European Commission

Directorate General Economic and Financial Affairs

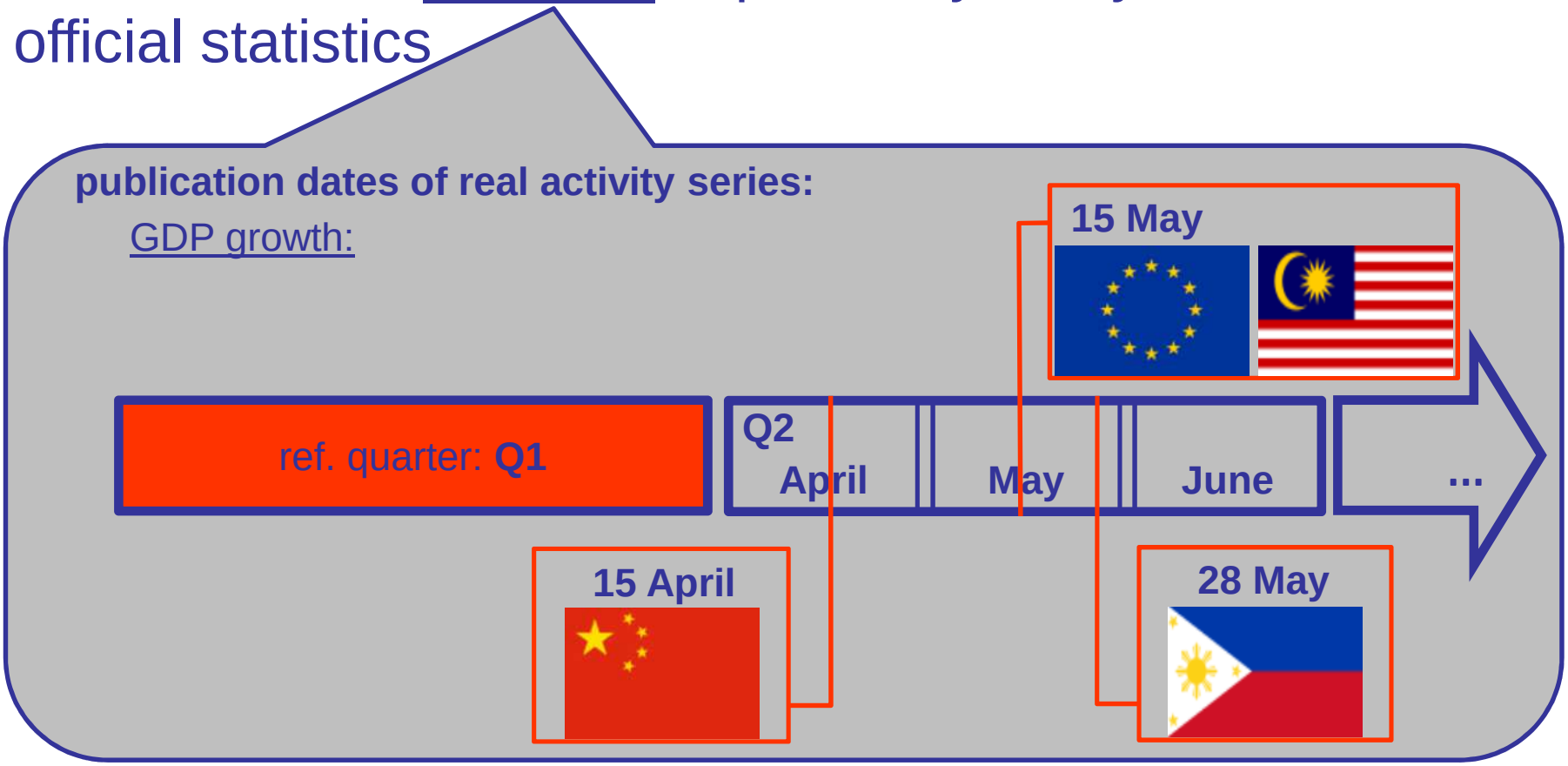
Bridge Models for Now- and Short-term Forecasting Euro-area GDP Growth and Evaluation of Forecasts

Andreas Reuter
Business and consumer surveys and short-term forecast (DG ECFIN A4.2)

International Workshop on Short-term Statistics, 18-20 May 2015, Beijing

Why do we need (short-term) forecasting models?

…to overcome **time-lag** imposed by delayed release of official statistics



Main types of (short-term) forecasting models

"classical" BMs

- using few (ca. 2-5), hand-selected predictor variables

pooling models

- running many (50 or more) bi-variate BMs
- averaging their forecasts

"factor-based" BMs

- creating artificial series (=factors), summarising common variation in large data-set (50 or more variables)
- using the series as predictors in a "classical" BM.

BRIDGE MODELS (BM)

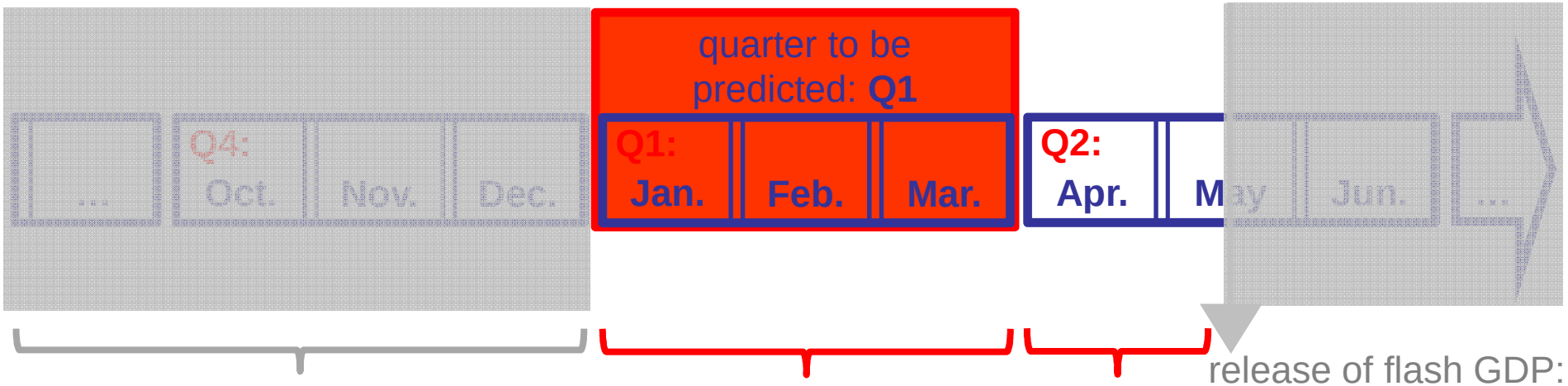
- use timely-available indicators to forecast (not yet published) target variable
- since timely indicators are monthly, bridge models put into relation **(they "bridge" between)** monthly predictors and a quarterly target variable (GDP)

quarterly averaging of a given predictor's monthly readings *(if incomplete: first forecasting missing monthly observations of predictor)*

= *models operated by EU Commission*

European Commission

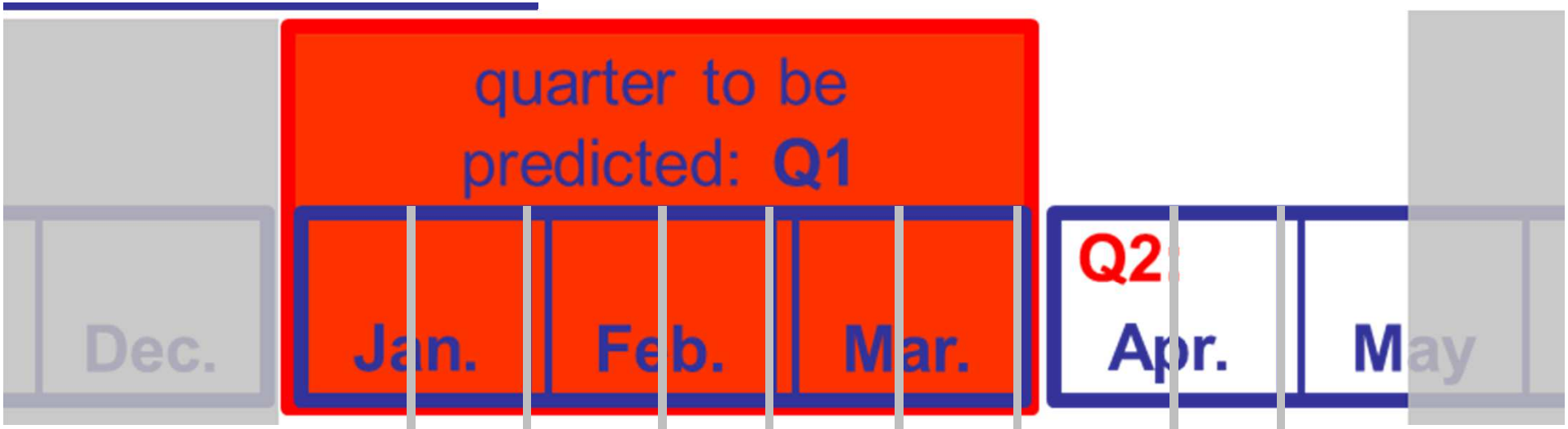
When to conduct nowcasts?



genuine forecast:
predicting a quarter t+1, while being in t

"nowcast":
predicting a quarter t, while being in quarter t

"backcast":
predicting a quarter t, while being in quarter t+1



"Nowcasts should be conducted whenever important, nowcast-relevant data is released!"

depends on predictor variables used:

financial series:

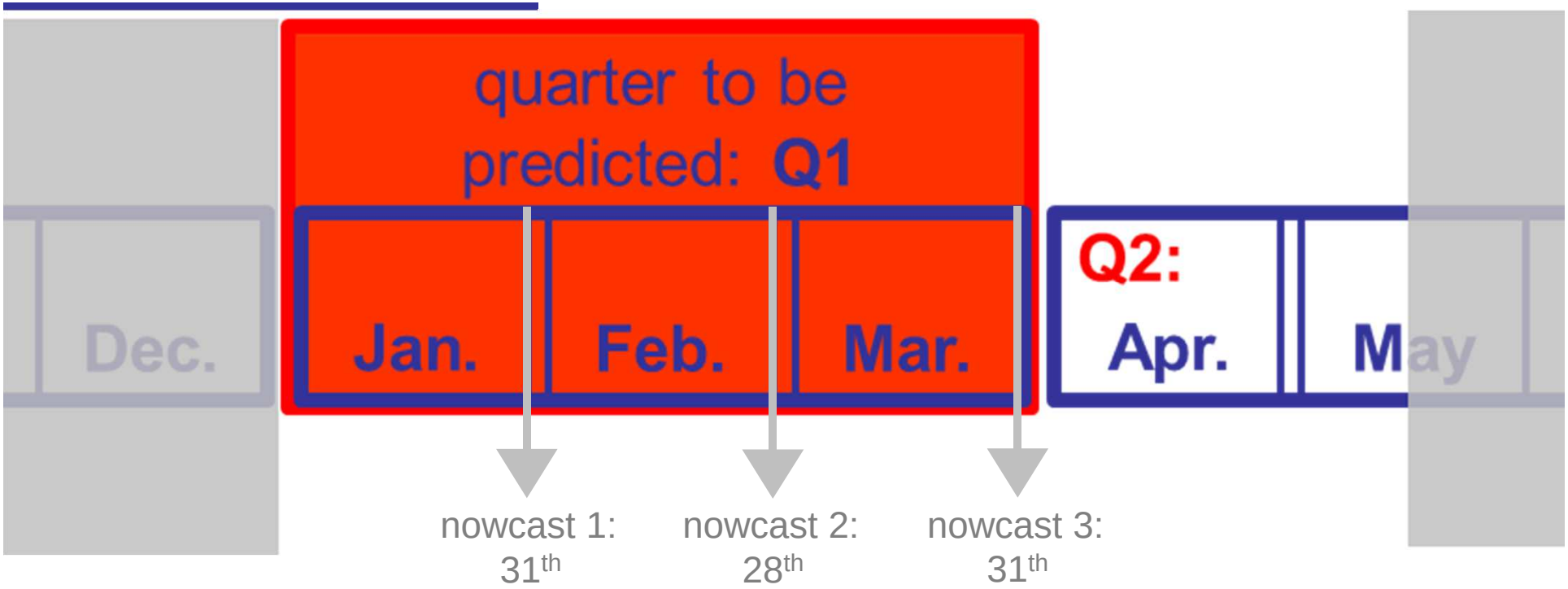
- stock market indices
- commodity prices
- interest rates
- etc.

- daily "release"
- convertible to monthly frequency (averaging)
- average on basis of first 15 days of a month is quite reliable

Business & consumer surveys:

real activity series:

European Commission

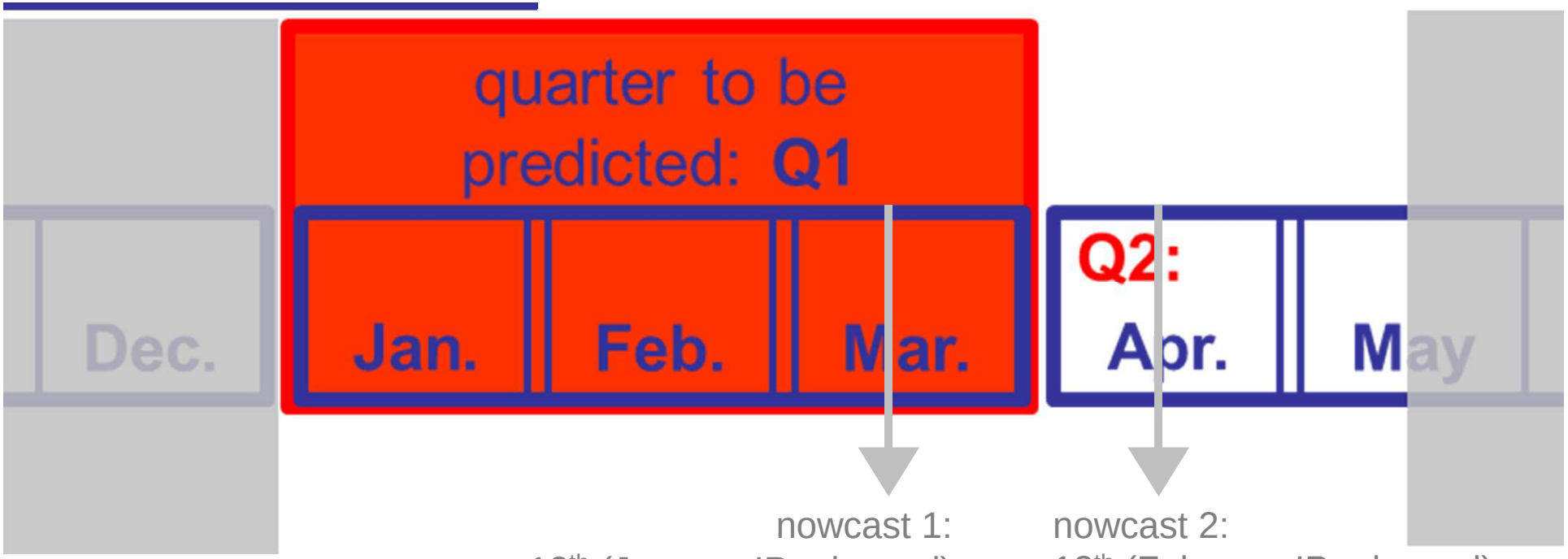


Business & consumer surveys:

- EU harmonised BCS programme (ESI, industrial confidence, etc.)
- Markit Economics' Purchasing Managers Index
- etc.

- monthly release
- available at end of reference month

European Commission



nowcast 1:
12th (January IP released)

nowcast 2:
12th (February IP released)

real activity series:

- industrial production index
- retail sales
- etc.

- mostly: monthly frequency
- available <u>after</u> reference month
  e.g. industrial production (IP) 42 days

The EU's nowcasting models

a) the linear model

history:

- launched in early 2009
- should be easy to administer >>"classical" bridge model
- should rely on survey data since:
  - real activity data published quite late
  - financial data lowly correlated with GDP
- testing predictive power of survey variables with broadest scope (i.e. sectoral confidence indicators, ESI)

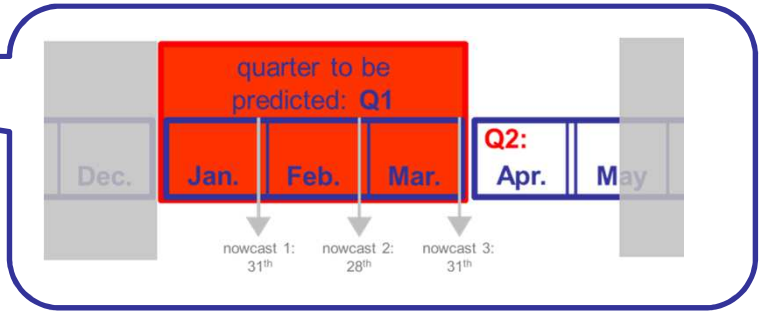
only industry and services confidence significant

industry "classically" contributing most to GDP variation

services being largest sector of economy

nowcasting dates:

- end of month 1, end of month 2, etc.



equation:

$$GDPqoq_t = c + \beta_1 INDU_t + \beta_2 INDU_{t-1} + \beta_3 SERV_{t-1} + u_t$$

in-sample period:

- period: 1996q1 to q_{t-1}
- predictors considered in terms of quarterly averages

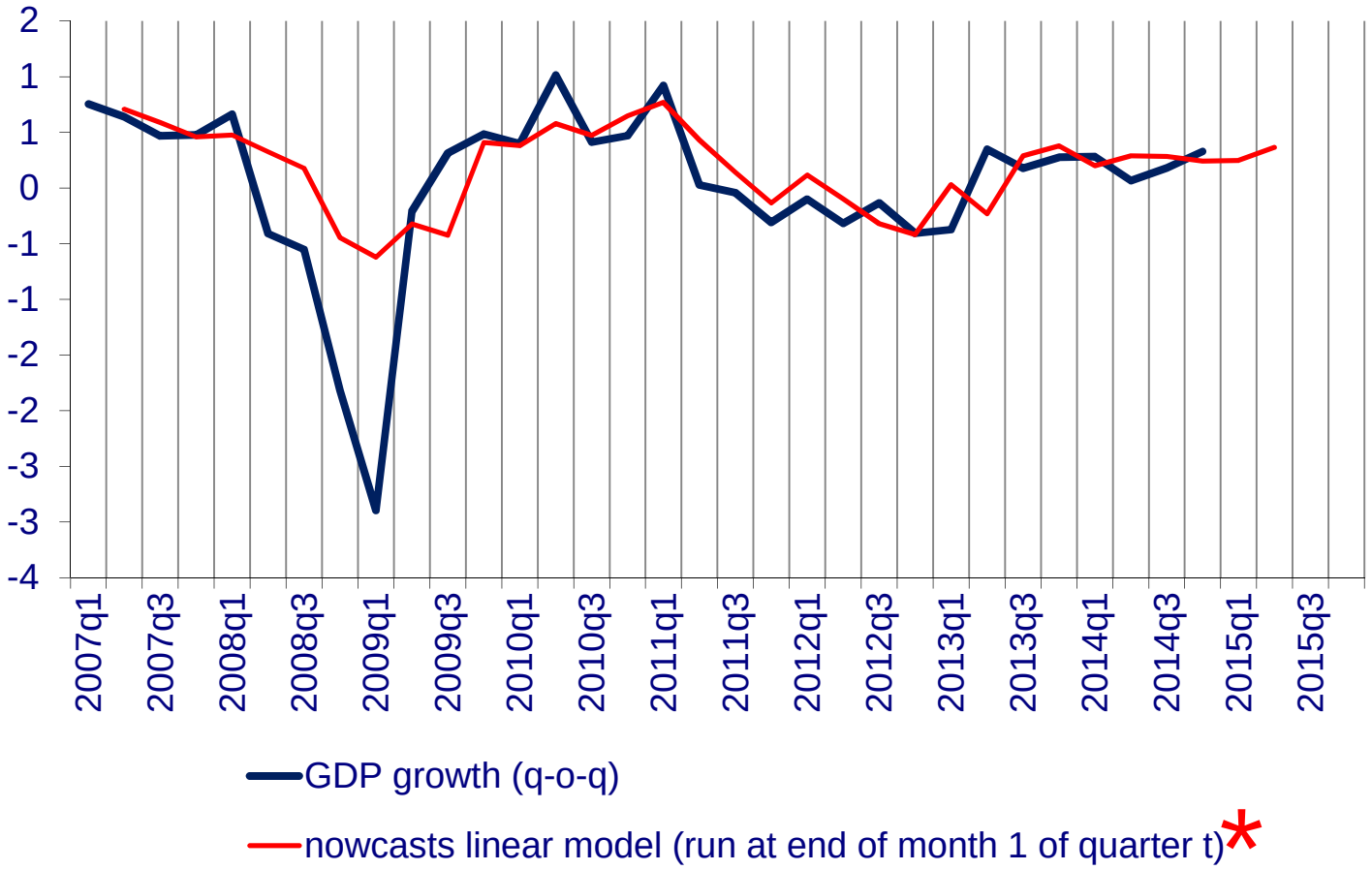
out-of-sample period:

- period: q_t
- predictors considered in terms of quarterly averages ***of available observations in q_t***

examples:

- nowcast of q1, while being at end of January: $INDU_{q1} = INDU_{January}$
- nowcast of q1, while being at end of February: $INDU_{q1} = avg.\ (INDU_{January}, INDU_{February})$
- etc.

Historical nowcasts:



* Nowcasts for period 2007q2 to 2009q3, based on pseudo real-time estimation.

b) the non-linear model

history:

- launched in 2010
- should remedy deficiencies of linear model in grasping sharpness of 2009-downturn

equation:

- assumption: relationship between survey data and GDP usually linear, but getting non-linear in case of extremely positive / negative GDP growth rates

fictitious example:

- increase in confidence indicator from +30 to +35 associated with 0.5% growth
- increase in confidence indicator from +35 to +40 associated with 1.0% growth

solution: rendering linear model quadratic

$$GDPqoq_t = c + \beta_1 INDU_t^2 + \beta_2 INDU_{t-1}^2 + \beta_3 SERV_{t-1}^2 + u_t$$

problem:

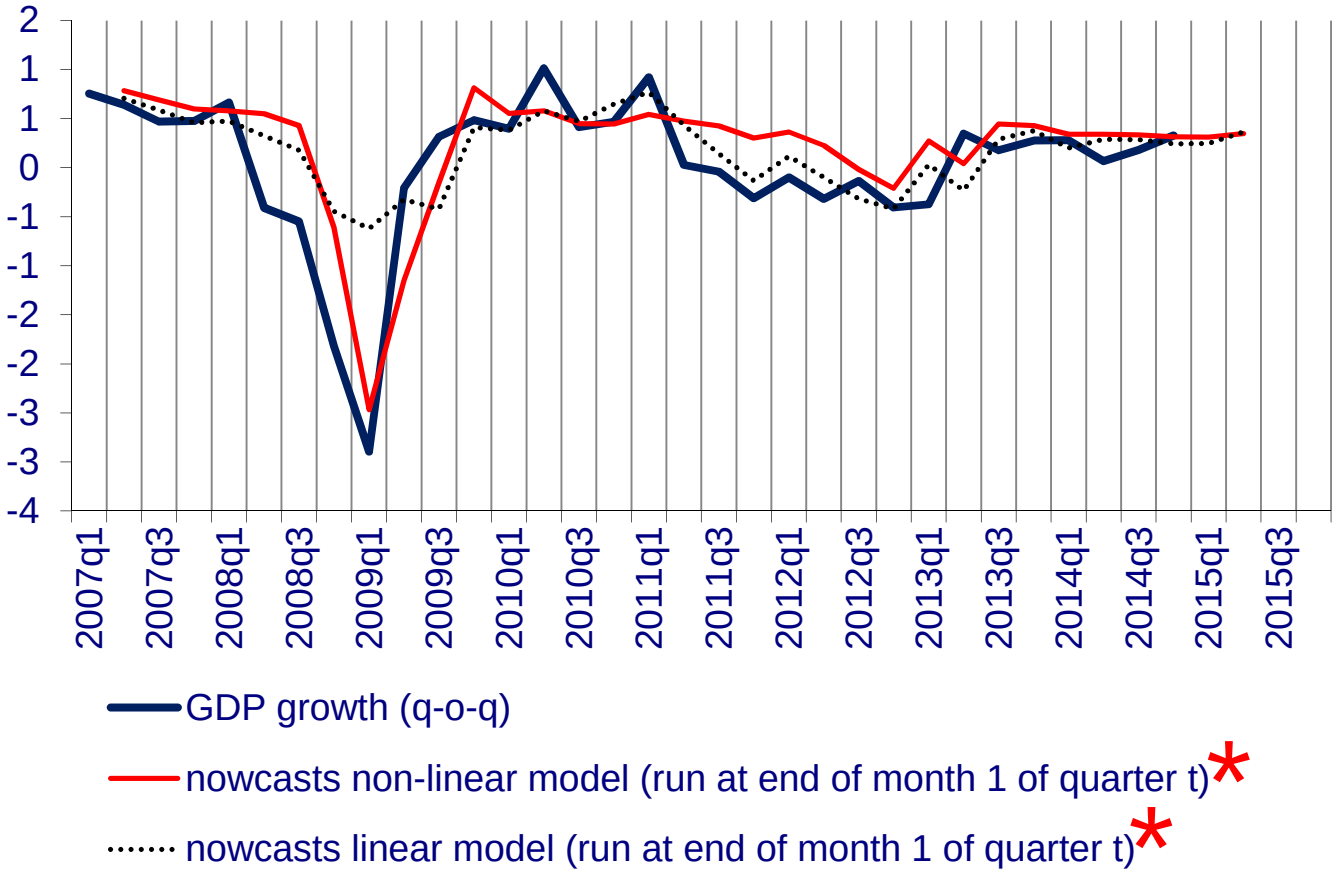
- confidence indicators can take on negative values
- quadratics would render all readings positive

$$GDPqoq_t = c + \beta_1 INDU_t * \beta_4 |INDU_t| + \beta_2 INDU_{t-1} * \beta_5 |INDU_{t-1}| + \beta_3 SERV_{t-1} * \beta_6 |SERV_{t-1}| + u_t$$

nowcasting dates in-/ out-of-sample period:

- identical to linear model

historical nowcasts:



*Nowcasts for period 2007q2 to 2009q3, based on pseudo real-time estimation.

c) factor-based bridge model

history:

- launched in 2015
- advantages:

examples:

- stock market indices
- commodity prices
- interest rates
- etc.

- daily "release"
- convertible to monthly frequency (averaging)
- average on basis of first 15 days of a month is quite reliable

(unexp vance

- also "me variables may be included

- **financial market variables** can be included
- factor-based bridge model can be run/updated every 15 days

nowcasting dates:

- end of month 1, **mid of month 2**, end of month 2, etc.

factor-based bridge modelling explained:

assumption:

- hidden (latent) structure in the vast data-set
- main variation patterns observed in a large data-set of potentially relevant variables can be summarised in a few artificially created variables (=factors)

constructing "factors":

- new variable (=factor 1) is built as linear combination of all series.
- constructed such that it extracts maximum variance from the variables in the data-set.
- fitted values (= "factorscores") are saved; they are the actual "factor"
- new factor 2 is built.
  - i) must be uncorrelated to factor 1
  - ii) extracts maximum remaining variance of the series
- … procedure repeated several times

generating nowcasting equation:

- plugging all factors of chosen factor-set into regression explaining GDP
- applying general-to-specific procedure, deleting factors insignificant at 5% level

in the case of our model:

- procedure repeated 6 times, instructing programme to extract:
  - max.6 factors (factor-set 1)
  - max.5 factors (factor-set 2)
  - etc.
- regressing GDP (separately) on factor-set 1, factor-set 2, etc.
- identify equation having lowest AIC value and continue working with the corresponding factor-set

Procedure repeated before every nowcast: factors can change over time + amount of retained factors can change from nowcast to nowcast

equation:

$$GDPqoq_t = c + \beta_1 FACTOR\ 1_t + \beta_2 FACTOR\ 2_{t-1} + \ldots + u_t$$

in-sample period:

- period: 1996q1 to q_{t-1}
- predictors considered in terms of quarterly averages

out-of-sample period:

- period: q_t
- factors have quarterly frequency (i.e. no necessity of averaging monthly observations across quarter)

the data-set used:

- comprises 29 variables*

financial variables:

- Dow Jones 65 composite average (US)
- Dow Jones 30 composite average (US)
- HWWI commodity price index (worldwide)

proxy of global environment

survey variables:

- ESI + experimental ESI
- sectoral confidence indicators
- industry survey questions:
  - past production
  - production expectations
  - past orders (quarterly)
- services survey questions:
  - past business situation
  - past demand
  - demand expectations
- retail trade survey questions:
  - Expected business activity
  - past demand
  - demand expectations
- construction survey questions:
  - past building activity
  - employment expectations
- consumer survey questions:
  - future financial position
  - expected econ. situation
- Ifo world economic survey (quarterly)

real activity variables:

- unemployment rate
- industrial production index (total)
- industrial production index (manufacturing)
- industrial production index (basic metals)
- industrial production index (intermediate goods)
- industrial production index (capital goods)
- exports (intra-euro area)
- imports (intra-euro area)

rationale:

- industry contributing most to GDP volatility
- early availability of IP (just 42 publication delay)

*Initial data-set comprised ca. 100 variables, but resulted in worse nowcasting performance.

special "bridging" technique for out-of-sample period – the blocking approach:

- in linear / non-linear model: predictors considered in terms of quarterly averages ***of available observations in q_t*** → e.g. forecasting Q1 at end of February: using average of variable's readings from January & February
- blocking approach: creates 3 quarterly series from a given monthly variable

	variable x
Jan. 2014	0.10
Feb. 2014	0.35
Mar. 2014	0.70
Apr. 2014	0.60
May 2014	0.33
Jun. 2014	0.27
Jul. 2014	0.31
Aug. 2014	0.87
Sep. 2014	0.91

→

	quarterly variable x 1
2014q1	0.10
2014q2	0.60
2014q3	0.31

	quarterly variable x 2
2014q1	0.35
2014q2	0.33
2014q3	0.87

	quarterly variable x 3
2014q1	0.70
2014q2	0.27
2014q3	0.91

- variables featuring value for nowcast quarter$_t$ are included coincidentally
- other variables are included in terms of their lag

nowcast-evaluation: pseudo out-of-sample exercise

- nowcast has no/little value if reliability of model is unknown

Quality assessment: What would model have predicted in the past? Was it right?

in an ideal world…

- select "evaluation horizon" (years for

in a practitioner's world…

mean absolute error (MAE):

$$\frac{\sum \left| \text{actual GDP growth} - \text{predicted GDP growth} \right|}{\text{number of nowcasts conducted}}$$

root mean squared error (RMSE):

$$\left(\frac{\sum \left(\text{actual GDP growth} - \text{predicted GDP growth} \right)^2}{\text{number of nowcasts conducted}} \right)^{0.5}$$

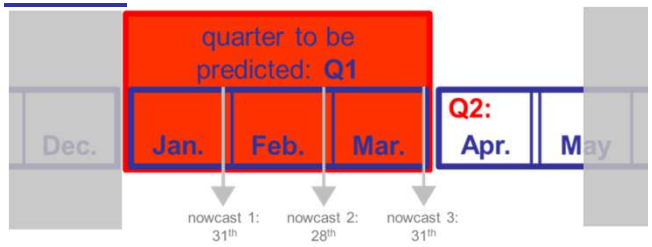
- conducting forecasts and comparing them to observed GDP

  "average" errors:
  - mean absolute error (MAE)
  - root mean squared error (RMSE)

determining MAE/RMSE from pseudo out-of-sample exercise

date:	we estimate:			date:	we estimate:		
15th Jan.	$Q4_f$	$Q1_f$		15th Aug.	$Q3_f$		
30th Jan.	$Q4_f$	$Q1_f$		30th Aug.	$Q3_f$		
15th Feb.		$Q1_f$		15th Sep.	$Q3_f$		
28th Feb.		$Q1_f$		30th Sep.	$Q3_f$		
15th Mar.		$Q1_f$		15th Oct.	$Q3_f$		
30th Mar.		$Q1_f$		30th Oct.	$Q3_f$		
15th Apr.		$Q1_f$	$Q2_f$	15th Nov.		$Q4_f$	
30th Apr.		$Q1_f$	$Q2_f$	30th Nov.			
15th May			Q2.	15th Dec.		$Q4_f$	
30th May				30th Dec		$Q4_f$	
15th Jun.				15th Jan.		$Q4_f$	
30th Jun.				30th Jan.		$Q4_f$	$Q1_f$
15th Jul.			f	15th Feb.			$Q1_f$
30th Jul.			$Q2_f$ $Q3_f$	30th Feb.			$Q1_f$
				15th Mar.			$Q1_f$
				…			…

simulating performance of linear model:



"tailored" data-set contains:

- IPI, etc. until May
- surveys until July
- …

"tailored" data-set contains:

- IPI, etc. until September
- surveys until October
- …

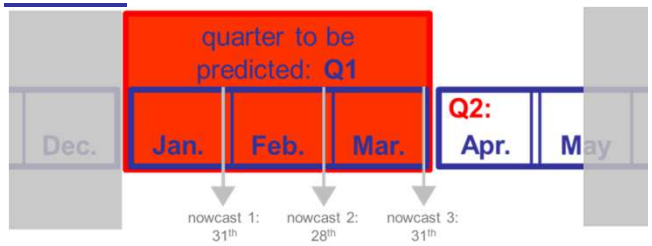
European Commission

determining MAE/RMSE from pseudo out-of-sample exercise

date:	we estimate:			date:	we estimate:		
15th Jan.	$Q4_f$	$Q1_f$		15th Aug.	$Q3_f$		
30th Jan.	$Q4_f$	$Q1_f$		30th Aug.	$Q3_f$		
15th Feb.		$Q1_f$		15th Sep.	$Q3_f$		
28th Feb.		$Q1_f$		Sep.	$Q3_f$		
15th Mar.		$Q1_f$		Oct.	$Q3_f$		
30th Mar.		$Q1_f$		Oct.	$Q3_f$		
15th Apr.		$Q1_f$	$Q2_f$	15th Nov.		$Q4_f$	
30th Apr.		$Q1_f$	$Q2_f$	30th Nov.		$Q4_f$	
15th May			$Q2_f$	15th Dec.		$Q4_f$	
30th May			$Q2_f$	30th Dec		$Q4_f$	
15th Jun.			$Q2_f$	15th Jan.		$Q4_f$	$Q1_f$
30th Jun.			$Q2_f$	30th Jan.		$Q4_f$	$Q1_f$
15th Jul.			$Q2_f$ $Q3_f$	15th Feb.			$Q1_f$
30th Jul.			$Q2_f$ $Q3_f$	30th Feb.			$Q1_f$
				15th Mar.			$Q1_f$
				…			…

model performance in month 2 of the quarter:
calculation of RMSE/MAE

simulating performance of linear model:



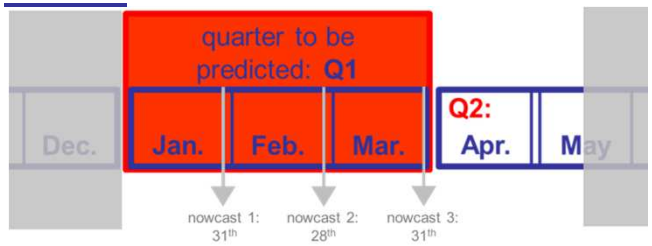
European Commission

determining MAE/RMSE from pseudo out-of-sample exercise

date:	we estimate:			date:	we estimate:		
15th Jan.	$Q4_f$	$Q1_f$		15th Aug.	$Q3_f$		
30th Jan.	$Q4_f$	$Q1_f$		30th Aug.	$Q3_f$		
15th Feb.		$Q1_f$		15th Sep.	$Q3_f$		
28th Feb.		$Q1_f$		30th Sep.	$Q3_f$		
15th Mar.		$Q1_f$		15th Oct.	$Q3_f$		
30th Mar.		$Q1_f$		30th Oct.	$Q3_f$		
15th Apr.		$Q1_f$	$Q2_f$	15th Nov.		$Q4_f$	
30th Apr.		$Q1_f$	$Q2_f$	30th Nov.		$Q4_f$	
15th May			$Q2_f$	15th Dec.		$Q4_f$	
30th May			$Q2_f$	30th Dec		$Q4_f$	
15th Jun.			$Q2_f$	15th Jan.		$Q4_f$	$Q1_f$
30th Jun.			$Q2_f$	30th Jan.		$Q4_f$	$Q1_f$
15th Jul.			$Q2_f$ $Q3_f$	15th Feb.			$Q1_f$
30th Jul.			$Q2_f$ $Q3_f$	30th Feb.			$Q1_f$
				15th Mar.			$Q1_f$
				…			…

model performance in month 3 of the quarter: calculation of RMSE/MAE

simulating performance of linear model:



European Commission

performance of EU nowcasting models:

out-of-sample period: 2010q1-2014q3			auto-regressive model	linear model	non-linear model	factor model
quarter $_t$	month 1	end	0.37	0.26	0.35	0.29
	month 2	mid				0.24
		end	0.34	0.27	0.35	0.24
	month 3	mid				0.23
		end	0.34	0.28	0.36	0.24
quarter $_{t+1}$	month 1	mid				0.23
		end				0.24
	month 2	mid				0.21

Observations:

- all models (except for non-linear model) are always better than AR benchmark
- linear model is best one at end of month 1
- factor-model is best one at end of months 2 and 3:
  - it can achieve this good performance even 15 days before the end of the month

out-of-sample period: 2008q1-2009q4			auto-regressive model	linear model	non-linear model	factor model
quarter $_t$	month 1	end	2.10	1.06	0.78	1.25
	month 2	mid				1.03
		end	1.43	0.88	0.60	1.05
	month 3	mid				0.69
		end	1.43	0.78	0.60	0.61
quarter $_{t+1}$	month 1	mid				0.50
		end				0.50
	month 2	mid				0.47

Observations:

- all models perform significantly worse in the crisis-period
- non-linear model is now the best-performing one in months 1 and 2
- in month 3, non-linear and factor model perform equally well
- in months 2 and 3, factor model performs particularly badly:
  - probably due to inclusion of lagged terms (helpful in smooth times, deterimental in volatile times)

Thanks for your attention!