Constructing Coincident and Leading Indices of Economic Activity for the Brazilian Economy

João Victor Issler   Hilton Notini   Claudia Rodrigues

EPGE-FGV

May, 2010

Introduction (General)

- In the past, business-cycle research focused on econometric models trying to capture the main features of either GDP or of the four coincident variables that the NBER is said to follow (employment, income, industrial production, and sales); see Stock and Watson (1988a, b, 1989, 1991, 1993a), Hamilton (1989), Kim and Nelson (1998), Chauvet (1998), Harding and Pagan (2003), Hamilton (2003), and Chauvet and Piger (2008), *inter-alia*.
- Arguably, these models missed a key variable that should be included in them – the NBER decisions on U.S. turning points. Although this information is usually available with a considerable lag, there is no reason not to include it *ex-post* on econometric models building coincident indices of economic activity.
- This point was forcefully made in Issler and Vahid (2006).

Introduction (General)

- There has been a recent trend to incorporate NBER dating-committee decisions into different econometric models of business cycles.
- Some of these contributions were independent. Still, they all recognize that one should not discard the informational content of these decisions in constructing econometric models.
- Birchenhall et al. (1999), Dueker (2005), Issler and Vahid (2006), and Chauvet and Hamilton (2006), *inter-alia*.

Introduction (Brazil)

- Empirical business-cycle research in Brazil has shown the usefulness of heuristic methods such as the one proposed by *The Conference Board* vis-a-vis econometric-based methods.
- Duarte, Issler and Spacov (2004), Issler, Notini and Rodrigues (2008), and Hollauer, Issler, and Notini (2009).

This Paper

1. Proposes a new coincident index for Brazil based on the method in Issler and Vahid (2006). Compares it with the basic TCB index.
2. It is shown to outperform the basic TCB index in dating recessions.
3. Ingredients of Issler and Vahid:
    1. Use NBER decisions and *only current* coincident series to *uncover* weights given by the NBER on coincident series in dating recessions.
    2. Instrumental-variable Probit Regression to estimate weights of cyclical components of the four coincident variables (employment, income, industrial production, and sales).
    3. Only cyclical components of coincident variables are allowed in the Probit regression (filters out noise).
4. The model is *structural*, but not necessarily the best forecasting model for NBER decisions or for the probability of a recession. Reason: serial correlation in Probit regression error.

The Method: Canonical Correlations

- $x_t = (I_t, Y_t, E_t, S_t)'$ is the set of coincident variables (income, output, employment and sales).
- z_t is the vector of m $(m \geq 4)$ "predictors" of x_t: lags of x_t as well as lags of the leading variables.
- Four independent linear combinations $A(x_t) = (\alpha_1' x_t, \alpha_2' x_t, \alpha_3' x_t, \alpha_4' x_t)$, with the property that $\alpha_1' x_t$ is the linear combination of x_t that is most (linearly) predictable from z_t, $\alpha_2' x_t$ is the second most predictable linear combination of x_t from z_t after controlling for $\alpha_1' x_t$, and so on.
- These linear combinations are orthogonal to each other and to all the corresponding linear combinations of the z_ts, $\Gamma(z_t) = (\gamma_1' z_t, \gamma_2' z_t, \gamma_3' z_t, \gamma_4' z_t)$.

The Method: Canonical Correlations

- The regression R^2s between $\alpha_i' x_t$ and $\gamma_i' z_t$ for $i = 1, 2, 3, 4$, which we denote by $\left(\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_4^2\right)$, are the squared canonical correlations between x_t and z_t. We call $(\alpha_1' x_t, \alpha_2' x_t, \alpha_3' x_t, \alpha_4' x_t)' = (c_{1,t}, c_{2,t}, c_{3,t}, c_{4,t})'$ the "basis cycles" in x_t.
- We are interested only on the significant basis cycles (linear combinations that are predictable from past information). They are associated with significant canonical correlations. We throw out the insignificant ones.
- Significant=Signal.
- Insignificant=Noise.

The Method: Identifying Assumption

- y_t^* – the unobserved state of the economy
- $\{c_{1t}, \cdots, c_{kt}\}$ – the significant basis cycles of the coincident series at time t.
- We assume that:

$$\mathbb{E}(y_t^* - \beta_0 - \beta_1 c_{1t} - \cdots - \beta_k c_{kt} \mid I_{t-1}) = 0. \quad (1)$$

- This implies that there must be a linear combination of y_t^* and $\{c_{1t}, \cdots, c_{kt}\}$ that is unpredictable from the information before time t.
- In the common-feature literature: there is a common cycle between the unobserved state of the economy y_t^* and the significant basis cycles $\{c_{1t}, \cdots, c_{kt}\}$.

The Method: Instrumental-Variable Probit Regression

$$\text{NBER}_t = \begin{cases} 1 & \text{if } \mathbb{E}(y_t^* \mid I_{t+h}) < 0 \\ 0 & \text{otherwise.} \end{cases},$$

since the NBER decides on a recession or not using future information. Using equation (1), we obtain:

$$\begin{aligned} \mathbb{E}\left(y_t^* \mid I_{t-1}\right) &= \beta_0 + \beta_1 \mathbb{E}\left(c_{1t} \mid I_{t-1}\right) + \cdots + \beta_k \mathbb{E}\left(c_{kt} \mid I_{t-1}\right) \\ &= \beta_0 + \beta_1 c_{1t} + \cdots + \beta_k c_{kt} + \omega_t, \text{ where } \mathbb{E}\left(\omega_t \mid I_{t-1}\right) = 0, \end{aligned}$$

and obviously ω_t is correlated with c_{it}, $i = 1, \cdots, k$.

The Method: Instrumental-Variable Probit Regression

We can always write

$$\mathbb{E}\left(y_t^* \mid I_{t+h}\right) = \mathbb{E}\left(y_t^* \mid I_{t-1}\right) + \xi_t + \xi_{t+1} \cdots + \xi_{t+h},$$

where ξ_{t+i} is the "surprise" associated with new information arriving in period $t+i$.
Then,

$$\begin{aligned} \mathbb{E}\left(y_t^* \mid I_{t+h}\right) &= \beta_0 + \beta_1 c_{1t} + \cdots + \beta_k c_{kt} + u_t, \\ u_t &= \omega_t + \xi_t + \xi_{t+1} \cdots + \xi_{t+h}. \end{aligned}$$

$\mathbb{E}\left(u_t \mid I_{t-1}\right) = 0$, has a "forward" $MA\left(h\right)$ structure, and is correlated with c_{it}.

Complete Statistical Model

$$
\begin{aligned}
\mathrm{NBER}_t &= \begin{cases} 1 & \text{if } E(y_t^* \mid I_{t+h}) < 0 \\ 0 & \text{otherwise.} \end{cases} \\
E(y_t^* \mid I_{t+h}) &= \psi_0 + \psi' \underset{4\times 1}{x_t} + u_t \qquad (2) \\
x_t &= \underset{4\times m}{\Pi} \underset{m\times 1}{z_t} + \varepsilon_t,
\end{aligned}
$$

where u_t may be correlated with ε_t, u_t and ε_t are jointly normal, and Π has rank k.

The Coincident Series

Y_t,: industrial production. (IBGE)
available from 1980:1 to 2007:11

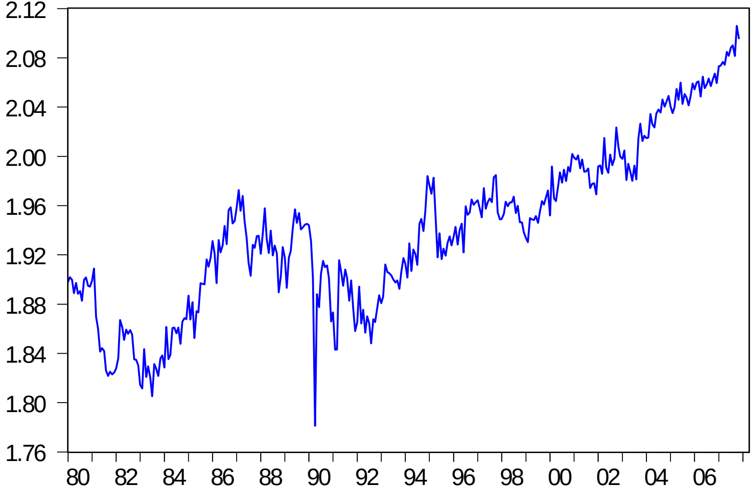
S_t : Brazilian production of corrugated paper: proxy for sales (ABPO)
available from 1980:1 to 2007:11

N_t,: number of persons (10 years or older) that have a job. (IBGE)
available from 2002:3 on (back-casted)

I_t : income proxied by the labor income series, .(IBGE)
available from 2002:2 on (back-casted)

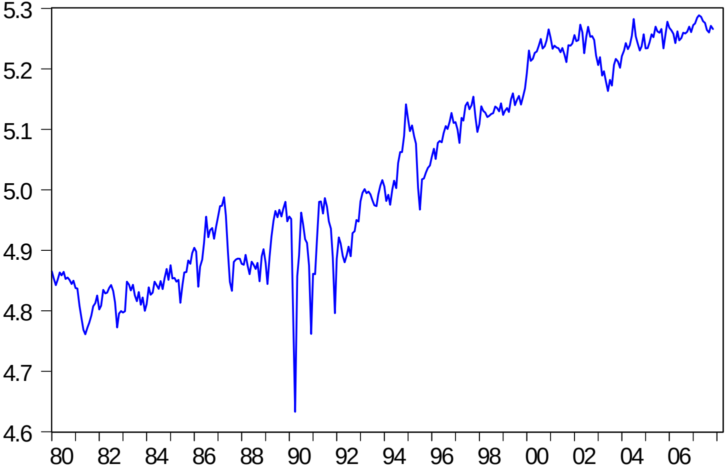
Coincident Series Graph: Industrial Production

Industrial Production - In log and Seasonally Adjusted



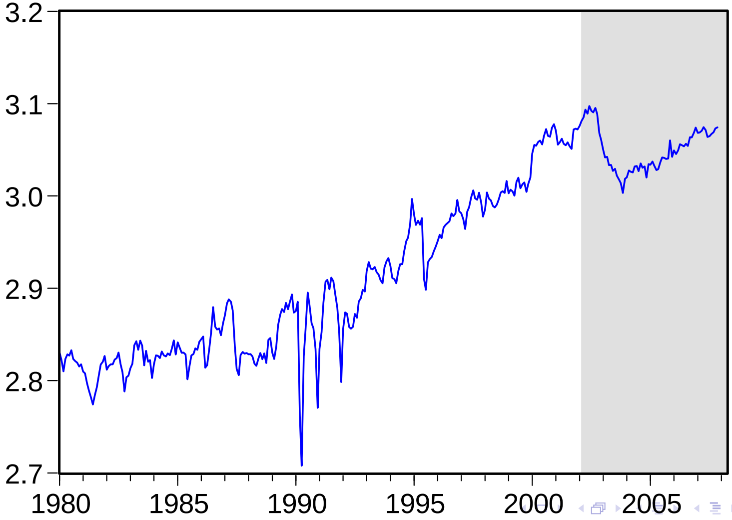
Coincident Series Graph: Sales

Sales - In log and Seasonally Adjusted



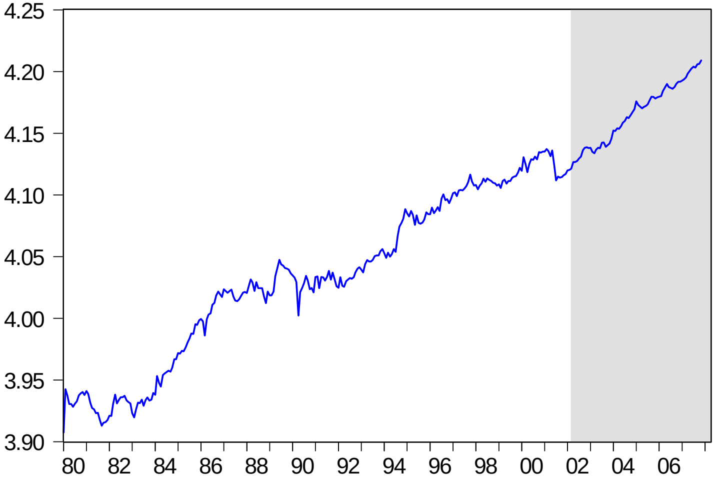
Coincident Series Graph: Income

Income - In log and Seasonally Adjusted; actual is shaded



Coincident Series Graph: Employment

Employment - In log and Seasonally Adjusted; actual is shaded



TCB Coincident Index

The TCB coincident index is an equally-weighted linear combination of four coincident series (income (I_t), output (Y_t), employment (N_t), and sales (S_t)) once we control for the fact that the growth rate of these series have different variances:

$$\Delta \ln (CI_t) = \frac{1}{4}\left[\frac{\Delta \ln (I_t)}{\sigma_{\Delta \ln(I)}} + \frac{\Delta \ln (Y_t)}{\sigma_{\Delta \ln(Y)}} + \frac{\Delta \ln (N_t)}{\sigma_{\Delta \ln(N)}} + \frac{\Delta \ln (S_t)}{\sigma_{\Delta \ln(S)}}\right], \qquad (3)$$

where $\sigma_{\Delta \ln(I)}$, $\sigma_{\Delta \ln(Y)}$, $\sigma_{\Delta \ln(N)}$, and $\sigma_{\Delta \ln(S)}$ are respectively the standard deviations of income, output, employment, and sales growth. It is straightforward to construct the level series $\ln (CI_t)$ or CI_t once we posses $\Delta \ln (CI_t)$. Is also straightforward to construct (3) with recursive weights.

Canonical correlation analysis found 4 significant basis cycles.

IV Probit regression results:

Two Stage Conditional Maximum Likelihood Estimates

Regressor	Est. Coeff.	Robust P-Value	Normalized Coeff.
$\Delta \ln Y_t$	-7.2230	0.1634	0.3669
$\Delta \ln S_t$	-0.9008	0.8960	0.0458
$\Delta \ln E_t$	-9.5572	0.0068	0.4855
$\Delta \ln I_t$	-2.0035	0.8551	0.1018
Constant	–	–	–

Standardized IVCI and Standardized TCB Indices

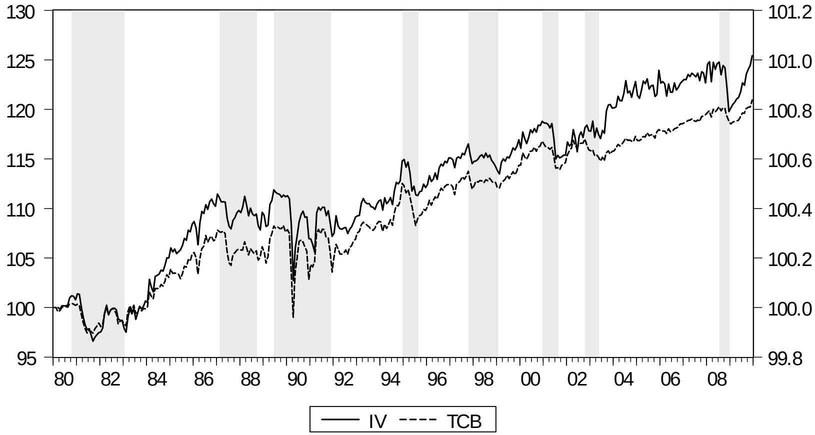
- Standardized IVCI:

$$IVCI_t = 0.10 \times \ln I_t + 0.37 \times \ln Y_t + 0.49 \times \ln N_t + 0.05 \times \ln S_t . \quad (4)$$

- Standardized TCB:

$$TCB_t = 0.26 \times \ln I_t + 0.24 \times \ln Y_t + 0.35 \times \ln N_t + 0.15 \times \ln S_t . \quad (5)$$

The Coincident Indices



IV and TCB Indices (Codace's Recessions Shown)

Measuring the Accuracy in which a Coincident Index Predicts the "State of the Economy"

How to evaluate the a potential coincident index:
Quadratic Probability Score, labelled as QPS: originally proposed by Diebold and Rudebusch (1999).

$$QPS = \frac{\sum_{t=1}^{T} (P_t - R_t)^2}{T},$$

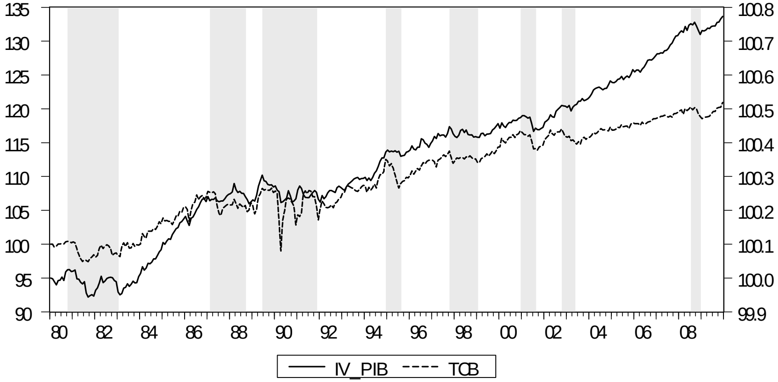
P_t is equal to 1 when a recession is dated by the dating committee, and 0 otherwise;
R_t is equal to 1 when a recession is dated by the coincident index, and 0 otherwise (Bry-Boschan);
T is the total number of sample observations;
QPS ranges between zero and one. It is an estimator of the frequency of erroneous dating by the coincident index vis-a-vis the decisions of the dating committee.

QPS Results

QPS Statistics vis-a-vis CODACE's Dating of Recessions			
Coinc Index	Whole Sample	Recession Only	Expansion Only
IVCI	0.1056	0.0873	0.1154
TCB	0.1806	0.4603	0.0299

QPS Statistics vis-a-vis GDP's Dating of Recessions			
Coinc Index	Whole Sample	Recession Only	Expansion Only
IVCI	0.0694	0.0246	0.0928
TCB	0.1889	0.4754	0.0422

GDP replaces Income (only measures labor income)



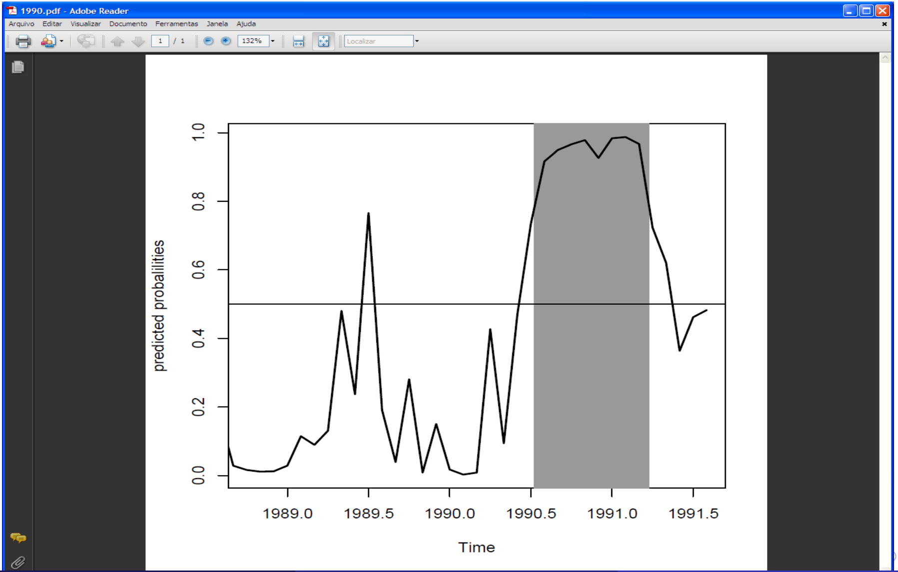
IVCI vs TCB (Codace's Recessions shown)

QPS Results (Robustness Analysis)

QPS Statistics vis-a-vis CODACE's Dating of Recessions			
Coinc Index	Whole Sample	Recession Only	Expansion Only
IVCI	0.1500	0.3254	0.0556
TCB	0.1611	0.4206	0.0214

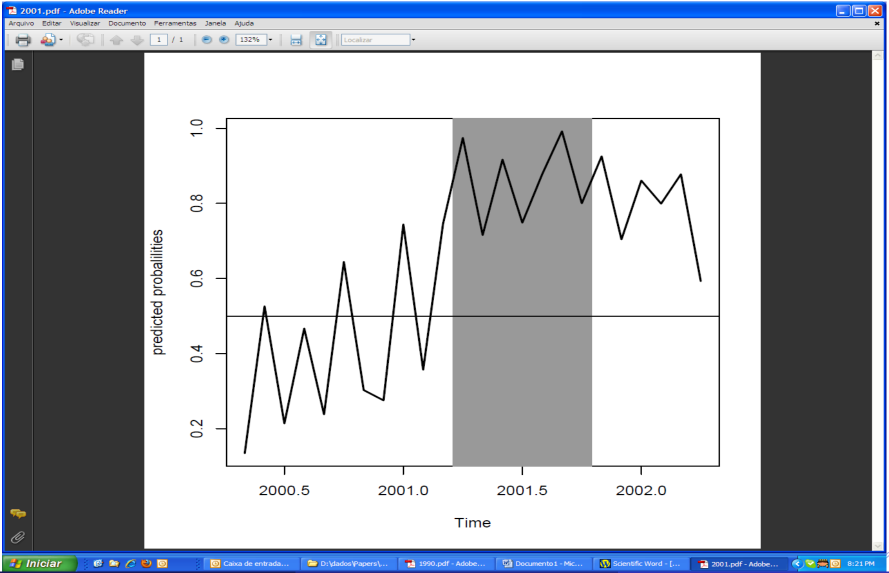
QPS Statistics vis-a-vis GDP's Dating of Recessions			
Coinc Index	Whole Sample	Recession Only	Expansion Only
IVCI	0.1028	0.2541	0.0253
TCB	0.1583	0.4180	0.0253

- 1990 Recession



Real-Time Analysis – USA

- 2001 Recession



Real-Time Analysis – USA

- 2007 Recession