Department of Economics

A GDP Nowcast for the UK

UCL Macro Monitor

May 2026

Corrado Calemma Tiziano di Biase Thomas Lazarowicz

1. Overview

This document describes a model for producing real-time nowcasts of UK quarterly gross domestic product (GDP) growth using a panel of mixed-frequency macroeconomic indicators. It builds on the Dynamic Factor Model (DFM) framework of Antolin-Diaz et al. (2024) – the same methodology underlying the current nowcasting model at the Federal Reserve Bank of New York – as an independent application to UK data.[1] The model extracts three latent common factors and a stochastic GDP trend from 22 time series observed at monthly and quarterly frequency, producing nowcasts that we update weekly as new data are released. The set of variables included in the model is deliberately small, covering the sameinformation set as comparable commercial nowcasts, supplemented by Euro Area survey and activity indicators that lead UK developments through trade and confidence channels. Calibration of the variable selection, identification scheme, and prior hyperparameters is documented in Section 7 below.

2. Outlier detection and removal

The data entering the model is subject to the outlier detection and removal procedure of Antolin-Diaz et al. (2024). This allows for the detection and removal of outliers in the raw data before implementation of any first differencing required to ensure stationarity of the series.[2] Outliers are specified as additive t-distributed shocks that are serially and cross-sectionally independent with constant degrees of freedom and scale parameters. More formally the relation between observed series $\tilde{y}_{it}$, outlier shocks o_{it}, and transformed series y_{it} is

$$y_{it} = \Delta_i \left(\tilde{y}_{it} - o_{it} \right),$$

where Δ_i is, with some abuse of notation, either the first difference operator or the identity operator, depending on the series, as detailed in Table 1. The outlier component o_{it} is defined as:

$$o_{it} = \sqrt{\psi_{it}} z_{it},$$

where $\nu_i \psi_{it}^{-1} \sim \chi^2_{\nu_i}$ and $z_{it} \sim \mathcal{N}(0, \sigma^2_{o,i})$. The outlier component is extracted and removed in a separate step of the Gibbs sampling algorithm, following Jacquier et al. (2004).

3. State-Space Representation

After removing outliers and transforming, the model has a state-space representation. For $i = 1, \ldots, n$ indicators, the measurement equation is:

$$y_{it} = c_i + \underbrace{\lambda_{i,\tau}\, \tau_t}_{\text{trend}} + \underbrace{\sum_{k=1}^{r} \sum_{j=0}^{s} \lambda_{i,k,j}\, f_{k,t-j}}_{\text{common factors}} + e_{it} \tag{1}$$

where τ_t is the GDP trend, a persistent component capturing the slow-moving permanent level of output, and $f_{k,t}$ for $k = 1, \ldots, r = 3$ are latent common cyclical factors shared across the 22 indicators. e_{it} is the idiosyncratic component. The trend loading $\lambda_{i,\tau}$ and the factor loadings $\lambda_{i,k,j}$ are subject to the identifying restrictions described in Section 4.[3]

[1] We are grateful to the authors for providing us with their code, which in turn draws heavily on code developed by Fulcrum Asset Management.

[2] See Table 1 below for a list of variable transformations employed.

[3] The distributed lags in the measurement equation are observationally equivalent to stacking $(f_t', \ldots, f_{t-s}')'$ into the state with contemporaneous loadings, so the choice does not affect identification. We follow Antolin-Diaz et al. (2024) because it makes each indicator's lead–lag profile explicit.

The trend follows a random walk; the three factors follow a stationary VAR(2):

$$\tau_t = \tau_{t-1} + \sigma_\tau\, \eta_{\tau,t} \tag{2}$$

$$f_t = \Phi_1 f_{t-1} + \Phi_2 f_{t-2} + \eta_t, \qquad \eta_t \sim \mathcal{N}(0,\, \Sigma_{\eta,t}) \tag{3}$$

with a stationarity constraint requiring all eigenvalues of the companion matrix to lie inside the unit circle. The factor innovation covariance $\Sigma_{\eta,t}$ is diagonal: each factor has its own independent univariate stochastic volatility process, so the three factors do not share innovation variance. Each idiosyncratic component, e_{it}, follows an independent AR(2):

$$e_{it} = \rho_{i1}\, e_{i,t-1} + \rho_{i2}\, e_{i,t-2} + \sigma_{u,i,t}\, s_{it}\, u_{it} \tag{4}$$

The shocks $\eta_{\tau,t}$, η_t and u_{it} are independently and identically distributed $\mathcal{N}(0,1)$. Equations (5)–(6) specify that the innovation log-variances of the factors and the idiosyncratic terms follow random walks (i.e. stochastic volatility)[4], allowing the model to adapt to changes in macroeconomic volatility (e.g. the 2008 financial crisis, the COVID-19 pandemic):

$$\Delta \ln \sigma^2_{\eta,k,t} = \gamma_{\eta,k}\, \nu_{\eta,k,t} \qquad\qquad \Delta \ln \sigma^2_{u,i,t} = \gamma_{u,i}\, \nu_{u,i,t} \tag{5}$$

The scalar s_{it} is an outlier scale drawn from a Student-t distribution with $\nu_o = 8$ degrees of freedom: it inflates the innovation variance for observations deemed outlying, protecting the estimated factors against data spikes[5]. The trend innovation variance σ^2_τ is held constant.

Quarterly GDP is treated as a partially observed monthly series. The observed quarterly growth rate is linked to the latent monthly states via the Mariano and Murasawa (2003) triangular aggregation weights [1/3, 2/3, 1, 2/3, 1/3] embedded in the measurement matrix, so that the 21 monthly indicators inform the quarterly target in real time as data arrive within the quarter.

4. Identification and Factor Structure

The model rests on two sources of identification. First, the factors follow a stationary VAR(2) with stochastic volatility; assuming that the SV processes in the VAR innovations are governed by three independent random walks in log space identifies the factors up to scaling and permutation. Second, scaling is fixed by a *unit diagonal* restriction on the loading matrix: three indicators are designated to normalise the scale of the three factors, each loading by construction with unit weight on its own factor. Specifically, GDP normalises the first factor, GfK Consumer Confidence normalises the second, and Industrial Production normalises the third; aside from these three diagonal entries, every variable loads freely on every factor. This resolves the sign and scale indeterminacy of the factors, which is essential for reliable inference, but leaves the rotation meaningfully unrestricted, which is innocuous for the forecasting application; causal interpretation of individual factors lies outside the scope of this work, and the cyclical signal extracted from the panel is invariant to rotations of the factor basis.

[4]With random-walk log-variances the transformed series are not unconditionally stationary, their unconditional variance being time-varying. This nonstationarity is intentional and does not impede estimation: the autoregressive dynamics remain stationary, and conditional on the volatility path the process is a stable Gaussian VAR, which the Kalman filter and simulation smoother extract regardless of whether the log volatility follows a random walk.

[5]The two outlier components capture different phenomena: o_{it} removes transient additive noise in the observed series, while s_{it} gives the persistent idiosyncratic process fat-tailed innovations.

The trend τ_t is identified separately through a key restriction: only GDP has a non-zero loading on the trend, fixed at unity ($\lambda_{\text{GDP},\tau} = 1$, $\lambda_{i,\tau} = 0$ for all $i \neq$ GDP). As in the single-factor specification of Antolin-Diaz et al. (2024), the trend is a slow-moving permanent component of the GDP series alone, identified through GDP's own dynamics rather than from co-movement across the panel. An exclusion restriction on lagged trends (no variable loads on τ_{t-j} for $j \geq 1$, reflecting the random-walk specification) and a zero GDP intercept (absorbed by the standardisation of the series) complete the identification of the trend block.

5. Estimation

The model, given its structure and Bayesian framework, benefits from a sequential estimation approach. Parameters and latent states are estimated via a Gibbs sampling algorithm, iterating over blocks of parameters and states conditional on the others. This allows the problem of sampling to be reduced from an exceedingly high-dimensional posterior into multiple manageable pieces. Latent states conditional on parameters are extracted using multiple Kalman Filtering steps and simulated using standard simulation smoothing steps like the one proposed by Durbin and Koopman (2002). Log-volatility paths, which can only be extracted from a non-Gaussian state-space model, are sampled following Kim et al. (1998), i.e. approximating the non-Gaussian distribution of interest (in this case a log-chi-squared) with a ten-component Gaussian mixture. The outlier components are sampled following Jacquier et al. (2004), drawing ψ_{it} conditional on z_{it} and z_{it} conditional on ψ_{it} (z_{it} is Gaussian, so this step reduces to a standard Kalman filtering exercise with simulation smoothing). Conditional on the latent states, all the constant coefficients are then estimated with standard Bayesian regression approaches. For instance, the factor VAR parameters are sampled using a multivariate regression with a Minnesota prior, while the factor loadings are sampled equation-by-equation with a univariate regression, again with a Minnesota prior. We take 8,000 Gibbs iterations, of which the first 2,000 are discarded as burn-in (6,000 retained); the real-time estimation procedure is documented in full in Section 7.

6. Adaptations for UK Data

Data and variable selection. The production model uses 22 indicators (one quarterly GDP target plus 21 monthly series), chosen to match the information content of comparable commercial UK nowcasts and to incorporate Euro Area indicators with established lead properties for UK activity through trade and confidence channels. The full panel and the stationarity transformation applied to each series are reported in Table 1. All transformed series are standardised to mean zero and unit variance before estimation.

Prior calibration. The prior hyperparameters of the model are calibrated to UK data by minimising post-COVID nowcast RMSE in pseudo-out-of-sample backtests; the calibration grid, ranking, and robustness checks are reported in Section 7. The production configuration is summarised below.

- *Trend innovation variance.* Inverse-Gamma prior with scale $S_\tau = 0.01$ and degrees of freedom $\nu_\tau = 1$. This is a deliberately weakly informative prior: with the unit-diagonal and no-trend identification restricting trend exposure to GDP alone, the trend's role is reduced to capturing the slow-moving level of the GDP series, and a flatter prior leaves the data free to determine its smoothness.

Table 1: Variable panel and stationarity transformations.

Variable	Category	Freq.	Transform
UK GDP	Target	Q	QoQ %
Monthly GDP	Activity	M	MoM %
Industrial production	Activity	M	MoM %
Manufacturing output	Activity	M	MoM %
Manufacturing turnover	Activity	M	MoM %
Retail sales	Trade	M	MoM %
Total goods exports	Trade	M	MoM %
Total goods imports	Trade	M	MoM %
New car registrations	Trade	M	12m diff
Claimant-count unemployment rate	Labour	M	1st diff (pp)
GfK Consumer Confidence	UK survey	M	Level
PMI Manufacturing	UK survey	M	Level
PMI Services	UK survey	M	Level
PMI Construction	UK survey	M	Level
CBI Distributive Trades (retailing balance)	UK survey	M	Level
Lloyds Business Barometer	UK survey	M	Level
EA Business Climate Indicator	Euro Area	M	Level
EA Construction Output	Euro Area	M	Level
EA Consumer Confidence	Euro Area	M	Level
EA Industrial Production (excl. construction)	Euro Area	M	Level
EA Manufacturing PMI	Euro Area	M	Level
EA Services PMI	Euro Area	M	Level

Frequency: M = monthly, Q = quarterly. Transform: *Level* = no transformation; *MoM %* = month-on-month percentage change; *QoQ %* = quarter-on-quarter percentage change; *1st diff* = first difference (in percentage points for rates); *12m diff* = twelve-month difference.

- *Idiosyncratic stochastic-volatility prior.* Tight inverse-Gamma on the volatility-of-volatility ($S = 10^{-4}$, $\nu = 1$). Without this anchor, the SV processes can absorb variation that should be attributed to the factors, causing the factor innovation variances $\Sigma_{\eta,t}$ to drift toward zero and producing unstable posterior estimates of the loadings.

- *Factor VAR coefficients.* Minnesota prior with overall tightness $\tau_{\text{VAR}} = 0.9$, lag decay $\alpha = 2$, and sum-of-coefficients shrinkage $\mu = 1$.

- *Factor loadings.* Minnesota prior with tightness $\tau_\lambda = 0.8$ and decay $\alpha = 2$.

- *Outlier component.* Student-t scale with $\nu_o = 8$ degrees of freedom, common across variables.

The relatively loose VAR coefficient and loading priors ($\tau_{\text{VAR}} = 0.9$, $\tau_\lambda = 0.8$) reflect the fact that with three SV-driven factors and only 22 indicators, the panel is informative enough that aggressive shrinkage is not required and would bias the posterior away from the cyclical signal we seek to extract.

Nowcasting cycle. In production, full re-estimation (including model parameters) occurs after each new GDP outturn is released by the Office for National Statistics (approximately a month and a half after the end of the reference quarter), and weekly updates incorporate the weekly data flow to update the factors conditional on the model parameters using the partial re-estimation procedure documented in Section 7.

7. Backtesting

This section describes the backtesting procedure used to develop and validate the model.

7.1 Real-time construction and evaluation

Our estimation procedure uses two types of model run. A *full* re-estimation calls the Gibbs sampler described above on the full dataset (that is, the vintage of data available at time t), producing a joint posterior over parameters and latent states. A *partial* re-estimation uses the parameter posterior from the most recent full re-estimation and updates only the latent states based on data releases since the last full re-estimation. Full re-estimation occurs with each GDP outturn release; partial re-estimation occurs weekly to incorporate the latest data flow. The motivation for the split is threefold. First, following a quarterly re-estimation schedule means that parameters change at the pace of the slowest signal in our model (quarterly GDP), while the latent state changes at the pace of each release. Second, holding parameters fixed between GDP releases means that we can easily isolate the news content of each release and produce an updated nowcast at each release without conflating parameter re-estimation and news. Third, partial re-estimation still allows us to utilise the model's outlier detection algorithm to decide whether a release contains news or is likely a historical outlier (without running a full re-estimation).

Mini-Gibbs procedure. Partial re-estimation is non-trivial under the Bayesian framework adopted here. Two issues prevent a standard Kalman filter and simulation smoother combination from being used directly. First, the object held fixed between full re-estimations is not a single point estimate of the parameters but a posterior distribution. Second, the outlier component introduces a non-linearity that makes the joint sampling of factors, innovations, and outlier scales infeasible even when the parameters are known with point precision. We resolve both by replacing the analytical Kalman recursion with a nested Monte Carlo procedure, which we refer to as a "mini-Gibbs" procedure.

From the most recent full re-estimation, the parameter posterior is approximated by D independent draws. For each draw, parameters are held fixed at the drawn values and a reduced Gibbs sampler is run on the latent states alone – factors, trend, idiosyncratic components, log-volatilities, and outlier scales – via a recursive sequence of Kalman filtering and simulation smoothing steps. The output is D independent posterior trajectories of the latent states conditional on the new vintage. The full posterior of the latent states under the partially re-estimated model is obtained by aggregating across the D parameter draws.

This procedure is, in principle, no faster than a single full re-estimation, however unlike the full re-estimation this process is easily parallelisable. In production we use $D = 100$ parameter draws and $R = 1{,}000$ inner iterations.

7.2 Backtest design

To evaluate the model, we construct a real time historical vintage dataset. Each vintage is a snapshot of the macroeconomic information set as it appeared on any given day: each observation is the value of the respective series that was available on a given day without the benefit of subsequent revisions. The full sample of vintages used for backtesting and calibration covers 2018Q1 to 2026Q1. For evaluation purposes we produce these vintages on any given Wednesday rather than for every given release.

At each weekly vintage the model produces a density forecast of quarterly GDP growth at multiple horizons. For evaluation, we focus on horizon $h = 0$, the nowcast of the current target quarter, where the target quarter is the next quarter for which a GDP outturn has not yet been released, and summarise the posterior by its mean.

Two RMSE metrics are reported:

- *Last h=0 RMSE*: for each target quarter, the h=0 forecast from the last vintage before the GDP release for that quarter is extracted and squared error against the realised outturn is computed. This produces one observation per target quarter. This is the metric most directly comparable to externally published nowcasts, which report a single best-before-release forecast per quarter.

- *Average h=0 RMSE*: for each target quarter, the squared error is the average of the squared error across all weekly vintages whose target is that quarter, weighting forecasts made when the information set is sparse equally with those made when it is rich.

Each metric is reported on three sub-samples: the full panel of target quarters in 2018Q1–2026Q1, a subsample excluding the COVID period (2020-2021), and the post-COVID subsample (2022Q1 onward). We exclude the COVID subsample from our headline assessment because the 2020Q2–Q3 outturns – a 20% contraction followed by a 17% rebound – are large enough to dominate full-sample RMSE on their own and are structurally distinct from the cyclical fluctuations the model is designed to track. The headline metric throughout this document is post-COVID Last h=0 RMSE.

7.3 Variable panel and factor count

We use this backtesting exercise to select our nowcast specification. The first calibration choice is over variable selection and the factor count, r. We consider five sets of variables and up to three factors for each for a combination of 15 cells. We consider the following sets of variables:

- 22 indicators (one quarterly + 21 monthly) A parsimonious set of macroeconomic variables for the UK, supplemented by Euro Area survey and activity indicators. Full list in the model description.

- The same set of 22 indicators detailed above augmented with nine UK financial market series (FTSE All-Share, three sterling exchange rates, three gilt yields, Brent crude); 31 variables total.

- 24 indicators corresponding to the panel of Anesti et al. (2022), with the closest available substitutes for series we do not have in our pipeline (gilt yields proxy for the term spread).

- A broader panel of 39 indicators, all surveys and financial data detailed above, with the addition of labour market variables and house price series.
- An exhaustive pool of 65 indicators from the data pipeline; included as an upper bound on panel size.

Within each set of variables, factor counts of one, two, and three were tested. For each (variables, factor) combination, the model is estimated for each point on the 256-cell hyperparameter grid described below:

- Trend innovation prior scale: $S_\tau \in \{0.005, 0.010, 0.020, 0.030\}$
- Trend innovation prior degrees of freedom: $\nu_\tau \in \{1, 5, 7, 10\}$
- Factor VAR Minnesota tightness: $\tau_{\text{VAR}} \in \{0.5, 0.7, 0.9, 1.1\}$
- Outlier Student-t degrees of freedom: $\nu_o \in \{4, 8\}$
- Loading Minnesota tightness: $\tau_\lambda \in \{0.5, 0.8\}$

Each cell of the grid represents an independent backtest using a screening MCMC budget of 4,000 Gibbs draws (1,000 burn-in) per re-estimation. The screening budget is appropriate for ranking configurations on RMSE but not for producing final-quality posteriors. For the parameterisation (number of factors, hyper-parameter combination) achieving the best post-COVID Last h=0 RMSE for each variable set based on this screening, the model is then re-estimated using the production budget of (8,000 / 2,000) and the resulting post-COVID Last h=0 RMSE is recorded. For each variable set, Table 2 reports the resulting RMSEs for each variable set and the associated number of factors.

Table 2: Best post-COVID Last $h = 0$ RMSE per variable set.

	Factors	Post-COVID RMSE
`22`	3	**0.218**
`anesti`	3	0.265
`39`	2	0.298
`22 + financial`	2	0.315

Notes: Best post-COVID Last $h = 0$ RMSE for each variable set, across the 256-cell hyper-parameter grid described in the text and $r \in \{1, 2, 3\}$ factors. For each variable set, the leading specification is chosen based on 4,000 Gibbs draws (1,000 burn-in), with the RMSEs reported based on a full 8,000 draw (2,000 burn-in) run.

As Table 2 shows, the compact `v22` panel with three factors dominates with a post-COVID RMSE of 0.218. The production specification is therefore 22 indicators with three common cyclical factors plus a stochastic GDP trend. We report in Table 3 the performance of our chosen specification across other sample periods.

Beyond the five main dimensions described in the grid above, we also consider a broader 836-configuration sweep, varying the Minnesota lag-decay parameters α_{VAR} and α_λ, the

Table 3: Production model nowcast RMSE by sub-sample (v22, three factors).

	Post-COVID (2022Q1–2026Q1)	Ex-COVID (excl. 2020–21)	Full sample (2018Q1–2026Q1)
Last h=0 RMSE	**0.218**	0.198	2.060
N (target quarters)	17	25	33

Notes: Last h=0 RMSE (QoQ %, one observation per target quarter) for the production specification at the production MCMC budget (8,000 draws, 2,000 burn-in). Post-COVID is the headline sub-sample. The full sample is dominated by the extreme 2020 outturns: the Q2 and Q3 2020 nowcast errors alone exceed 20 annualised percentage points, inflating the full-sample RMSE by an order of magnitude.

level of the factor and idiosyncratic stochastic-volatility priors, the volatility-of-volatility scales, and the idiosyncratic AR coefficient priors $\rho_{i,1}, \rho_{i,2}$. Within sensible neighbourhoods of the production values, perturbing these axes has small impact on post-COVID Last h=0 RMSE relative to the five main dimensions; they are therefore held fixed at the values listed in the model description's prior calibration table for the main grid above.

7.4 Computational resources

The sweep over calibration grid for each (variables, factors) pair was run on UCL's Myriad HPC cluster. Each cell of the grid – one (variables, factor count, hyperparameter configuration, seed) tuple – is an independent job: a single backtest covering 2018Q1-2026Q1 on a single CPU core with single-threaded BLAS. To ensure reproducibility all jobs are run on a fixed node class with identical CPUs.

Per-job resources. An eight-year weekly backtest comprises 32 full re-estimations and roughly 420 partial-re-estimation (mini-Gibbs) updates, taking 20–24 wall-clock hours per cell.

7.5 Robustness

The chosen configuration is subjected to three robustness checks.

Seed sensitivity and convergence. The top five configurations from the 22 variable three factor sweep are re-run across a panel of RNG seeds, with full posterior parameter draws retained at each re-estimation, to verify that the cross-cell ranking is not an artefact of a single seed. Geweke p-values and effective sample sizes for the leading factor loading are computed for each (configuration, seed) pair and indicate convergence.

Local sensitivity around the winner. A fine-grained local sweep replicates the best performing configuration at neighbouring values of S_τ with all other hyperparameters held fixed, to confirm that 0.010 is not a knife-edge optimum and to characterise the local response of post-COVID RMSE to perturbations of the trend prior.

7.6 Benchmark comparison

As a headline external benchmark for our UK GDP nowcasts we use a private sector nowcast provider, which publishes a single best-before-release forecast per target quarter using a related but different dynamic factor model. Forecasts from both models are aligned to the same target quarters, release dates, and information sets. On post-COVID Last h=0 RMSE – the metric on which the benchmark's published performance is most directly comparable with our model – the production model achieves a nearly identical RMSE.

7.7 Summary of the production specification

Our backtesting and calibration procedure set out above chooses the following variables, factor structure and hyperparameter set:

- *Variable panel*: 22 indicators (one quarterly GDP target plus 21 monthly indicators), as documented in Table 1.
- *Factor count*: three common cyclical factors plus a stochastic GDP trend.
- *Identification*: As documented in the model description (only GDP loads on the trend; three diagonal-unity normalisers fixed at GDP, GfK Consumer Confidence, and Industrial Production).
- *Hyperparameters*: $S_\tau = 0.010$, $\nu_\tau = 1$, $\tau_{\text{VAR}} = 0.9$, $\nu_o = 8$, $\tau_\lambda = 0.8$. All other priors as documented in the model description.
- *MCMC*: 8,000 Gibbs iterations with 2,000 burn-in (6,000 retained); no thinning. Mini-Gibbs at partial re-estimation: 100 parameter draws $\times$ 1,000 inner iterations.
- *Performance*: post-COVID Last h=0 RMSE of 0.218, almost identical to the external benchmark on the same target quarters.

References

Anesti, N., Galvão, A. B., & Miranda-Agrippino, S. (2022). Uncertain kingdom: Nowcasting gdp and its revisions. *Journal of Applied Econometrics*, *37*(1), 42–62.

Antolin-Diaz, J., Drechsel, T., & Petrella, I. (2024). Advances in nowcasting economic activity: The role of heterogeneous dynamics and fat tails. *Journal of Econometrics*, *238*(2), 105632.

Durbin, J., & Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, *89*(3), 603–616.

Jacquier, E., Polson, N. G., & Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, *122*(1), 185–212.

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, *65*(3), 361–393.

Mariano, R. S., & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, *18*(4), 427–443.