Constructing Coincident Economic Indicators for Emerging Economies

Cem Çakmaklı Sumru Altuğ

Conference on
'POLICY ANALYSIS IN THE POST GREAT RECESSION ERA'

Istanbul, 16-17 October 2014

Summary of the Paper

- We construct a (statistical) monthly coincident indicator for measuring economic conditions for Emerging Market (EM) economy (Turkey),
- We take the uncertainty due to unobserved components and missing observations ,and parameter uncertainty into account by conducting Bayesian inference,
- We estimate the density of the indicator using Bayesian semiparametric inference, rather than only point prediction.
- Structural changes, time-varying volatility and extreme observations can be handled within the framework

Emerging Markets

- The recovery after the Great Recession was/is sluggish for developed countries,
- This was/is not the case for the Emerging Markets, i.e. recovery of the world economy was/is driven by the EM economies,
- Increasing attention on Emerging Markets, but lack of accurate (statistical) measures of timely economic conditions;
- Emerging market characteristics
  - Data is available as original observations
  - It is more like a patchwork rather than combination of mixed frequency data, a characteristic of typical EM data

Earlier contributions

- Descriptive approach: The coincident economic indicator of Department of Commerce (of The Conference Board (TCB) after 1995) using *ad hoc* combination of key macroeconomic variables: industrial production, income, trade and sales, employment ,
- Statistical approach: ,
  - Stock and Watson (1989, NBER): Dynamic factor models using variables used by TCB,
  - Chauvet (1996, IER) and Diebold and Rudebusch (1996, ReStat): Dynamic factor models together with two regimes of parameters (Markov process) to capture changes in recessions and expansions but only approximate inference,
  - Kim and Nelson (1998, ReStat): Exact inference.

Earlier contributions - Statistical Approach

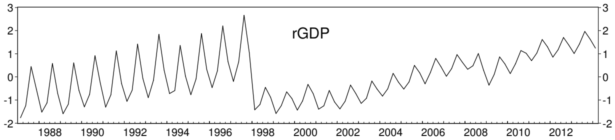
- Monthly coincident indicator using a mixed frequency dataset: Mariano and Murasawa (2003, JAE) ,
- Weekly coincident indicator using a mixed frequency dataset: Aruoba etal (2009, JBES) ,
- Nowcasting using a mixed frequency dataset: Banbura etal (2012, Handbook of Economic Forecasting) and many other papers ,
- This paper falls into the first category

Earlier contributions - Turkey

- Monthly leading indicator using combination of monthly series: Atabek etal (2005, EMFT) ,
- Nowcasting, backcasting GDP: Akkoyun and Gunay (2012, CBRT, WP),
- Monthly coincident indicator using monthly and quarterly series using dynamic factor model on seasonally adjusted series: electricity production, industrial production, imports, employment and GDP: Aruoba and Sarikaya (CBR, 2013).

Analysis of some datasets

- real GDP: Two series:
  - 1987-1997
  - 1998- at quarterly frequency available as original observation



- Non-agricultural employment: Three series:
  - 1988-1999 at semi-annual frequency
  - 2000-2004 at quarterly frequency
  - 2005- at monthly frequency

available as original observations (seasonally adjusted series are available only for recent years)

Our dataset

Variable name	Mnemonic	Frequency	Release
Industrial production index	IP	Monthly	1986M1:2014M5
Import quantity index	MQ	Monthly	1986M1:2014M5
Export quantity index	XQ	Monthly	1986M1:2014M5
Real Gross Domestic Product	rGDP	Quarterly	1987Q1:1997Q1 1998Q1:2014Q1
Real Final Consumption	rFC	Quarterly	1987Q1:1997Q1 1998Q1:2014Q1
Total employment less agricultural employment	NAE	Biannual Quarterly Monthly	1987B1:1999B2 2000Q1:2004Q4 2005M1:2014M5
Trade and services turnover index	TST	Quarterly	2005Q3:2014Q1
Retail sales volume index	RSV	Monthly	2010M1:2014M5

Dynamic factor model - Latent factors

- Consider the **monthly** coincident factor with $AR(p)$ dynamics

$$\Phi(L)C_t = \mu_t + \sigma_{C,t}\eta_{C,t}$$

  We model μ_t and $\sigma_{C,t}$ in a time varying manner using Bayesian semiparametric inference: Density estimation for the coincident factor

- Consider seasonality factor with monthly dynamics

$$S_{1,t} = -\sum_{s=1}^{11} S_{1,t-s} + \sigma_{S_{1,t}}\eta_{S_{1,t}}$$

- Consider seasonality factor with quarterly dynamics

$$S_{2,t} = -S_{2,t-3} - S_{2,t-6} - S_{2,t-9} + \kappa_t\sigma_{S_{2,t}}\eta_{S_{2,t}}$$

  where κ_t takes the value 1 if quarterly variables are observed and 0 if they are not observed.

Temporal aggregation

- Coincident variables load on the (single) coincident factor as

$$X_t^i = \lambda_i^1 C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \lambda(L) C_t + \psi(L) X_t^i + \varepsilon_t^i$$

- We work with the logarithm of the level series (deterministicly detrended)
- All the parameters in the measurement equation are allowed to change with the change of releases
- Stock variables

$$\tilde{X}_t^i = \begin{cases} \lambda_1^i C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \lambda(L) C_t + \psi(L) X_t^i + \varepsilon_t^i & \text{if observed} \\ \text{NA} & \text{otherwise} \end{cases}$$

- Flow variables

$$\tilde{X}_t^i = \begin{cases} \sum_{s=0}^{D_i - 1} \left(\lambda_1^i C_{t-s} + \lambda_2^i S_{1,t-s} + \lambda_3^i S_{2,t-s} + \lambda(L) C_{t-s} + \psi(L) X_{t-s}^i + \varepsilon_{t-s}^i \right) & \text{if observed} \\ \text{NA} & \text{otherwise} \end{cases}$$

where $D_i = 3$ for quarterly variables and $D_i = 6$ for semi-annual variables

Estimating the density of the coincident indicator

- Consider the **monthly** coincident factor with $AR(p)$ dynamics

$$\Phi(L)C_t = \mu_t + \sigma_{C,t}\eta_{C,t}$$

  Let denote $\theta_t = \{\mu_t,\ \sigma_{C,t}\}$

- A finite mixture model consists of latent states, $s_t = 1, \ldots, K$ where, conditional on these states, the parameters take specific values, θ_{S_t}.
- Dummy variables with certain time dependence (Markov process) to be estimated along with model parameters, θ_{S_t}.
- As the number of states tend to ∞ we have an infinite mixture model (iMM).
- In practice the number of states cannot exceed the number of observations, i.e. each observation has a specific parameter in the model (Semi-parametrics in frequentist econometrics)

Statistical background

- Dirichlet process as a generalization of Dirichlet distribution

$$f(\theta_t|\theta_{-t}) \equiv G_t = \frac{\alpha}{\alpha + T - 1} G_0(\Lambda) + \sum_{i=1, i \neq t}^{T} \frac{\delta^j(\theta_i)}{\alpha + T - 1},$$

  where α is the precision parameter and G_0 is the base prior.
- An iMM with Markov structure can be generated using a hierarchical structure where the base prior itself follows a Dirichlet process,

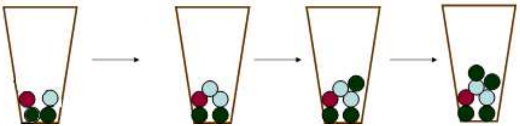
$$G_0(\Lambda) \sim DP(\gamma, H).$$

  We estimate the distribution of the coincident indicator using infinite Mixture Model (iMM), i.e. the number of mixtures and mixture parameters are estimated jointly.

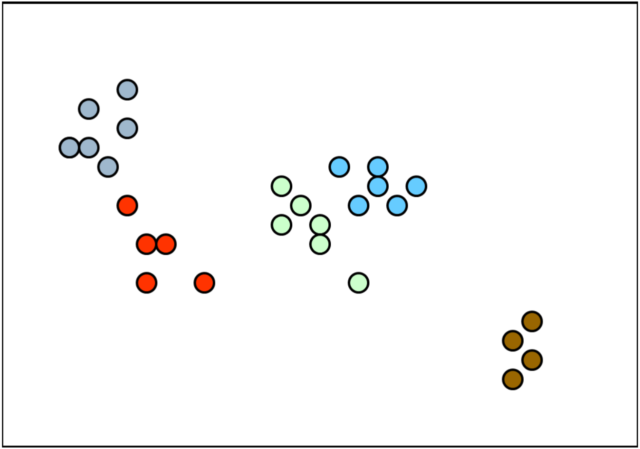
Polya's Urn Scheme: Blackwell and MacQueen (1973, AS)

- Suppose a continuous sample of colored balls of size α and with the distribution of the colors of these balls as $G_0(.)$

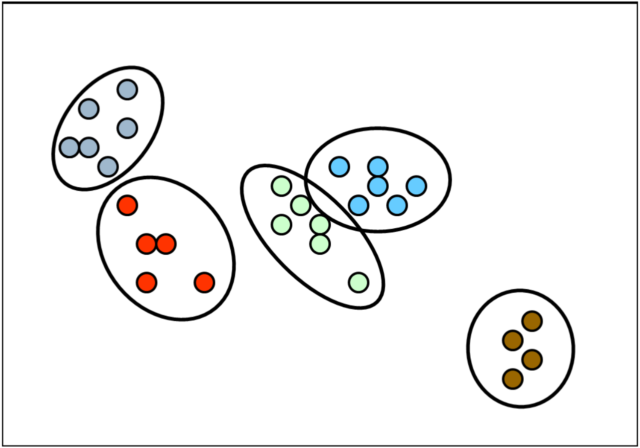
1. A ball is sampled from the urn and upon observing the color of the sampled ball another ball of exactly the same color is added to the urn along with the sampled ball.
2. In the next round, either a ball with a new color from the initial sample of α will be drawn or the ball that is identical to the previously drawn ball will be drawn...
3. Repeating this T times results in a T sample of balls with J different colors, where the distribution of this combination of J colors are distributed according to a Dirichlet distribution



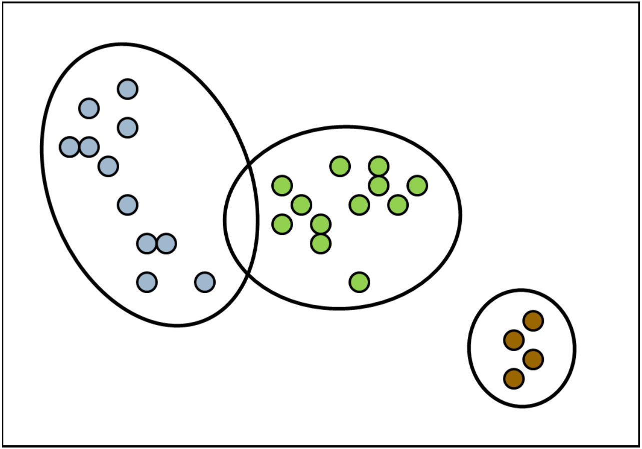
Suppose we observe the following data (error terms)



Should we group the data as



or as



Model in a compact form

$$
\begin{aligned}
\tilde{X}_t^{i,S} &= \lambda_1^i C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \varepsilon_t^i \\
\tilde{X}_t^{i,F} &= \sum_{s=0}^{D_i-1} \left(\lambda_1^i C_{t-s} + \lambda_2^i S_{1,t-s} + \lambda_3^i S_{2,t-s} + \varepsilon_t^i\right) \quad && \varepsilon_t^i \sim N(0, \sigma_i^2) \\
\Phi(L) C_t &= \eta_{C,t} && \eta_{C,t} \sim N(0, \theta_t) \\
\theta_t | G &\sim DP(\alpha, G_0) && G_0 \sim DP(\gamma, H) \\
S_{1,t} &= -\sum_{s=1}^{11} S_{1,t-s} + \eta_{S_{1,t}} && \eta_{S_{1,t}} \sim N(0, \sigma_{S_{1,t}}^2) \\
S_{2,t} &= -S_{2,t-3} - S_{2,t-6} - S_{2,t-9} + \kappa_t \eta_{S_{2,t}} && \eta_{S_{2,t}} \sim N(0, \sigma_{S_{2,t}}^2)
\end{aligned}
$$

Bayesian Inference

- Takes uncertainty about parameters when estimating unobserved components and vice versa
- Bayesian semiparametrics allow for explicit density estimation in the state space framework
- We can also make a density forecast (nowcast/backcast) of the missing observations as we estimate the density of the coincident factor (not in this paper)

Prior specifications

- Base prior

$$H(\mu_0, a, \nu, S) = N - IG(\mu_0, a, \nu, S) \mapsto \mu|\sigma_H^2 \sim N(\mu_0, \frac{\sigma_H^2}{a}) \text{ and } \sigma_H^2 \sim IG(\nu, S),$$

  Parameters are set such that the base prior is standard normal.

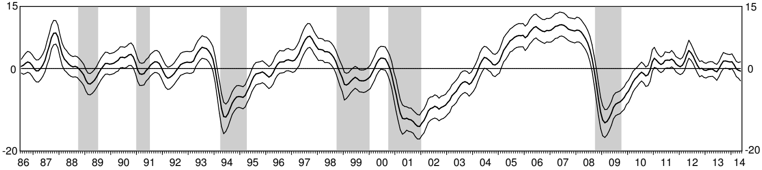
- Precision: $\alpha : f(\alpha) \propto (1 - \frac{\alpha - \alpha_{\min}}{\alpha_{\max} - \alpha_{\min}})^p$ where $p = 0.8$ and $\alpha \in [0.1, 0.5]$
- Persistence of the coincident factor: $vec(\Phi) \propto N(0, I * 10)$
- Factor loadings: $\lambda_{i,k} \propto N(1, 10)$ for $i = 1, \ldots, N$ and $k = 1, 2, 3$
- Variance of the measurement error: $R : f(\sigma_i^2) \propto \frac{1}{\sigma_i^2}$
- Variance of the state error: Set to identity matrix except the coincident factor for identification.

MCMC Algorithm

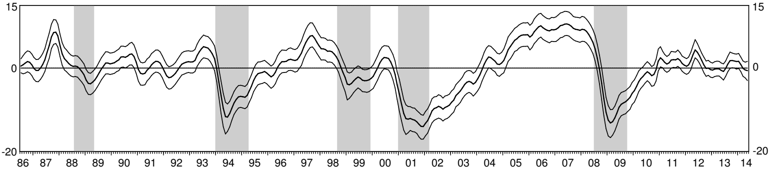
1. Initialize the parameters by drawing θ_t from the base prior $G(\Lambda)$. At step (m) of the iteration
2. Sample $f_t^{(m)} = (C_{1:T}^{(m)}, S_{1,1:T}^{(m)}, S_{2,1:T}^{(m)})'$ from $p(f_t|y_{1:T}, \theta_{1:T}^{(m)}, \Phi^{(m)})$ using Carter and Kohn (1994) or another alternative.
3. Sample $\Phi^{(m)}$ from $p(\Phi| C_{0:T}^{(m-1)}, \theta_{1:T}^{(m-1)})$.
4. Sample $\sigma_i^{2,(m)}$ from $p(\sigma_i^2|y_{1:T}, f_t^{(m-1)})$ using the observations that are not missing.
5. Sample $\lambda_i^{(m)} = (\lambda_{i,1}^{(m)}, \lambda_{i,2}^{(m)}, \lambda_{i,3}^{(m)})'$ from $p(\lambda_i|y_{1:T}, f_t^{(m-1)})$ using the observations that are not missing.
6. Sample $\theta_{1:T}^{(m)}$ from $p(\theta_t| C_{1:T}^{(m-1)}, \theta_{-t}^{(m-1)}, \Phi^{(m)})$ for $t = 1, 2, \ldots, T$ using slice sampler of Neal (2003, Ann.Stat.) and Kalli etal (2001, Stat&Comp).
7. Sample $\alpha^{(m)}$ from $p(\alpha|\theta_{1:T}^{(m)})$.
8. Repeat (2)-(6) M times.

Estimated components: Coincident factor (deviation from the trend)

Coincident factor and GDP based recession dates (BBQ algorithm yoy)



Coincident factor and IP based recession dates (BBQ algorithm yoy)



We obtain level series using the filter. We use a modified version of the trick in Stock and Watson (1989). Observe that the coincident indicator is a function of the data

$$C_{t|t} = W(1)X_t$$

Taking the expectations of both side

$$\begin{aligned} E[C_{t|t}] &= W(1)E[X_t] \\ \alpha_0 + \alpha_1 t &= W(1)\bar{X}_t \end{aligned}$$

We can obtain this weight, W, from the filter using the steady state conditions and an approximation through the relation between $c_{t|t} = W\tilde{X}_t$.

In compact form the model (conditional on the parameters) can be written as

$$\begin{aligned} X_t &= \mu + HS_t + \varepsilon_t \qquad && \varepsilon_t \sim N(0, R) \\ S_t &= FS_{t-1} + \eta_t && \eta_t, \sim N(0, Q), \end{aligned}$$

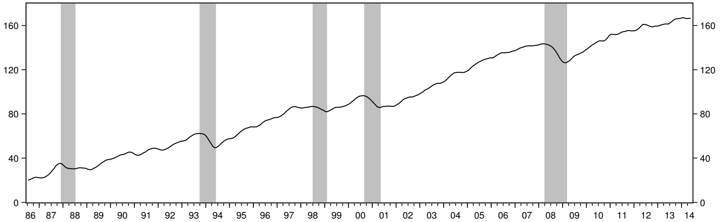
Then;

$$\begin{aligned} S_{t|t} &= W\tilde{X}_t \\ S_{t|t} &= S_{t|t-1} + K_t \zeta_{t|t-1} \\ S_{t|t} &= FS_{t-1|t-1} + K_t \zeta_{t|t-1} \\ S_{t|t} &= FS_{t-1|t-1} + K_t(\tilde{X}_t - HS_{t|t-1}) \\ &\vdots \\ S_{t|t} &= (I - (I - KH)F)K\tilde{X}_t \end{aligned}$$

thus $W = (I - (I - KH)F)K$ and we use the 1^{th} row of the matrix W

Estimated components: Coincident factor (level)

Coincident factor and indicated recession dates



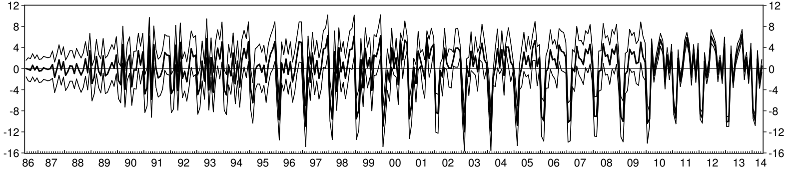
- BBQ algorithm based on yoy growth rates indicate recessions later than that of indicated by the coincident factor.
- This delay can also be observed using the series published by TurkStat
  yoy changes: 1999Q1-1999Q4, 2001Q2-2001Q4, 2008Q4-2009Q3
  qoq changes: 1998Q3-1999Q3, 2001Q1-2001Q2, 2008Q2-2009Q1
- The model can capture recessions timely.

Estimated components: Coincident factor

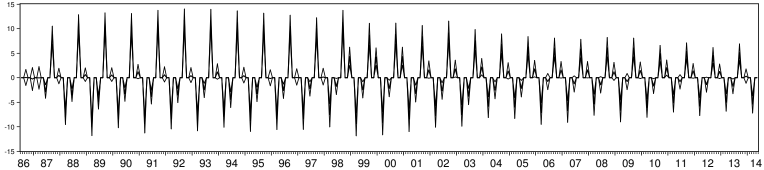
- The coincident factor is able to track the recessions: 1994 financial crisis, 1998 Russia crisis, 2001 crisis and 2008-09 great recession.
- It captures the rapid growth period of 2000's.
- It indicates a sluggish growth in the post great recession era.

Estimated components: Seasonal factors

Monthly factor



Quarterly factor

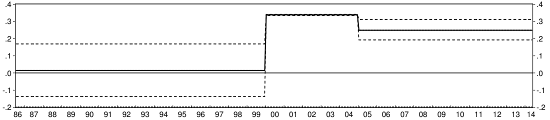


Estimated components: Seasonal factors

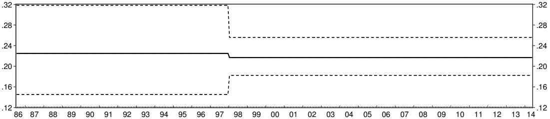
- The monthly factor is more variable with larger bands indicating the uncertainty surrounding the magnitude of the seasonality.
- Quarterly factor shows decreasing (in magnitude) pattern.
- Seasonality can be captured nicely.

Factor loadings for coincident factor

Factor loadings of non-agriculture employment series



Factor loadings of GDP series



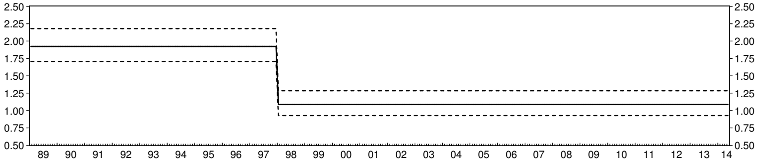
- Other loadings of remaining variables on coincident factors are stable.

λ_C

	5% LB	Median	5% UB
IP	1	1	1
rFC	0.19	0.22	0.26
XQ	0.30	0.52	0.74
MQ	1.49	1.75	2.06
RT	0.94	0.96	0.97
THC	0.28	0.55	0.83

Factor loadings for seasonal factors

Factor loadings of GDP series for quarterly seasonal factor

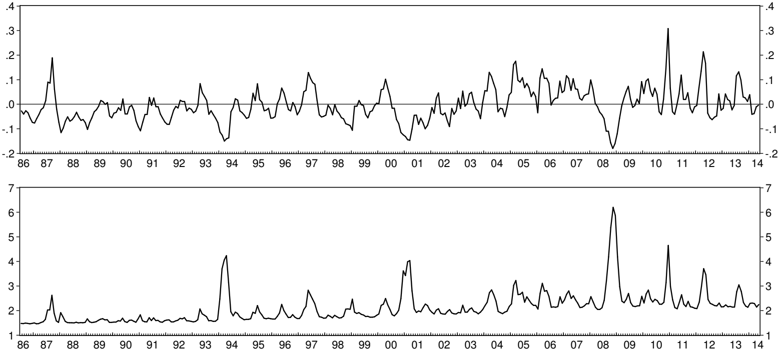


	λ_{S_1}			λ_{S_2}		
	5% LB	Median	5% UB	5% LB	Median	5% UB
IP	0.75	0.89	1.04	0	0	0
rFC	-0.17	-0.08	0.01	1.12	1.31	1.53
XQ	0.71	0.99	1.27	-0.84	-0.49	-0.17
MQ	0.80	1.06	1.32	-0.34	-0.03	0.26
RT	1.78	1.78	1.79	1.57	1.58	1.58
THC	-0.50	0.19	0.96	1.08	1.49	1.97

- Monthly series load more on monthly seasonal factor (e.g. XQ, MQ)
- Quarterly series load more on quarterly seasonal factor (e.g. rFC, THC)

Effects of semiparametric structure

iMM allows for time variation in the intercept and the variance



- Intercept is volatile often with changing signs (always negative during recessions).
- Volatility is increasing during recessions and for some period it increases enormously capturing extremes.

Conclusion

- A monthly coincident indicator is constructed.
- Structural changes and mixed frequencies are taken into account.
- Common seasonality is modelled explicitly.
- The density of the coincident factor is estimated along with other model parameters.

Findings

- Change in the data structure affects factor loadings dramatically.
- Monthly and quarterly common seasonal factors are sufficient for capturing seasonal patterns.
- Semiparametric structure allows for changes in volatility and intercept.

Future work

- Coincident indicator at a higher frequency,
- Density nowcasting using semiparametric structure,
- ...