A New Index Outperforms the Purchasing Managers' Index

Rolando F. Peláez
University of Houston-Downtown

Several recent papers have documented the effects on domestic and overseas financial markets of unexpected changes in the Purchasing Managers' Index (PMI). This paper introduces a new composite index of diffusion indices from the Institute for Supply Management. The proposed index dominates the PMI as an indicator of the growth rates of real GDP and of industrial production. Moreover, as a stand-alone indicator of business cycle turning points, the new index is superior in timeliness to the Conference Board's composite coincident index and equal to it in consistency.

Introduction

The manufacturing *Report on Business* of the Institute for Supply Management (ISM) is an important indicator of economic activity (Niemira and Zukowski, 1998). The Purchasing Managers' Index (PMI) is the centerpiece of the report. Market participants await its release with some anxiety due to its history of destabilizing bond and equity prices on the first business day of each month. Of 24 regularly scheduled announcements, Fleming and Remolona (1997) rank the ISM survey seventh in order of importance for trading activity in the Treasury bond market. Ederington and Lee (1996) also report a statistically significant effect of the ISM survey on interest rates in both the U.S. Treasury and Eurodollar markets. Other studies also document the informational content of the PMI, e.g., Klein and Moore (1988), Kauffman (1999), Harris (1991), Bretz (1990), and Dasgupta and Lahiri (1992).

Given the PMI's importance, it is worth considering if a better index could be constructed using only ISM diffusion indices that are released at the same time as the PMI. This paper develops such an index. The new index dominates the PMI as an indicator of the growth rates of real GDP and of industrial production.

The Purchasing Managers' Index

Since 1931 the National Association of Purchasing Managers (NAPM) has surveyed a sample of its members and produced diffusion indices of economic activity

0747–5535/02/1600 & 1300/0041/$2.50

for the manufacturing sector.[1] The survey covers a stratified sample of about 400 industrial companies; over time the response rate has averaged about two-thirds. Around the third week of each month, respondents compare activity in the current month to the previous month; however, responses are limited to *better, worse*, or *unchanged*. Adding the percentage reporting *better,* and one-half of the percentage reporting *unchanged,* yields diffusion indices bounded by 0 and 100; these are then seasonally adjusted with factors from the U.S. Department of Commerce.[2] Diffusion indices measure the breadth of change, but cannot measure its magnitude with precision.

In 1980 an economist with the U.S. Department of Commerce introduced the PMI as an equally weighted index of five diffusion indices: new orders, production, employment, supplier deliveries, and inventories. Two years later, NAPM and the Department of Commerce introduced a re-weighted PMI (Bretz, 1990). The weights were selected to maximize the correlation of the index with GDP growth, and the index was re-calculated from January 1948 (Torda, 1985). Unlike other indices that are re-weighted periodically to enhance their usefulness, the PMI's weights have remained fixed: new orders (0.3), production (0.25), employment (0.2), vendor performance (0.15), and inventories (0.1).

The release of the PMI precedes the releases of all major measures of economic activity. For example, the PMI precedes by three weeks the first release of the Conference Board's composite coincident index, which initially reflects only three of four components and is revised twice in subsequent months. In contrast, the PMI itself is not revised, but like other seasonally adjusted series, the seasonal factors undergo revision in January of each year. The PMI for a given quarter is available one month before the advance release of GDP and a full three months before the final revision.

Financial markets react to unanticipated changes in the PMI partly from expectations that the Federal Reserve may change its policy stance. According to the policy anticipations hypothesis, the Federal Reserve may change its target for the Fed funds rate in response to macroeconomic news. The hypothesis argues that as the Fed's objectives change, so does the market's reaction to the arrival of unanticipated information, see, e.g., Poole (1988) and Santomero (1991).

Alternative Models

This section compares the informational value, for real GDP growth, of the PMI and of four other models consisting of ISM indices,

$$G_t = \alpha_0 + \alpha_1 PMI_t + e_t, \tag{1}$$

---

[1] The National Association of Purchasing Managers was recently renamed the Institute for Supply Management.

[2] The price index is not seasonally adjusted.

$$G_t = \alpha_0 + \alpha_1 NOI_t + \alpha_2 PI_t + \alpha_3 EMI_t + \alpha_4 INV_t + \alpha_5 SD_t + v_t\,, \tag{2}$$

$$G_t = \alpha_0 + \alpha_1 PI_t + \alpha_2 EMI_t + \alpha_3 PRICE_t + \alpha_4 NOI_{t-3} + e_t, \tag{3}$$

$$G_t = \beta_0 + \beta_1 NOI_t + \beta_2 EMI_t + \beta_3 PRICE_t + \beta_4 INV_{t-2} + \beta_5 \Delta SD_{t-2} + e_t, \tag{4}$$

$$G_t = \gamma_0 + \gamma_1 PI_t + e_t, \tag{5}$$

Model (1) regresses the quarterly growth rate of GDP, (G_t), on the PMI. Model (2) includes the components of the PMI without *a priori* restrictions on the coefficients. Model (3) includes the contemporaneous production and employment indices, a three-quarter lag of the new orders index, and an ISM diffusion index of industrial materials prices which is not a component of the PMI. Model (4) includes the contemporaneous values of three ISM indices (new orders, employment, and prices); two other indices (inventories and the first difference of supplier deliveries) enter lagged two quarters. Finally, model (5) includes only the production index (PI) that, as shown earlier, enters the PMI with a weight of 0.25.

Table 1—OLS Estimation Results (1955Q1—2003Q4)

The regressand, G_t, is the annualized growth rate of real GDP. The ISM indices are quarterly averages of monthly observations (t-statistics in parentheses, and P-values in brackets)

	Models				
Regressors	M1	M2	M3	M4	M5
PMI_t	0.36 (12.8)				
PI_t		0.18 (1.5)	0.26 (4.6)		0.32(11.0)
NOI_t		0.16 (1.6)		0.20(4.2)	
EMI_t		0.10 (1.2)	0.20 (2.8)	0.27(4.5)	
INV_t		-0.05 (-0.9)			
SD_t		-0.05 (-1.7)			
$Price_t$			-0.06 (-3.4)	-0.05(-2.8)	
NOI_{t-3}			-0.10 (-3.2)		
INV_{t-2}				-0.11(-3.2)	
ΔSD_{t-2}				-0.10(-3.0)	
Const.	-15.6 (-10.4)	-15.7 (-9.7)	-11.8 (-6.6)	-12.7(-6.4)	-14.6(-9.0)
Adj. R^2	0.40	0.51	0.53	0.54	0.44
DW	1.9	2.2	2.35	2.3	2.3
AR 1-5[a]	0.50[0.77]	0.66[0.66]	2.1[0.07]	1.59[0.16]	0.9[0.5]
ARCH 1-4[b]	2.43[0.05]*	0.38[0.83]	0.57[0.68]	0.80[0.53]	0.8[0.6]
Hetero test[c]	1.15[0.32]	1.60[0.11]	1.06[0.39]	1.43[0.17]	0.02[1.0]
Reset test[d]	5.79[0.02]*	0.82[0.37]	3.79[0.05]	1.92[0.17]	1.9[0.2]
Normality[e]	10.2[0.006]**	13.4[0.001]**	20.8[0.000]**	12.4[0.002]*	13.0[0.002]
Instability[f]	1.44**	2.11*	0.62	1.53	1.21*
AIC[g]	2.13	1.98	1.90	1.89	1.88

[a]Breusch (1978) and Godfrey (1978) F-test for serial correlation of residuals
[b]Engle's (1982) ARCH test (F-stat.) for autoregressive conditional heteroscedasticity of residuals
[c]White's (1980) test (F-stat.) for heteroscedasticity of residuals
[d]Ramsey's (1969) RESET test (F-stat.) for specification error
[e]Doornik-Hansen (1994) test for normality of residuals (Chi^2 stat.)
[f]Hansen's (1992) parameter instability test
[g]Akaike (1981) information criterion

Table 1 shows estimation results and tests of model adequacy. During 1955Q1—2003Q4, model (1) is inferior to its rivals judging by adjusted R^2 and by the Akaike (1981) information criterion (AIC)—a procedure for selecting undominated parsimonious models. Because a low value of AIC is preferred, models (4) and (5) are preferred to the alternatives. Model (2), which includes the PMI's components without *a priori* restrictions, outperforms the PMI; yet due to extreme collinearity the coefficients are not significant.[3] The F-test does not reject the null hypothesis that in model (2) the new orders, employment, and inventories indices contribute nothing to the explanation of real GDP growth [$F(3,190) = 1.13$, P-value = 0.34] in the presence of the production and supplier deliveries indices.

The price diffusion index enters models (3) and (4) with a negative coefficient because it proxies for inflation. It is less clear why the coefficient of lagged new orders (NOI_{t-3}) should be negative; one possible explanation involves inventory corrections following large changes in new orders. The inventories index indicates the direction of change in manufacturing inventories; a rising index signaling rising inventories. Although the contemporaneous inventories index is redundant, it is worth considering if it contains useful information when lagged. For example, other things equal, the growth rate of GDP should vary inversely with the lagged level of inventories, as the burden of adjustments to inventories falls on future GDP growth. The estimation results for model (4) in Table 1 confirm this; the coefficient of the lagged inventories index (INV_{t-2}) is negative and statistically significant.

The supplier deliveries index (SD) is a measure of vendor performance; a higher value indicates that delivery is slower than in the previous period; conversely, the index decreases during economic slowdowns, indicating faster deliveries due to reduced backlogs. Because a diffusion index is a crude growth rate, the first difference of supplier deliveries proxies for acceleration. The lagged first difference of the supplier deliveries index enters model (4) with a significant and negative coefficient, i.e., after accounting for the effects of the other regressors, the growth rate of GDP in the current quarter tends to decrease if the supplier deliveries index accelerated two quarters earlier. Model (5) has a higher R-squared and a lower SIC than the equally parsimonious model (1).

Residual autocorrelation is evident only in model (3). This invites modeling and, in fact, including one lag of the dependent variable yields a better fit. Autoregressive conditional heteroscedasticity is present in the residuals from model (1). For the other models, tests of the residuals reveal the absence of serial correlation, heteroscedasticity, and autoregressive conditional heteroscedasticity (ARCH); however, consistent with the presence of outliers, the Doornik-Hansen (1994) chi-squared test rejects the null of normality.

---

[3] Some of the diffusion indices are more highly correlated with each other than with the growth rate of GDP.

Out-of-Sample Forecasts of Real GDP Growth Rates

Market participants are not interested in the theoretical underpinnings or elegance of the PMI, only the predictive ability of the index and its timeliness possess utility. Therefore, if an index equal in timeliness to the PMI but superior in predictive accuracy were available, it would obviously have greater utility.

In order to assess predictive accuracy, ex ante one-quarter-ahead forecasts are obtained with rolling regressions after reserving a subset of the sample for model initialization. Each model is first estimated sampling 1955Q1—1965Q4 in order to obtain out-of-sample forecasts for 1966Q1. Re-estimation then proceeds recursively; at each step the estimation sample includes one more observation, i.e., at the next step, the estimation sample is 1955Q1—1966Q1 and the forecast period is 1966Q2, and so on. This way, the out-of-sample forecasts do not use data or coefficients from the forecast period.

Root mean squared forecast error (RMSFE), mean absolute forecast error (MAFE), and Theil's (1961) inequality coefficient are commonly used measures of forecast accuracy. All three are zero if forecasts are perfect. Formally, let F1 - F5 denote the ex ante forecasts obtained with models (1) through (5). Table 2 shows that F3, F4, and F5 outperform F1 and F2 according to all three measures.

Table 2—Forecast Statistics for One-Step-Ahead Forecasts F1, F2, F3, F4, and F5 Obtained with Models 1, 2, 3, 4, and 5, 1966Q1 through 2003Q4

Statistic	F1	F2	F3	F4	F5
Number of Forecasts	152	152	152	152	152
MFE	0.21	0.50	0.22	0.27	0.42
P-value for Forecast Bias	0.37	0.02*	0.28	0.20	0.05*
MAFE	2.13	2.07	1.92	1.91	2.01
RMSFE	2.84	2.67	2.50	2.53	2.61
Theil's U	0.69	0.64	0.60	0.61	0.63
Jarque-Bera Normality Test	24.1	30.9	41.2	38.2	14.3
P-value for Error Normality	0.000**	0.002**	0.000**	0.000**	0.000**

MFE and MAFE are the mean forecast error and the mean absolute forecast error, respectively. RMSFE is the square root of the mean squared forecast error. Theil's U statistic, or inequality coefficient, assesses predictive accuracy relative to a naïve no-change model. It is unitary when the MSFE equals the mean square error of naïve no-change forecasts, and it is less than 1.0 if predictions are more accurate than naïve forecasts. P-values are the significance levels of test statistics, i.e., the probability of getting a number at least as large under the zero null. The Jarque-Bera (1980) statistic tests for normality of the forecast errors. * and ** denote significance at the 0.05 and 0.01 levels, respectively

It is important to determine if forecasts are biased, i.e., if the mean forecast error (MFE) differs significantly from zero. Regressing the forecast error on a constant, and letting the t-test establish if the constant differs significantly from zero satisfies this requirement. The test results (P-values for forecast bias) shown in Table 2 do not allow rejection of the null of unbiased forecasts for models 1, 3, and 4.

Although forecasts from model (4) exhibit the lowest MAFE, it is important to test formally whether the difference between F1 and F4 is statistically significant. Granger and Newbold (1977) show that if the forecast errors from one model are correlated with those from another, the usual variance ratio or F test is inappropriate for this task. They develop a test of one-step-ahead forecasts that is valid if the loss function is quadratic and if the forecast errors are zero-mean, Gaussian, and serially uncorrelated. More recently, Harvey, Leybourne, and Newbold (1997) introduce a variation of the Granger-Newbold test that, while being easier to implement, suffers the same limitations concerning error loss and forecast errors. Diebold and Mariano (1995) develop a better measure of forecast performance; one applicable to more general loss functions, and that remains valid even if forecast errors are non-Gaussian, non-zero mean, serially correlated, and contemporaneously correlated with those from another model. Let the differential loss in period t from using F1 versus F2 be,

$$D_t = g(e_{1t}) - g(e_{2t}) \,, \tag{5}$$

where $g(e_{1t})$ and $g(e_{2t})$ denote the losses in period t from the forecast errors of the respective models. The Diebold-Mariano procedure tests whether the mean differential loss (MDL) is zero,

$$\mathrm{MDL} = \frac{1}{T}\sum_{t=1}^{T}[g(e_{1t}) - g(e_{2t})] \tag{6}$$

Table 3—Testing for Forecast Accuracy with the Diebold-Mariano Test

Hypothesis	P-Val. (1966-2003)	P-Val. (1980-2003)
F1 = F2 vs. F1 better than F2	0.90	0.06*
F1 = F2 vs. F2 better than F1	0.10	0.94
F1 = F3 vs. F1 better than F3	0.98	0.72
F1 = F3 vs. F3 better than F1	0.02**	0.28
F1 = F4 vs. F1 better than F4	0.99	0.92
F1 = F4 vs. F4 better than F1	0.003***	0.08*
F1 = F5 vs. F1 better than F5	0.98	0.10
F1 = F5 vs. F5 better than F1	0.02**	0.90

* 0.1 level of significance, ** 0.05 level of significance, and *** 0.01 level of significance

Under the null hypothesis of equal forecast accuracy, MDL is zero. Table 3 shows the results of the Diebold-Mariano (DM) test of forecast accuracy. The difference between F4 and F1 is statistically significant during 1966Q1—2003Q4, as well as during the shorter period, 1980Q1—2003Q4. For example, the DM test rejects the null that F1 is as accurate as F4 (against the alternative that F4 is better than F1) at the 0.01 level for 1966-2003, and at the 0.1 level for 1980-2003. Model (3) dominates model (1) during 1966-2003, but is only equal to it during the more recent 1980-2003 period. Models (2) and (5) outperform model (1) during 1966-2003, but the reverse occurs during 1980-2003. In short, all of the models outperform the PMI-based model during the extended forecast period, but only model (4) does so during both forecast evaluation periods. Therefore, it is worth considering an index based on model (4). The next section explains the construction of the index and tests its usefulness for dating business cycle turning points.

The Proposed Index

This section develops two versions of the new index; both are based on model (4). One has variable weights; the second has fixed weights and is the preferred version. Recursive coefficients were obtained with rolling regressions after first initializing the model over the sample 1955Q1—1965Q3. This procedure avoids ex post bias, i.e., the coefficients at each point t reflect only information available up to that point. The sum of the recursive regression coefficients from model (4) is,

$$S_t = \beta_{1t} + \beta_{2t} + \beta_{3t} + \beta_{4t} + \beta_{5t}\,. \tag{7}$$

The components' weights (WC_{it}) now evolve over time with the evolving recursive coefficients,

$$WC_{it} = \frac{\beta_{it}}{S_t}\,. \tag{8}$$

Figures 1 and 2 show the estimated weights. The weights are unstable during the first few years because of the small sample size, outliers, and high multicollinearity, all of which make it impossible to estimate the coefficients with much precision. Figure 1 shows that the weights of the inventories and supplier delivery indices dropped rapidly at first and finally stabilized in the mid 1980s; however, the weight of the price index shows a more gradual decline. Figure 2, shows a similar pattern, although the weights of the employment and new orders indices continued to evolve, albeit at a slower pace. In the new index, NI, the components' weights are lagged one quarter, so that at each point they reflect only past information,

$$\begin{aligned} NI_t = {} & WNOI_{t-1} \times NOI_t + WEMI_{t-1} \times EMI_t + WPRICE_{t-1} \\ & \times PRICE_t + WINV_{t-1} \times INV_{t-2} + WSD_{t-1} \times \Delta SD_{t-2}\,. \end{aligned} \tag{9}$$

The construction of the variable-weighted index is cumbersome. Beside the advantage of computational simplicity, an index with constant weights may be more desirable, given that in small samples and in the presence of multicollinearity OLS estimators are sensitive to small changes in the data, resulting in unstable weights. The simplest way to obtain a set of constant weights is by averaging the variable weights over some suitable period. Because, as shown earlier in Figures 1 and 2, the individual weights became more stable after 1984, the means of the variable weights during 1985Q1—2003Q4 were selected to form a constant-weighted index (CWI) that can be calculated back to 1948Q1,

$$CWI_t = 1.08 \times NOI_t + 1.14 \times EMI_t - 0.2 \times PRICE_t - 0.51 \times INV_{t-2} - 0.5 \times \Delta SD_{t-2}\,. \tag{10}$$

Figure 1—Variable Weights, 1966Q1 to 2003Q4

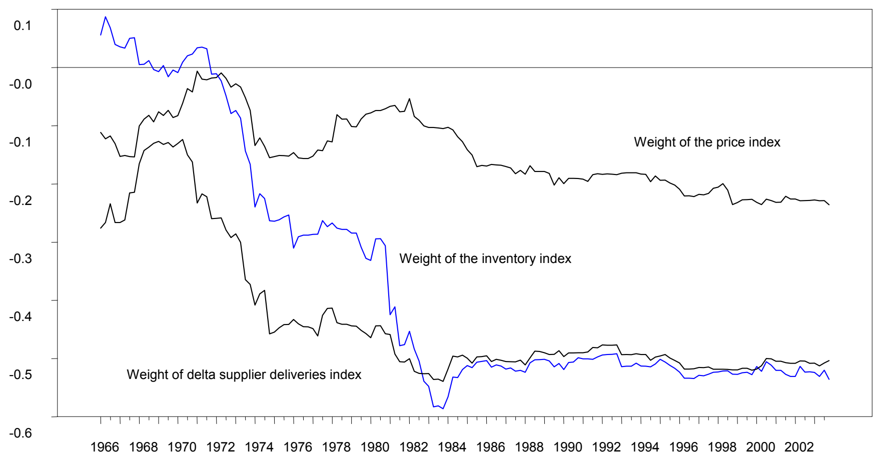

Figure 2—Variable Weights, 1966Q1 to 2003Q4

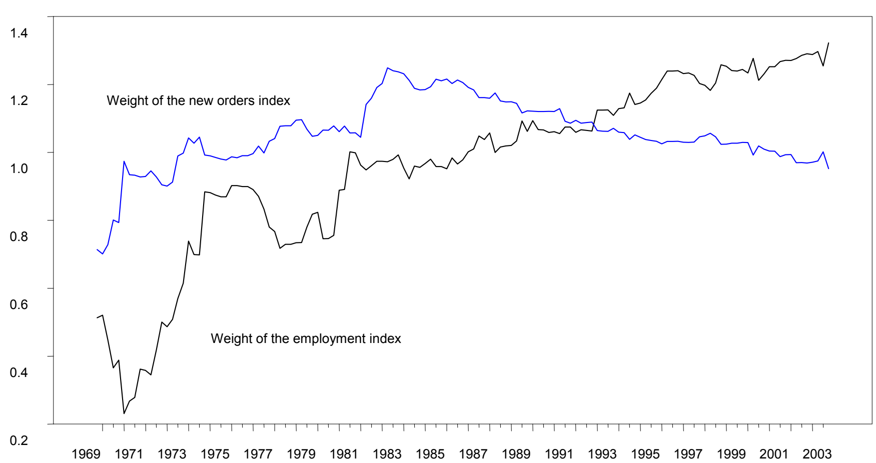

Figure 3—PMI and CWI 1949Q1 to 2003Q4

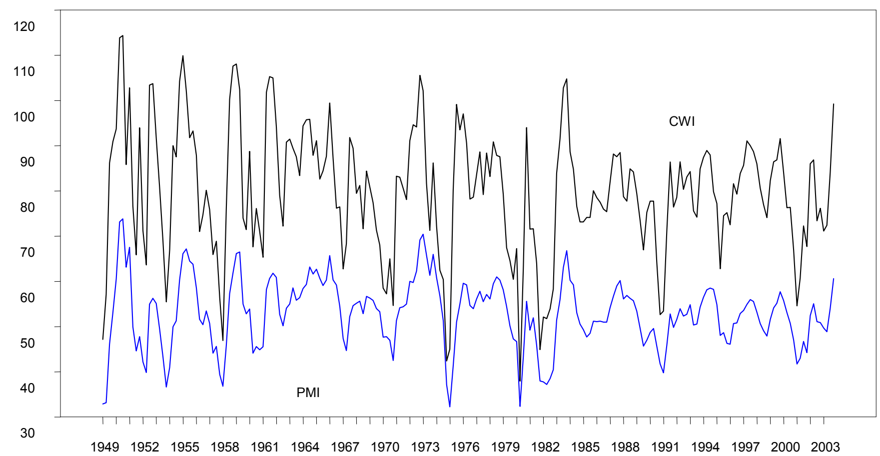

Figure 4—Estimated Probability of Recession 1960Q1 to 2003Q4

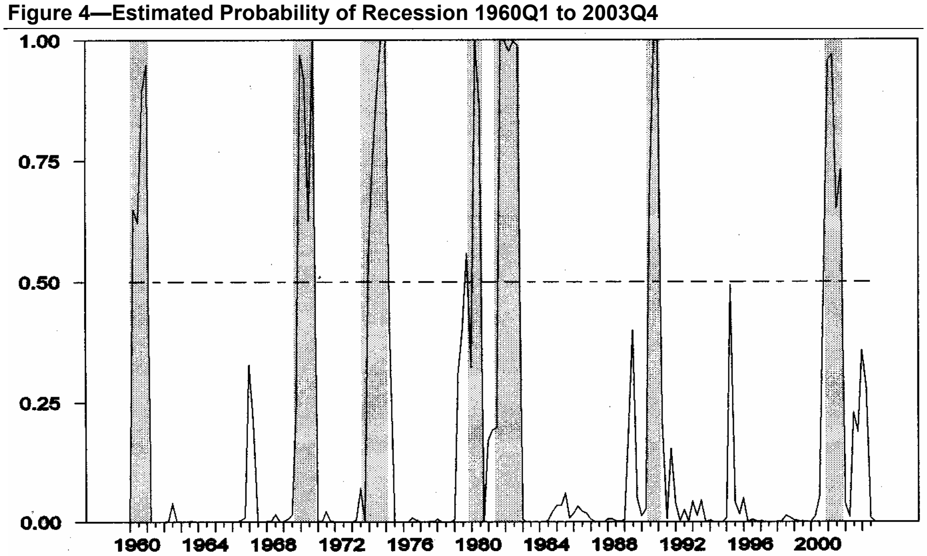

Figure 3 plots the PMI and the CWI; the latter exhibits greater variability, and shows a slight leading tendency relative to the PMI. The CWI is not bounded by 100 because not all of its components are similarly bounded. The regression results in Table 4 show that, in terms of R^2, the CWI does better than the PMI in accounting for the quarterly variation of GDP over the extended sample period 1948Q1—2003Q4, as well as over the shorter sample, 1948Q1—1984Q4. Likewise, Table 5 shows that the CWI is superior to the PMI in explaining the quarterly variation of the Federal Reserve's industrial production index.

Table 4—Regressions of the Growth Rate of GDP on the CWI and the PMI (t-statistics in parentheses)

	Adj. R^2	DW	Sample Period	T
(6.1) $G_t = -13.4 + 0.21CWI_t$	0.53	2.3	1949Q1—2003Q4	220
(-12.3) (15.7)				
(6.2) $G_t = -15.2 + 0.35PMI_t$	0.40	1.9	1949Q1—2003Q4	220
(-10.4) (12.9)				
6.3) $G_t = -14.2 + 0.22CWI_t$	0.55	2.3	1949Q1—1984Q4	144
(-10.4) (13.2)				
(6.4) $G_t = -16.1 + 0.37PMI_t$	0.44	1.9	1949Q1—1984Q4	144
(-10.1) (12.2)				

Table 5—Regressions of the Growth Rate of the Industrial Production Index on the CWI and the PMI (t-statistics in parentheses)

	Adj. R^2	DW	Sample Period	T
(7.1) $G_t = -35.7 + 0.49CWI_t$	0.70	1.6	1949Q1—2003Q4	220
(-20.2) (22.4)				
(7.2) $G_t = -42.2 + 0.86PMI_t$	0.62	1.4	1949Q1—2003Q4	220
(-17.5) (19.1)				
(7.3) $G_t = -37.5 + 0.51CWI_t$	0.71	1.6	1949Q1—1984Q4	144
(-16.6) (18.6)				
(7.4) $G_t = -44.2 + 0.89PMI_t$	0.63	1.5	1949Q1—1984Q4	144
(-14.3) (15.7)				

Dating Business Cycle Turning Points with the CWI

Dating business-cycle turning points, i.e., determining ex post when the economy shifted from one stage of the cycle to another is difficult, albeit less so than predicting ex ante when the shift will occur. In dating turning points, the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER) studies the behavior of several monthly indicators. The two most important are economy-wide employment and personal income less transfer payments in real terms. Two other indicators focus on manufacturing, i.e., industrial production and manufacturing and trade sales in real terms. Those four indicators compose the Conference Board's composite coincident index (CCI). Other indicators also play a role, "but there is no fixed rule about which other measures contribute information to the process" (Hall et al., 2002). In short, the committee eschews formulaic approaches, favoring judgment instead.

In order to assess the signaling value of the CWI, consider the following models that convert observations on rival indices (PMI, and the growth rate of the CCI) into conditional probabilities of recession,

$$Z_t = \alpha_0 + \alpha_1 CWI_t + v_t \tag{11}$$

$$Z_t = \alpha_0 + \alpha_1 GCCI_t + v_t \tag{12}$$

$$Z_t = \alpha_0 + \alpha_1 CWI_t + \alpha_2 GCCI_t + v_t \tag{13}$$

$$Z_t = \alpha_0 + \alpha_1 PMI_t + \alpha_2 GCCI_t + v_t \ . \tag{14}$$

Z_t is unitary during recessions, otherwise zero. The logit model transforms the binary dependent variable into probabilities bounded by zero and one. If the estimated probability exceeds 0.5 the economy is presumed to be contracting; otherwise it is expanding. The models signal a turning point when the estimated probability crosses 0.5 in either direction. During the 176 quarters from 1960Q1 through 2003Q4, the economy experienced seven recessions and several growth cycles. It would be interesting to see how well these simple models succeed in identifying the state of the economy and signaling that a turning point has occurred.

Equations (11) and (12) compare the signaling value of the CWI and of the growth rate of the coincident index (GCCI), whereas equations (13) and (14) compare the informational value of the CWI and the PMI in the presence of the GCCI. Table 6 shows the estimation results. Based on pseudo-R^2, equations (11) and (12) are virtually indistinguishable. However, equation (12) yields 167 correct cases as compared to 163 for equation (11), but this small advantage in dating ability is illusory. The problem is that the ISM's diffusion indices are available in real time, yet the revised data on the CCI are available with a lag of roughly three months; therefore, a decision-maker does not have the luxury of estimating both models in real time. Table (6) also shows that the coefficient of the CWI in equation (13) is statistically significant in the presence of GCCI, but the coefficient of the PMI is not significant in equation (14). Equation (13) correctly signals the state of the economy in 171 of 176 quarters, but the price of this impressive result is a loss in timeliness relative to equation (11).

Figure 4 shows business cycle reference points (shaded areas bounded by peak and trough dates established by the NBER's Business Cycle Dating Committee) and the probabilities of recession estimated with equation (13). A peak marks the end of an expansion and the beginning of a recession. Between trough and peak, the economy expands. Four features stand out. First, there are no false negatives. Second, the model was two quarters late in recognizing the recession of 1973, in which the oil embargo and price shock played a major role. Third, the model identifies periods of economic slowdowns that did not ripen into recessions, e.g., 1966, 1985-86, 1993, 1995, and 2002. This is worth noting, as slowdowns during which growth rates remain positive may cause as much suffering as mild recessions. Finally, the model clearly signals the 2001 recession that humbled most of the professional forecasters and the Federal Reserve alike.

Table 6—Results of Estimating Equations 11-14 by Logit, 1960Q1—2003Q4, (t-statistics in parentheses)

Regressors	Equation (11)	Equation (12)	Equation (13)	Equation (14)
CWI_t	-0.27(-5.5)		-0.19(-3.2)	
$GCCI_t$		-1.16(-5.4)	-0.87(-3.5)	-1.12(-4.6)
PMI_t				-0.02(-0.3)
Constant	17.9(3.4)	-0.4(-1.2)	13.1(3.1)	0.63(0.2)
Pseudo-R^2	0.57	0.58	0.67	0.58
Cases Correct	163	167	171	168
Log Likelihood	-33.2	-32.3	-24.7	-32.3

In January 2001, when most professional forecasters predicted strong growth for 2001, the CWI-based model in equation (11) signaled that the economy was dangerously close to recession. For example, on January 3, 2001, the ISM released the *Report on Business* containing diffusion indices for the previous month; at that point, estimation of equation (11) would have yielded a probability of recession of 0.46 for the fourth quarter of 2000, which suggests that the recession may have started earlier than March 2001. Upon the release on April 2, 2001, of the *Report on Business* the estimated probability of recession for 2001Q1 obtained with equation (11) was 0.96. Underscoring the importance of this result, it would take the Federal Reserve another five to seven months to wake up to the recession. As the economy contracted, monetary policy was driven by the assumption that the economy kept growing, albeit very slowly. The minutes of the FOMC meetings of March 20, May 15, June 26-27, and August 21 do not mention a recession or downturn. Instead, the staff forecasts prepared for those meetings posited that the economy was growing very slowly. Consistent with this view, the *Monetary Policy Report to the Congress* of July 18 placed the central tendency of the Federal Reserve's forecasts for the increase in real GDP over the four quarters of 2001 at a range of 1¼ percent to 2 percent. In fact, real GDP growth in 2001 was only 0.05 percent.

The first mention of a possible mild downturn finally appeared in the minutes of the FOMC meeting of October 2, albeit incorrectly attributing the downturn to the September 11 event. "The information reviewed at this meeting suggested that the attacks of September 11 might well have induced a mild downturn in economic activity after several months of little movement in the level of economic activity."

The recession ended in November, and only then did the staff forecast for the FOMC meeting of that month finally predict a recession extending into 2002. "The staff forecast prepared for this meeting emphasized the continuing wide range of uncertainty surrounding the outlook in the wake of the September 11 attacks. The mild downturn in economic activity in the third quarter was seen as likely to deepen over the remaining of the year and to continue for a time next year" *(*minutes*,* meeting of Nov. 6, 2001).

Professional forecasters did not fare better. In the fourth quarter of 2000, the forecasters in the Survey of Professional Forecasters (SFP) conducted by the Federal Reserve Bank of Philadelphia predicted strong growth for 2001. Having failed to predict the recession, once it began they failed to recognize it and continued to predict strong economic growth even as GDP growth declined. As Stock and Watson (2003) point out, "The SPF one-quarter-ahead forecast of 2001Q1 growth was 3.3 percent, whereas GDP actually fell by 0.6 percent; the one-quarter-ahead forecast of 2001Q2 growth was 2.2 percent, but GDP fell by 1.6 percent; and the one-quarter-ahead forecast of 2001Q3 growth was 2.0 percent, while GDP fell by 0.3 percent." Finally, in the fourth quarter of 2001, the SPF forecasters predicted that growth would be negative in that quarter; instead, GDP growth was a strong 2.7 percent (Stock and Watson, 2003).

Conclusions

The proposed CWI index is better than the PMI in accounting for the growth rates of GDP and of industrial production over long time spans. As a stand-alone indicator of business cycle turning points, the CWI rivals the CCI which, developed over many years at the NBER, is generally considered the best index of economic activity. The CWI has an important three-month advantage in timeliness over the final release of the CCI, however. As an illustration of the value of this lead time, consider the hypothetical case of a trader, who upon estimating equation (11) on January 3, 2001 and finding that the probability of recession in 2000Q4 was 0.46, proceeded to sell equities and buy bonds. Eight months later, in August 2001, with the Fed still undecided, the S&P 500 index had dropped by about 8 percent and the yield on the 10-year Treasury note had fallen by about 30 basis points, and this was just the beginning.

References

1. Akaike, H., "Likelihood of a Model and Information Criteria," *Journal of Econometrics*, 16 (May 1981), pp. 3-14.

2. Bretz, R.J., "Behind the Economic Indicators of the NAPM Report on Business," *Business Economics* (July 1990), pp. 42-50.

3. Breusch, T.S., "Testing for Autocorrelation in Dynamic Linear Models," *Australian Economic Papers*, 17 (December 1978), pp. 334-355.

4. Dasgupta, S., and K. Lahiri, "A Comparative Study of Alternative Methods of Quantifying Qualitative Survey Responses Using NAPM Data," *Journal of Business and Economic Statistics* (October 1992), pp. 391-400.

5. Diebold, F.X., and R.S. Mariano, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13 (July 1995), pp. 253-263.

6. Doornik, J.A., and H. Hansen, "A Practical Test for Univariate and Multivariate Normality," discussion paper (1994), Nuffield College.

7. Ederington, L.H., and J.H. Lee, "The Creation and Resolution of Market Uncertainty: The Impact of Information Releases on Implied Volatility," *The Journal of Financial and Quantitative Analysis* (October 1996), pp. 513-539.

8. Engle, R.F., "Autoregressive Conditional Heteroscedasticity, With Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50 (July 1982), pp. 987-1007.

9. Fleming, M.J., and E.M. Remolona, "What Moves the Bond Market?" *Economic Policy Review, Federal Reserve Bank of New York* (December 1997), pp. 31-50.

10. Godfrey, L.G., "Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables," *Econometrica*, 46 (November 1978), pp. 1303-1313.

11. Granger, C.W.J., and P. Newbold, *Forecasting Economic Time Series* (New York: Academic Press, 1977).

12. Hall, R., M. Feldstein, J. Frankel, R. Gordon, N. Mankiw, and V. Zarnowitz, "The NBER's Business-Cycle Dating Procedure," National Bureau of Economic Research (December 7, 2002).

13. Hansen, B., "Testing for Parameter Instability in Linear Models," *Journal of Policy Modeling,* 14 (August 1992), pp. 517-533.

14. Harris, E.S., "Tracking the Economy with the Purchasing Managers' Index," *Quarterly Review*, Federal Reserve Bank of New York (Fall 1991), pp. 61-69.

15. Harvey, D.I., S.J. Leybourne, and P. Newbold, "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting,* 16 (June 1997), pp. 281-291.

16. Kauffman, R.G., "Indicator Qualities of the NAPM Report on Business," *The Journal of Supply Chain Management* (Spring 1999), pp. 29-37.

17. Klein, P.A., and G.H. Moore, "N.A.P.M. Business Survey Data: Their Value as Leading Indicators," *Journal of Purchasing and Materials Management* (Winter 1988), pp. 32-40.

18. Niemira, M.P., and G.F. Zukowski, *Trading the Fundamentals* (New York: Academic Press, 1998).

19. Poole, W., "Monetary Policy Lessons of Recent Inflation and Disinflation," *Journal of Economic Perspectives,* 2 (Summer 1988), pp. 73-100.

20. Ramsey, J.B., "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," *Journal of the Royal Statistical Society,* 31(1969), pp. 350-371.

21. Santomero, Anthony M., "Money Supply Announcements: A Retrospective," *Journal of Economics and Business,* 43 (February 1991), pp. 1-23.

22. Stock, J.H., and M.W. Watson, "Evidence on Structural Instability in Macroeconomic Time-Series," *Journal of Business and Economic Statistics,* 14 (January 1996), pp. 11-30.

23. Torda, T.S., "Purchasing Management Index Provides Early Clue on Turning Points," *Business America*, U.S. Department of Commerce (June 24, 1985), pp 11-13.

24. Theil, H., *Economic Forecasts and Policy* (Amsterdam: North Holland, 1961).

25. White, H., "A Heteroscedastic-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica,* 48 (May 1980), pp. 817-838.

Data Appendix

All diffusion indices are from the manufacturing *Report on Business* of the Institute for Supply Management. The monthly indices were averaged in order to obtain quarterly observations. GDP, seasonally adjusted, in chained (1996) dollars is from the U.S. Department of Commerce, Bureau of Economic Analysis. The industrial production index, seasonally adjusted, is from the Federal Reserve.