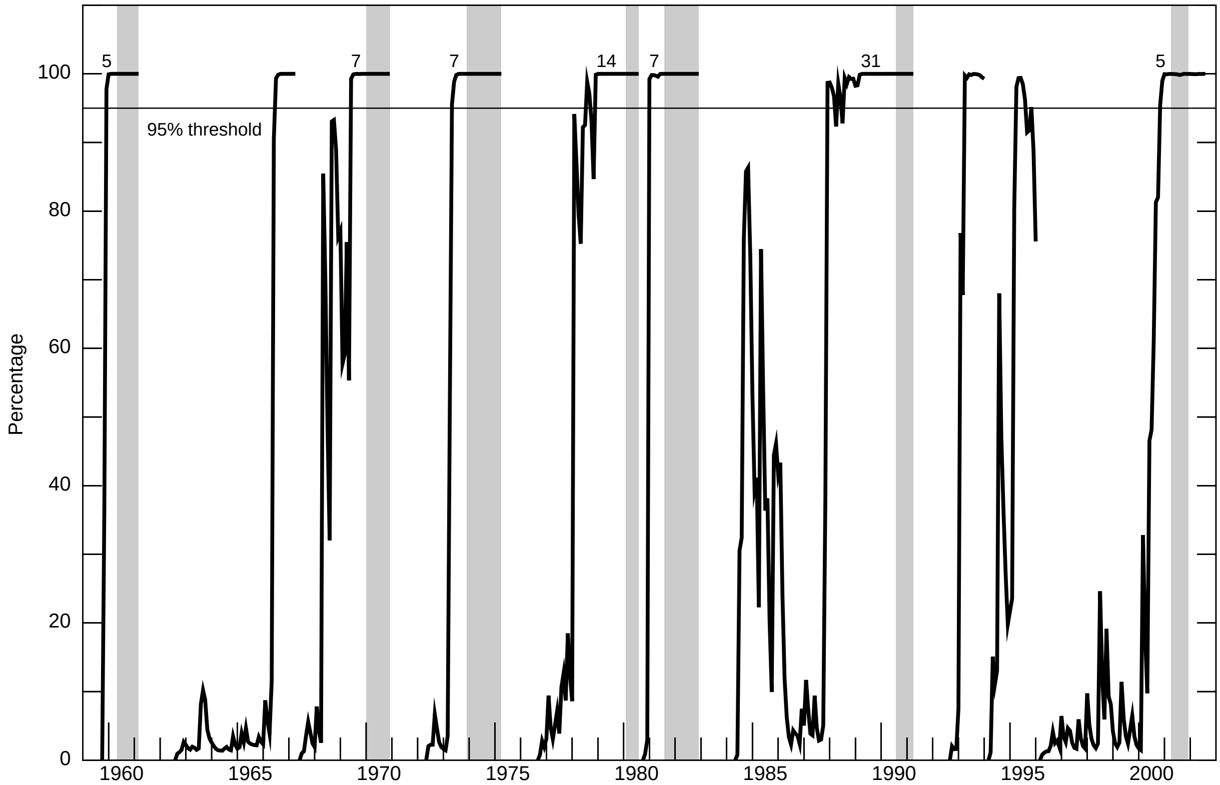


Note: The height of the narrow vertical bars represents the number of false signals from the respective rules of thumb for each of the real-time data vintages (vintages from March 1992 to July 2002).

Sources: BIS; Author's calculations.

Figure 2

Probability of imminent recession using Neftçi's model

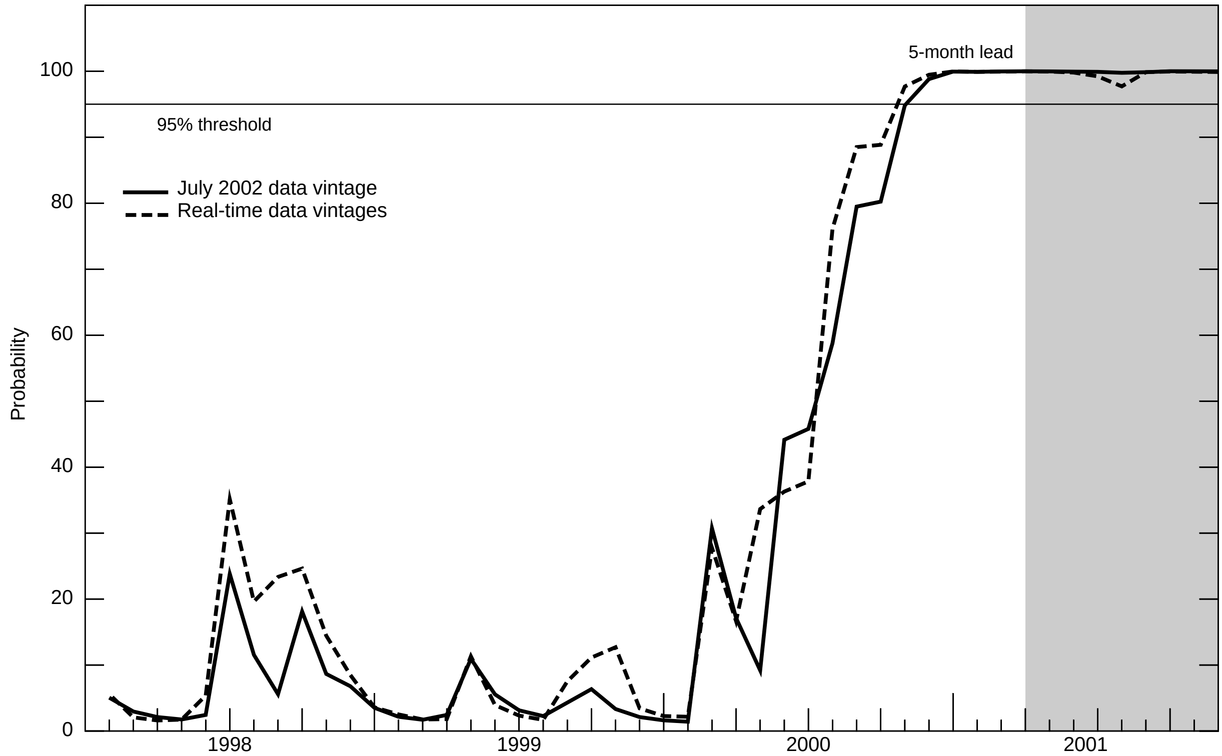


Note: NBER recessions are represented by the shaded bars. The numbers next to the bars indicate the number of months prior to the peak when the probability is above the 95% threshold level of confidence. The sample runs from January 1959 to July 2002.

Sources: NBER; BIS; Author's calculations.

Figure 3

Real-time performance of Neftçi's model

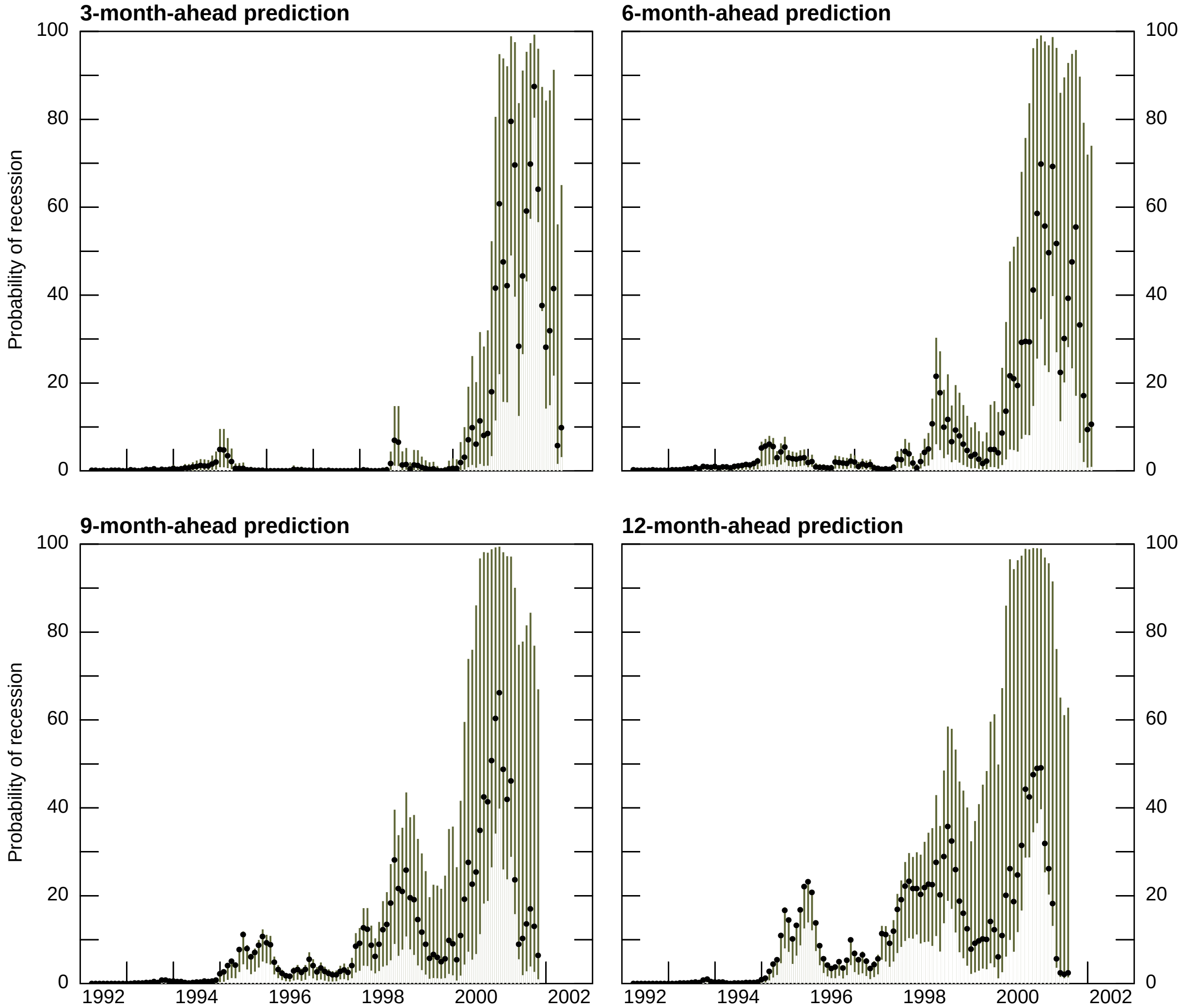


Note: The NBER Business Cycle Dating Committee dated the peak of the expansion as March 2001 and the trough as November. 2001.

Sources: NBER; BIS; Author's calculations.

Figure 4

Performance of probit model



Note: The dark hash lines indicate the probability of recession using the July 2002 vintage of data. The vertical lines represent the range of probability estimates (minimum to maximum) from the model using all relevant real-time data vintages (vintages from March 1992 to July 2002).

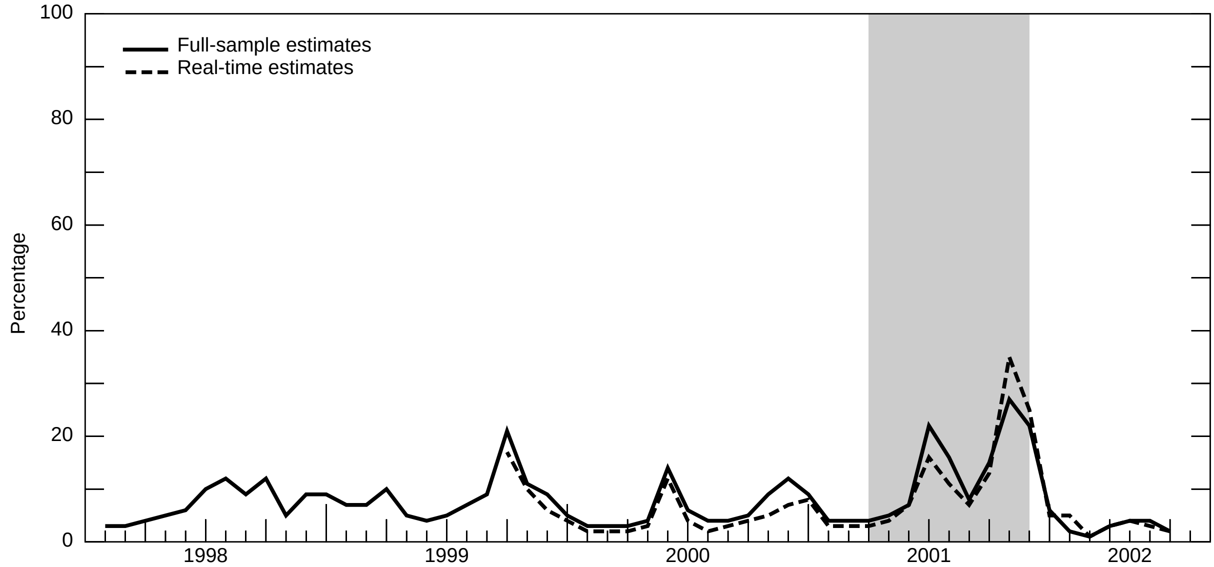
Sources: BIS; Author's calculations.

Figure 5

Performance of Stock and Watson's experimental recession index

Stock and Watson's leading recession index (XRI)

Real-time estimates versus full-sample estimates in latest recessionary period

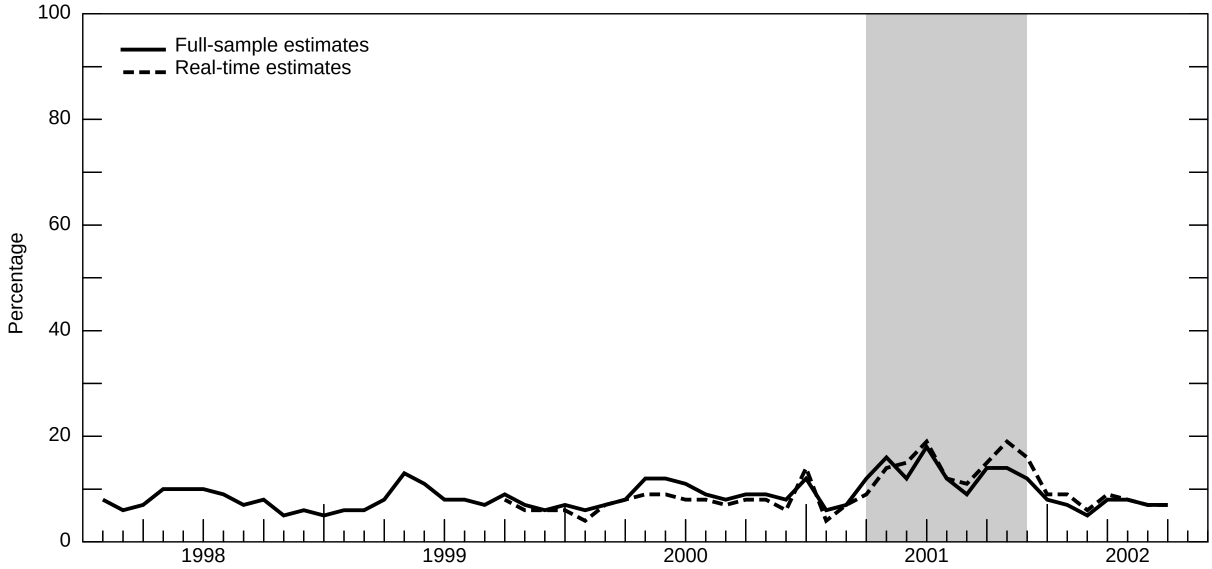


Sources: NBER; BIS; Author's calculations.

Figure 6

Performance of Stock and Watson's alternative experimental recession index

Stock and Watson's non-financial leading recession index (XRI-2)
Real-time estimates versus full-sample estimates in latest recessionary period



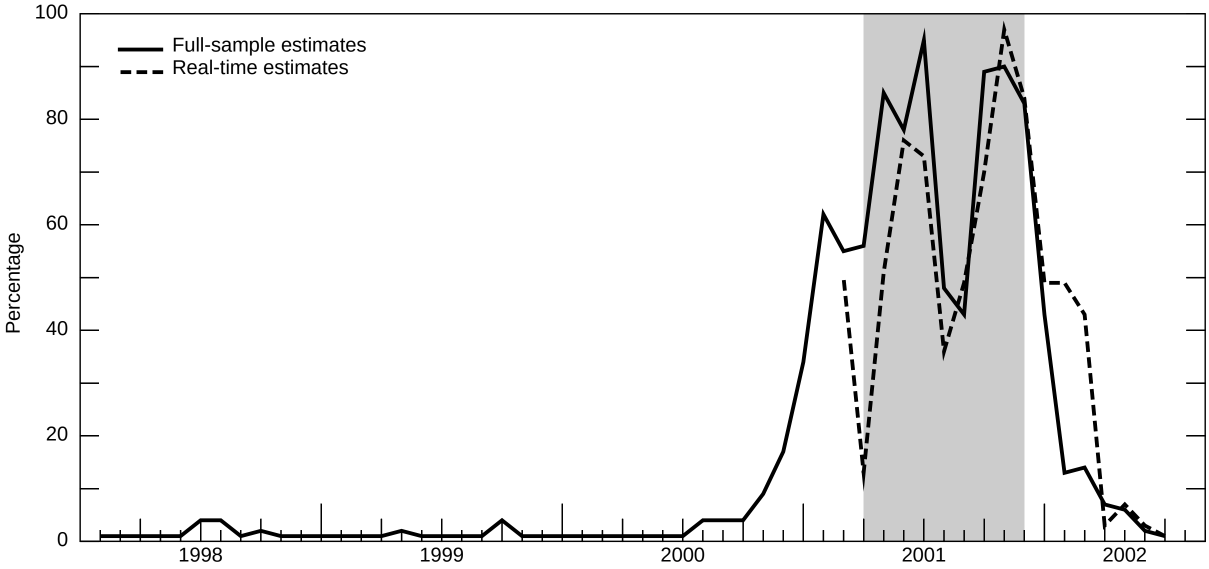
Sources: NBER; BIS; Author's calculations.

Figure 7

Performance of Stock and Watson's experimental coincident recession index

Stock and Watson's coincident recession index (XRI-C)

Real-time estimates versus full-sample estimates in latest recessionary period



Note: The real-time coincident recession index (XRI-C) was first published in the February 2001 *Stock and Watson Indicator Report.*

Sources: NBER; BIS; Author's calculations.

Previous volumes in this series

No	Title	Author
147 February 2004	Whither monetary and financial stability? The implications of evolving policy regimes	Claudio Borio and William R. White
146 January 2004	Consumer credit scoring: do situational circumstances matter?	R B Avery, P S Calem & G B Canner
145 January 2004	Are changes in financial structure extending safety nets?	William R White
144 October 2003	Transparency versus constructive ambiguity in foreign exchange intervention	Priscilla Chiu
143 October 2003	The Balassa-Samuelson effect in central Europe: a disaggregated analysis	Dubravko Mihaljek and Marc Klau
142 October 2003	Three episodes of financial fragility in Norway since the 1890s	Karsten R Gerdrup
141 September 2003	Financial strains and the zero lower bound: the Japanese experience	Mitsuhiro Fukao
140 September 2003	Asset prices, financial imbalances and monetary policy: are inflation targets enough?	Charles Bean
139 September 2003	External constraints on monetary policy and the financial accelerator	Mark Gertler, Simon Gilchrist and Fabio M Natalucci
138 September 2003	Public and private information in monetary policy models	Jeffery D Amato and Hyun Song Shin
137 September 2003	The Great Depression as a credit boom gone wrong	Barry Eichengreen and Kris Mitchener
136 September 2003	The price level, relative prices and economic stability: aspects of the interwar debate	David Laidler
135 September 2003	Currency crises and the informational role of interest rates	Nikola A Tarashev
134 September 2003	The cost of barriers to entry: evidence from the market for corporate euro bond underwriting	João A C Santos and Kostas Tsatsaronis
133 September 2003	How good is the BankScope database? A cross-validation exercise with correction factors for market concentration measures	Kaushik Bhattacharya
132 July 2003	Developing country economic structure and the pricing of syndicated credits	Yener Altunbaş and Blaise Gadanecz
131 March 2003	Optimal supervisory policies and depositor-preference laws	Henri Pagès and João A C Santos
130 February 2003	Living with flexible exchange rates: issues and recent experience in inflation targeting emerging market economies	Corrinne Ho and Robert N McCauley
129 February 2003	Are credit ratings procyclical?	Jeffery D Amato and Craig H Furfine
128 February 2003	Towards a macroprudential framework for financial supervision and regulation?	Claudio Borio

All volumes are available on request from the BIS and are also on our website www.bis.org.