The OECD Weekly Tracker of GDP level

High-level summary

The OECD Weekly Tracker of GDP levels replaces the previous versions.

A new version of the Weekly Tracker estimates the level of weekly GDP in almost real time. It replaces both the original version, which proxied year-on-year GDP growth, and the Counterfactual Tracker. It uses Google Trends data, a similar, but slightly modified, modelling methodology and a new data pre-processing method. It is easier to interpret and so more informative than the previous versions. It also has a longer time coverage (2004 onwards) and is more robust to outliers, while remaining consistent with the previous two Trackers.

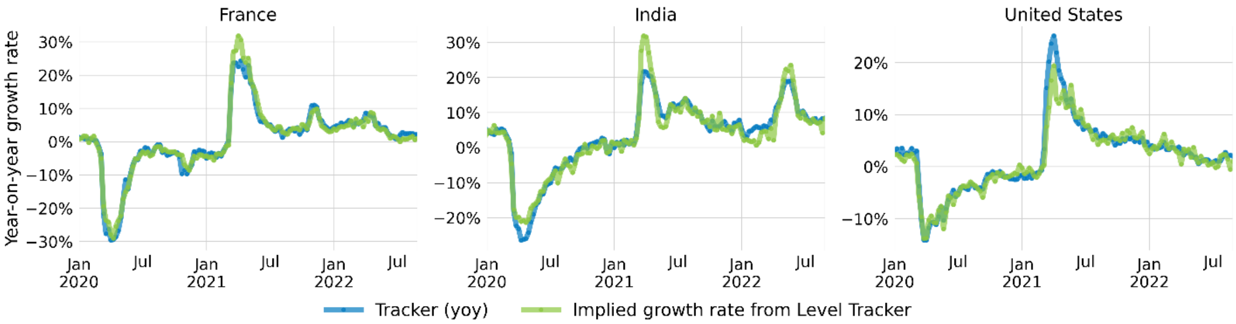
It is more informative about the current state of the economy than the Tracker of year-on-year GDP growth. Weekly series of year-on-year growth rates can be difficult to interpret, as they can be driven by variations occurring either in the current level of GDP or a year earlier. This problem was made conspicuous by the strong base effects which appeared a year after the first lockdowns, in March 2021, when sharp upturns in the Tracker were caused by the contraction from a year earlier. The new Tracker will thus make it easier to detect turning points in real time.

The Tracker of the level of GDP also replaces the Counterfactual Tracker. This second version was introduced in 2021 to address the base effect problem. It estimated weekly GDP relative to the pre-crisis trend. The latter was proxied by OECD projections from November 2019, which only covered a period to the end of 2021, when the publication of the Counterfactual Tracker had to be interrupted. In contrast, the Tracker of GDP levels has a longer time coverage, from 2004 to the present day, with no expiration date.

Building the Tracker of GDP level required a key technical challenge in high-frequency economics to be tackled, namely the issue of seasonal adjustment. Seasonal patterns in weekly economic series are difficult to filter out. That is the reason why several papers including the original Tracker paper have used 52-week differencing, which conveniently removes seasonal patterns (Lewis, Mertens and Stock, 2020[1]; Lourenço and Rua, 2020[2]; Bilek-Steindl et al., 2020[3]; Delle Monache, Emiliozzi and Nobili, 2021[4]). The Tracker of GDP level uses a new approach to seasonality based on machine learning, which allows GDP to be modelled with the Google Trends series levels rather than growth rates.

The new Tracker provides information consistent with the previous versions. Growth rates derived from the weekly GDP level series are broadly concordant with the original Tracker (Figure 1). This is a strong result given the differences in the two modelling approaches.

Figure 1. The Weekly Tracker of GDP level is consistent with the original Tracker



Source: OECD Weekly Tracker

The Weekly Tracker of GDP level is also more accurate. Forecast simulations in pseudo-real time yield better results than with the Tracker of GDP growth. This results from the Level Tracker algorithm's larger dataset, which includes observations from 2004 (whereas the training sets of growth rate models start in 2005). This also relates to the fact that an extreme event such as the historic contraction in the second quarter of 2020 will cause two, not one, outliers in the GDP year-on-year growth series (at the time of their occurrence, and through base effects one year later), which are both hard to predict. Predicting levels rather than growth rates thus mechanically reduces forecasting error.

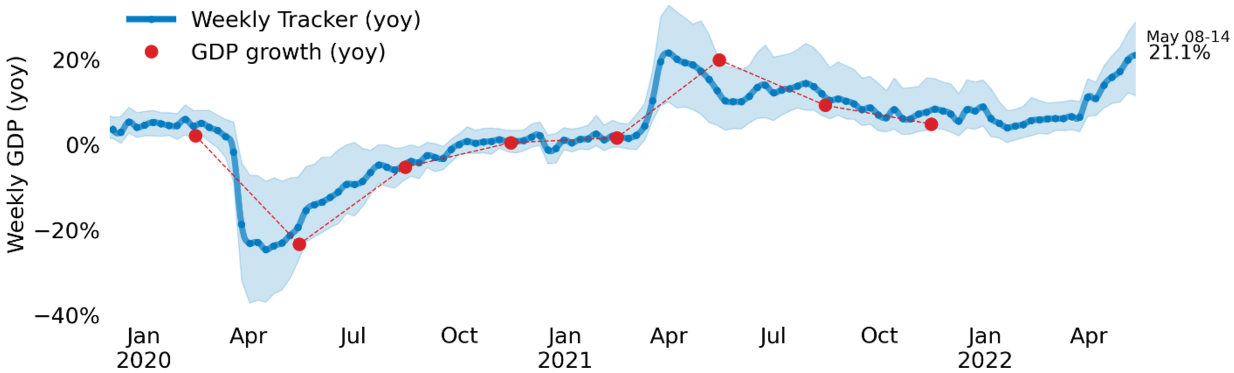
The GDP level Tracker

Among the technical challenges raised by the new generation of high-frequency indicators is seasonal adjustment (Guha and Ng, 2019[54]). This was already a concern to national statisticians working with lower frequency data, but which became thornier when it came to factoring in weekly or daily patterns such as holidays, national days, or religious celebrations. The most commonly adopted solution was to transform data using either yearly (Lewis, Mertens and Stock, 2020[1]; Lourenço and Rua, 2020[2]; Bilek-Steindl et al., 2020[3]; Delle Monache, Emiliozzi and Nobili, 2021[4]) or quarterly growth rates (Eraslan and Götz, 2020[55]; Adam et al., 2021[56]). The Tracker used the former, which helped articulate a quarterly and a weekly model. However, the limitations of this approach became conspicuous around the anniversary of the first lockdowns, when major base effects blurred the interpretability indices based on yearly growth rates, including the New York Federal Reserve Board Weekly Economic Index or the Tracker. This was clear in India for instance, where the strong upturn in March-April 2021, which mechanically results from the sharp contraction occurring a year earlier, obfuscates the economic effect of a second wave of the pandemic (Figure 2, Panel A). More generally, high-frequency series taken in yearly (or quarterly) growth rates are difficult to interpret whenever volatility in the base year (or quarter) is large, as it becomes hard to tell whether the growth rate is driven by current or past movements.

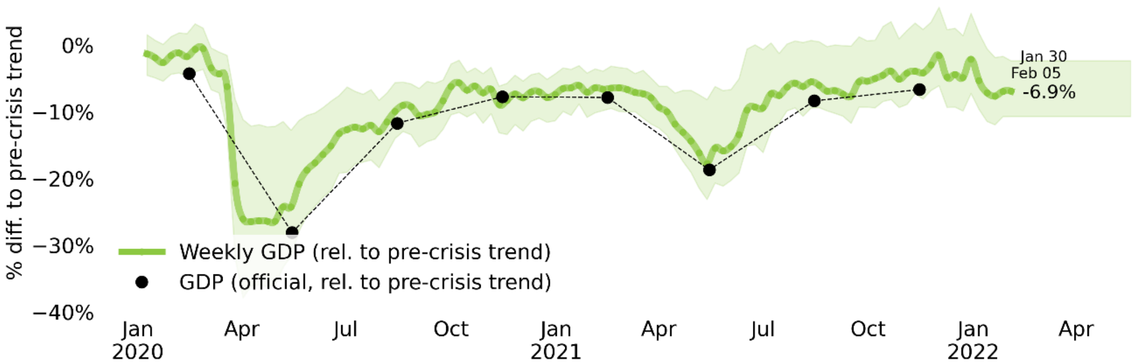
The OECD Counterfactual Weekly Tracker (Figure 2, Panel B) was introduced in 2021 to tackle the base effect problem. It relies on year-on-two-year growth rates. Shifting the base by one year in the past allowed the computation of weekly GDP in 2021 relative to the same week in 2019, when weekly volatility was low. The year-on-two-year growth rates are used to provide series of weekly GDP relative to the pre-crisis trend, proxied by the quarterly GDP projections from the November 2019 OECD Economic Outlook. This solution was no longer valid after 2021, as the weekly volatility in the base year (2020) was high again in 2022 and the pre-COVID OECD Economic Outlook forecast only covered a period to the end of 2021. As a result, the publication of the Counterfactual Weekly Tracker was interrupted in early 2022. This note introduces a perennial replacement to the Counterfactual Tracker: the GDP level Tracker (Figure 2, Panel C).

Figure 2. Three generations of Weekly Trackers, India

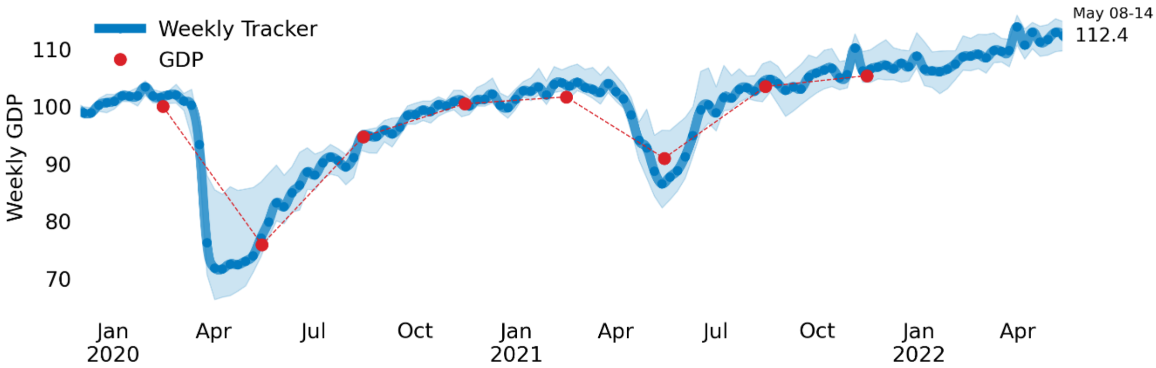
Panel A. The Weekly Tracker of GDP growth



Panel B. The Counterfactual Tracker



Panel C. The Weekly Tracker of GDP level



Note: On Panel A, the Weekly Tracker is an estimate of year-on-year growth in weekly GDP. The Counterfactual Tracker on Panel B represents the percent difference between weekly GDP and the pre-crisis trend, proxied by OECD forecasts from the November 2019 OECD Economic Outlook. Lastly, Panel C shows the Weekly Tracker of GDP level, which is an index of GDP level where 2019 Q4 = 100. With the Tracker of GDP growth on Panel A, the base effect in March-April 2021 obfuscates the effect of the second wave, which is clearly visible on the other two panels.

Source: OECD Weekly Tracker (Woloszko, 2020); OECD Quarterly National Accounts; OECD Economic Outlook n106 (Nov. 2019).

Methodology

A third generation model will replace the former two and aims at providing a perennial solution to the base effect problem. The “GDP level Tracker” provides estimates of weekly GDP levels, expressed as an index where 2019 Q4 = 100:

$$LT^w \equiv \frac{Y^w}{Y^{2019\,Q4}} * 100 + \sigma_w \tag{2.7}$$

It uses a new approach to high-frequency seasonality based on machine learning, which allows GDP to be modelled from the level of the Google Trends series rather than growth rates. It is easier to interpret and so more informative than the previous versions. It also has a longer time coverage (2004 onwards) and it is more robust to outliers, while remaining consistent with the previous two Trackers. This section describes the modelling approach used to produce the Tracker of the level of Weekly GDP. The model is similar to that of the original Tracker, except that it does not use the growth rate transformation, which was applied to both GDP series and search volume indices. The following paragraphs explain how GDP level models can be derived from GDP growth models, then formally introduce the Level Tracker model, and the new seasonality adjustment method based on machine learning.

From growth rate to level models

Official seasonally adjusted GDP for country i at time t (t can be a given quarter or a week) can be written as an exponential model:

$$Y_{it} = e^{\gamma_i + \alpha_i t + D_{it}} \tag{2.8}$$

GDP in country i is expressed as a process with a scale γ_i, an average growth rate α_i, plus idiosyncratic variations noted D_{it}. This notation implies no loss of generality: it simply fits with the fact that GDP levels are non-stationary, which is why economists usually focus on their growth rates.

Taking the growth rate of this quantity gives the basis of a GDP growth model with fixed effects:

$$\frac{d\log(Y_{it})}{dt} = \alpha_i + \frac{dD_{it}}{dt} \tag{2.9}$$

In discrete time, this yields:

$$\frac{Y_{it}}{Y_{it-4}} - 1 \approx 4\alpha_i + D_{it} - D_{it-4} \tag{2.10}$$

Fixed effects in a GDP growth model capture the term $4\alpha_i$ which corresponds to country-specific average time trends. The core idea behind the GDP level Tracker is to rather model D_{it}, by using a specification derived from the primitive of equation 2.9:

$$\log(Y_{it}) = \gamma_i + \alpha_i t + D_{it} \tag{2.11}$$

The derivative of the time-trend $\gamma_i + \alpha_i t$ is equal to α_i, and the derivative of D_{it} in continuous time is equal to the demeaned GDP growth rate. In other words, *a model of GDP growth with country dummies corresponds to a model of log GDP with country-specific time trends*.

This approach avoids the main limits of year-on-year GDP growth models, which clearly appear from equation 2.9:

- ***Outlier duplication***. Outliers appear twice, successively with a positive sign (D_{it}) and a negative sign ($-D_{it-4}$), which can degrade predictive performance, since extreme values are hard to predict.
- ***Ambiguity.*** Variations in the GDP growth rate can either be attributed to D_{it} or $-D_{it-4}$, which blurs the interpretability of resulting series and makes it more difficult to detect turning points.

However, the better interpretability of the level model comes at a cost. Whereas the growth rate model (equation 2.9) was based on the hypothesis that GDP is difference stationary, thus allowing for shock persistence, the level model implies that GDP is trend stationary. The additional parameter γ_i in the level model implies that GDP reverts to its trend, which may not be true.[1] The key feature of the Level Tracker is to relax this constraint by adding a degree of freedom and including a time variable in the machine learning model described in the next paragraph.

The GDP level model based on Google Trends

The Tracker's GDP growth model can be noted as:

$$4\alpha_i \; + \; D_{it} - D_{it-4} = f_{yoy}\left(\frac{S_{it}}{S_{it-4}} - 1\right) + \; \widehat{\alpha_\iota} \tag{2.12}$$

The country dummy captures the time trend $4\alpha_i$, which is equal to the average GDP growth rate for country i. The central modelling task consists in modelling the demeaned GDP growth by estimating f_{yoy}, which is an arbitrarily complex function of the year-on-year growth rate in the search volume indices.

The approach introduced here consists in modelling D_{it} directly instead of $D_{it} - D_{it-4}$ and using country-specific time trends instead of intercepts:

$$\log(Y_{it}) = \; g(s_{it}, t) + b_i + a_i t + \epsilon_{it} \tag{2.13}$$

The log search volume indices are noted s_{it} and are pre-processed using a methodology described in the next section. The linear term $b_i + a_i t$ captures the average trend over the training period, while the time variable in g allows for local deviations from it, such as persistent shocks.

This model is estimated in two steps: (1) estimating the time trends ($\hat{b}_i + \hat{a}_i t$) with a linear regression, and (2) estimating the relationship g from the following model:

$$D_{it} = g(s_{it}, t) + \epsilon_{it} \tag{2.14}$$

The term D_{it} is the detrended log GDP and has a straightforward interpretation as the percent deviation of GDP from its average trend: $D_{it} = \log(Y_{it}) - \hat{b}_i + \hat{a}_i t \cong \frac{Y_{it}}{\hat{b}_i + \hat{a}_i t} - 1$.

In practice, the model described in equation 2.14 is estimated from quarterly[2] official GDP series and quarterly aggregates of the search volume indices. The training set includes data pooled from 46 countries, as in the original version of the Tracker. The function g is estimated using a neural network, by regressing the detrended log GDP on the seasonally adjusted log search volume indices. It is then applied to the weekly seasonally adjusted search volume indices S_{iw}, and the Weekly Tracker of GDP level for week w is obtained as:

$$LT_{iw} = \exp[\; \hat{g}(S_{iw}, t) + \hat{b}_i + \hat{a}_i w \;] \tag{2.15}$$

A difference with the year-on-year tracker model is that the training set is slightly larger (around 3300 against around 3100 observations) as it includes observations from 2004, for which it was not possible to compute the year-on-year growth rates.

[1] A vast literature on the subject spans the past four decades and includes for instance (Fleissig and Strauss, 1999[105]) (Papell and Prodan, 2004[104]) (Camacho, 2011[106])

[2] Except for Canada, the United Kingdom, Mexico and Norway for which official monthly GDP series are used.

Machine learning for seasonal adjustment

The search indices are assumed to be well explained by a multiplicative model including a scaling factor δ_i, a time trend $\eta_i t$, a seasonal term Z_t and an idiosyncratic component S_{it}:

$$SVI_{it} = e^{\delta_i + \eta_i t + Z_{it} + S_{it}} \tag{2.16}$$

The seasonal component is assumed to have a 52-week cyclicality, whereby its value is equal to some constant z_{iw} in each week w:

$$Z_{it} = \sum_{w=1}^{52} z_{iw} * 1_{t=w} \tag{2.17}$$

It is easily seen that $Z_t - Z_{t-52} = 0$, which is the reason why the 52-week log-differencing is a form of seasonal adjustment:

$$\log(SVI_{it}) - \log(SVI_{it-52}) = 52\,\eta_i + S_{it} - S_{it-52} \tag{2.18}$$

The time trend is reduced to a constant, the seasonal components disappear, but the quantity of interest is blurred by the introduction of minus the one-year lagged idiosyncratic shock.

The Weekly Tracker of GDP level uses a seasonal adjustment based on machine learning. For each search volume index and each country, the z_w are estimated by fitting a penalised regression model on the detrended log search volume indices:

(1) The trend is fitted by regressing each log search index on time and an intercept: $\log(SVI_{it}) = \delta_i + \eta_i t + \epsilon_{it}$

(2) The z_{iw} are estimated using a Ridge regression on the detrended search volume indices:

$$\log(SVI_{it}) - \widehat{\delta_\iota} - \widehat{\eta_\iota} t = \sum_{w=1}^{52} z_{iw} * 1_{t=w} \tag{2.19}$$

(3) The trend and estimates $\widehat{z_{\iota w}}$ are subtracted from the log search volume indices:

$$S_{iw} = \log(SVI_{it}) - \widehat{\delta_\iota} - \widehat{\eta_\iota} t - \widehat{Z_{\iota w}} \tag{2.20}$$

The use of machine learning helps to reduce the level of noise in the estimation. Without regularization, the estimates of $\widehat{z_{\iota w}}$would be equal to the average value of the detrended search volume index over each week number w:

$$\widehat{z_{\iota w}}_{OLS} = \frac{1}{18}\sum_{t=1}^{N}\left(\log(SVI_{it}) - \widehat{\delta_\iota} - \widehat{\eta_\iota} t\right) * 1_{t=w} = z_{iw} + \frac{1}{18}\sum_{t=1}^{N} S_{it} * 1_{t=w} \tag{2.21}$$

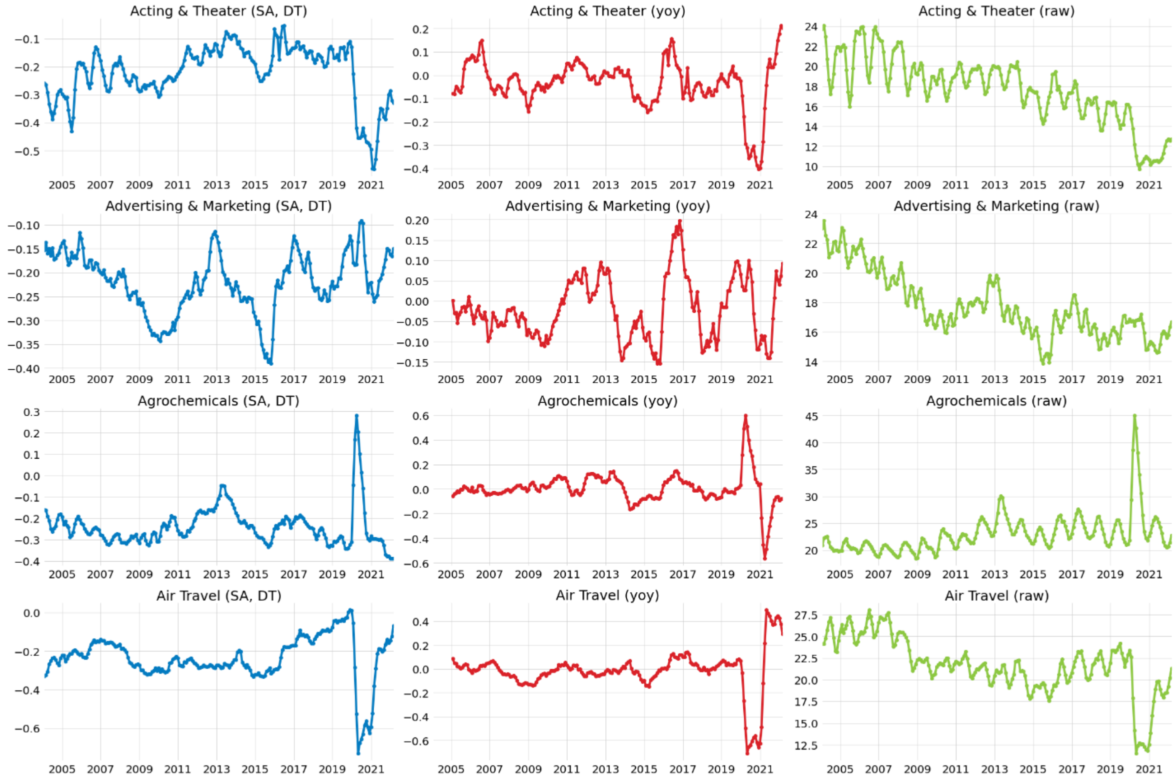
18 is the number of years in the sample (from 2004 to 2022), and it appears that the noise term is equal to the average idiosyncratic shocks occurring in week number w over the years. This noise can be large in the presence of major shocks, and subtracting the OLS estimates of the z_{it} can thus delete precious information about large shocks.

Further, the estimation of each 52 $z_{iw}$for each country and each search index is made more precise by the availability of weekly search volume indices back to 2004. Google Trends series are either available at a monthly frequency back to 2004 or at a weekly frequency over the past 5 years. That is the reason why the original Weekly Tracker was only made available over the past five years. This constraint was circumvented by looping over 5-year segments and splicing the 5-year weekly series together, thus emulating the approach found in (Borup and Schütte, 2020[10]).

Figure 3~~Figure 3~~ compares the resulting seasonally adjusted and detrended (SA, DT) log search indices with the year-on-year log differences and shows that the two series share similar statistical properties, except for the shock symmetry present in the year-on-year log differences.

Figure 3. Selected search volume indices (United States)

Seasonally adjusted and detrended, year-on-year-growth rate, raw

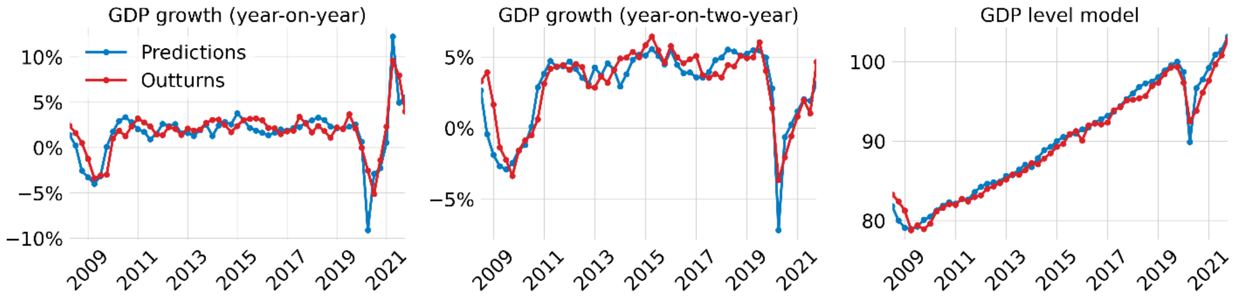


Note: The first column shows selected detrended and seasonally adjusted search volume indices for the United States. The same search indices are shown in year-on-year growth rate in the second column, and in level in the third column.
Source: Secretariat's computations and Google Trends.

The GDP level Tracker achieves higher performance

This section uses forecast simulations and shows that the Tracker of GDP level is more accurate than the previous two versions. Forecast simulations emulate the conditions faced by a forecaster at each quarter and provide out-of-sample performance metrics. The lower predictive error of the Level Tracker Is explained by two factors.

Figure 4. Forecast simulations results, United States



Note: Out-of-sample forecast simulations results form three models for the United States.
Source: OECD Quarterly National Accounts; Secrétariat's computations.

First, it is trained on a larger dataset. Its training set is 7% larger than that of the GDP growth tracker (which starts in 2005) and 14% larger than that of the Counterfactual Tracker (which starts in 2006). To see how this impacts performance, it is useful to compare the Root Mean Squared predictive Error (RMSE) of the GDP Growth, Counterfactual and Level Trackers to that of autoregressive models trained to predict GDP growth, GDP year-on-two-year growth and GDP level (Figure 4). On average across 46 countries, the relative RMSE of the Level Tracker is 8% lower than that of the GDP growth Tracker, and 20% lower than that of the Counterfactual Tracker (Table 1).

Table 1. Forecast simulations: comparisons with autoregressive models

RMSE relative to autoregressive model, G7 countries and average across 46 countries

	GDP Growth Tracker		Counterfactual Tracker		GDP Level Tracker	
	2004-2022	*2020-2021*	*2004-2022*	*2020-2021*	*2004-2022*	*2020-2021*
Canada	0.67	0.52	0.62	0.41	0.51	0.38
France	0.45	0.37	0.35	0.24	0.26	0.20
Germany	0.80	0.58	0.92	0.50	0.66	0.41
Italy	0.47	0.34	0.56	0.27	0.39	0.26
Japan	0.74	0.25	1.01	0.31	0.82	0.35
United Kingdom	0.48	0.40	0.43	0.36	0.25	0.19
United States	0.58	0.42	0.55	0.27	0.38	0.22
Average (46 countries)	*0.78*	*0.50*	*0.90*	*0.56*	*0.72*	*0.47*

Note: The out-of-sample Root Mean Square Errors obtained from forecast simulations are divided by those of an autoregressive model (AR 4). The performance metrics related to the Counterfactual Tracker are derived from the year-on-two-year growth rate model (equation 2.4).
Source: Secretariat's computations.

Second, the Weekly Tracker of GDP level is more robust to outliers. This relates to the fact that extreme events such as the historic contraction in the second quarter of 2020, which are particularly hard to predict, appear twice in GDP year-on-year growth series. Of course, this effect does not appear in Table 1 because the same line of reasoning applies to the three benchmark autoregressive models. Alternatively, the accuracies of the GDP growth and GDP level models can also be compared by deriving the year-on-year GDP growth rate from the out-of-sample predictions of the GDP level model. Using the predicted GDP level $\widehat{Y_i^q}$ and the official GDP level Y_i^{q-4}, the implied GDP growth rates are given by[3]:

[3] At any quarter q, GDP in $q-4$ is known, so the best estimate for GDP growth uses the model prediction at the numerator and the published outturn from a year earlier at the denominator.

$$\widehat{y_{LM}^{q}} = \frac{\widehat{Y_i^q}}{Y_i^{q-4}} - 1 \tag{2.20}$$

Table 2. Forecast simulations: predicting GDP growth

Normalised RMSE of GDP growth predictions, G7 countries and average across 46 countries

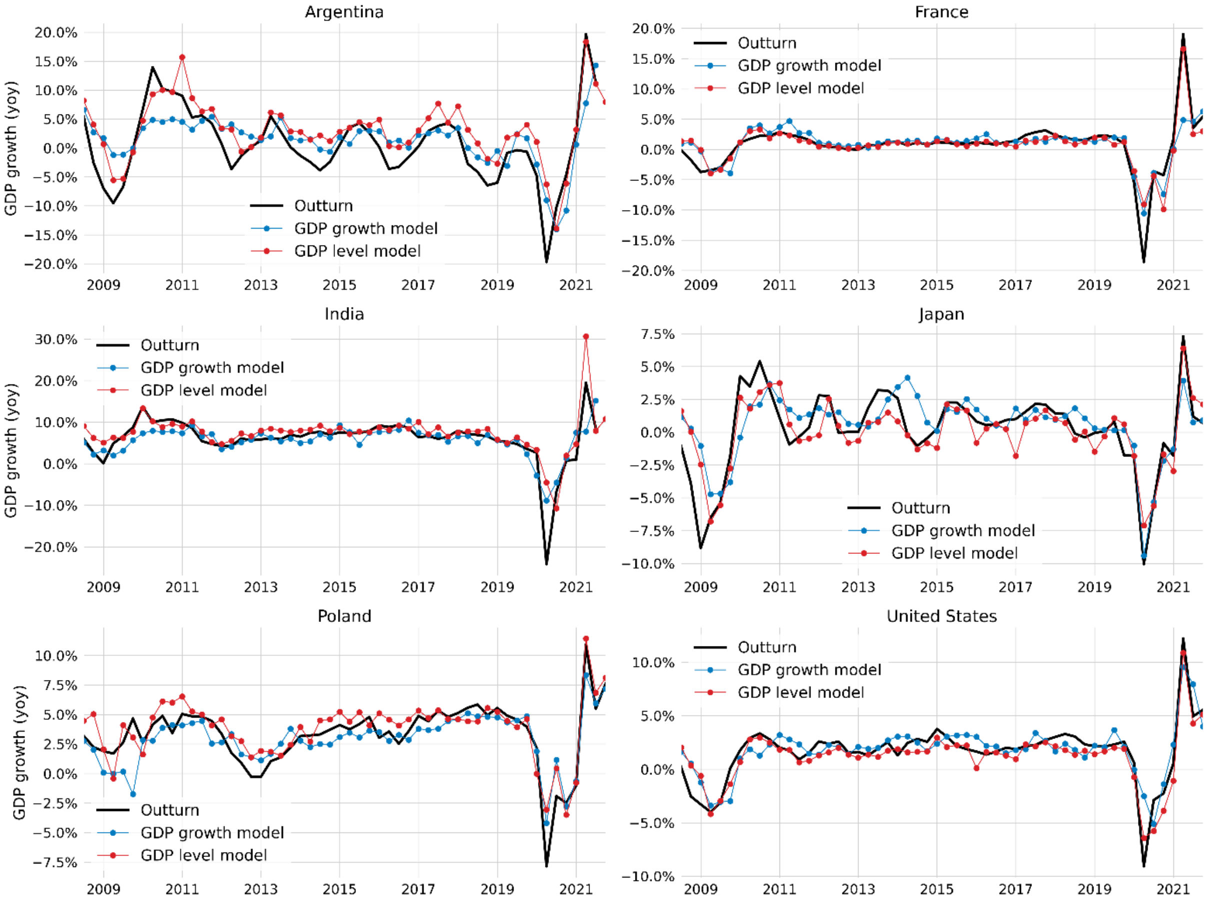
	GDP Growth Tracker		Counterfactual Tracker		GDP Level Tracker	
	2004-2022	*In 2021*	*2004-2022*	*In 2021*	*2004-2022*	*In 2021*
Canada	0.62	1.27	0.56	0.35	0.52	0.57
France	0.66	1.78	0.53	0.22	0.49	0.51
Germany	0.65	1.15	0.56	0.42	0.53	0.35
Italy	0.54	1.13	0.56	0.36	0.45	0.65
Japan	0.69	0.57	0.71	0.26	0.57	0.44
United Kingdom	0.61	1.49	0.56	0.33	0.42	0.33
United States	0.59	0.85	0.47	0.31	0.43	0.41
Average (46 countries)	0.66	1.05	0.57	0.43	0.56	0.55

Note: This table compares the performance of the three models at predicting GDP growth. It shows the RMSEs normalised by the standard deviations of official GDP series in each country for the three models. The GDP Growth Tracker predictions are compared to official year-on-year GDP growth rates. The predictions from the year-on-two-year model which is used to derive the Counterfactual Tracker are compared against official year-on-two-year GDP growth rates. And the predictions of the GDP Level Tracker are converted to growth rates using equation 2.20 and then compared to official year-on-year GDP growth rates.
Source: Secrétariat's computations.

Table 2 compares the RMSE of the GDP growth and Counterfactual Trackers to that of the GDP growth estimates implied by the Level Tracker predictions. The RMSEs are normalised by the standard deviations of the GDP growth series, which amounts to comparing performance to that of a "conservative oracle" which always predicts the period average. The handicap of the original Tracker from outlier duplication becomes conspicuous, as its normalised RMSE over 46 countries is 0.66 against 0.57 and 0.56 for the Counterfactual and Level Tracker respectively. Over 2021, the average RMSE across countries of GDP growth Tracker is around twice higher than these of the Counterfactual and Level Tracker.

In countries where the contraction in March 2020 was particularly severe, the peak observed a year later in the year-on-year series is symmetrically high and predictions from the GDP growth model fall far below (by around 19% points in France, Figure 5). By contrast, the year-on-year growth rate in the predictions derived from the GDP level models are substantially more accurate in the first quarter of 2021. The gap in performance between the two models is smaller in countries where the 2020 trough and 2021 peaks were less extreme, such as Poland or the United States. Predicting levels rather than growth rates thus mechanically reduce forecasting error.

Figure 5. Forecast simulations: comparison with the GDP growth model



Note: Pseudo-real time projections of the GDP growth from the GDP growth tracker model (blue) and the GDP level tracker model (red). The latter are derived using equation 1.1.
Source: Secrétariat's computations.

Comparing the weekly series shows that the Level Tracker is broadly consistent with the previous versions while being slightly better aligned with official GDP series. Growth rates derived from the weekly GDP level series are broadly concordant with the original Tracker (Figure 6). This is a strong result given the differences in the two modelling approaches. Moreover, based on the root mean squared error of their quarterly averages, the Level Tracker is 5% closer to the official series than the GDP growth Tracker. A comparison with the Counterfactual Tracker yields similar conclusions. The GDP Level Tracker is thus both broadly consistent with and more accurate than the previous versions.

Figure 6. Comparing the level and growth rate Trackers

Tracker (yoy) and year-on-year growth rates implied from the GDP Level Tracker

Note: The implied growth rate is computed as the 52-week growth rate in the Level Tracker series.
Source: OECD Weekly Tracker