

Covid-19 forecasting and implications for macroeconomic forecasting

Jennifer L. Castle, Magdalen College, University of Oxford
with Jurgen A. Doornik and David F. Hendry

Start-up Workshop: Model invariance and constancy in the face of large shocks to the Norwegian macroeconomic system

March 22nd, 2022

Delighted to be joining team on Model invariance and constancy in the face of large shocks to the Norwegian macroeconomic system.

English version of website

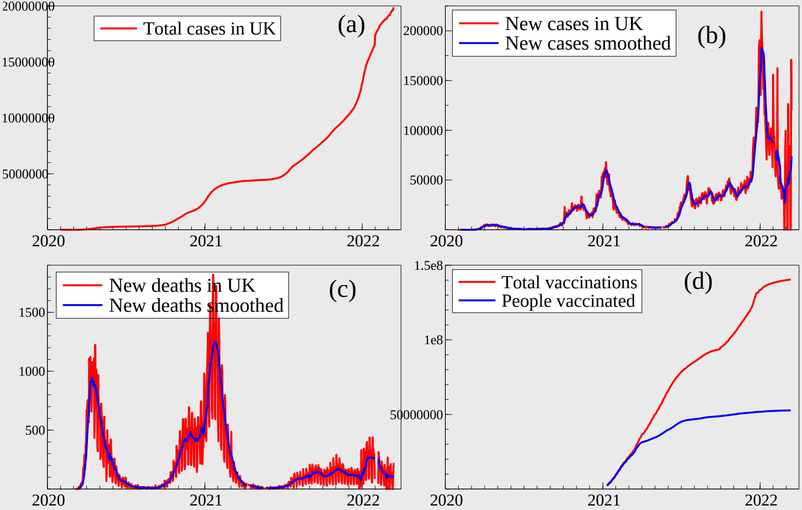
(1) **Introduction**

(2) Forecasting Covid-19 confirmed cases and deaths

(3) Forecasting UK unemployment during the pandemic

(4) Conclusions

UK data for Covid-19 to 15 March 2022   Source: OWID



- Explosive roots then slowing, followed by further explosive roots: **highly non-stationary**
- Abrupt shifts due to measurement errors and policy responses: **structural breaks and outliers**

Methodologies for reporting data are also non-stationary:

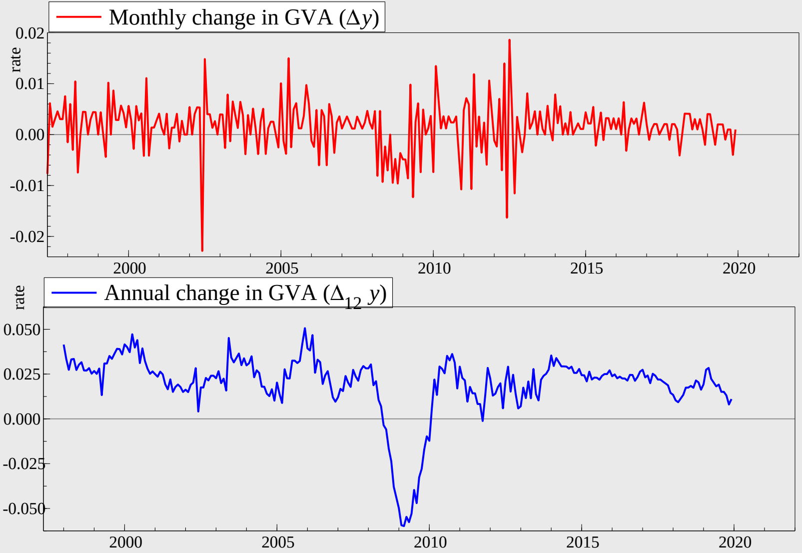
- ramping up of testing and now reduced testing: **stochastic trends**
- changing definitions and reporting errors (inclusion of care home cases, definition of covid-related deaths): **distributional shifts**

Changing seasonality also present.

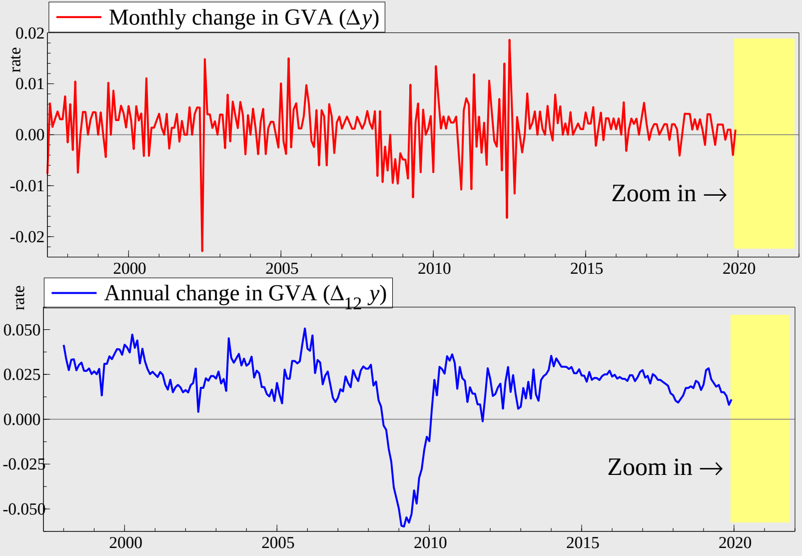
Two sources of non-stationarity interact with underlying data, reporting process, and changing seasonality.

Challenging data to forecast — requires methods that are adaptable to breaks and measurement changes.

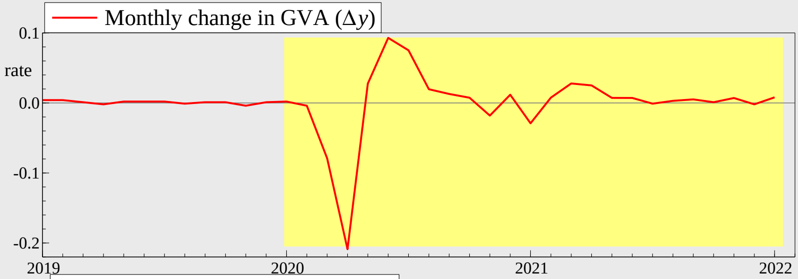
UK Gross Value Added

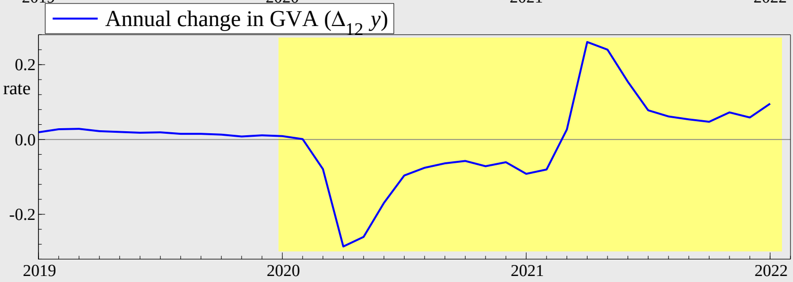


UK Gross Value Added

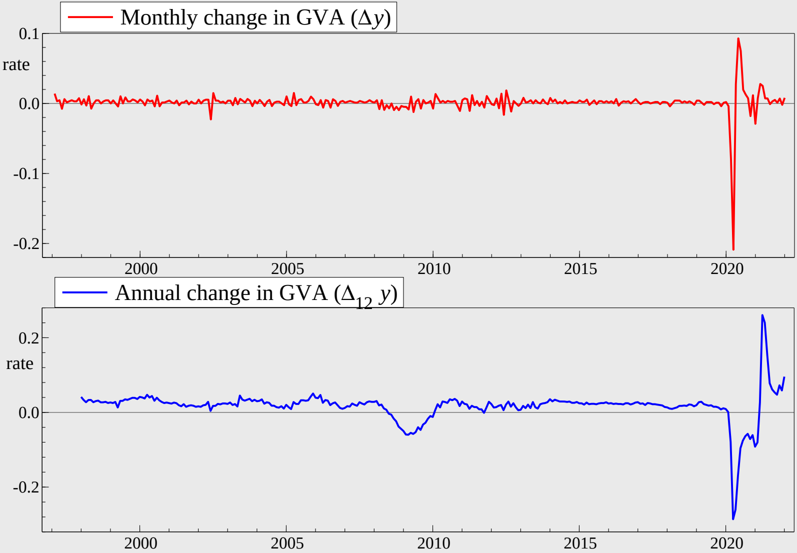


UK Gross Value Added





UK Gross Value Added



Macroeconomic data also highly non-stationary

Large shocks to macroeconomy from pandemic resulting in structural breaks.

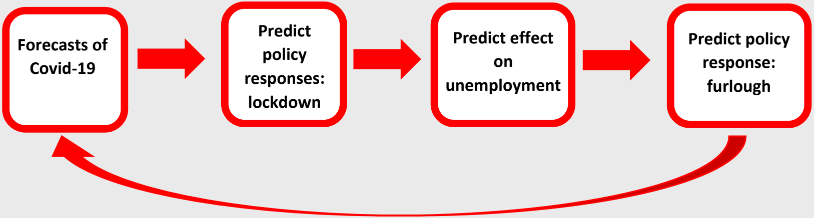
Two-way interaction between pandemic and economy:

- Confirmed cases & deaths grow exponentially $\Rightarrow$ health policy imposes lockdown $\Rightarrow$ fall in output $\Rightarrow$ fall in cases and deaths with lag $\Rightarrow$ relax lockdowns.

Endogenity between health and economic systems, linked by behavioural responses and economic/political policy responses as well as institutional design.

Creates challenges for forecasting both Covid-19 and macroeconomic outcomes, but also suggests similitude in forecasting approaches.

Given close links between health and economic response to pandemic, forecasting devices that are robust to structural breaks should perform well for both pandemic and economic outcomes.



Economic and health responses operate at different frequencies:

- 7 day ahead forecasts for Covid-19. Huge uncertainty in long horizon forecasts.
- 1-3 month ahead forecasts for unemployment. Data available with substantial lag.

Tying together health and economic systems is difficult, but commonalities in data and forecasting procedures to explore.
E.g. combine these methods with Gunnar and Ragnar's work on smooth transition modelling and forecasting in next talk.

Two main approaches to forecasting

- Structural models
  - E.g. SIR models in epidemiology
  - E.g. DSGE models in economics
- Time series models

  E.g. AR; Theta: Assimakopoulos and Nikolopoulos (2000); Cardt: Castle, Doornik, and Hendry (2019); Growth curves: Harvey and Kattuman (2020)

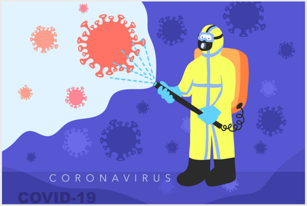
Neither aims to model DGP which depends on host, social, environmental, policy variables, etc. Process too complex to model.

Theoretical models supported by available evidence are crucial to policy-making in economics, epidemiology, ...

- But: history of more simple data-based devices ‘out-forecasting’ ‘structural’ models.
- Shifts in distributions from past lead to systematic mis-forecasting

Our approach is purely statistical: captures trends and breaks, but not causes.

(1) Introduction
(2) **Forecasting Covid-19 confirmed cases and deaths**
(3) Forecasting UK unemployment during the pandemic
(4) Conclusions

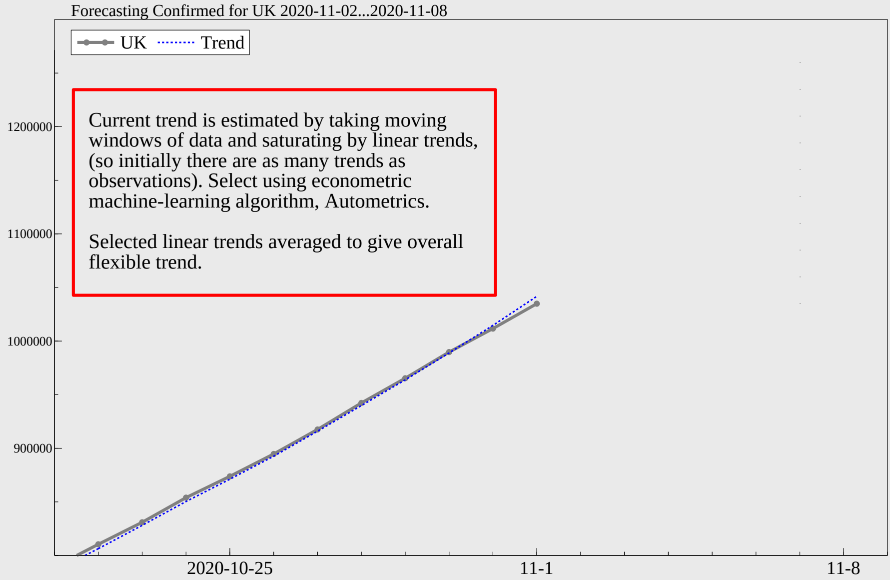


We forecast cumulative daily **confirmed cases** and **deaths**
Data source: Johns Hopkins University/CSSE – thanks!

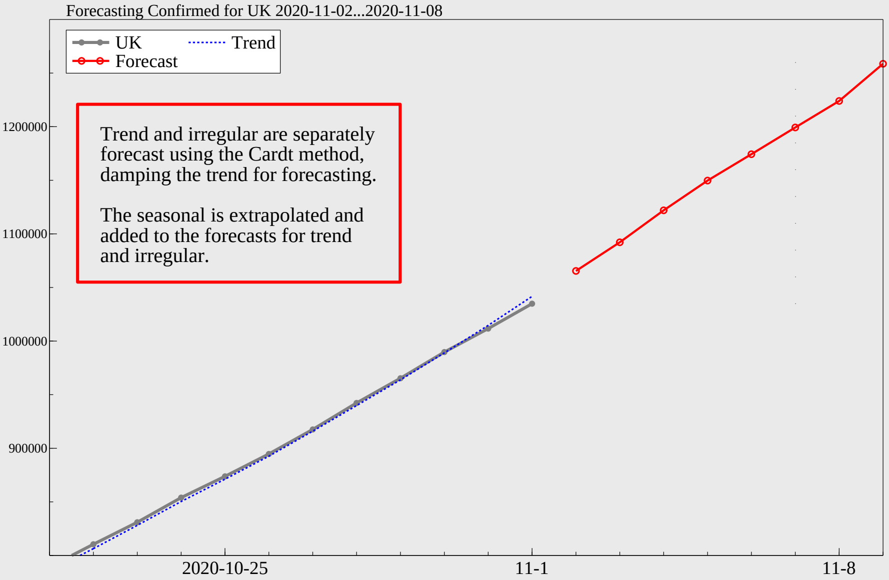
- Started mid-March 2020,
- cumulative daily confirmed cases and deaths,
- $\approx$ 50 countries, $\approx$ 50 US states, $>$ 300 England local authorities, 4 models each: 3200 model estimates each time.
- about 4 times a week,
- Short-term forecasting of the Coronavirus Pandemic (*IJF*),
- Statistical — (*Journal of Clinical Immunology and Immunotherapy*).

Forecasts are hosted at **www.doornik.com/COVID-19**

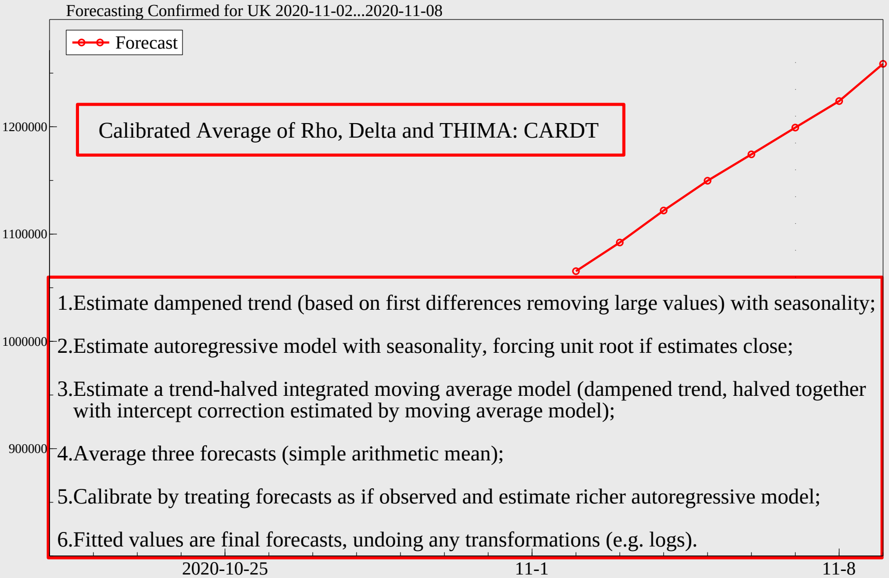
A closer look at our forecasts



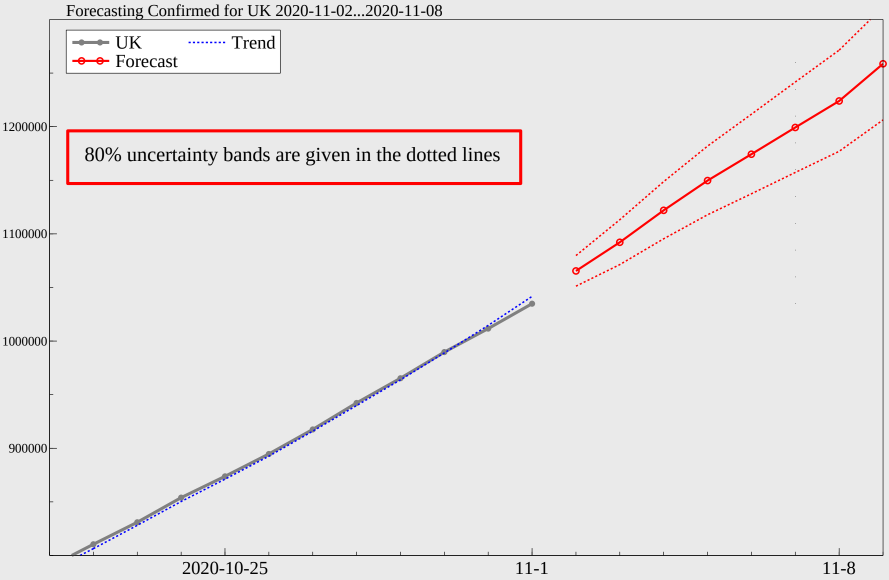
A closer look at our forecasts: current



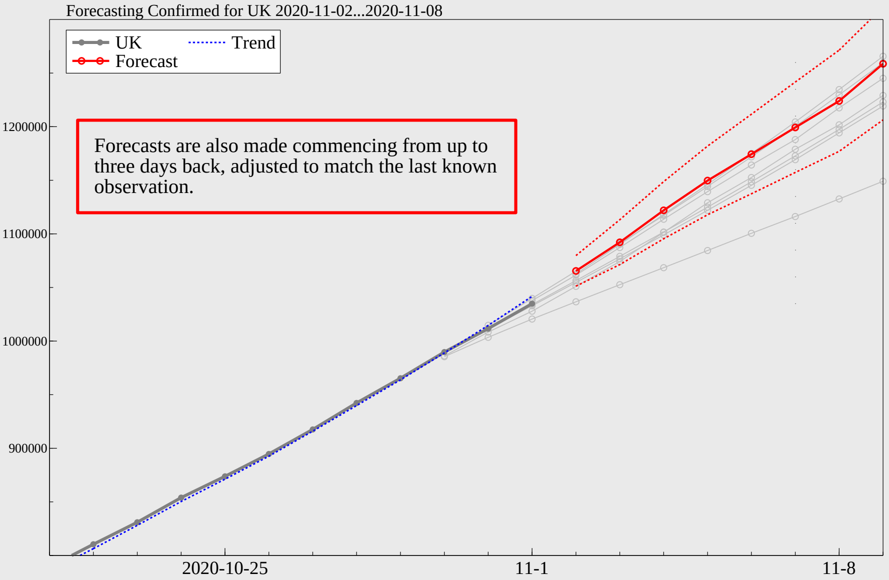
A closer look at our forecasts: cardt forecasts



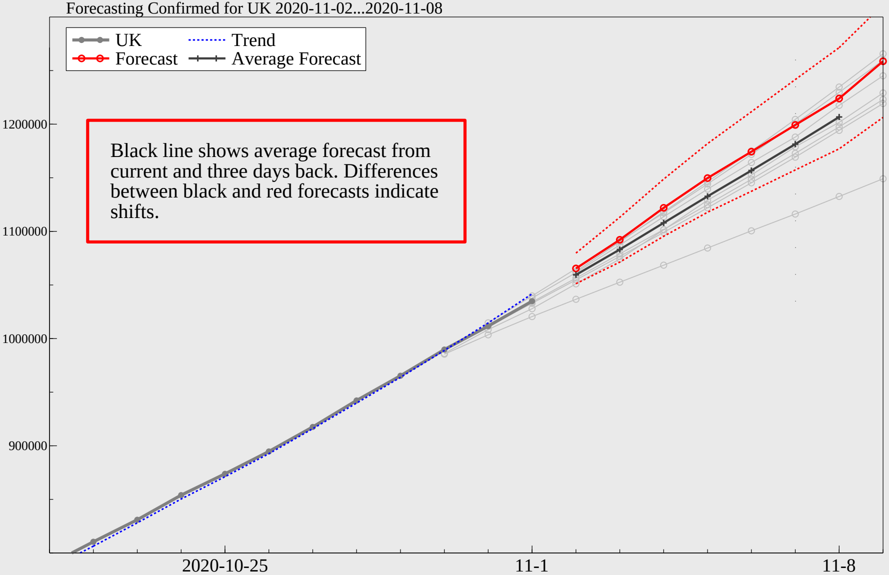
A closer look at our forecasts: current with 80% intervals



A closer look at our forecasts: three days back

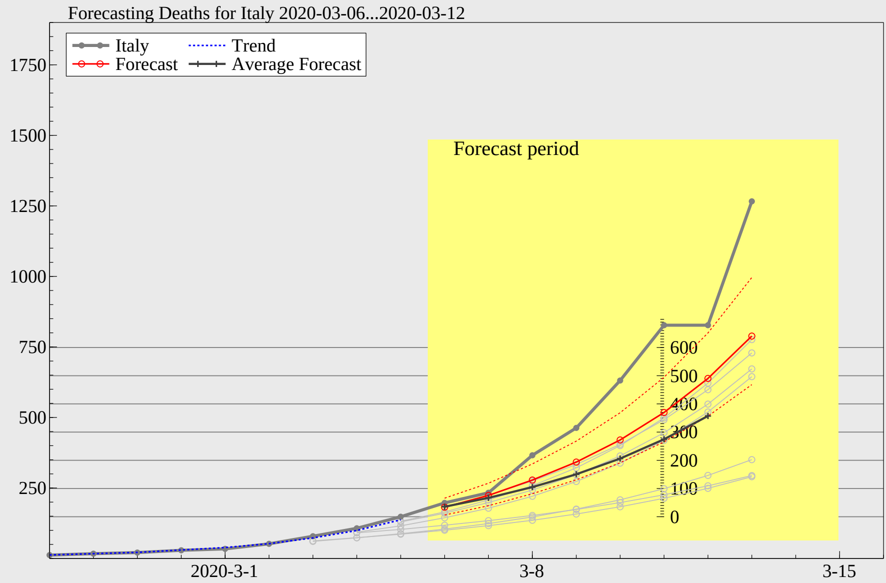


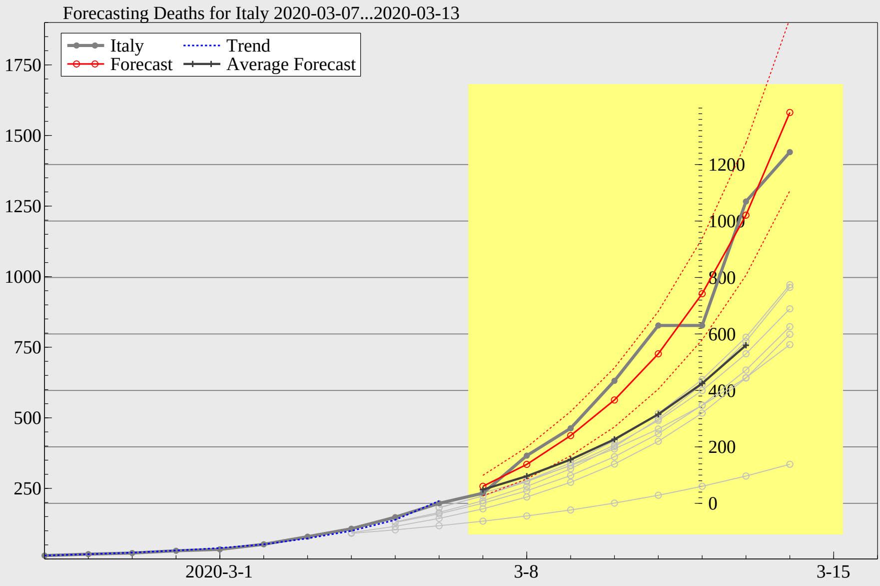
A closer look at our forecasts: with average (adjusted)



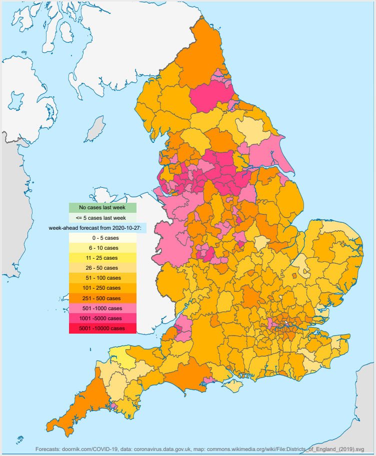
Black line shows average forecast from current and three days back. Differences between black and red forecasts indicate shifts.

Recovery from a break




Forecasting Deaths for Italy 2020-03-07...2020-03-13
Italy
Trend
Forecast
Average Forecast
1750
1500
1250
1000
750
500
250
1200
1000
800
600
400
200
0
2020-3-1
3-8
3-15

Local authority forecasts slideshow

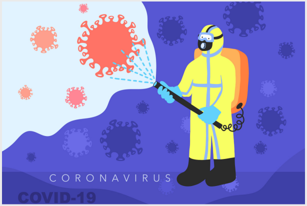


(1) Introduction

(2) Forecasting Covid-19 confirmed cases and deaths

(3) **Forecasting UK unemployment during the pandemic**

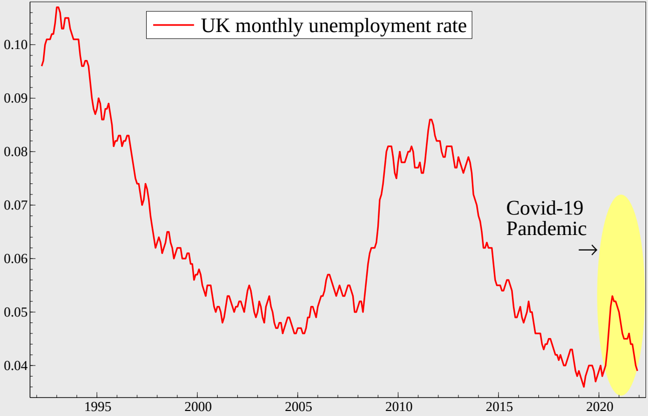
(4) Conclusions



Monthly unemployment data

Unemployment data from labour force survey undertaken by International Labour Organisation. Three-monthly survey of 85,000 individuals.
Monthly frequency reported as mid-month of three-month average time period.
Last obervation in December 2021 is average of unemployment rates over November 2021 – January 2022. Generates monthly lead and lag persistence.

Monthly unemployment data



Given 20% fall in GDP in one month (April 2020) from pandemic the effect on unemployment rates very modest, reaching a peak of 5.3% in September 2020.

Structural break in relationship between output & unemployment.

U is outcome of supply and demand for labour, aggregated across all prospective workers, with labour demand derived from demand for goods and services. **Implies highly complex DGP**.

Instead use a profits proxy, R; assumes unemployment falls when hiring labour is profitable, and increases if not profitable.

$$\boxed{R_t = R_{l,t} - \Delta_{12} y_t - \Delta_{12} p_t}$$

$R_{l,t}$: long-term interest rate;
$\Delta_{12} y_t$: annual change in log Gross Value Added;
$\Delta_{12} p_t$: annual Consumer Price Index inflation rate.

U is outcome of supply and demand for labour, aggregated across all prospective workers, with labour demand derived from demand for goods and services. **Implies highly complex DGP**.

Instead use a profits proxy, R; assumes unemployment falls when hiring labour is profitable, and increases if not profitable.

$$\boxed{R_t = R_{l,t} - \Delta_{12} y_t - \Delta_{12} p_t}$$

$R_{l,t}$: long-term interest rate;
$\Delta_{12} y_t$: annual change in log Gross Value Added;
$\Delta_{12} p_t$: annual Consumer Price Index inflation rate.

Autoregressive distributed lag general model includes:

- Profits proxy; nominal wage inflation; output gap; wage share
- Sufficient dynamics (13 lags on monthly data)
- Saturation estimators for outliers and breaks; Seasonals
- Non-linear transformations

Methodology:

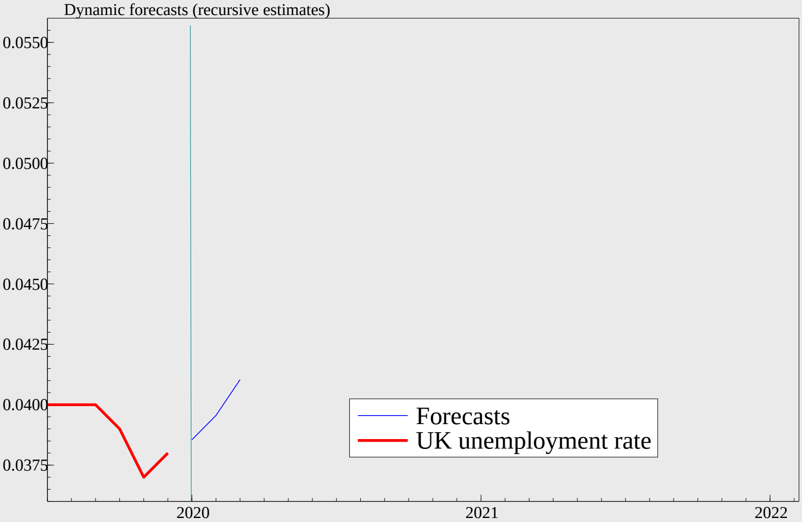
- Set regressors, constant and seasonals as retained and select saturation estimators and non-linearities at $\alpha = 0.0001$.
- **621** (81 fixed) regressors with 215 observations [2002(2)-2019(12)].
- Then select over regressors at $\alpha = 0.001$.
- Transform to stationary representation with ecm.
- Re-select at $\alpha = 0.001$ from equilibrium correction representation including lagged ecm and 12 lags of differenced regressors.

Methodology:

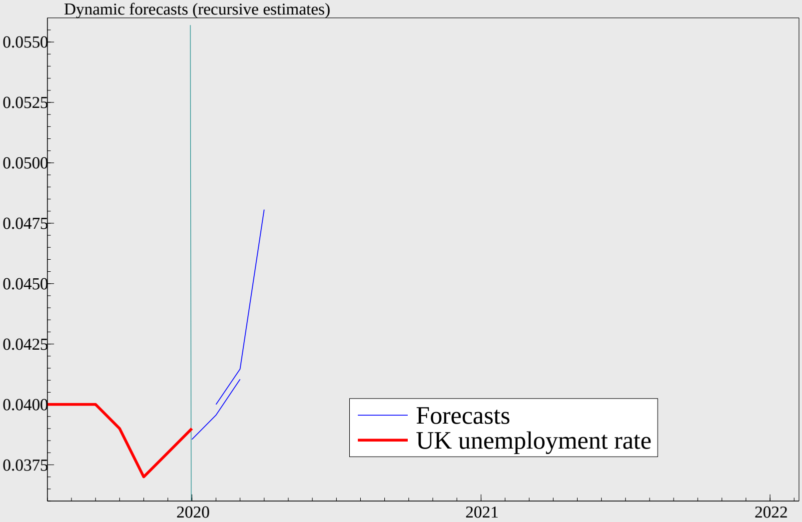
- Set regressors, constant and seasonals as retained and select saturation estimators and non-linearities at $\alpha = 0.0001$.
- **621** (81 fixed) regressors with 215 observations [2002(2)-2019(12)].
- Then select over regressors at $\alpha = 0.001$.
- Transform to stationary representation with ecm.
- Re-select at $\alpha = 0.001$ from equilibrium correction representation including lagged ecm and 12 lags of differenced regressors.

Congruent, well-specified model passing all diagnostic tests.
No indicators or non-linearities retained.
Similar impact of R to annual model of excess demand for unemployment reported in Hendry (2001) over 1865–1991.
Long-run equilibrium unemployment rate of $\approx 5\%$, matching mean over last 150 years, yet here on non-overlapping monthly data.

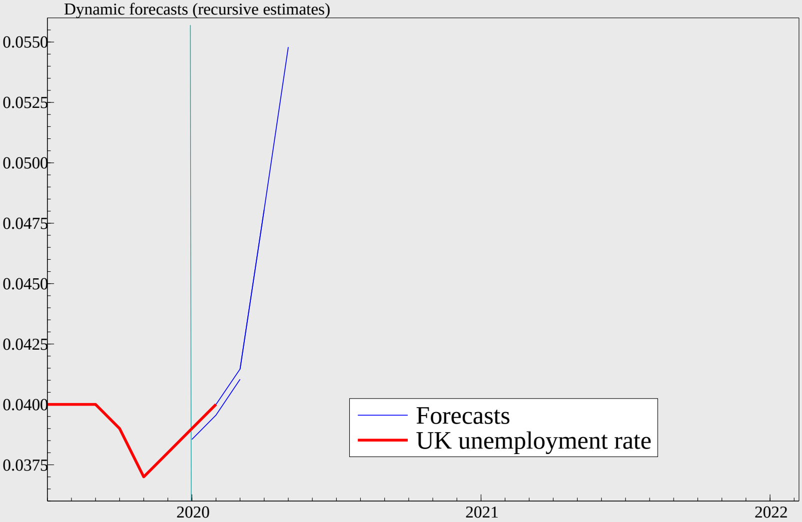
3 month ahead forecasts from econometric model



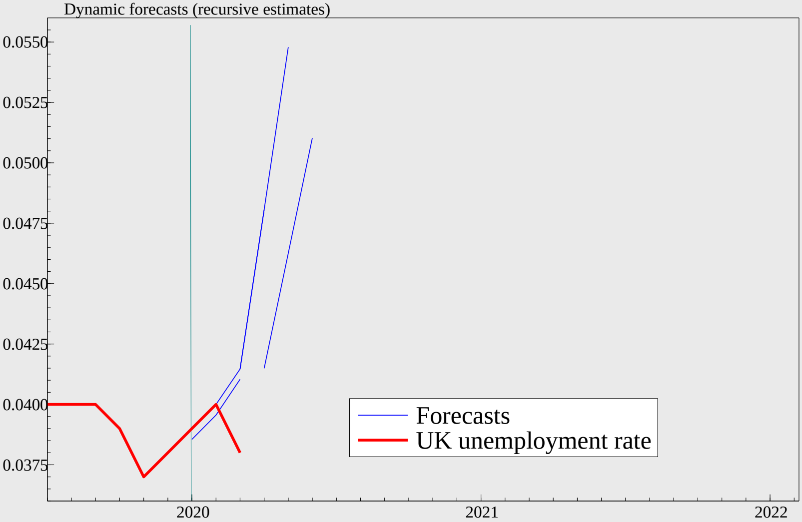
3 month ahead forecasts from econometric model



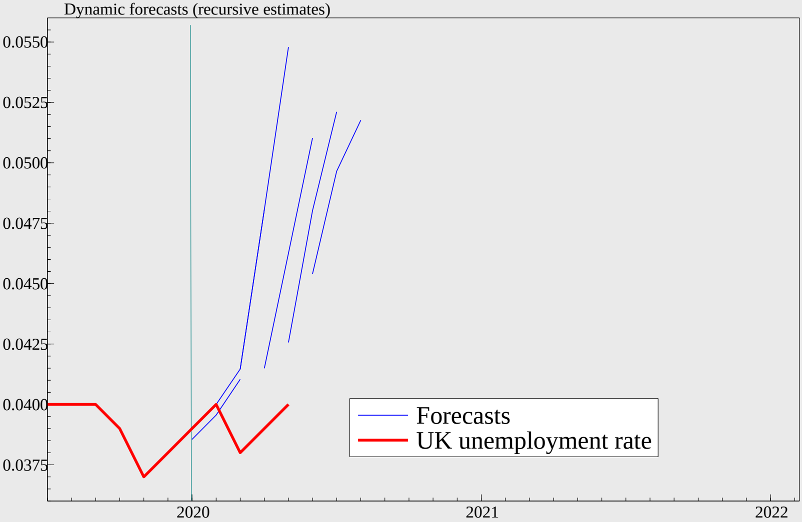
3 month ahead forecasts from econometric model

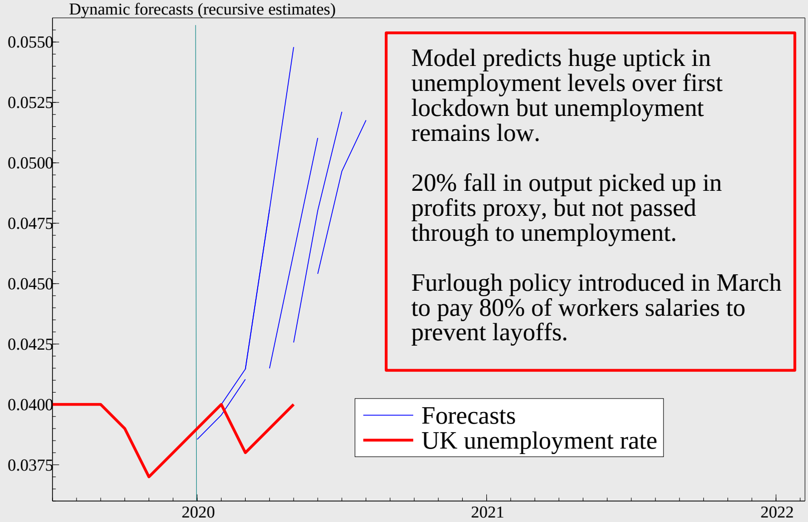


3 month ahead forecasts from econometric model

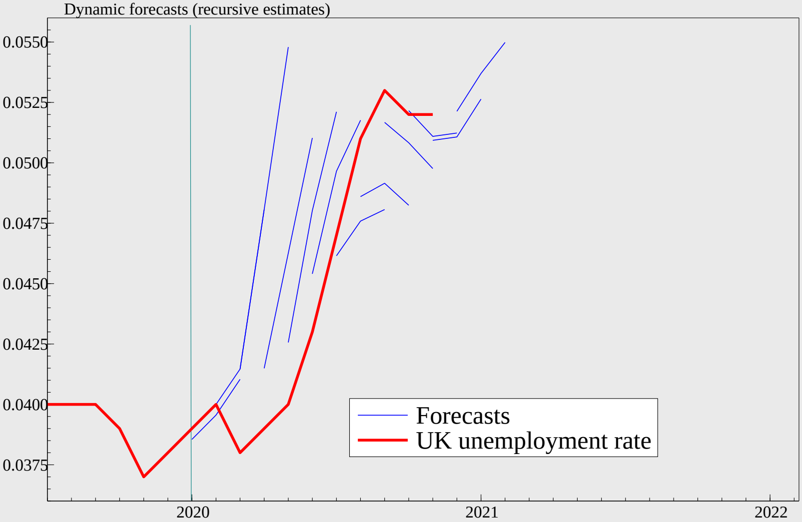


3 month ahead forecasts from econometric model

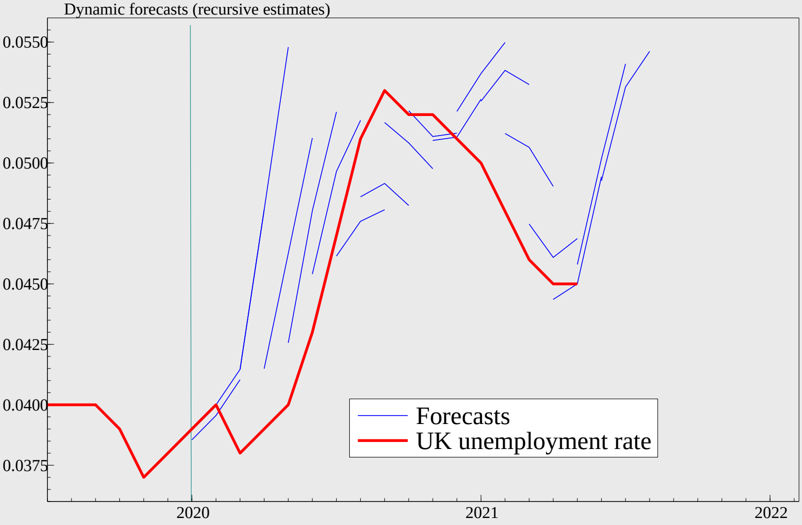



Dynamic forecasts (recursive estimates)
0.0550
0.0525
0.0500
0.0475
0.0450
0.0425
0.0400
0.0375
2020
2021
2022
Model predicts huge uptick in unemployment levels over first lockdown but unemployment remains low.
20% fall in output picked up in profits proxy, but not passed through to unemployment.
Furlough policy introduced in March to pay 80% of workers salaries to prevent layoffs.
Forecasts
UK unemployment rate

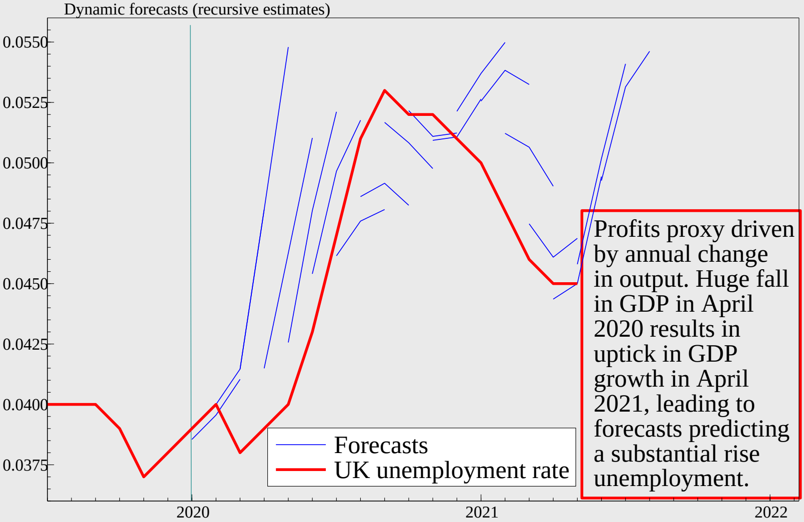
3 month ahead forecasts from econometric model



3 month ahead forecasts from econometric model

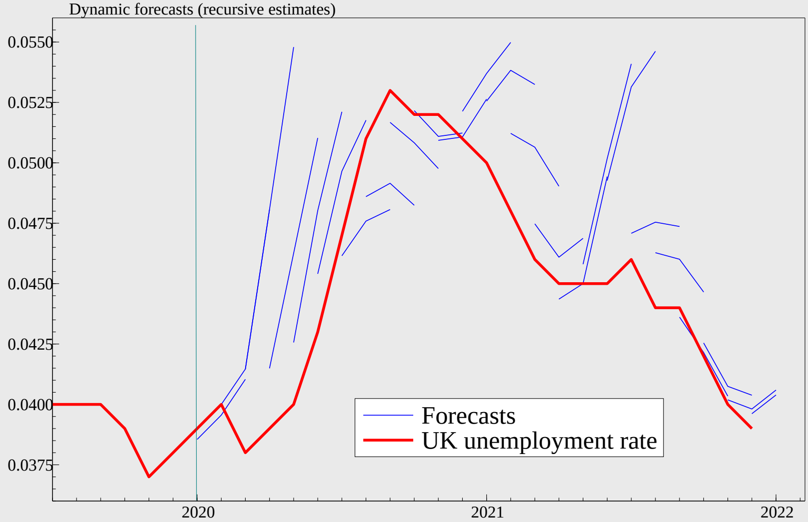


3 month ahead forecasts from econometric model



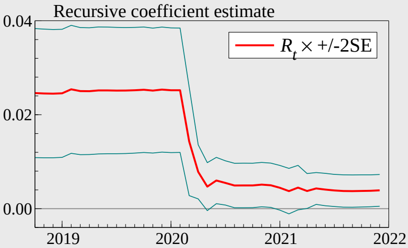
Profits proxy driven by annual change in output. Huge fall in GDP in April 2020 results in uptick in GDP growth in April 2021, leading to forecasts predicting a substantial rise unemployment.

3 month ahead forecasts from econometric model



Failure of superexogeneity

Forecasts are *ex post*, they include contemporaneous data (ΔR_t).
Would usually give an information advantage over *ex ante* forecasts.

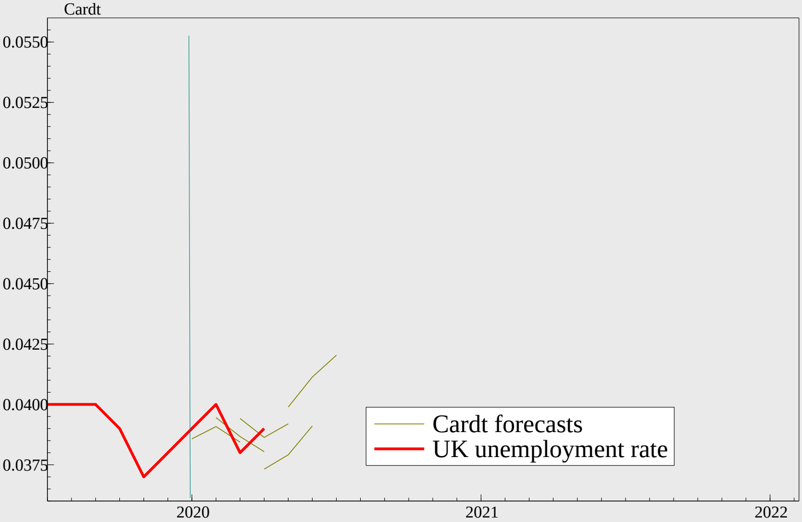


Test by including step dummies interacting with R over pandemic period and select. One significant interaction term retained (March 2020), full sample R coefficient close to 2019(12).

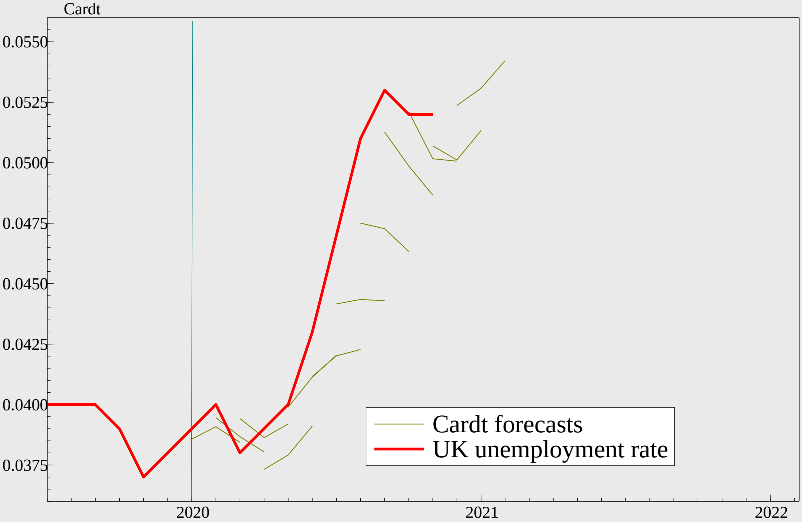
Changing relationship between profits proxy and unemployment rate due to policy intervention.

A form of Multiplicative Indicator Saturation (MIS) – successfully detect induced shifts in estimated models following policy intervention.

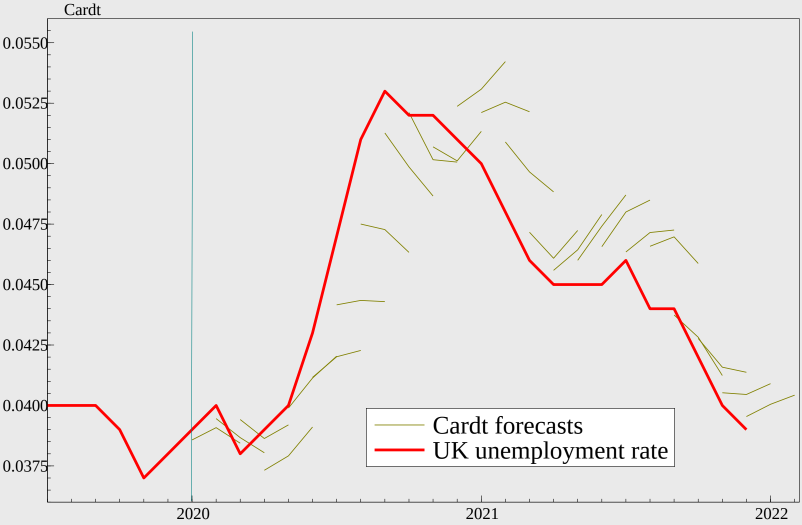
How our statistical forecasts fare: Cardt



How our statistical forecasts fare: Cardt



How our statistical forecasts fare: Cardt



Robust forecasts using the economic model

If structural break from pandemic caused equilibrium mean to shift, robust version of economic model can improve forecast performance.

The model of unemployment is:

$$\widehat{\mathrm{Ur}}_t = \widehat{\beta}_0 + \widehat{\beta}_1 \mathrm{Ur}_{t-1} + \widehat{\mathbf{\Gamma}}' \mathbf{X}_t \quad (1)$$

where $\mathbf{X}_t$ includes the profits proxy and lags of the wage share and nominal wage inflation.

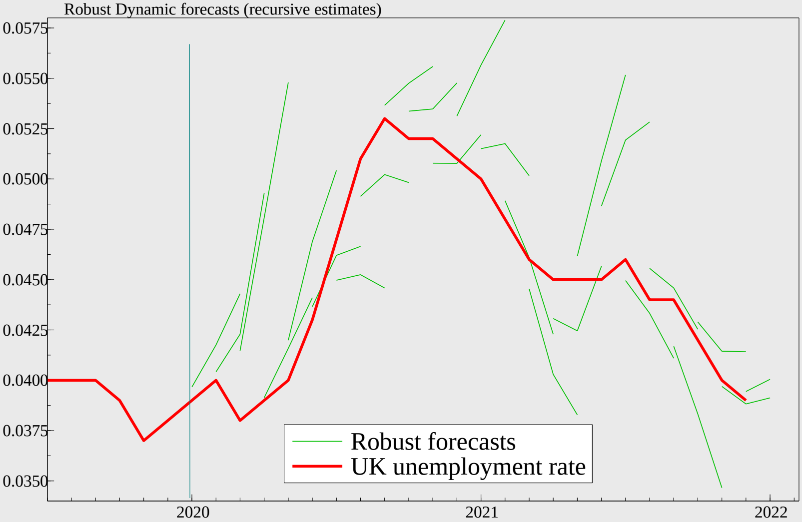
Robust forecasts are obtained by differencing the model after the parameters are estimated:

$$\widehat{\mathrm{Ur}}_{T+1|T} = \mathrm{Ur}_T + \widehat{\beta}_1 \Delta \mathrm{Ur}_T + \widehat{\mathbf{\Gamma}}' \Delta \mathbf{X}_T \quad (2)$$

thereby removing the intercept and embedded means in the exogenous regressors.

In practice a local average could replace Ur_T for a smoother forecast device, see Martinez, Castle, and Hendry (2021).

3 month ahead Robust forecasts of UK unemployment



How the forecasts perform

	Model		Robust		Cardt	
H	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1	0.19	3.41	*0.17*	*3.14*	**0.16**	**2.84**
2	0.40	6.98	*0.38*	*6.78*	**0.30**	**5.26**
3	*0.60*	*10.87*	0.63	11.66	**0.43**	**7.67**

Table: Forecast evaluation: bold indicates smallest root mean square forecast error (RMSE) and mean absolute percentage error (MAPE), and italics denote second smallest. H is forecast horizon. Evaluated over 2020(1)-2021(12), giving 24 1-step ahead forecasts.

	Model		Robust		Cardt	
H	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1	0.19	3.41	*0.17*	*3.14*	**0.16**	**2.84**
2	0.40	6.98	*0.38*	*6.78*	**0.30**	**5.26**
3	*0.60*	*10.87*	0.63	11.66	**0.43**	**7.67**

Table: Forecast evaluation: bold indicates smallest root mean square forecast error (RMSE) and mean absolute percentage error (MAPE), and italics denote second smallest. H is forecast horizon. Evaluated over 2020(1)-2021(12), giving 24 1-step ahead forecasts.

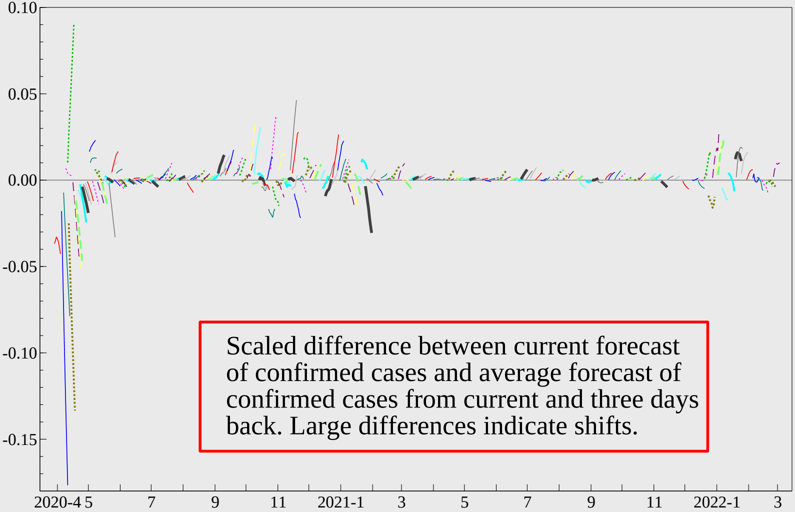
- Cardt dominates forecast performance at all horizons.

	Model		Robust		Cardt	
H	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1	0.19	3.41	*0.17*	*3.14*	**0.16**	**2.84**
2	0.40	6.98	*0.38*	*6.78*	**0.30**	**5.26**
3	*0.60*	*10.87*	0.63	11.66	**0.43**	**7.67**

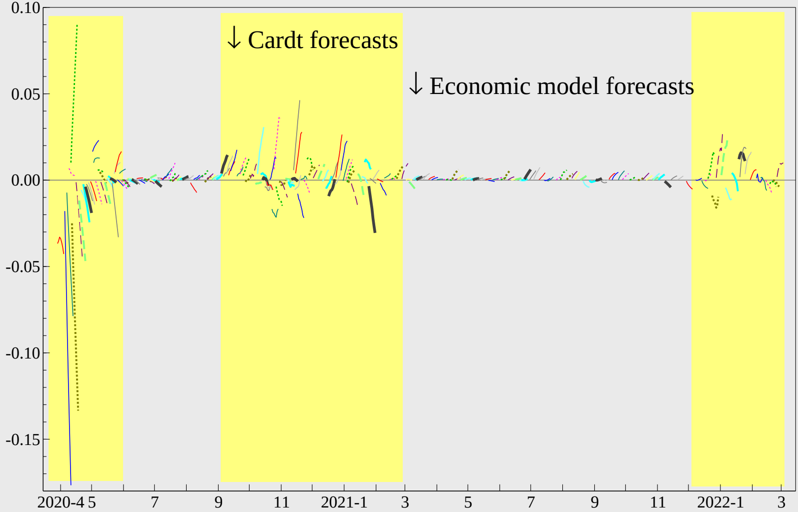
Table: Forecast evaluation: bold indicates smallest root mean square forecast error (RMSE) and mean absolute percentage error (MAPE), and italics denote second smallest. H is forecast horizon. Evaluated over 2020(1)-2021(12), giving 24 1-step ahead forecasts.

- Cardt dominates forecast performance at all horizons.
- Robust is preferred to model based forecasts for shorter horizons, where the robust device is designed for 1-step ahead forecasts.
- Model based forecasts poorest despite conditioning on contemporaneous information.

Can we do better by switching using Covid-19 forecasts?



Can we do better by switching using Covid-19 forecasts?



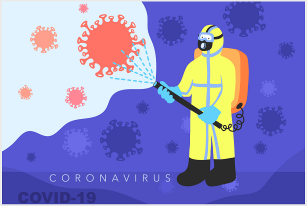
Can we improve our forecasts?

	Model		Robust		Cardt		Switching		Average	
H	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1	0.19	3.41	0.17	3.14	0.16	2.84	0.16	2.80	**0.15**	**2.56**
2	0.40	6.98	0.38	6.78	0.30	5.26	**0.29**	**5.11**	0.30	5.17
3	0.60	10.87	0.63	11.66	**0.43**	**7.67**	0.47	8.58	0.46	8.25

Table: Switching denotes 0/1 weights on Model and Cardt with weight 1 on Cardt for 2020(3)(4)(5)(9)(10)(11)(12),2021(1)(2)(12). Average has equal weights on Model, Robust and Cardt.

- Cardt, switching and average all have similar forecast performance.
- All substantially improve on model forecasts.

(1) Introduction
(2) Forecasting Covid-19 confirmed cases and deaths
(3) Forecasting UK unemployment during the pandemic
(4) **Conclusions**



Forecasting under spotlight during pandemic.
But data is messy: need methods that can handle non-stationary data subject to breaks and revisions.

Cannot avoid forecast errors, but want to recover quickly afterwards

Extrapolative statistical forecasts:

- can supplement explanatory models — often perform better as soon as breaks occur.
- but cannot be used to assess policy interventions or undertake scenario analysis.

Need both types of forecasting model.

Large forecast errors for econometric model relative to those of Cardt show success of furlough scheme in maintaining employment during lockdown.

Little evidence that using Covid-19 forecasts to signal structural breaks in macroeconomy and switch to more adaptive forecasts helps.

References I

Assimakopoulos, V. and K. Nikolopoulos (2000).
The theta model: a decomposition approach to forecasting.
International Journal of Forecasting 16(4), 521 – 530.

Castle, J. L., J. A. Doornik, and D. F. Hendry (2019).
Some forecasting principles from the M4 competition.
Economics papers 2019-w01, Nuffield College, Oxford University.

Clements, M. P. and D. F. Hendry (1998).
Forecasting Economic Time Series.
Cambridge: Cambridge University Press.

Harvey, A. C. and P. Kattuman (2020).
Time series models based on growth curves with applications to forecasting coronavirus.
Harvard Data Science Review Special issue 1.

Hendry, D. F. (2001).
Modelling UK inflation, 1875-1991.
Journal of Applied Econometrics 16, 255–275.

Martinez, A. B. and Castle, J. L. and Hendry, D. F. (2021).
Smooth Robust Multi-Step Forecasting Methods.
Advances in Econometrics Forthcoming.

Model of aggregate unemployment to 2019(12)

$$\widehat{U}_t = \underset{(0.0005)}{0.0018} + \underset{(0.007)}{0.976}U_{t-1} + \underset{(0.007)}{0.025}R_t - \underset{(0.007)}{0.030}R_{t-2} - \underset{(0.006)}{0.021}\Delta_{12}w_{t-2}$$

$$+\underset{(0.004)}{0.024}\,(w-p-y+l)_{t-2} + \text{seasonals}$$

$\widehat{\sigma} = 0.09\%$; $F_{ar}(7,191) = 1.26$; $F_{arch}(7,201) = 0.43$; $\chi^2(2) = 0.44$;
$F_{hetero}(21,193) = 1.04$; $F_{reset}(2,196) = 1.73$; $T = 2002(2) - 2019(12)$

Solved long-run solution:

$$\boxed{\widehat{d} = U - \underset{(0.016)}{0.049} + \underset{(0.268)}{0.46}\,R + \underset{(0.364)}{1.32}\,\Delta_{12}w - \underset{(0.468)}{1.49}\,(w-p-y+l)}$$

Climate Econometrics

UNIVERSITY OF OXFORD

$$\widehat{U}_t = \underset{(0.0005)}{0.0018} + \underset{(0.007)}{0.976}U_{t-1} + \underset{(0.007)}{0.025}R_t - \underset{(0.007)}{0.030}R_{t-2} - \underset{(0.006)}{0.021}\Delta_{12}w_{t-2}$$
$$+\underset{(0.004)}{0.024}(w - p - y + l)_{t-2} + \text{seasonals}$$
$$\widehat{\sigma} = 0.09\%;\ F_{ar}(7, 191) = 1.26;\ F_{arch}(7, 201) = 0.43;\ \chi^2(2) = 0.44;$$
$$F_{hetero}(21, 193) = 1.04;\ F_{reset}(2, 196) = 1.73;\ T = 2002(2) - 2019(12)$$

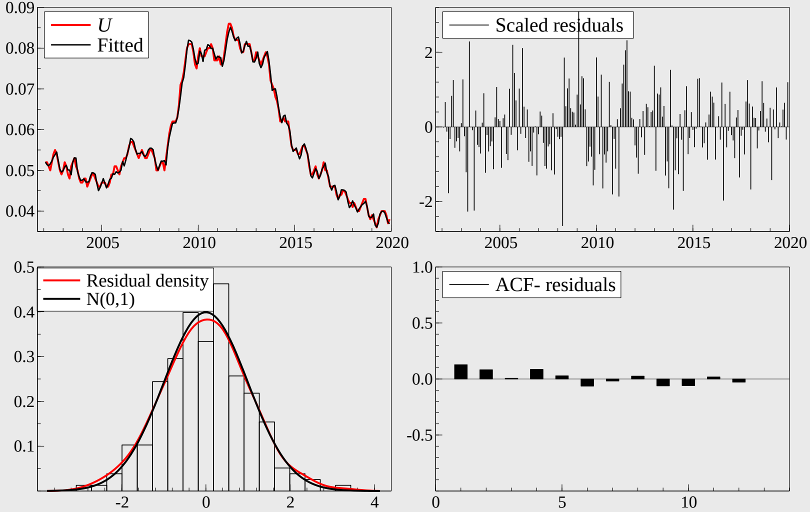
Solved long-run solution:

$$\boxed{\widehat{d} = U - \underset{(0.016)}{0.049} + \underset{(0.268)}{0.46}R + \underset{(0.364)}{1.32}\Delta_{12}w - \underset{(0.468)}{1.49}(w - p - y + l)}$$

Long-run equilibrium unemployment rate of $\approx 5\%$, matching mean over last 150 years, yet here on non-overlapping monthly data.

No indicators or non-linearities retained.

Model fit


U
Fitted
0.09
0.08
0.07
0.06
0.05
0.04
2005
2010
2015
2020
Scaled residuals
2
0
-2
2005
2010
2015
2020
Residual density
N(0,1)
0.5
0.4
0.3
0.2
0.1
-2
0
2
4
ACF- residuals
1.0
0.5
0.0
-0.5
0
5
10

$$\widehat{\Delta U_t} = \underset{(0.0002)}{+0.0003} + \underset{(0.067)}{0.20}\,\Delta U_{t-1} + \underset{(0.009)}{0.024}\Delta R_t - \underset{(0.002)}{0.013}\widehat{d}_{t-1} + \text{seasonals}$$

$$\widehat{\sigma} = 0.09\%;\ R^2 = 0.65;\ F_{ar}(7,193) = 1.54;\ F_{arch}(7,201) = 1.06;$$

$$\chi^2(2) = 0.85;\ F_{hetero}(17,197) = 0.94;\ F_{reset}(2,198) = 1.93;$$