GDP revisions are not cool.

The statistical agencies' trade-off.

S. Asimakopoulos [1] M. Lalik [2] J. Paredes [2] J. Salvado [2]

[1] Brunel University London

[2] European Central Bank

October 19, 2023

Outline

Motivation

How do Statistical Agencies manage the interplay between revisions to GDP and it's subcomponents (i.e. private consumption, government consumption, investment, net trade, and inventories)? Do they revise GDP to the same extent as the expenditure components or is there any trade-off ("loss function")?

- Do the revisions fulfill desirable statistical properties? (e.g. mean zero, small volatility etc.)
- Are the revisions "news" or "noise" (predictable)?
- Are SAs treating GDP as a "King" (i.e. not revise it too much over time, otherwise the initial releases would be seen as unreliable, but being more open to revisions to the expenditure components)?
- Is the underlying cross-correlations among revisions to GDP **and its components** of any use to forecast initial GDP releases?

Literature

Properties of GDP revisions and how they affect forecasting:

- i) mean zero; ii) small volatility; iii) small revision-to-announcement ratio; and iv) small noise-to-signal ratio.
- Mankiw, Runkle and Shapiro (1984), Mankiw and Shapiro (1986), Mork (1987), Croushore and Stark (2003), Faust, Rogers and Wright (2005), Aruoba (2008), Croushore (2011).

Incorporating GDP vintages and revisions to VAR approach

- Garrat, Lee, Mise, and Shields (2008), Clements and Galvao (2012), Clements and Galvao (2013) and Carriero, Clements, and Galvao (2015).

Statistical Agencies trade-off:

- Eurostat (2015): "There are two important requirements for quarterly national accounts. Quarterly national accounts must be: available as soon as possible after the end of the reference period; and as accurate as is feasible to require as little subsequent revision as possible."

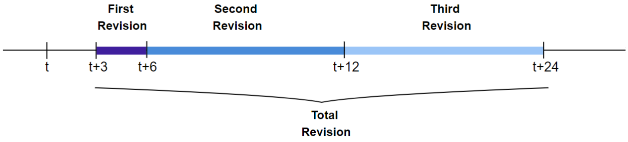
Data

We construct a unique dataset of revisions to the GDP growth and the revisions to the contributions of expenditure components. The dataset is based on national accounts data published by Eurostat, for Germany, France, Italy and Spain over the period 2002-2022.

"Snapshot" approach: Following the approach of Giannone, Henry, Lalik and Modugno (2012) and using monthly snapshots of quarterly data, we monitor the appearance of new information to construct vintages of "first releases" (y^t_{t+3}), "second releases" (y^t_{t+6}), "third releases" (y^t_{t+12}), and "final releases" (y^t_{t+24}).

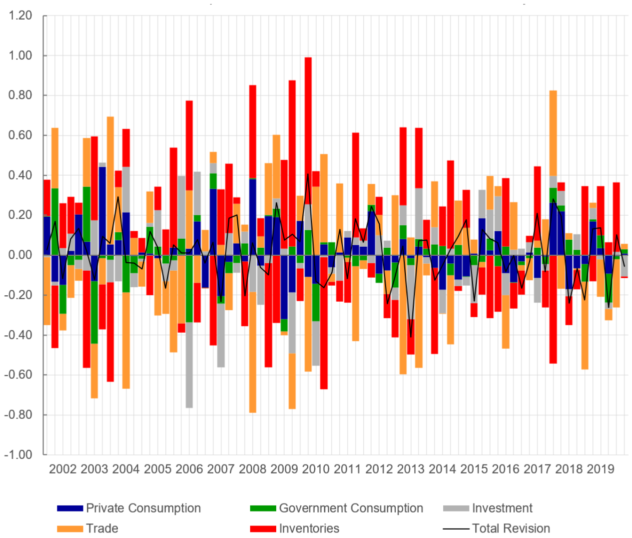
Finally, this dataset is used to derive the revisions that will be studied:

- First revision (r^t_1) = second release - first release
- Second revision (r^t_2) = third release - second release
- Third revision (r^t_3) = final release - third release
- Total revision (r^t_T) (Cumulative revision) = final release - first release



Data

Example: r_3^t for Germany



Properties of data revisions

We start by assessing the statistical behaviour of revisions based on the following properties:

- **(P1)**: $E(r_T^t) = 0$
- **(P2)**: $var(r_T^t)$ is small
- **(P3)**: Revision-to-announcement ratio $(\overline{r_T^t}/\overline{y_{t+3}^t})$ is relatively small
- **(P4)**: Noise-to-signal ratio $(\sigma_{r_T^t}/\sigma_{y_{t+24}^t})$ is relatively small

Properties of data revisions

Unconditional properties of total revisions: Germany and France

Table 1a: Summary Statistics of Total Revisions

	GDP growth rate	Private consumption contribution	Government consumption contribution	Investment contribution	Trade contribution	Inventories contribution
Panel A: Germany						
Mean of total revisions ($\overline{r_T^t}$)	**0.051**	**0.044**	0.010	-0.006	0.003	0.000
Standard deviation of total revisions ($\sigma_{r_T^t}$)	0.221	0.217	0.122	0.181	0.346	0.419
Revisions-to-announcements ratio ($\overline{r_T^t}/\overline{y_{t+3}^t}$)	0.205	0.606	0.158	-0.110	0.048	-0.035
Noise-to-signal ratio ($\sigma_{r_T^t}/\sigma_{y_{t+24}^t}$)	0.313	0.699	1.059	0.505	0.461	0.653
Corr. with initial ($\rho_{r_T^t, y_{t+3}^t}$)	-0.192	**-0.228**	**-0.527**	**-0.233**	**-0.535**	**-0.439**
Panel B: France						
Mean of total revisions ($\overline{r_T^t}$)	-0.006	-0.015	-0.001	**0.036**	0.001	-0.027
Standard deviation of total revisions ($\sigma_{r_T^t}$)	0.187	0.133	0.063	0.120	0.227	0.240
Revisions-to-announcements ratio ($\overline{r_T^t}/\overline{y_{t+3}^t}$)	-0.023	-0.079	-0.014	1.010	-0.024	3.733
Noise-to-signal ratio ($\sigma_{r_T^t}/\sigma_{y_{t+24}^t}$)	0.412	0.628	0.923	0.570	0.596	0.607
Corr. with initial ($\rho_{r_T^t, y_{t+3}^t}$)	**0.194**	**-0.251**	**-0.382**	-0.010	**-0.280**	**-0.413**

Note: N = 74, Boldface denotes significance at the 10% level

Properties of data revisions

Unconditional properties of total revisions: Italy and Spain

Table 1b: Summary Statistics of Total Revisions

	GDP growth rate	Private consumption contribution	Government consumption contribution	Investment contribution	Trade contribution	Inventories contribution
Panel C: Italy						
Mean of total revisions ($\overline{r_T^t}$)	0.000	-0.029	-0.015	0.008	0.027	0.008
Standard deviation of total revisions ($\sigma_{r_T^t}$)	0.184	0.173	0.110	0.161	0.377	0.476
Revisions-to-announcements ratio ($\overline{r_T^t}/\overline{y_{t+3}^t}$)	-0.140	-0.699	-0.803	-0.149	1.711	-0.487
Noise-to-signal ratio ($\sigma_{r_T^t}/\sigma_{y_{t+24}^t}$)	0.307	0.524	1.088	0.457	0.854	1.021
Corr. with initial ($\rho_{r_T^t, y_{t+3}^t}$)	0.141	0.118	**-0.452**	**-0.277**	**-0.519**	**-0.585**
Panel D: Spain						
Mean of total revisions ($\overline{r_T^t}$)	-0.006	-0.016	-0.004	0.028	0.001	-0.015
Standard deviation of total revisions ($\sigma_{r_T^t}$)	0.160	0.286	0.232	0.255	0.430	0.253
Revisions-to-announcements ratio ($\overline{r_T^t}/\overline{y_{t+3}^t}$)	-0.014	-0.078	-0.046	-9.799	0.014	-0.463
Noise-to-signal ratio ($\sigma_{r_T^t}/\sigma_{y_{t+24}^t}$)	0.283	0.610	1.229	0.542	0.814	1.217
Corr. with initial ($\rho_{r_T^t, y_{t+3}^t}$)	-0.086	-0.160	**-0.713**	-0.186	**-0.398**	**-0.600**

Note: N = 74, Boldface denotes significance at the 10% level

Properties of data revisions

Unconditional properties of total revisions.

- **(P1)** is satisfied for the GDP growth rate for all countries except of Germany. Additionally, we also detect a significant positive revision for the private consumption in Germany and investment contribution in France.
- **(P2)** - **(P4)** are not supported by the data, which shows that the volatility of revisions is considerable in comparison to the volatility of the series themselves.
- We find statistically significant *negative* correlations between total revisions of the components to growth and it's initial announcements. This implies that sub-components contributions might cancel each other over time.

News vs noise

- Following Mankiw and Saphiro (1986) and Aruoba (2008), data revisions are sometimes characterized as “news” or “noise”.
- Data revisions are “news” when they add new information, and “noise” when they reduce measurement error.
- Under the “news” hypothesis the initial announcement is an efficient forecast that reflects all available information and subsequent estimates reduce the forecast error, incorporating new information. This way the revision is correlated with the final value but uncorrelated with the data available when the estimate is made.
- However, under the “noise” hypothesis the initial announcement is an observation of the final series, measured with error. This means that the revision is uncorrelated with the final value but correlated with the data available when the estimate is made.

News vs noise tests

To test the news and noise hypotheses we estimate the two equations below:

$$y^t_{t+24} = \alpha_1 + \beta_{news} * y^t_{t+3} + \epsilon^1_t \tag{1}$$

$$y^t_{t+3} = \alpha_2 + \beta_{noise} * y^t_{t+24} + \epsilon^2_t \tag{2}$$

The first equation tests the "news" hypothesis under the β_{news}=1 and α_1=0 joint F-test. Whereas, the second equation tests the "noise" hypothesis under the β_{noise}=1 and α_2=0 joint F-test. These two hypotheses can be both rejected, where it can be concluded that the data are not informative regarding the news/noise dichotomy.

News vs noise

Joint F-statistics (robust Wald statistic) for news and noise

Germany	GDP	PCR	GCR	ITR	BTR	SCR
News	4.22**	1.91	14.68***	1.08	9.98***	4.72**
Noise	2.29*	16.77***	12.87***	2.39*	0.75	1.07
France	**GDP**	**PCR**	**GCR**	**ITR**	**BTR**	**SCR**
News	1.66	1.69	2.69*	4.08**	2.94**	5.54***
Noise	14.82***	7.37***	21.08***	23.52***	4.25**	2.25
Italy	**GDP**	**PCR**	**GCR**	**ITR**	**BTR**	**SCR**
News	0.37	1.11	9.73***	3.82**	12.30***	18.59***
Noise	4.21**	22.24***	25.10***	0.70	3.53**	4.62**
Spain	**GDP**	**PCR**	**GCR**	**ITR**	**BTR**	**SCR**
News	0.33	1.37	19.82***	1.16	6.51***	9.56***
Noise	0.63	7.25***	4.11**	5.22***	7.77***	19.74***

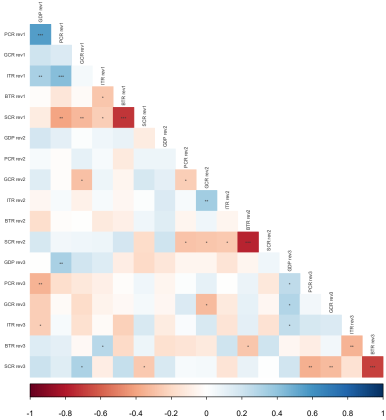
Notes: The results above show the joint F-statistics (robust Wald statistic) for each regression and each component. 1%, 5% and 10% significance levels are denoted by ***, ** and * respectively. Where we have defined as PCR: Private Consumption Contribution; GCR: Government Consumption Contribution; ITR: Investment Contribution; BTR: Trade Contribution; SCR: Inventories Contribution.

A theoretical framework for the statistical agencies' trade-off problem

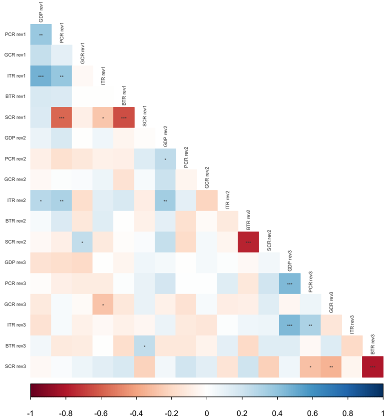
- Statistical agencies face a trade-off: They want to be reliable, but also timely in their publications as users demand speedy information.
- Our hypothesis is that the SAs prioritize GDP growth: as to avoid big fluctuations over time and maintain the reliability of this series, SAs compensate the news through at least one expenditure component when updating any other component.
- This *redistribution* effect in the contributions to GDP growth on the expenditure side is hinted by the correlations matrices among components: there is a statistically significant correlation among some components for the same vintage: **Revisions of the contribution of inventories has a significant negative correlation with the revisions in the trade contribution across countries and vintages.**

Revisions correlation matrix

Germany and France



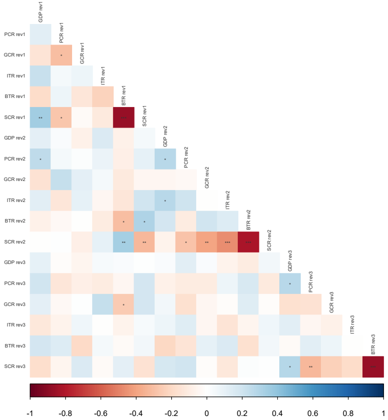
a) Germany



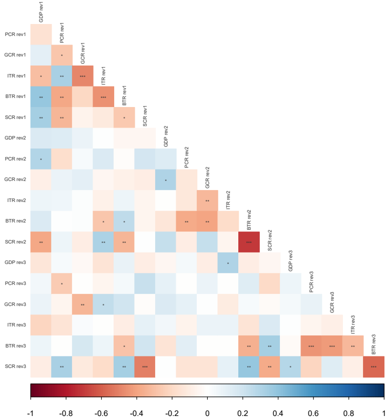
b) France

Revisions correlation matrix

Italy and Spain



c) Italy



d) Spain

Empirical approach - OLS

- The correlation matrix suggests that there is at least one candidate that SAs use to balance the incorporation of new information: Inventories.
- For each vintage and country we test whether the revision of the contribution of inventories to GDP growth can be explained by the revision of the rest of contributions.

For each revision j = 1, 2, 3, T (total) and country i = DE, FR, IT, ES, the following equation is estimated:

$$SCR_{j,i}^{t} = constant_{j,i} + \delta_1 * PCR_{j,i}^{t} + \delta_2 * GCR_{j,i}^{t} + \delta_3 * ITR_{j,i}^{t} + \delta_4 * BTR_{j,i}^{t} + u_{j,i}^{t} \qquad (3)$$

We also estimate the same equation, incorporating one explanatory variable at a time to check the relative importance of each of them.

Empirical results - OLS

	Germany				France			
	r_1^t	r_2^t	r_3^t	r_T^t	r_1^t	r_2^t	r_3^t	r_T^t
Private Consumption	-0.646***	-0.923***	-0.703***	-0.650***	-0.744***	-0.633***	-0.613***	-0.560***
	(0.080)	(0.118)	(0.118)	(0.112)	(0.144)	(0.138)	(0.138)	(0.143)
Government Consumption	-0.853***	-0.956***	-0.600***	-0.455**	-0.243	-0.048	-1.067***	-0.419
	(0.104)	(0.184)	(0.173)	(0.193)	(0.325)	(0.306)	(0.271)	(0.286)
Investment Contribution	-0.873***	-0.751***	-0.810***	-0.928***	-0.319*	-0.347*	-0.393**	-0.276*
	(0.103)	(0.199)	(0.138)	(0.139)	(0.176)	(0.180)	(0.147)	(0.158)
Trade Contribution	-0.969***	-0.931***	-0.975***	-1.004***	-0.863***	-0.842***	-0.862***	-0.844***
	(0.041)	(0.059)	(0.068)	(0.071)	(0.105)	(0.070)	(0.058)	(0.080)
Constant	0.013	0.019	0.014	0.030	-0.021*	-0.008	0.003	-0.025
	(0.011)	(0.014)	(0.017)	(0.024)	(0.011)	(0.011)	(0.011)	(0.018)
Observations	74	74	74	74	74	74	74	74
R^2	0.915	0.838	0.800	0.785	0.655	0.689	0.804	0.640

	Italy				Spain			
	r_1^t	r_2^t	r_3^t	r_T^t	r_1^t	r_2^t	r_3^t	r_T^t
Private Consumption	-0.884***	-0.789***	-0.734***	-0.602***	-0.898***	-0.851***	-0.962***	-0.695***
	(0.114)	(0.085)	(0.085)	(0.117)	(0.098)	(0.064)	(0.064)	(0.072)
Government Consumption	-1.219***	-1.050***	-1.165***	-1.086***	-0.840***	-0.734***	-0.937***	-0.711***
	(0.200)	(0.136)	(0.224)	(0.185)	(0.096)	(0.081)	(0.117)	(0.089)
Investment Contribution	-0.894***	-0.844***	-0.864***	-0.816***	-1.012***	-0.860***	-0.972***	-0.780***
	(0.084)	(0.095)	(0.145)	(0.124)	(0.121)	(0.081)	(0.077)	(0.096)
Trade Contribution	-1.054***	-0.958***	-1.030***	-1.066***	-0.778***	-0.930***	-0.984***	-0.775***
	(0.035)	(0.041)	(0.053)	(0.053)	(0.073)	(0.044)	(0.043)	(0.065)
Constant	0.001	-0.003	0.004	0.010	0.008	-0.006	-0.007	-0.006
	(0.008)	(0.010)	(0.015)	(0.020)	(0.011)	(0.009)	(0.012)	(0.017)
Observations	74	74	74	74	74	74	74	74
R^2	0.934	0.929	0.885	0.881	0.678	0.873	0.891	0.697

Note: Standard errors in parenthesis; ***:$p < 0.01$,**:$p < 0.05$,*:$p < 0.1$

Empirical results: V-VAR and CV-VAR models

We can get the following V-VAR representation:

$$\begin{bmatrix} y_t^{t-3} \\ r_1^{t-6} \\ r_2^{t-12} \\ r_3^{t-24} \end{bmatrix} = c + \Gamma \begin{bmatrix} y_{t-3}^{t-6} \\ r_1^{t-9} \\ r_2^{t-15} \\ r_3^{t-27} \end{bmatrix} + \epsilon_t \quad (4)$$

whereas the CV-VAR model will be:

$$\begin{bmatrix} y_t^{t-3} \\ r_{1,j}^{t-6} \\ r_{2,j}^{t-12} \\ r_{3,j}^{t-24} \end{bmatrix} = b + \Delta \begin{bmatrix} y_{t-3}^{t-6} \\ r_{1,j}^{t-9} \\ r_{2,j}^{t-15} \\ r_{3,j}^{t-27} \end{bmatrix} + \phi_t \quad (5)$$

where $j = [1,2,3,4,5]$ indicates the different expenditures components. In the CV-VAR model each revision variable is a 5x1 vector.

Empirical results

Standard V-VAR vs CV-VAR approach

Forecast comparison of GDP initial announcements 1 and 4 periods ahead

	Germany		France		Italy		Spain	
	h=1	h=4	h=1	h=4	h=1	h=4	h=1	h=4
pre-Covid	0.913	1.045	0.859*	1.042	1.134	0.966	0.891*	0.870**
Covid	0.987**	0.971	0.997	1.052	0.990	0.948	0.998	1.027
full sample	0.934*	1.021	0.899	1.045	1.093	0.960	0.923	0.920

Notes: The results above show the ratio of the average RMSFEs when comparing the forecasts between the standard V-VAR model without the components and our real-time CV-VAR model with components. A value less than one indicates that the CV-VAR approach performs better on average than the alternative. The Diebold-Mariano statistic is also applied and we show the statistical significance with asterisks, ***:$p < 0.01$,**:$p < 0.05$,*:$p < 0.1$.

Empirical results: CV-VAR forecasting revisions

Germany

	1st rev.	2nd rev. h=1	3rd rev.	1st rev.	2nd rev. h=4	3rd rev.
pre-Covid	0.9032	1.0182	0.9823	0.9903	1.0077	0.9984
Covid	0.7344	1.0742**	0.6881***	0.6342	1.1519	0.8278
full sample	0.7624	1.0581	0.8512**	0.6791	1.1043	0.9223

France

	1st rev.	2nd rev. h=1	3rd rev.	1st rev.	2nd rev. h=4	3rd rev.
pre-Covid	1.0137	1.2763	1.0056	1.0355*	0.9716	1.0036
Covid	0.8162*	0.5745	0.7364	0.4084	0.1642	0.4699
full sample	0.8659	0.6789	0.8539	0.4883	0.18525	0.5894

Italy

	1st rev.	2nd rev. h=1	3rd rev.	1st rev.	2nd rev. h=4	3rd rev.
pre-Covid	1.0935	1.0965	0.9434	1.0706*	0.9951	1.0020
Covid	0.6938	0.9758	1.1359	0.1476	0.1681	0.6810
full sample	0.7827	1.0573	1.0197	0.1721	0.2911	0.7657

Spain

	1st rev.	2nd rev. h=1	3rd rev.	1st rev.	2nd rev. h=4	3rd rev.
pre-Covid	1.2598	0.9521	1.1008	1.0655*	0.9433	1.0047
Covid	0.8887	0.5558*	0.4440	0.3300	0.1374	0.0552
full sample	0.8912	0.5749	0.7613	0.3317	0.1430	0.1205

Notes: The results above show the ratio of the average RMSFEs when comparing between the direct forecasts of GDP revisions via the standard V-VAR model without the components and the indirect forecast of GDP revisions via the components revisions using our real-time CV-VAR model. A value less than one indicates that the CV-VAR approach performs better on average than the alternative. The Diebold-Mariano statistic is also applied and we show the statistical significance with asterisks, ***:$p < 0.01$,**:$p < 0.05$,*:$p < 0.1$.

Conclusions

According to our hypothesis, statisticians consider that the initial GDP growth estimates are accurate and have no incentive to revise it. They internalise the trade-off between publication timeliness and reliability as minimising GDP revisions.

We empirically prove it and show that new incoming information will change other GDP sub-components, for example, on the expenditure side, which cancel each other at an aggregate level.

If those expenditure side items are to be used for forecasting GDP, these would need to be included in one model in order to account for the important existing cross-correlation across them. Forecasting them in different models in isolation and then adding up them to form the GDP aggregate would not seem to be a good strategy.