**International Symposium on Forecasting**

June 2008
Nice

# Business Cycle Indicators in VARs: A Quarterly Forecasting Model of the Italian Economy

**Carmine Pappalardo**
**Ciro Rapacciuolo**
**Anna Ruocco**

# Aim

To build a **small-scale econometric model** for performing short-term forecasts of the Italian economy. The model should provide forecast of quarterly GDP and its components from the demand side.

GDP forecast is obtained to **aggregating predictions** of the various components. A supply-side GDP prediction is also obtained, supplementing that from the demand side.

Use a set of **quantitative and qualitative** short-run indicators to get predictions for each variable in a VAR (**Vector Autoregression**) framework.

The model focuses on **real developments** in the economy, abstracting from prices, financial and monetary developments. It is not able to perform simulations for policy purposes.

# Outline

- **Review on main Bridge-Equations approaches**
- **Modelling strategy**
- **Model structure**
- **Empirical results** (forecast evaluation, encompassing properties, role to NA revisions)
- **Conclusions**

# Bridge-equation approaches

# Bridge-equation models(1)

Official releases of quarterly national accounts are published with some delay. Italian NSI publishes the first official estimate of real GDP about 8 weeks after the end of the reference quarter. This delay stands in quite some contrast to the need for **timely and reliable** information on the state of the economy.

To infer the most recent developments on economic activity, forecasters pay attention to **short-run information** from related monthly indicators: *hard* (such as industrial production and retail trade data), *soft* (business surveys), financial variables, composite indicators.

Incorporating monthly indicators into forecasting models (**bridge**) may lead to considerable improvements in current-quarter predictions (*now-cast*). Sedillot and Pain (2003) provide a survey on indicator models of real GDP growth adopted in selected OECD countries.

# Single-equation models(2)

**SE models** are typically derived from unrestricted Autoregressive Distributed Lag models (ARDL[$p$,$q$]). Real GDP growth is regressed on survey data or other monthly indicators aggregated to a quarterly frequency. The unrestricted model has the form:

$$\rho(L)\Delta y_t = \sum_j \delta_j(L) x_{j,t} + \varepsilon_t$$

Currently used to provide a forecast for growth in the **current quarter** (*nowcast*) when indicators are available but GDP is not (**conditional forecast**). A conditional forecast would be expected to produce a more accurate outcome than unconditional models.

**Drawbacks**: Reduced length of the forecast horizon (up to 1 step ahead). Solution as in Grassman and Keyreman, to sequentially lag available information to increase the forecast horizon.

# VAR models

Vector autoregression (VAR) models combine variable indicators and GDP growth. Used to **circumvent the delays** from waiting for new quarterly data on the indicator variables. The VAR model can be written as follows:

$$y_t = \nu + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \varepsilon_t \tag{1}$$

A single VAR model can be used to provide multi-period forecasts when data for both GDP and the indicator variables are not available (*unconditional forecast*) whilst the SE bridge model can be run only for one quarter at a time (as data for the indicator become available).

**Drawbacks**: poor predictive accuracy since do not include all recent available information for the prediction on the current quarter (*nowcast*).

**Solutions**: to estimate conditional VAR models (BVAR) can be used to provide a forecast when (some) indicators are available but GDP is not (Robertson and Tallman, 1999).

# Combining monthly and quarterly data

Increasing studies address explicitly the means of **incorporating new monthly information** that becomes available within the quarter for which growth projections are being made. This methodology has been adopted to face the limits of both SE and VAR approaches.

- **single-equation**: The missing high frequency information for the current quarter (and, potentially, for the following quarter) is projected using **monthly auxiliary models** (univariate models, multivariate models with surveys). This approach allows us to obtain predictions based on incomplete monthly information.

  Studies which deal with the efficient treatment of new monthly information that becomes available within the quarter are in Rathjens and Robins (1993), Ingenito and Trehan (1996), Stark (2000). Zheng and Rossiter (2006) for Canada. Rünstler and Sédillot (2003) was the first research for the euro zone.

- **VAR framework**: Salazar and Weale (1999) present a technique for estimating a VAR on monthly data, making use of interpolated estimates of GDP and correcting for the impact of measurement error. Test whether the model estimated from the interpolated monthly data contains information absent from the quarterly VAR.

  Reynaud and Sherrer (2002) employ VAR model in the **block recursive** form. This representation allows to perform, for a given period, multiple forecasts (up to 5 for each quarter) based on additional information (*innovations*). To compare those sequential forecast provides a measure of the new additional short-term information.

  Camba-Mendez *et alii* (2001) propose an **Automatic Leading Indicator** (ALI) model: VAR augmented with factors extracted from a dynamic factor model to summarise the information content of a pooling of variables. Results show that the forecasting performance of ALI model is significantly better than that of traditional model selection criteria with VARs. This study is carried out using quaterly data for France, Germany, Italy and the United Kingdom.

# Models for GDP and Components

**Empirical studies** concerning the short-term forecasting of GDP and its components: Trevor, Thorp (1988; Australia); Baffigi *et al.* (2004) for the euro zone; Parigi and Schlitzer (1995), Bovi *et al.* (2000; BEEF) for Italy; Irac and Sédillot (2002; OPTIM) for France.

- The majority of these models are based on **SE** including *soft* and financial **indicators**, other than the quantitative ones. Also include factors from a pooling of variables (OPTIM).

- **Reduced forecast horizon** because based on quarterly aggregates of monthly indicators.

- To get multi-step forecast, they **require indicators forecasting**.

# This model

# VAR Approach

**Basic idea** is to get short-run forecasts of quarterly GDP and its components by means of a set of short-run indicators in a VAR context.

The approach relies on **unconditional** VAR models. This allow us to face the drawback of the limited forecasting horizon performing multi-step forecasts (up to 4 steps ahead).

**Do not include financial** variables for prediction nor factors from dynamic factor models (as in Camba-Mendez *et al.* and in Irac and Sèdillot in OPTIM model for France.)

(One of) the main limit(s) is the poor **predictive accuracy for the current** quarter, since the *nowcast* is obtained do not using the (even partial) available information.

To account for this drawback and evaluate the lack of predictive power of the *nowcast*, we define a **set of conditional single equation** models for the current quarter, so to get conditional forecasts. The bridge models are specified for the main NA variables of interest.

We evaluate the **predictive accuracy** of VAR models versus SE and univariate benchmark models in terms of both *ex-post* in-sample forecast evaluation and encompassing.

# Unconditional versus conditional

As predictions for each NA variable are obtained through small-scale VARs, this implies an additional implication in both estimation and forecasting: the **strong exogeneity condition**.

For *in-sample* statistical inference, only weak exogeneity is required to support the conditioning on the not-modeled variables. For *multi-step* forecasting, **strong** exogeneity is essential (Engle, Hendry and Richard, 1993).

One aspect of strong exogeneity is **Granger not-causality**. It is a necessary though not sufficient condition for strong exogeneity, such that rejection would reject the validity of the conditioning for forecasting.

Distortionary effects of **invalid conditioning** on multi-step forecast performance (Clemens, Hendry, 1999). One way round the difficulties entailed by conditioning is to require that all the variables are *endogenized* in any forecasting exercise using **VAR models**.

# Modelling strategy

# Indicators

Availability of indicator variables is limited by the necessity of having timely information closely related to the phenomena to predict. Some of them are **fairly standard** and include so called *hard* data.

**Soft indicators** are mainly derived by the ISAE surveys on manufacturing, construction and households sectors. Indicators record economic agents assessment on both the current development of economic variables and their short-term expectations.

**Do not use** financial variables and factor indicators from dynamic factor models. In our statistical approaches do not combine quarterly and monthly frequency data (as in Salazar and Weale, 1999).

**Selection crteria**: Variables are selected to satisfy high correlation and conventional diagnostic tests in quarterly models. The **leading properties** plays an important role in the single-equation framework.

# Statistical Treatment

Before being used in the model, all the predictors are treated so to be **consistent with NA aggregates** (chain-linked, seasonally and working day adjusted).

**First**, potential outliers are detected in raw data by using the Tramo-Seats procedure (Gómez and Maravall, 1997). **Secondly**, WD adjustment only applies to *hard* indicators available at monthly frequency. **Finally**, seasonal adjustment applies to quarterly averages of (eventually wd adjusted) monthly series.

This procedure also applies to **survey indicators** (except wd adjustment). A wide literature stressed the fact that seasonality in business survey data should be considered as spurious. And also investigate the potentially stationary nature of the seasonal component of qualitative variables (Franses, 1996, Proietti, 2000).

Following a consolidated practice, quantitative variables are expressed in logs. All variables are **differenced** to get stationarity.

# Functional Forms(1)

**VAR** models are specified using differenced (stationary) time series and are of the form:

$$\Delta y_t = \alpha + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \ldots + \beta_q \Delta y_{t-q} + \phi \mathrm{det}_t + \rho \mathrm{IC}_t + \varepsilon_t \quad (2)$$

where $\Delta y_t$ is a vector of variables, $\Delta = (1 - L)$ (DVAR), $\mathrm{det}_t$ is an additional deterministic component (additive dummies) and $\mathrm{IC}_t$ is the intercept correction term.

**Models do not include the long-run term**. Clements and Hendry (1995) show that, in general, this term does not significantly improve the multi-steps forecasts of the changes in the variables.

The **reduction process** is based on testing sequential joint restrictions on the parameters of the VAR equations until the null is rejected and, testing for correct specification.

**Differencing and intercept corrections** help to offset some forms of forecast failure, yet have no causal basis, but essentially due to not-modeled shifts in deterministic factors.

**Intercept corrections** (ICs) are data-based shift factors that change the mean of a model over the forecast period relative to its in-sample value.

Since ICs can eliminate recent deterministic shifts, they are a potential remedy for breaks that have occurred; a condition in which the **static** one-step-ahead forecasts systematically overestimate (underestimate) the realized values.

For multi-step forecasting, we find appropriate to include an adjustment reflecting the **average static prediction error** of the recent past (Busetti, 2001).

# Functional forms(2)

**Single-equations** specification are derived from an unrestricted Autoregressive Distributed Lag (ARDL[$p$,$q$]) model. NA variables are regressed on their past, on current and past values of related indicators aggregated to a quarterly frequency:

$$\Delta\mathbf{y}_t = \alpha + \sum_{i=0}^{p} \gamma_i \Delta\mathbf{y}_{t-i} + \sum_{j=1}^{q} \beta_j \Delta\mathbf{x}_{t-j} + \delta\mathbf{d}_t + \varepsilon_t$$

where $\Delta = (1 - L)$ is the filtering applied to the log of dependent variable and to the exogenous variables specific to each model, $\mathbf{d}_t$ are the deterministic components, $\varepsilon_t$ the idiosyncratic error term.

Reduced models take the form of **conditional** bridge equation models.

**Model reduction**: *i*) each general unrestricted model includes up to the 4th lag of the independent and dependent variables; *ii*) the *General to Specific* approach for model reduction is performed running Pc−Gets (Krolzig and Hendry, 2001).

We **do not make prediction** of indicators to compensate for the lack of monthly information in the current period.

SE models can be effectively used at the **end of the current quarter**, when the information on indicators is fully available but not on NA variables.

We use these models and, the univariate ones, essentially to **improve the evaluation** of 1-step-ahead forecast (*nowcast*) coming from VARs.

# Three main criteria

**Stationarity**. The parameters in the VAR must be such that the system is stationary and not explosive. This amounts to check that the eigenvalues of the autocorrelation matrix in a VAR(1) representation of the system are all smaller than one.

**Specification tests**. The error term of *in-sample* estimates must satisfy the usual properties (normality, no serial correlation, parameter stability, functional form).

**Granger causality**. Indicators included in the VAR must not be exogenous with respect to the variable to predict. Otherwise, VAR may reduce to a SE model (as in Baffigi *et al.* (2005)).

If the null of no-Granger causality cannot be rejected, we conclude that **indicators do not have power** to explain the growth of the reference variable.

# Model Structure

# Model Structure(1)

The goal is to predict (up to 4-steps-ahead) the development (*q-on-q*) of the **main NA aggregates**, chain-linked (2000=100), seasonally and working-day adjusted.

To account for the **lack of additivity**, we follow a two steps procedure. **First**, the forecast of each component is performed using VAR models. **Secondly**, the prediction of the main aggregate is obtained trough an auxiliary regression (*single-equation*) including the first step forecasts as regressors.

**Changes in stocks** are no more provided (at least for Italy). But its forecast plays an important to get a more accurate GDP predictions. The **difference** between the *official* GDP estimates and the summing up its components consists of two main aggregates: the statistical discrepancy and, the *true* series of changes in stocks.

**We model this difference**. A single-equation with indicators (related to the changes in stocks) and MA errors is estimated. We assume that the MA component captures the statistical discrepancy. An estimates of changes in stocks is obtained as the **structural fit** of the equation (net of the MA specification).

The model provides two GDP forecasts. From the **demand side**, it is obtained as the result of the prediction of its main components (**indirect forecast**). From the **supply side**, GDP is predicted **directly** using sectoral indicators.

Overall, the model consits of 12 VARs, 8 single-equation models with indicators, 1 univariate equation (Government Consumption).

## Model Structure

| Aggregates | VAR | SE | ARIMA |
|---|---|---|---|
| Consumption | 4 | 3 | 1 |
| Investment | 3 | 1 | |
| Exports | 2 | 1 | |
| Imports | 2 | 1 | |
| Stocks | | 1 | |
| GDP | 1 | 1 | |

# Consumption

| | Aggregates | Models | Lags | Indicators |
|---|---|---|---|---|
| $C_1$ | Not-Durables | VAR | 2 | Confidence climate |
| $C_2$ | Durables | VAR | 1 | Vehicles reg., conf. climate, orders (BS), IP cons. goods |
| $C_3$ | Semi-Durables | VAR | 3 | confidence climate |
| $C_4$ | Services | VAR | 2 | Vehicles reg., IP construction sector |
| CFTER | National | SE | | $C_1$, $C_2$, $C_3$, $C_4$ |
| CFAM | Private | SE | | CFTER |
| CCONS | Governmet | ARIMA | | |
| CFIN | Final | SE | | CFAM, CCONS |

# Investments

| | Aggregates | Models | Lags | Indicators |
|---|---|---|---|---|
| $I_1$ | Machinery and Equipment | VAR | 1 | Orders inv. goods (BS), IP inv. goods, business climate, US$/euro exch. rate |
| $I_2$ | Transport | VAR | 2 | IP, auto and comm. vehicles reg. |
| $I_3$ | Construction | VAR | 2 | IP construction sect. |
| INVFL | Gross Fixed | SE | | $I_1$, $I_2$, $I_3$ |

# Exports and Imports

| | Aggregates | Models | Lags | Indicators |
|---|---|---|---|---|
| Exports | | | | |
| $X_1$ | Goods | VAR | 1 | Nominal exch. rate, exports (vol.) |
| $X_2$ | Services | VAR | 2 | Nominal exch. rate, IP Germany and US |
| XBS | Exports | SE | | $X_1$, $X_2$ |
| Imports | | | | |
| $M_1$ | Goods | VAR | 2 | Domestic demand |
| $M_2$ | Services | VAR | 2 | Nominal exch. rate, imports of goods |
| MBS | Imports | SE | | $M_1$, $M_2$ |

# Changes in stocks

| | Aggregates | Models | Lags | Indicators |
|---|---|---|---|---|
| S | Stocks | SE | | Domestic demand, Imports of goods, IP$_{t-2}$ (investment goods), MA(4) |

# GDP

| | Aggregates | Models | Lags | Indicators |
|---|---|---|---|---|
| Demand side $GDP_d$ | Demand | SE | | CFIN, INVFL, XBS, MBS, S |
| Supply side $GDP_s$ | Supply | VAR | 2 | IP, IP construction sect, employment services sect. |

# Out-of-Sample Forecasting

# Empirical Results

- **Forecast evaluation**

- **Encompassing tests**

- **Pseudo real-time evaluation** (role of NA revisions)

# Forecast Evaluation

*Ex-post* in-sample forecast evaluation. We choose between a number of **forecasting schemes**. Estimated coefficients are fixed at their in-sample values for the whole *ex-post* forecast horizon; **updated recursively** as the period unfolds. In order to simulate the effect of successive estimations, we will also look at the results of **rolling estimates**.

The forecast evaluation exercise is carried out in the period [2005:q2-2007:q1]. For each model, a full set of **8 predictions** at several steps ahead ($h$=1,...,4) is obtained.

The analysis of forecast accuracy is carried out in terms of **RMSFE** and, using both the recursive and rolling schemes. The aim is to asses to what extent the smaller RMSFE could be attributed to sampling variability.

# GDP (RMSFE)

| (recursive) | current quarter | 1-step ahead | 2-steps ahead | 3-steps ahead |
|---|---|---|---|---|
| $GDP_d$ | 0.004 | 0.001 | 0.002 | 0.003 |
| $GDP_{d,stocks}$ | 0.005 | 0.001 | 0.001 | 0.001 |
| $GDP_s$ | 0.005 | 0.004 | 0.004 | 0.004 |
| ARIMA | 0.005 | 0.005 | 0.004 | 0.004 |
| SE | 0.003 | | | |
| $SE_{stocks}$ | 0.005 | | | |
| (rolling) | | | | |
| $GDP_d$ | 0.004 | 0.001 | 0.002 | 0.003 |
| $GDP_{d,stocks}$ | 0.006 | 0.001 | 0.001 | 0.002 |
| $GDP_s$ | 0.005 | 0.004 | 0.003 | 0.004 |
| ARIMA | 0.006 | 0.005 | 0.003 | 0.004 |
| SE | 0.003 | | | |
| $SE_{stocks}$ | 0.004 | | | |

# Main Results

**At 1-step-ahead**, conditional SE models outperform model forecasts (both from the demand and supply side).

As the forecast horizon increases (**multi-step context**), GDP model from the demand side (including inventories) performs significantly better than rival specifications.

The same applies in both recursive and rolling scheme for the main **GDP components** (except Consumption).

With regard to the **GDP components** and for $h>1$, the performance of forecasting models for Consumption and Imports do not significantly differ from the univariate benchmark. Models for Investments and Exports outperform the ARIMA benchmark (also at 1-step-ahead).

# Consumption (RMSFE)

| Consumption (recursive) | current quarter | 1-step ahead | 2-steps ahead | 3-steps ahead |
|---|---|---|---|---|
| VAR | 0.002 | 0.002 | 0.002 | 0.002 |
| ARIMA | 0.002 | 0.002 | 0.002 | 0.002 |
| SE | 0.002 | | | |
| **Consumption (rolling)** | | | | |
| VAR | 0.002 | 0.002 | 0.001 | 0.002 |
| ARIMA | 0.002 | 0.002 | 0.002 | 0.002 |
| SE | 0.002 | | | |

# Investments (RMSFE)

| Investment (recursive) | current quarter | 1-step ahead | 2-steps ahead | 3-steps ahead |
|---|---|---|---|---|
| VAR | 0.015 | 0.008 | 0.012 | 0.013 |
| ARIMA | 0.014 | 0.014 | 0.013 | 0.014 |
| SE | 0.008 | | | |
| Investment (rolling) | | | | |
| VAR | 0.015 | 0.009 | 0.012 | 0.013 |
| ARIMA | 0.015 | 0.014 | 0.014 | 0.012 |
| SE | 0.008 | | | |

# Exports (RMSFE)

| Exports (recursive) | current quarter | 1-step ahead | 2-steps ahead | 3-steps ahead |
|---|---|---|---|---|
| VAR | 0.018 | 0.013 | 0.009 | 0.011 |
| ARIMA | 0.020 | 0.019 | 0.019 | 0.019 |
| SE | 0.021 | | | |
| Exports (rolling) | | | | |
| VAR | 0.019 | 0.017 | 0.004 | 0.010 |
| ARIMA | 0.021 | 0.020 | 0.021 | 0.019 |
| SE | 0.021 | | | |

# Imports (RMSFE)

| Imports (recursive) | current quarter | 1-step ahead | 2-steps ahead | 3-steps ahead |
|---|---|---|---|---|
| VAR | 0.014 | 0.010 | 0.011 | 0.0010 |
| ARIMA | 0.011 | 0.010 | 0.009 | 0.011 |
| SE | 0.012 | | | |
| Imports (rolling) | | | | |
| VAR | 0.015 | 0.010 | 0.010 | 0.011 |
| ARIMA | 0.012 | 0.011 | 0.011 | 0.011 |
| SE | 0.013 | | | |

# Encompassing tests

**Basic question**: test whether the forecast *errors* of model $\mathbf{M}_A$ is explained by the forecasts from model $\mathbf{M}_B$. Under the null, the forecast errors are innovations, hence unpredictable by the competing model $\mathbf{M}_B$, which is encompassed by model $\mathbf{M}_A$.

Alternatively, if $\mathbf{M}_B$ is informative, errors from model $\mathbf{M}_A$ are predictable. To exclusively rely on $\mathbf{M}_A$ gets a loss of information.

We use the modified Diebold-Mariano test of equal predictive accuracy (Harvey, Leybourne, Newbold, 1998): if the forecast from $\mathbf{M}_A$ encompasses the forecast from model $\mathbf{M}_B$, then the covariance between $\mathrm{e}_{\mathrm{A}}$ and $(\mathrm{e}_{\mathrm{A}} - \mathrm{e}_{\mathrm{B}})$ is $\leq 0$. If the covariance is $> 0$ , the null of encompassing is rejected.

The test statistics is the following:

$$HLN = \frac{\bar{d}}{\sqrt{var(\bar{d})}}$$

where $d = e_{A,t}(e_{A,t} - e_{B,t})$ and $\bar{d} = n^{-}1 \sum_{t=1}^{n} d$. Under the null, HLN statistics has an asymptotic standard normal distribution.

# GDP - nowcast

| Recursive nowcast | DEM | SUP | SE | SE$_{sup}$ | ARIMA |
|---|---|---|---|---|---|
| DEM | | 0.394 (0.693) | 4.201 (0.000) | 8.653 (0.007) | −1.038 (0.299) |
| SUP | 3.353 (0.001) | | 3.801 (0.000) | 6.584 (0.000) | 1.872 (0.061) |
| SE | −0.954 (0.340) | 1.278 (0.201) | | 4.384 (0.000) | 1.573 (0.116) |
| SE$_{sup}$ | 1.460 (0.144) | 0.628 (0.529) | 1.855 (0.063) | | 0.784 (0.433) |
| ARIMA | 3.382 (0.001) | 3.239 (0.001) | 3.960 (0.000) | 5.639 (0.000) | |

(p-values in brackets)

# GDP - 3-steps-ahead

| Recursive | DEM | SUP | ARIMA |
|---|---|---|---|
| DEM | | 0.307 (0.759) | 0.059 (0.953) |
| SUP | 3.444 (0.001) | | 0.791 (0.429) |
| ARIMA | 3.678 (0.000) | 3.678 (0.000) | |

(p-values in brackets)

# Main Results

For the current quarter, the **informative content** of the predictions from the demand side encompasses that from the supply side; though GDP forecasts from both the demand and supply side show similar RMSFE. **Conditional SE** models encompass all the others and, SE$_{supply}$ encompasses SE.

VARs show a lack of predictive power for the *nowcast*. **Opportunity** to select several predictions for the current quarter to get a most accurate (efficient) GDP forecast.

Advantage of model combination seems also to emerge for **GDP components** (conditional SE models marginally encompasses all the others). Except Exports, for which Model forecasts encompasses all the others.

When $h>1$, GDP forecast from the **demand side encompasses** those based on supply indicators and the ARIMA benchmark.

This conclusion does not applies to **GDP components**. The encompassing tests are not conclusive to indicate the superior informative content of Model forecasts. It emerges the opportunity of forecast combining up to 1 step ahead. Model forecast encompasses the ARIMA benchmark only over longer forecasting horizons.

# Consumption

| Recursive nowcast | Model | SE | ARIMA |
|---|---|---|---|
| Model | | 2.552 (0.011) | 2.320 (0.020) |
| SE | 1.653 (0.098) | | 1.992 (0.046) |
| ARIMA | 4.008 (0.000) | 2.953 (0.003) | |

(p-values in brackets)

| steps ahead | 1 | 2 | 3 |
|---|---|---|---|
| Model vs. ARIMA | 2.013 (0.044) | 6.102 (0.211) | 0.788 (0.431) |
| ARIMA vs. Model | 3.611 (0.000) | 2.346 (0.019) | 4.161 (0.000) |

(p-values in brackets)

# Investments

| Recursive nowcast | Model | SE | ARIMA |
|---|---|---|---|
| Model | | 4.342 (0.000) | 2.857 (0.004) |
| SE | −0.546 (0.585) | | 3.787 (0.000) |
| ARIMA | 2.951 (0.003) | 4.368 (0.000) | |

(p-values in brackets)

| steps ahead | 1 | 2 | 3 |
|---|---|---|---|
| Model vs. ARIMA | −3.047 (0.002) | 3.035 (0.412) | 0.305 (0.761) |
| ARIMA vs. Model | 3.035 (0.002) | 2.368 (0.018) | 3.478 (0.001) |

(p-values in brackets)

# Exports

| Recursive nowcast | Model | SE | ARIMA |
|---|---|---|---|
| Model | | 0.881 (0.378) | 0.582 (0.561) |
| SE | 2.875 (0.004) | | 3.342 (0.001) |
| ARIMA | 2.050 (0.040) | 2.850 (0.004) | |

(p-values in brackets)

| steps ahead | 1 | 2 | 3 |
|---|---|---|---|
| Model vs. ARIMA | 1.227 (0.220) | −0.281 (0.779) | 0.519 (0.604) |
| ARIMA vs. Model | 4.687 (0.000) | 6.848 (0.000) | 4.612 (0.000) |

(p-values in brackets)

# Imports

| Recursive nowcast | Model | SE | ARIMA |
|---|---|---|---|
| Model | | 6.216 (0.000) | 6.570 (0.000) |
| SE | 1.679 (0.093) | | 3.591 (0.000) |
| ARIMA | −0.202 (0.840) | 0.305 (0.760) | |

(p-values in brackets)

| steps ahead | 1 | 2 | 3 |
|---|---|---|---|
| Model vs. ARIMA | 3.374 (0.001) | 3.933 (0.803) | 2.159 (0.314) |
| ARIMA vs. Model | 4.687 (0.000) | 0.249 (0.000) | 3.139 (0.002) |

(p-values in brackets)

# Short-run forecast and real-time data

NSIs often revise GDP data (and related components) because of **statistical and definitional changes**. **Statistical changes** stem from the availability of additional information and generally only concern the most recent quarters. **Definitional changes** (base year, methodological reforms such as changes in classification) are more pervasive, occur at discrete times, involves a retrospective change of the whole historical sample.

Our aim is to infer the **impact of NA data revisions** on the forecasting ability of our models. To this aim, we need: a real-time dataset representing the data availability at any given date in the past; a number of competing models.

**Wide literature**: Robertson, Tallman (1998), Evans (2005), Diron (2006), Busetti(2001), Golinelli, Parigi (2006), Giannone, Riechlin, Small (2006). We build the *real-time* dataset as in Golinelli-Parigi (2006).

We define each GDP series of the real-time dataset as $\mathbf{y}_t^{r,v}$ where $r = 1$, 2,...,T+$f$ is the release index; $v = 1$, 2,...,$f$ is the vintage index (f labels the latest available vintage); $t = 1$, 2, ...,T+$f$ is the period index.

In the **data matrix** rows are periods, columns are vintages (i.e. $\mathrm{y}^v$ is the GDP time series over the period 1 to T+$v$ published by the statistical agency in its issue $v$), and diagonals are releases (or *outturns* as in Golinelli-Parigi, 2006).

The sequence of observations $\mathbf{y}_{T+1}^{1,1}$, $\mathbf{y}_{T+2}^{1,2}$, $\mathbf{y}_{T+3}^{1,3}$,...,$\mathbf{y}_{T+f}^{1,f}$ is the **first release** of GDP data that have never been revised; the time series of the second release $\mathbf{y}^{2,v}$ contains data that have all been revised once, and so on.

# The real-time dataset

| | | vintages, $v$ | | | | |
|---|---|---|---|---|---|---|
| | period, $t$ | 1 | 2 | 3 | … | f |
| | 1 | $\mathbf{y}_1^{T+1,1}$ | $\mathbf{y}_1^{T+2,2}$ | $\mathbf{y}_1^{T+3,3}$ | … | $\mathbf{y}_1^{T+f,f}$ |
| | 2 | $\mathbf{y}_2^{T,1}$ | $\mathbf{y}_2^{T+1,2}$ | $\mathbf{y}_2^{T+2,3}$ | … | $\mathbf{y}_2^{T+f-1,f}$ |
| | 3 | $\mathbf{y}_3^{T-1,1}$ | $\mathbf{y}_3^{T,2}$ | $\mathbf{y}_3^{T+1,3}$ | … | $\mathbf{y}_3^{T+f-2,f}$ |
| | … | … | … | … | … | … |
| | T-1 | $\mathbf{y}_{T-1}^{3,1}$ | $\mathbf{y}_{T-1}^{4,2}$ | $\mathbf{y}_{T-1}^{5,3}$ | … | $\mathbf{y}_{T-1}^{f+2,f}$ |
| | T | $\mathbf{y}_T^{2,1}$ | $\mathbf{y}_T^{3,2}$ | $\mathbf{y}_T^{4,3}$ | … | $\mathbf{y}_T^{f+1,f}$ |
| (2005:4) | T+1 | $\mathbf{y}_{T+1}^{1,1}$ | $\mathbf{y}_{T+1}^{2,2}$ | $\mathbf{y}_{T+1}^{3,3}$ | … | $\mathbf{y}_{T+1}^{f,f}$ |
| (2006:1) | T+2 | .NA | $\mathbf{y}_{T+2}^{1,2}$ | $\mathbf{y}_{T+2}^{2,3}$ | … | $\mathbf{y}_{T+2}^{f-1,f}$ |
| (2006:2) | T+3 | .NA | .NA | $\mathbf{y}_{T+2}^{1,3}$ | … | $\mathbf{y}_{T+2}^{f-2,f}$ |
| … | … | … | … | … | … | … |
| (2007:1) | T+f | .NA | .NA | .NA | … | $\mathbf{y}_{T+f}^{1,f}$ |

We build this matrix in terms of g**rowth rates** instead of levels. **The first release** for each NA variable (GDP and components) is the sequence of the growth rates computed on the preliminary series published by the statistical agency.

As in Golinelli-Parigi (2005), the **indicators dataset** is not organized by vintages: some of them are not usually revised (business surveys); hard indicators are held fixed at the latest available vintage at the time this exercise has been performed.

**Number of competing models**. For each vintage and reference variables, we consider: the VAR models, the SE bridge models, the univariate ARIMA specifications. Their specifications are held fixed over the vintages and the sample periods.

To measure the effect of NA revisions on model's predictive accuracy, we carry out **two pseudo-real-time exercises**.

**First of all**, we take each model varying from one vintage to the other: **1-quarter-ahead forecasts for each vintage** are obtained from a recursive procedure.

**First exercise**: the one-quarter ahead growth rate predictions are compared with the **first release**. On the basis of the one-quarter ahead forecast errors, we compute a number of alternative measures of forecast evaluation for all models (ME, MAE, RMSFE).

**Second exercise**: the one-quarter ahead forecasts are compared with the **most recent** available vintage (*actual* series before a benchmark revision). This exercise allows to simulate what happen in the forecasting practices, when predictions are compared with the most recent data.

Additionally, we also consider the **predictive content of the first release** (never revised) vs. the latest available vintage. The difference between *actual* and the first release is the *forecast error* of the first release.

Given the methodological homogeneity across several vintages, the error embodies only statistical changes due to new information availability.

We take **preliminary growth rates as predictions** of the *actual* GDP data, and assess their forecasting ability with respect to a number of alternative models.

We find that the prediction of the *actual* available data could be significantly improved by **combining** the preliminary releases and 1-step-ahead forecasts from the Model.

# Forecasting ability comparison with respect to the 1st release

| | | RMSFE | MAE | MAPE | ME |
|---|---|---|---|---|---|
| $GDP_d$ | Model | 0.004 | 0.003 | 1.129 | -0.003 |
| | ARIMA | 0.004 | 0.004 | 6.732 | -0.001 |
| Cons | Model | 0.004 | 0.003 | 1.490 | 0.000 |
| | ARIMA | 0.003 | 0.002 | 1.122 | 0.000 |
| Inv | Model | 0.009 | 0.008 | 0.560 | 0.000 |
| | ARIMA | 0.013 | 0.012 | 0.786 | 0.001 |
| Exp | Model | 0.019 | 0.014 | 0.634 | -0.008 |
| | ARIMA | 0.021 | 0.017 | 1.149 | -0.007 |
| Imp | Model | 0.011 | 0.012 | 0.894 | -0.004 |
| | ARIMA | 0.010 | 0.008 | 0.695 | -0.002 |

## Forecasting ability comparison with respect to the latest available vintage

| | | RMSFE | MAE | MAPE | ME |
|---|---|---|---|---|---|
| GDP$_d$ | Model | 0.004 | 0.004 | 0.876 | -0.003 |
| | ARIMA | 0.005 | 0.005 | 2.005 | -0.001 |
| | First Release | 0.001 | 0.001 | 0.262 | 0.000 |
| Cons | Model | 0.003 | 0.004 | 1.451 | 0.000 |
| | ARIMA | 0.003 | 0.002 | 0.867 | 0.001 |
| | First Release | 0.001 | 0.001 | 0.275 | 0.000 |
| Inv | Model | 0.006 | 0.005 | 0.596 | 0.000 |
| | ARIMA | 0.013 | 0.012 | 1.122 | 0.000 |
| | First Release | 0.005 | 0.003 | 0.713 | -0.001 |
| Exp | Model | 0.019 | 0.015 | 0.732 | -0.008 |
| | ARIMA | 0.021 | 0.018 | 1.195 | -0.007 |
| | First Release | 0.004 | 0.003 | 0.187 | 0.000 |
| Imp | Model | 0.012 | 0.013 | 1.421 | -0.003 |
| | ARIMA | 0.010 | 0.008 | 0.919 | -0.001 |
| | First Release | 0.005 | 0.004 | 0.683 | 0.001 |

# Conclusions

**Forecast evaluation**: At 1-step-ahead, conditional SE models outperform Model forecasts (both from the demand and supply side). VARs outperform the others for longer forecast horizons ($h>2$).

**Forecast encompassing**: At 1-step-ahead, **conditional SE** models encompass all the others. **Opportunity** to select current predictions for combining and get the most accurate (efficient) GDP forecast. Advantage of model combination seems also to emerge for **GDP components**.

When $h>1$, GDP model from the **demand side** encompasses those based on supply indicators and the ARIMA benchmark. For **GDP components**, the encompassing tests are not conclusive: opportunity of forecast combining up to 2 steps ahead. Model forecast encompasses the ARIMA benchmark only for $h>2$.

**Role of NA revisions**: We find that the prediction of the *actual* available data could be significantly improved by **combining** the preliminary releases and 1-step-ahead forecasts from the Model.