Tracking Indian Growth in Real Time

Giovanni Veronese

Bank of Italy - Embassy of Italy in New Delhi

17 September 2009

The task

Data does not help

Tracking GDP in realtime

Forecasting in (still) a data poor environment

Bridge model approach

Real time evaluation

A large dataset approach

Conclusions

Mandate: develop tools to track the state of the economy

- Assessment of current state of the economy is key for good macro forecasting (starting point) and hence policy
- This holds even more in today's environment, characterized by extreme uncertainty
- Longer-term analysis largely a leap of faith

What do we know in India?

We have various indicators, unfortunately not the plethora available elsewhere

On Aug.30 2009 just before the Q2-2009 GDP data release we knew:

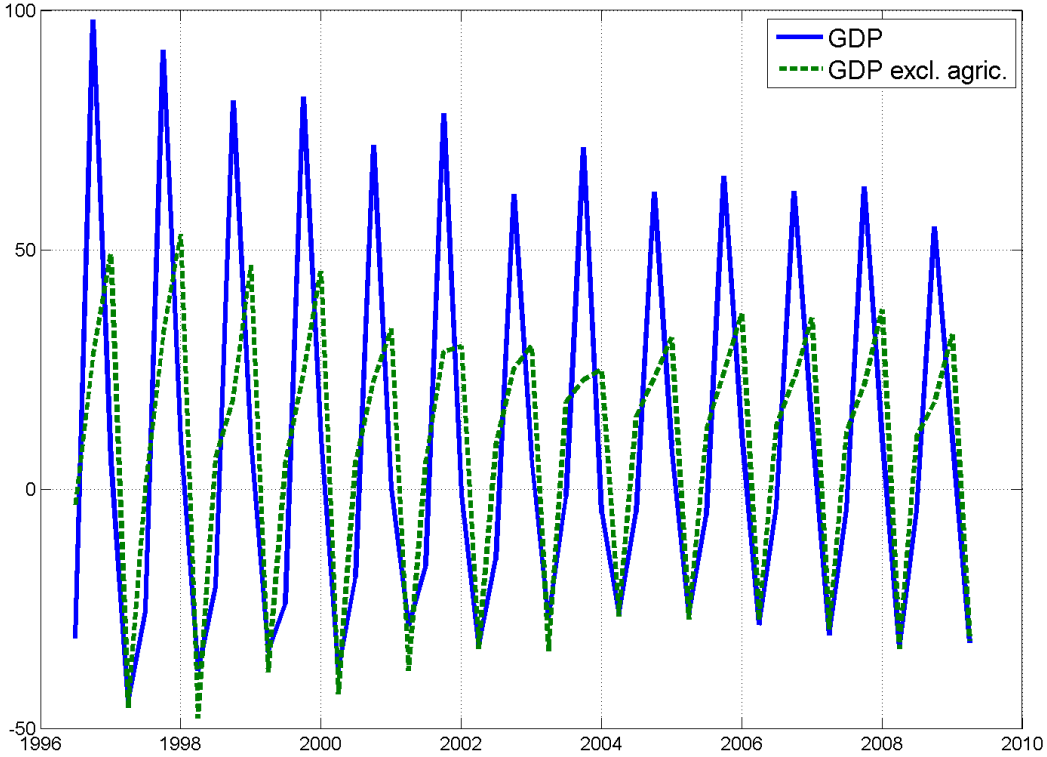
- GDP: Q1-2009
- IIP: July 2009
- Monetary variables: July 09
- Small set of activity series: car sales, trade, electricity (mainly July 09)
- Financial markets data: in real time

Fragmented picture:

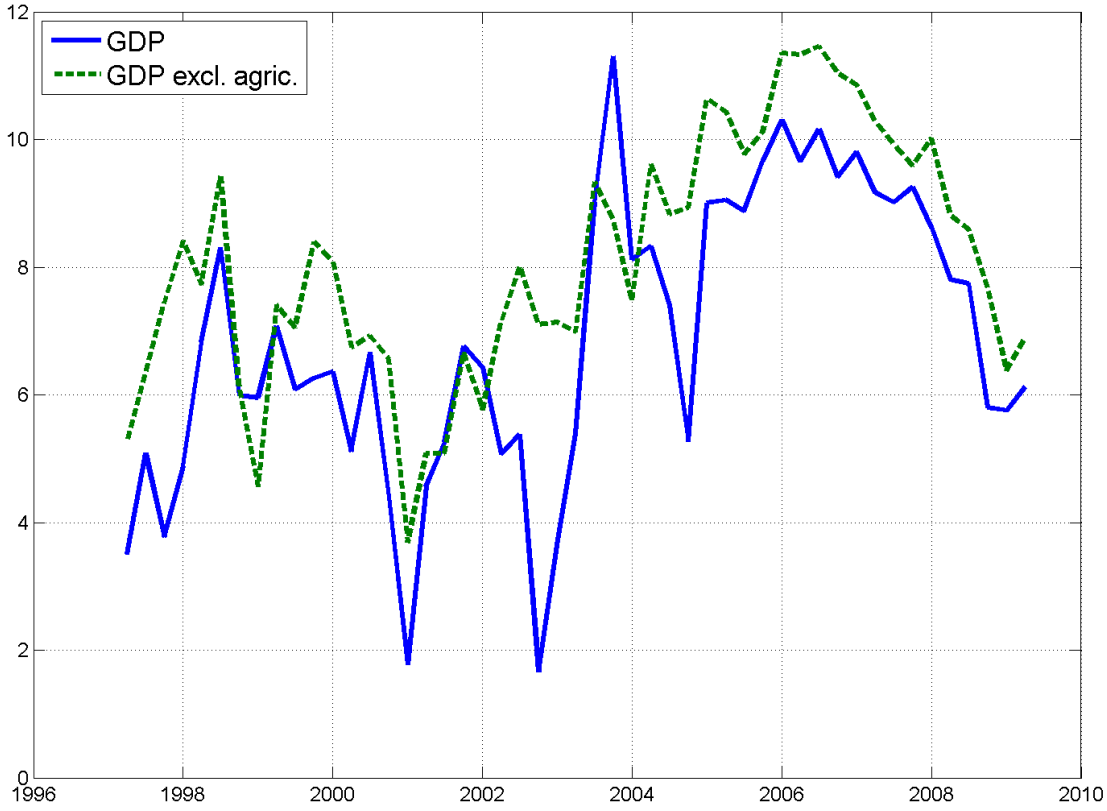
- $\rightarrow$ hard to gauge most recent conditions
- $\rightarrow$ even harder to see through the up/downs (smoothing)

This difficulty compounds with lack of official SA data

Unfortunately, both for GDP and GDP excluding agriculture, seasonality is a substantial source of variability in the growth rates,



Agriculture adds further volatility to growth



Agriculture driven mostly by weather/temperature conditions

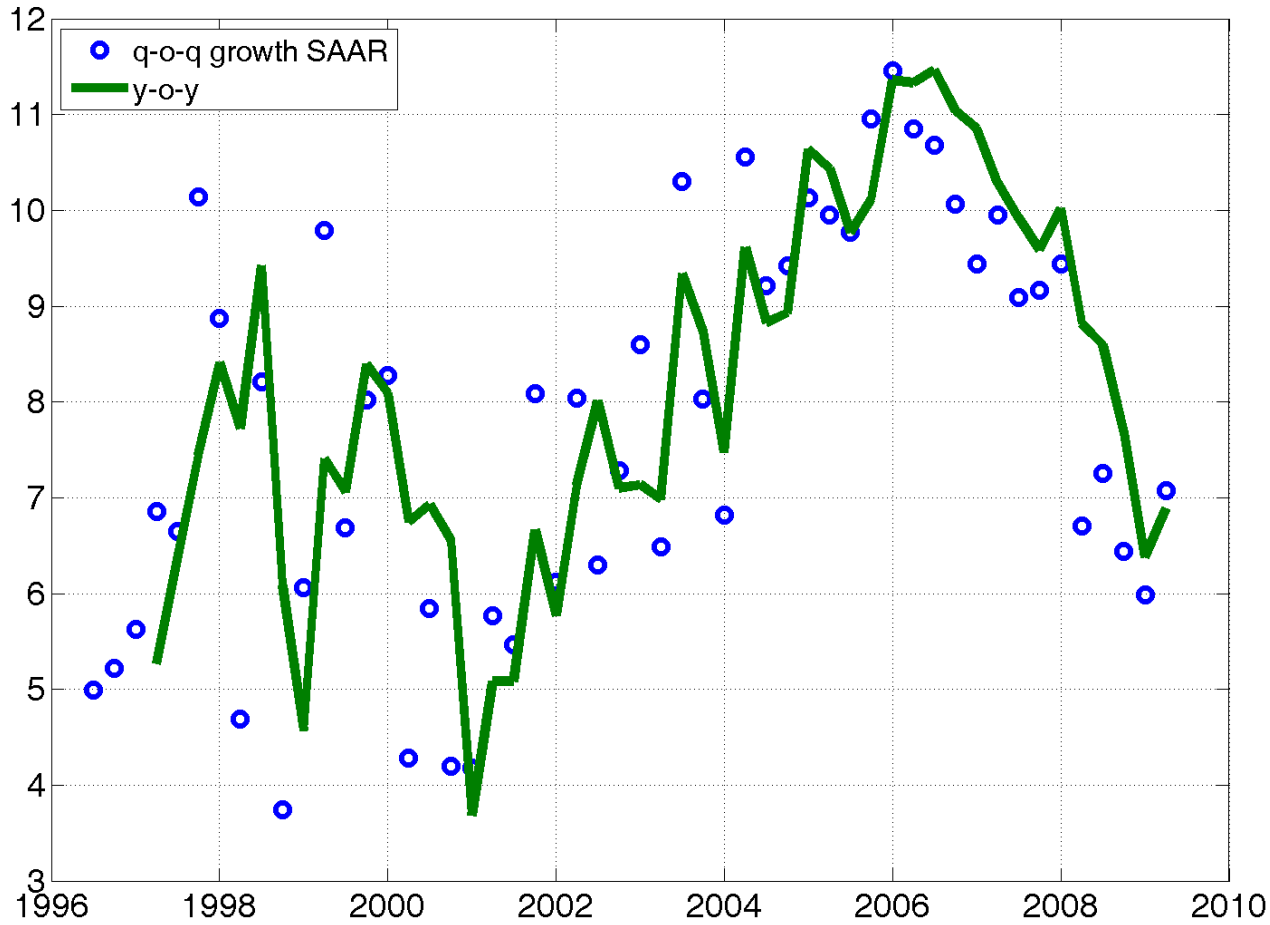
From now on we will focus on GDP excluding agriculture

Lack of SA data: look at year-on-year growth rates

But we would like to assess:

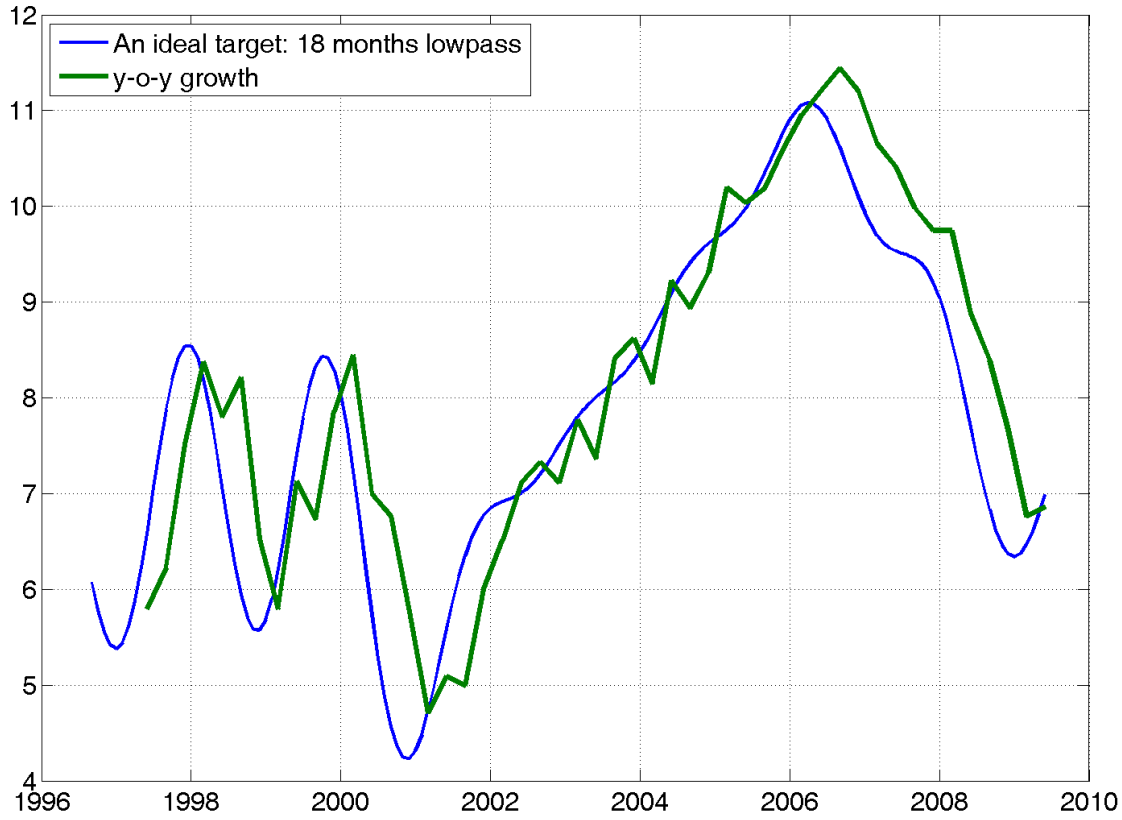
a) more recent dynamics → SA the data,

b) ideally, also be able to have a smoother picture



b): smoother signal from lowpass filter

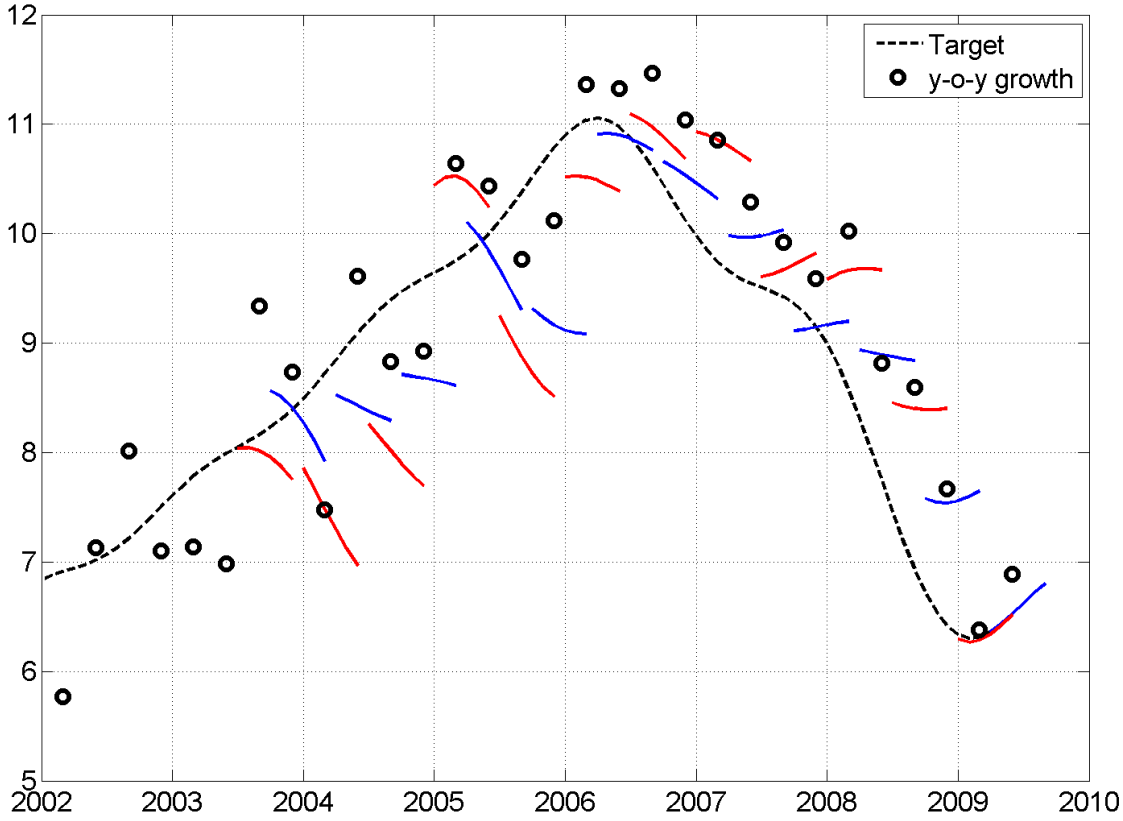
The signal leads the y-o-y, because it relies on bilater filter



In practice this is a challenge at the end of the sample

Good filtering needs good fore(now)-casting

Using only the *past* of GDP and univariate filtering techniques



But we may miss turning points and signal false ones too often.

What's out there

The literature on GDP tracking considers:

- Bridge Models: (Bank of Italy, ECB, many private analysts)
- Factor Models
    1. small scale: $\rightarrow$ state$-$space (Bank of Spain, Philadelphia Fed)
    2. large scale: $\rightarrow$ Eurocoin (Bank of Italy), ECB models, Swiss Central Bank
- Large scale BVAR models (ECB)

Accounting for more timely information

- Univariate methods: using seasonal adjustment may help somewhat,

  but we ignore information available

- how can we use this extra info: e.g. from monthly indicators?

Bridge model approach

- Bridge models are developed to link monthly releases with quarterly GDP growth
- Used by many institutions and private organizations
- Large literature on the subject (Baffigi et al. 2004)
- Often resort to hard (IIP) and soft data (surveys)

In India: little or no literature on the subject

It may have to do with lack of data (more on this later).

An aside: what is GDP

We should think of GDP as a black box.

Statistical offices (in India the CSO) access large amount of information, only partly observable to the public (e.g. NIPFP). The CSO processes this information set using NA methods

$$\Omega_t^{CSO} \rightarrow \boxed{\text{National Accounts}} \rightarrow \widehat{GDP}$$

$$\Omega_t^{NIPFP} \rightarrow \boxed{\text{Proxy Model}} \rightarrow \widetilde{GDP}$$

and $\Omega_t^{NIPFP} \subset \Omega_t^{CSO}$

However: Ω_t^{CSO} is a limited information set

In India:

- with a small set of indicators one can almost *reconstruct* the GDP data (excluding agriculture)
- relying on timely **monthly** indicators does the job
- estimation can be performed in real-time as they are released

National accounts information set

Sectors	Indicators	availability
1. agriculture	quarterly forecast crops	no
2. livestock	milk, cheese, eggs and wool production :	no
3. forestry	annual forecast split in four quarter	no
4. fishing	production fish	no
5. mining and quarrying	IIP mining, coal, petroleum	yes
6. manufacturing	IIP manufacturing	yes
7. electricity,gas and water supply	IIP electricity	yes
8. construction	cement, steel, coal production	partial
9. trade, hotels and restaurants	Gross Trading Index	partial
10. railways	net tonne Kms, passenger Kms	partial
11. other transport	commercial vehicles, ports, air freight	partial
12. communication	installed telephones	partial
13. banking and insurance	deposits , bank credits, WPI banking	partial
14. public administration	central govt revenue expenditure, CPI	partial

The main data releases

	Mar-09	Apr-09	May-09	Jun-09	Jul-09	Aug-09	Last release
GDP	O			O			31Aug2009
IIP	X	X	X	X	X		12Aug2009
Cement	X	X	X	X	X		26Aug2009
Steel	X	X	X	X	X		26Aug2009
Coal	X	X	X	X	X		26Aug2009
Railway	X	X	X	X	X		18Aug2009
Ports	X	X	X	X	X		26Aug2009
Turists	X	X	X	X	X		03Sep2009
Cars	X	X	X	X	X		11Aug2009
Electricity	X	X	X	X	X	X	02Sep2009
Phones	X	X	X	X	X		13Aug2009
Credits	X	X	X	X	X		14Aug2009
Deposits	X	X	X	X	X		14Aug2009
Govexp	X	X	X	X			28Aug2009

GDP reconstructed from sectoral VA

Bottom-up approach: try to mimick what CSO effectively does, but with smaller information set
and with some uncertainty on functional forms:

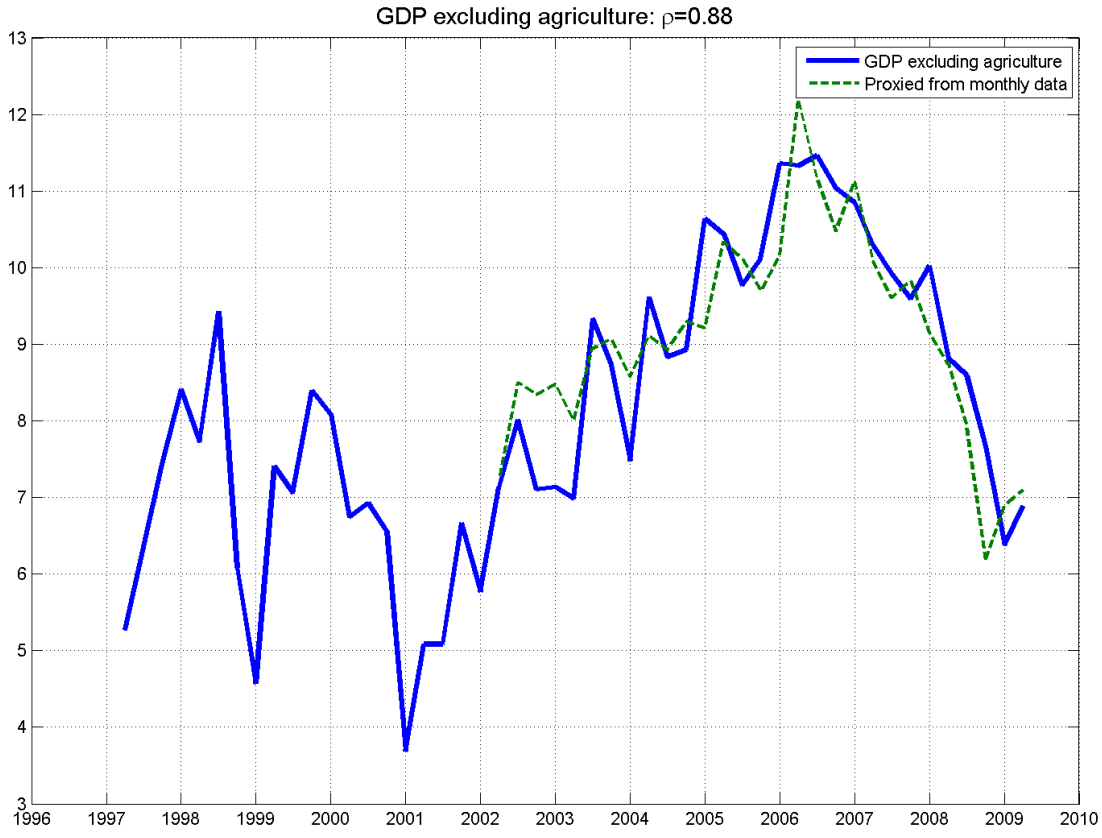
- $\widehat{VA}_t^s = VA_{t-4} * \sum_j \alpha_j \Delta(x_j^s)$
- x_j^s are proxy indicators (mainly monthly) for sector s
- $\Delta\widehat{GDP} = \sum_s \beta_s \Delta\widehat{VA}_t^s$

The essence of a bridge model:

Two step procedure:

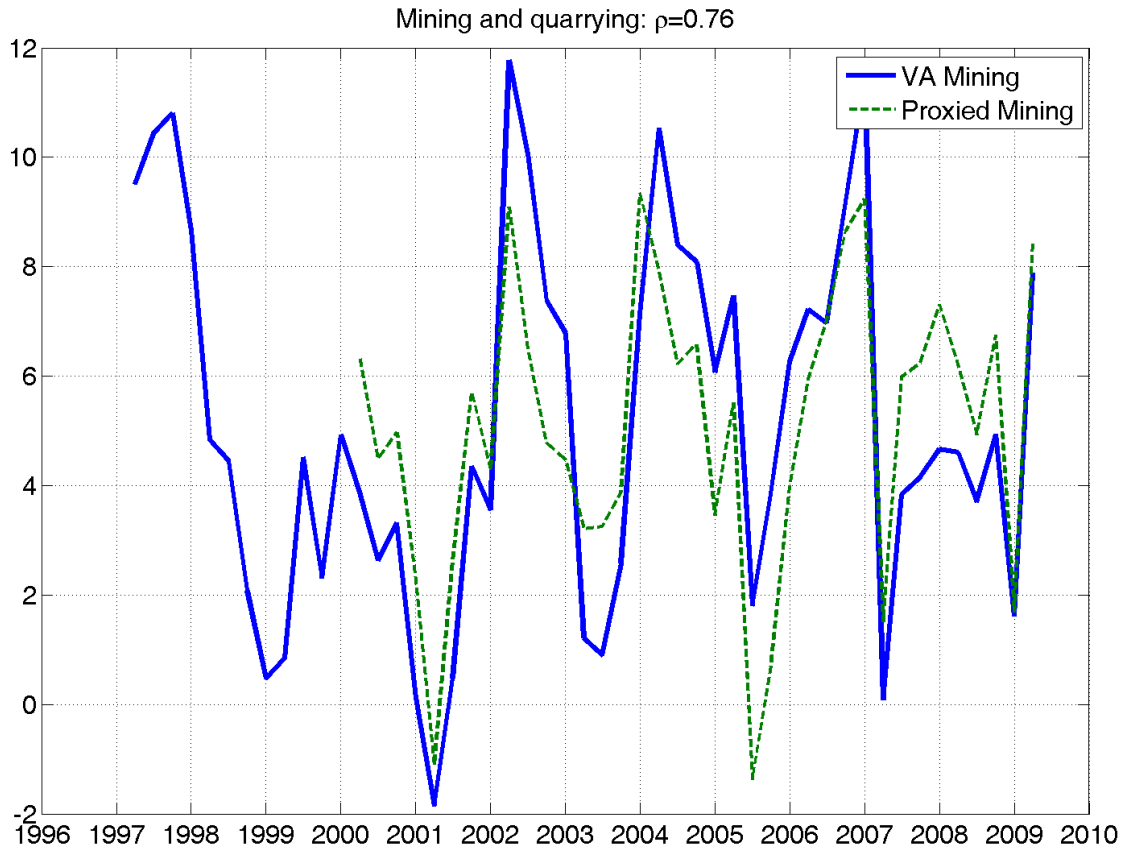
1. monthly step
   - for each sector we consider monthly indicators $x_{1,t}, \ldots, x_{k,t}$.
   - forecast the individual indicators with univariate methods,
   - aggregate the monthly variables to quarterly frequency: x_i^Q
2. quarterly step:
   - estimate bridge equations: $VA_{i,t}^Q = \mu_i + \sum_{s=0}^{q_i} x_{i,t-s}^Q + u_{i,t}$
   - aggregate sectoral forecast to obtain GDP forecast

In sample fit using a bottom-up approach



Caveat: comparisons using *final* data can be misleading

CSO may gradually revise their estimates using more detailed info, available after 1-2 years.



The *pseudo-real time* exercise (1)

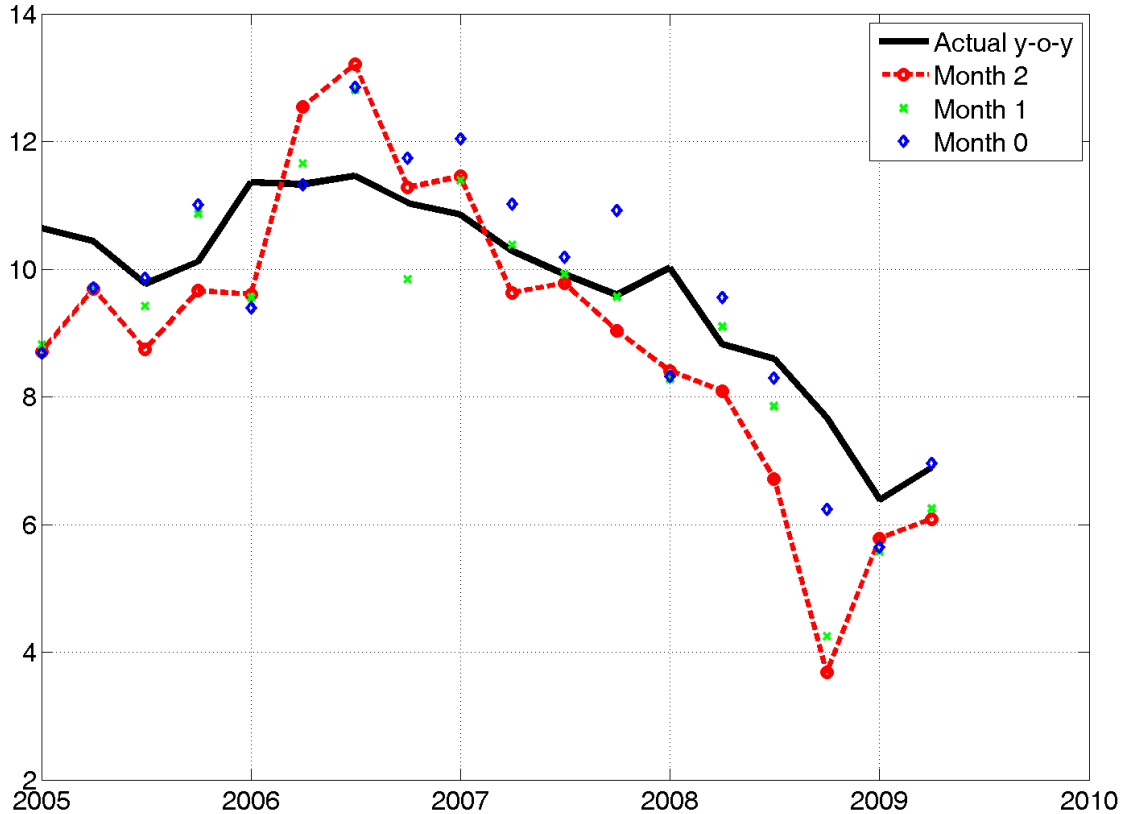
- At each step we replicate the information set available at the time of the GDP release
- We seasonally adjust at each step all the monthly indicators
- At each step we estimate our bridge models
- We take into account the actual release dates:

  (e.g. at end of Aug.2009 we have IIP up to June, but car sales up to July)

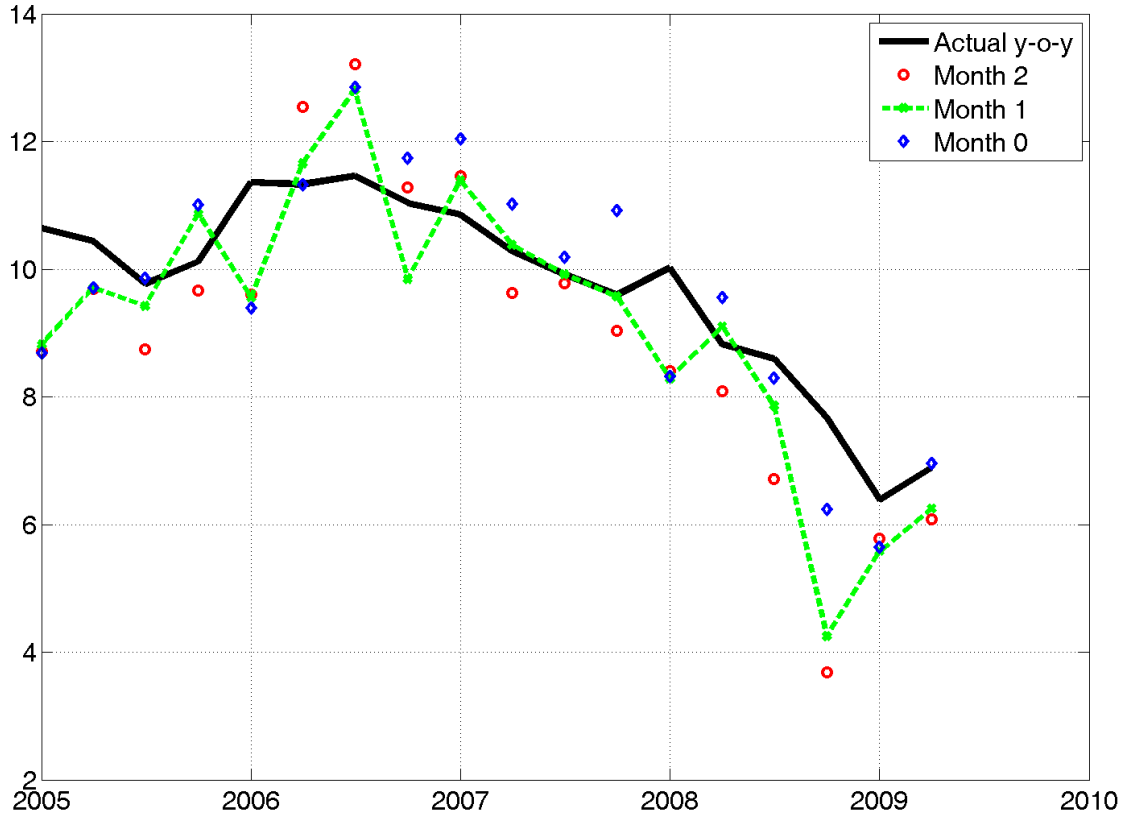
The *pseudo-real time* exercise (2)

- For each quarter we consider 3 consecutive forecasts (*nowcasts*):
  1. month 2: two months ahead of GDP release
  2. month 1: one month ahead of GDP release
  3. month 0: just ahead of the GDP release (our information set closest to the CSO one)

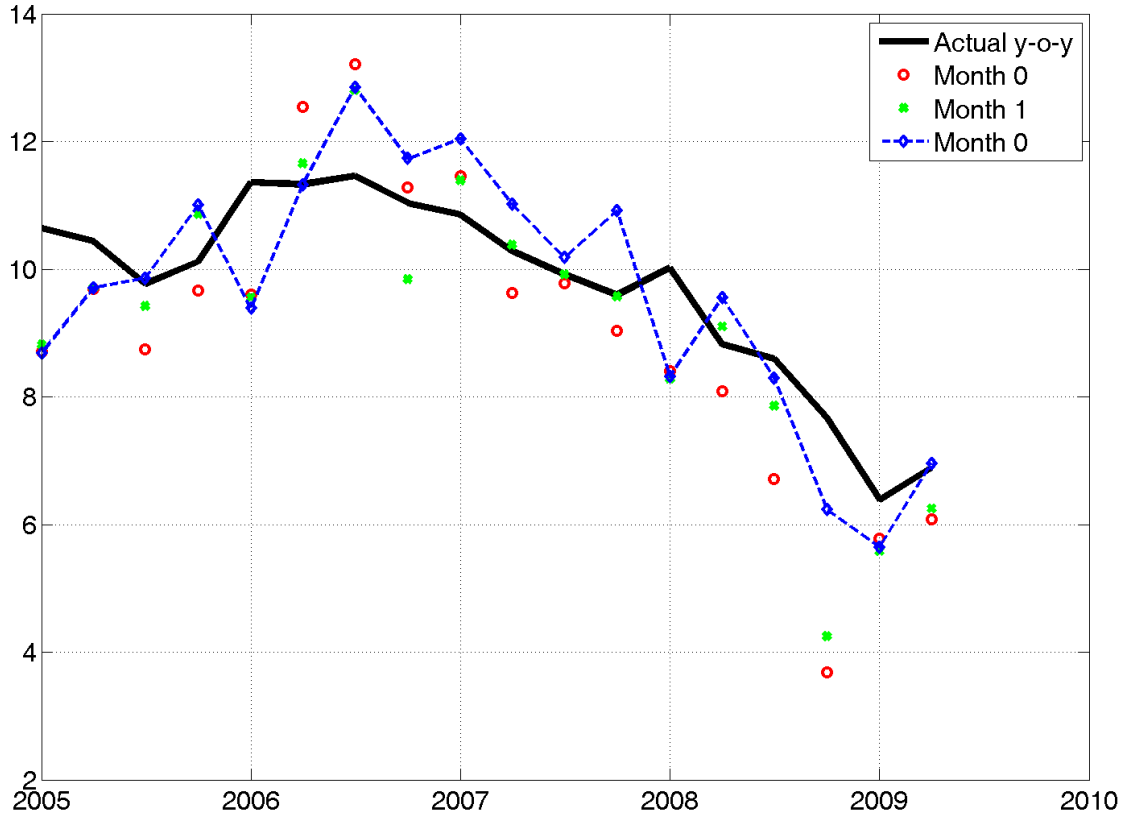
Forecasting: 2 months away from GDP release



Forecasting: 1 month away from GDP release



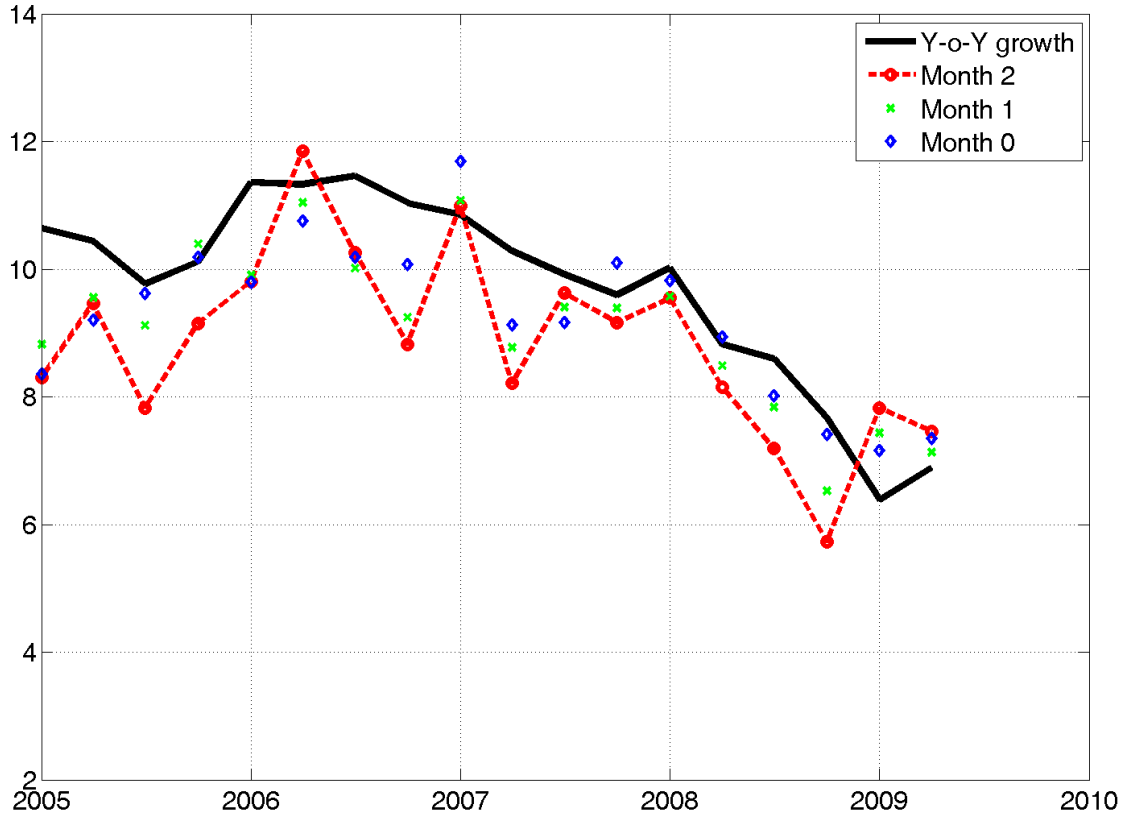
Forecasting: 0 months from GDP release (right before)



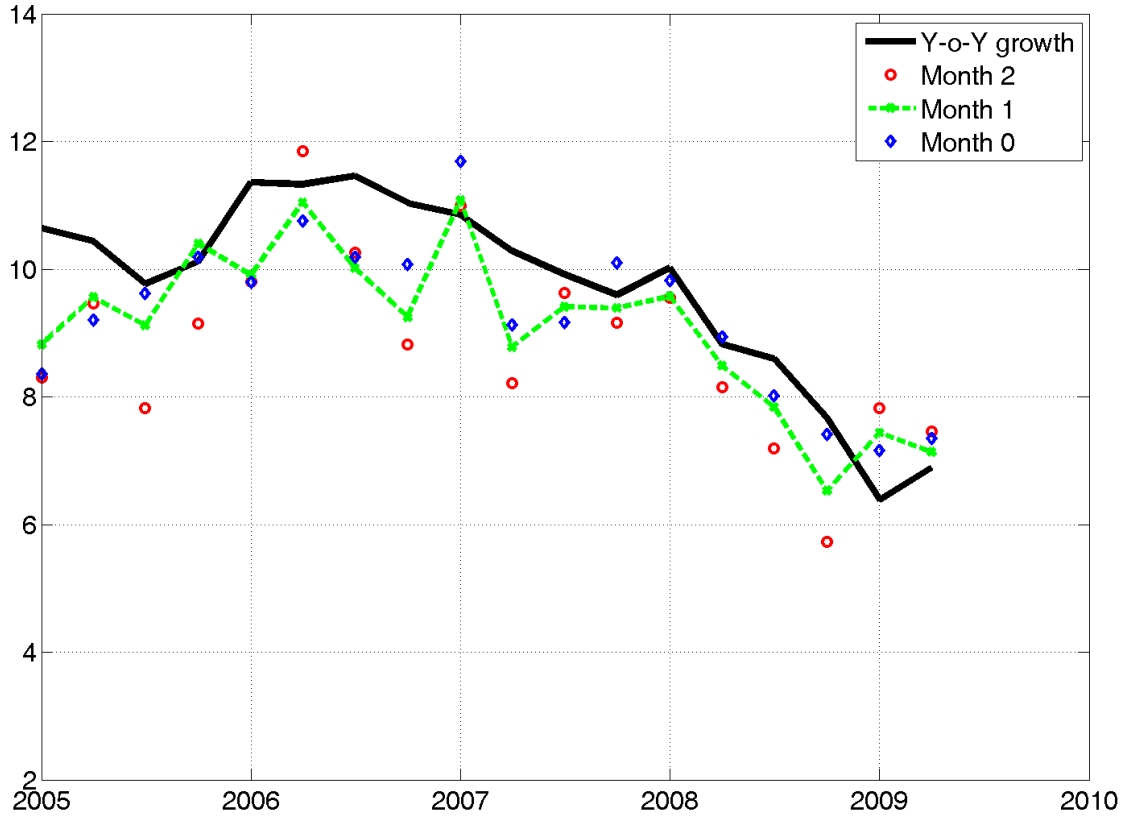
Can we obtain more timely info using q-o-q data?

- estimate bridge models on q-o-q SA monthly data
- use univariate forecasts implicit from the ARIMA model estimated
- expect some improvement in leading info
- but probably at a cost of greater volatility

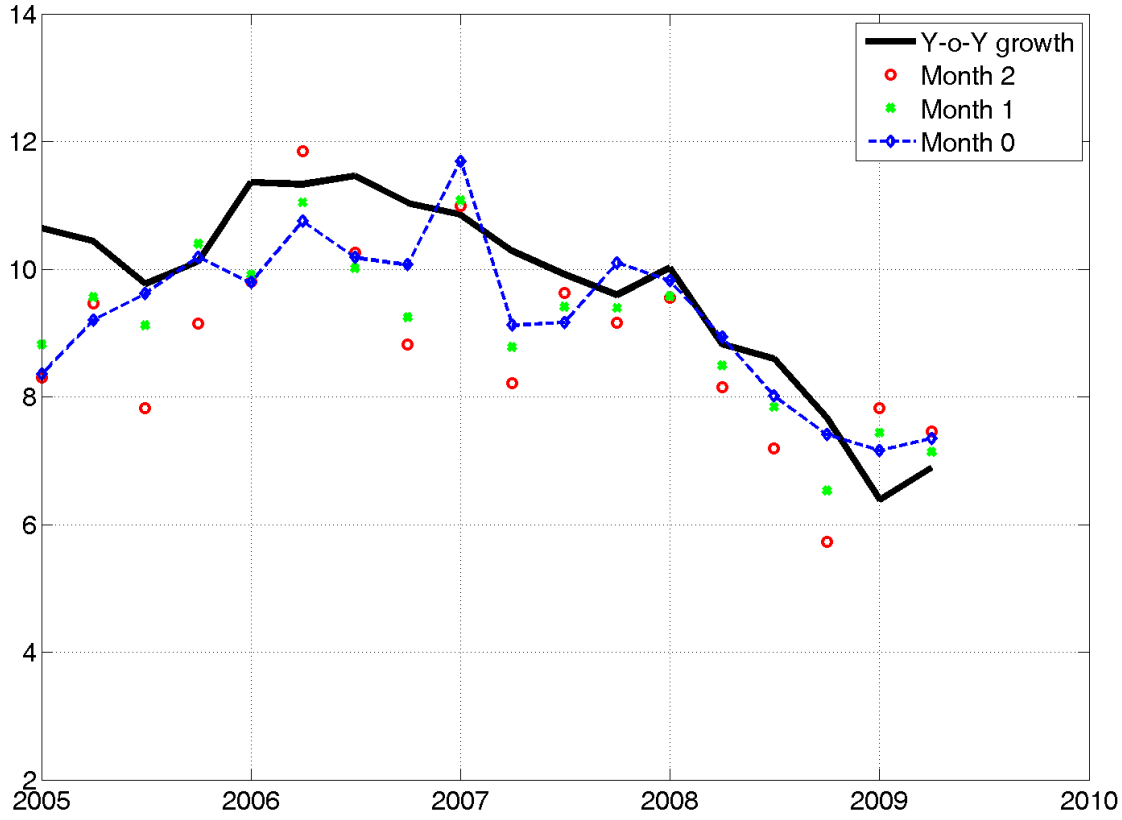
Using SA data: 2 months away from GDP release



Using SA data: 1 month away from GDP release



Using SA data: few days from GDP release



Forecasting GDP y-o-y: RMSE

Info set	Month 2	Month 1	Month 0	Final
GDP AR				1.08
GDP RW				0.88
Bridge top down				1.74
Bridge bottom up	1.42	1.33	1.16	1.02
Bridge bottom up SA	1.36	1.00	0.95	0.92

Out of sample for the quarters: 2004Q1 to 2009Q2
Bridge models: top down uses only GDP and the monthly indicators, bottom up takes into account sectoral composition of VA

Can we exploit more information?

We could also construct very large info set ($\Omega_t^{NIPFP} \supset \Omega_t^{CSO}$)

How to exploit it?

- Difficult in bridge model framework: collinearity issues
- Factor models: extract information from large dataset:

  we depart from the "mimimicking" the CSO approach to *fish* for potentially more informative series
- Pros: allow to deal in efficient way non synchronous flow of data, as missing obs on monthly indicators are predicted in optimal manner, using the entire cross-section
- Cons: ignore the zero/1 weights constraints on many variables

We are *bridging* with factors

Two step procedure:

1. monthly step
   - Consider large set of monthly indicators $X_t = x_{1,t}, \ldots, x_{N,t}$.
   - Extract small set of factors: $X_t = \Lambda f_t + \xi_t$
   - We can forecast each variable using the factor structure
2. quarterly/monthly step:
   - Version 1: Forecast GDP using quarterly factors
   - Version 2 (here): Project GDP on monthly factors mixed frequency to obtain monthly indicator

Factor models to extract cyclical component

Intuition: look for suitable linear combinations of variables: (averaging across sectors)

In Eurocoin *averaging* is designed to maximize common medium to long run features (more interesting for policy)

We attempt similar methods to obtain a monthly indicator of *underlying* GDP growth in India

We would like to think that: resorting to a wider information set may improve estimation of what the CSO will eventually revise its GDP to be (but it takes time to verify such a claim).

FYI: some statistical agencies are considering move towards factor models in quarterly national accounts estimation (**OECD 2009**)

Estimation of cyclical component of GDP growth c_t

$c_t = \beta(L) GDP_t^{qoq}$ ... so we need to address 3 problems:

1. How to reach the *end of the sample*, as $\beta(L)$ is two-sided
2. How to solve the *end-of-sample unbalacing* due to the varying timeliness of data releases
3. How to obtain a *monthly* indicator, as *observed* GDP is only available quarterly

use a large panel to obtain a reliable (monthly) estimate of MLRG at the end of the sample

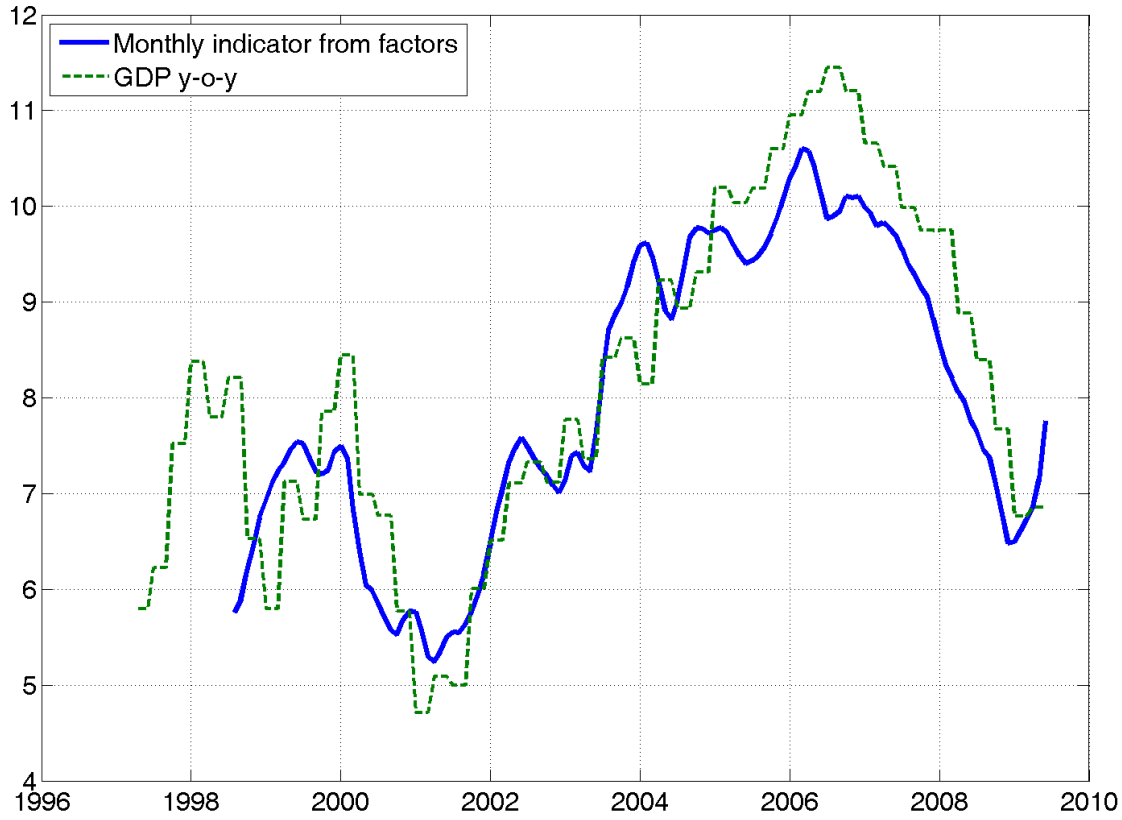
Summary of results for India

In sample fit of the model:

- by construction: it leads y-o-y GDP by approximately 1.5 qtrs, but
- we have lots of info on industry but
- very little info at monthly frequency of services sector, if it's there it's too short
- this probably is the cause of missing out the peak growth episode (driven by private services)
- foreign variables added: monthly business surveys from US and Europe (a smooth signal)
- unfortunately no Indian survey variable available: private industry still in infant stage (eg PMI), while RBI does not disclose their own

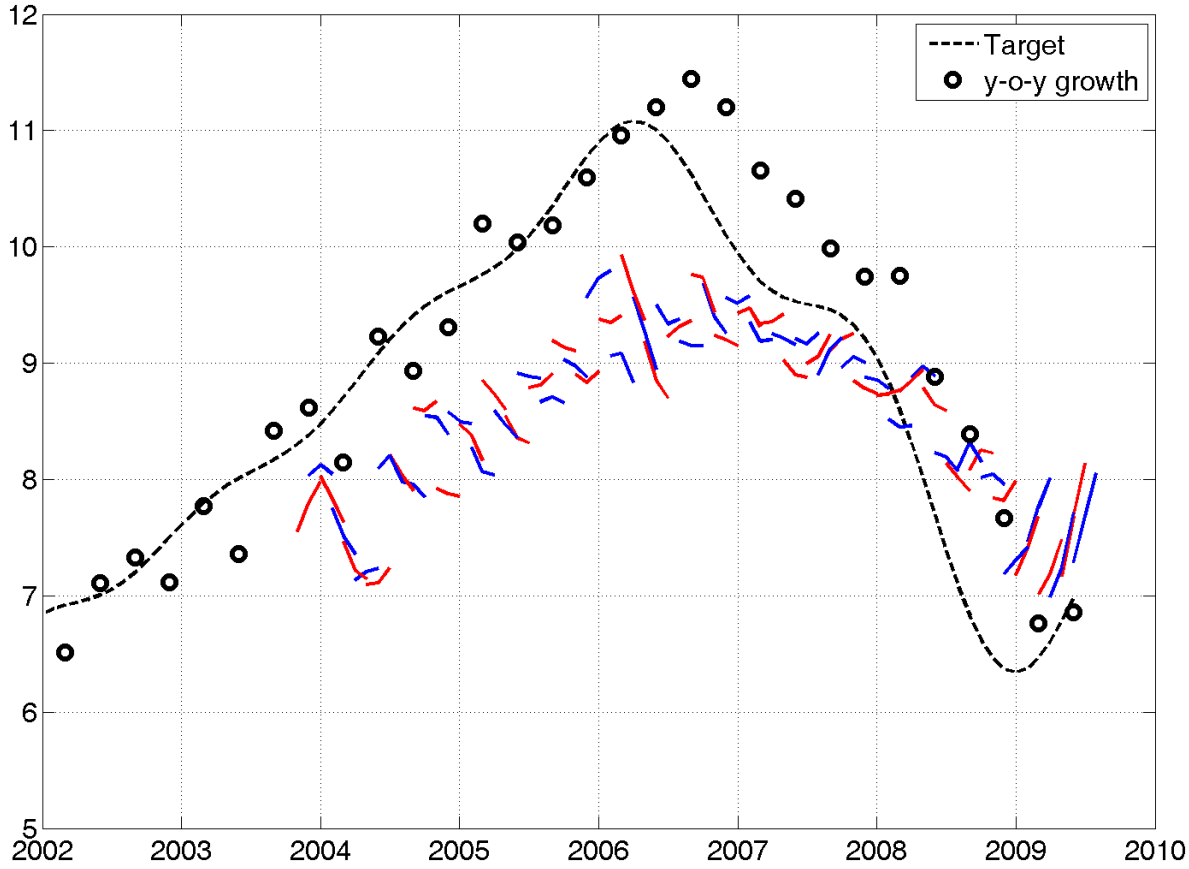
In sample: task is easier

We project quarterly GDP growth on a set of smooth monthly factors



Monthly dataset includes approximately 80 variables (India + Rest of world)

Out of sample: tough life



too many false signals: further calibrations may improve performance

Conclusions 1: tracking can be done effectively

- GDP tracking: first ever attempt to perform pseudo real-time assessment
- Small scale model seem more promising (and there are obvious reasons for this)
- Need to improve the modelling using seasonally adjusted data: may help in providing more timely information
- Large scale model: still overly ambitious, lack of enough indicators on services sector and surveys
- Expand the information set to embed quarterly information (e.g. from CMIE Prowess)

Conclusions 2: Policy implications

- Need to expand the statistical basis of GDP
- Growth pattern in some sectors depends almost entirely on few proxy variables
- Sometimes these proxy variables may have drawbacks (e.g. IIP) in measuring value added growth
- More effective short term analysis only with SA data