Nowcasting UK consumption and investment with monthly output components and real-time indicators

ANA BEATRIZ GALVÃO (WARWICK), RORY MACQUEEN (ESCOE) AND STEPHEN WRIGHT (BIRKBECK)

How do we improve early estimates of macroeconomic aggregates?

Initial releases are revised, often significantly. How can we minimise this?

- New data sources (five years of monthly sectoral GDP, real-time indicators)
- More frequent input-output tables
- MIDAS techniques mixed data
- Literature
  - Bayesian estimation of UMIDAS model (Carriero, Clark and Marcellino (2018) for priors)
  - Carreiro, Galvão & Kapetanios (2019) combining single-indicator forecasts
  - Real-time indicators (e.g. Kapetanios & Papaillas (2021))

	Investment Revisions		Consumption Revisions	
Initial Release:	Mean	standard dev	Mean	standard dev
1 (1st, M2)	0.55	2.22	-0.01	0.56
1 (QNA)	0.26	2.12	0.01	0.55
2	0.13	2.32	0.00	0.51
3	0.21	2.39	0.01	0.52
4	0.28	2.20	0.01	0.51
5	0.02	2.24	-0.04	0.52
6	0.09	2.19	-0.02	0.48
7	-0.01	1.90	0.00	0.48
8	0.07	2.05	0.00	0.46
9	0.08	2.03	-0.03	0.43
10	0.15	1.96	-0.02	0.41
11	0.12	1.87	-0.02	0.40
12	0.07	1.75	-0.02	0.37
13	0.18	1.26	0.00	0.29
14	0.00	1.03	-0.02	0.17
15	0.04	0.60	0.00	0.12

Two approaches in this project

Mixed-data sampling

1. Model
2. Data
3. Results

Input-output tables

1. Model
2. Data
3. Results
4. Future research

UMIDAS model for single indicator forecast

$$y_t = \alpha_{L,0} + \sum_{i=1}^{py} \alpha_{L,i}\, y_{t-i} + \sum_{j=0}^{px-1} \beta_{L,j+1}\, x^{(m)}_{t-(\frac{L+j}{m})} + \varepsilon_{t,L}$$

y_t = consumption, investment (2022Q1). Sampled quarterly (t = 1,2, … , T).

$x^{(m)}$ = Traditional survey indicators (GfK, OECD, CBI, Retail sales);
Monthly sector GVA;
VAT diffusion indices;
Vacancies;
Card spending.

$x^{(m)}$ sampled m times more frequently than y_t

e.g. for monthly data $m = 3$ and we observe x_t for $t = \frac{1}{3}, \frac{2}{3}, 1, \frac{4}{3}, \frac{5}{3}, 2, \ldots, T$

py = 1

px = 12

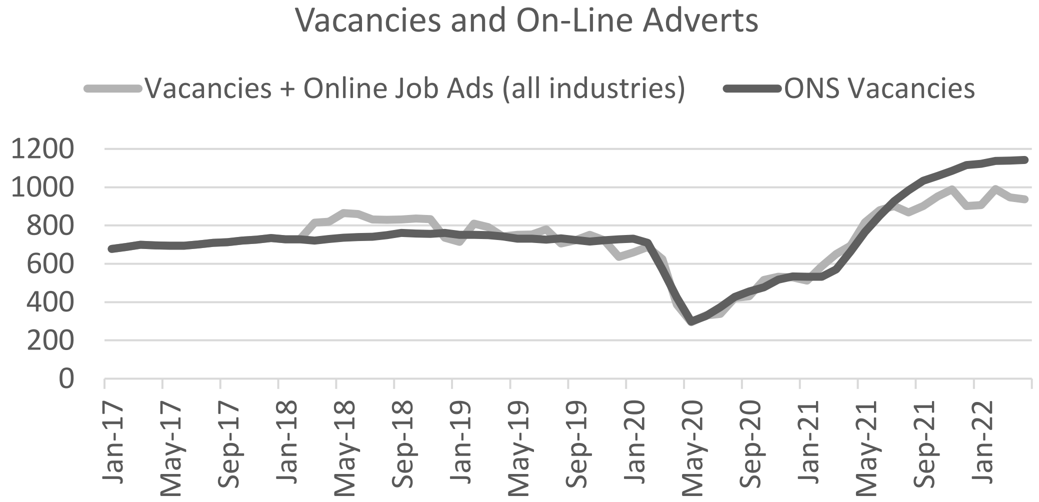
L = 0, except for monthly sector GVA (try both 0 and 1)

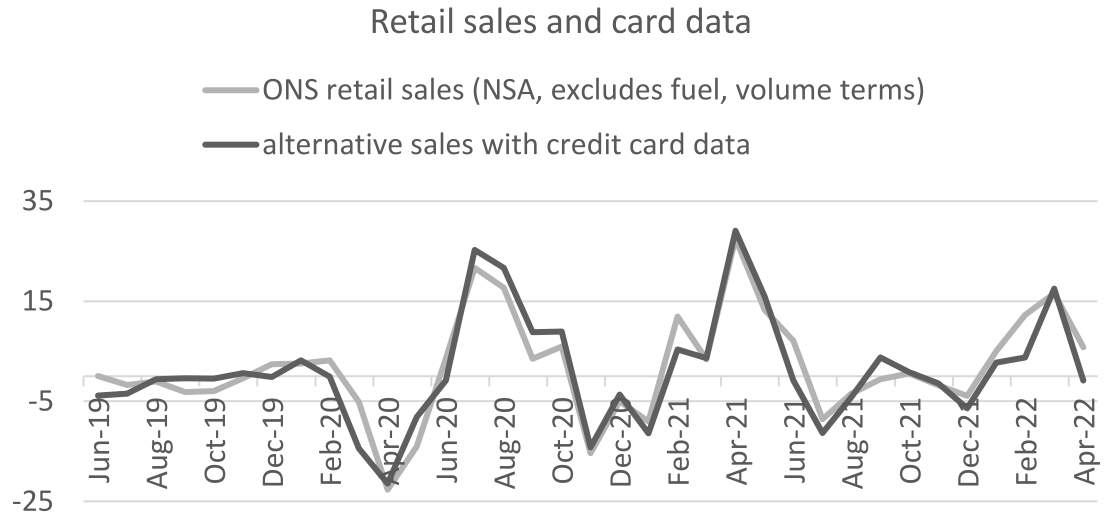
Benchmarks: first estimate; AR-only (β=0)

Augmenting series with weekly indicators

$$x_t = \beta_0 + \sum_{i=1}^{4} \beta_i \, z_{t-(\frac{i}{4})} + \epsilon_t$$

x_t = ONS vacancies (left); ONS Retail Sales (right)
z_t = Online job adverts (Adzuna, left); Card payments (BoE CHAPS, right)





Indicator-based models generally reduce bias during pandemic…

	2010Q1-2019Q4	
	Consumption	**Investment**
ONS first estimate	0.10	0.53
AR only	0.03	0.94
Monthly GVA (3mths)	0.03	0.93
Monthly GVA (2mths)	0.02	0.94
Survey indicators	0.03	0.92

	2020Q1-2021Q4	
	Consumption	**Investment**
ONS first estimate	0.54	1.32
AR only	11.93	2.16
Vacancies	-2.17	-3.81
Card spending	2.59	-1.36
Monthly GVA (3mths)	1.65	-0.79
Monthly GVA (2mths)	0.64	-1.64
Survey indicators	2.34	0.85
VAT	1.87	0.25

Before Covid:

- Less biased than ONS first estimates of consumption

During Covid:

- Three models are less biased than ONS first estimates
- Potential bias-reducing gains from using additional data compared with AR

…and reduce RMSFEs too

	2010Q1-2019Q4		Relative RMSFEs (AR only = 1)	
	Consumption	Investment	Consumption	Investment
ONS first estimate	0.69	2.04		
AR only	0.81	2.44		
Monthly GVA (3mths)	0.81	2.44	1.00	1.00
Monthly GVA (2mths)	0.81	2.44	1.00	1.00
Survey indicators	0.80	2.40	0.98	0.99

	2020Q1-2021Q4		Relative RMSFEs (AR only = 1)	
	Consumption	Investment	Consumption	Investment
ONS first estimate	1.45	2.86		
AR only	35.65	11.30		
Vacancies	9.16	6.18	0.26	0.55
Card spending	13.56	8.76	0.38	0.78
Monthly GVA (3mths)	11.74	6.89	0.33	0.61
Monthly GVA (2mths)	9.60	6.08	0.27	0.54
Survey indicators	12.62	10.04	0.35	0.89
VAT	11.96	9.06	0.34	0.80

Before Covid:

- Little gain from adding survey indicators
- Even sector GVA indices add little in this approach

During Covid:

- All statistical models deteriorate
- Gains over AR from traditional and real-time indicators much larger
- Vacancies data best

Input-output identity

$$\boldsymbol{y_T^e \equiv y_T^o\, R_T\, C_T + m_T\, Q_T + T_T}$$

$\boldsymbol{y_T^e}$ (1x6) = Total Final Expenditure aggregates: C(HH+NPISH), Gc, GFCF, Inventories, Valuables, Exports
$\boldsymbol{y_T^o}$ (1x20) = Gross Value Added in each domestic top-level SIC industry A-T
$\boldsymbol{R_T}$ (20x20) = diagonal matrix whose elements are the ratios between GVA by industry and GVA by product
$\boldsymbol{C_T}$ (20x6) = matrix whose element c_{ij} corresponds to the domestically added value of products in industry (row) *i* which have their final use in expenditure aggregate (column) *j* – for example: the expenditure in pounds sterling of consumption on agriculture products in 2018 – divided by the total GVA for products in industry (row) *I*
$\boldsymbol{m_T}$ (1x20) = value of imported products categorised by SIC industry A-T
$\boldsymbol{Q_T}$ (20x6) = the imports counterpart to $\boldsymbol{C_T}$, whose element q_{ij} corresponds to the imported value of products in sector (row) *i* which have their final use in expenditure aggregate (column) *j*, divided by the total GVA for products in industry (row) *i*
$\boldsymbol{T_T}$ = a 1x6 vector, each of whose elements is the total taxes minus subsidies on products in each expenditure category

Input-output model: not real-time feasible

$$\boldsymbol{y}_T^e \equiv \boldsymbol{y}_T^o \, \boldsymbol{R}_T \, \boldsymbol{C}_T + \boldsymbol{p}_T \, M_T + \boldsymbol{T}_T$$

$\boldsymbol{p}_T$ (1x6) = vector whose elements, which sum to 1, represent the share of total imports (whether consumed directly or indirectly) finding their final use in the corresponding final expenditure aggregate to $\boldsymbol{y}^e{}_T$;
M_T = total imports

Several reasons why this identity can't be used to give us perfect nowcasts:

1. Nominal vs. real
2. Annual vs. quarterly
3. Matrices only available with several years' lag
4. Measurement error in early vintages of GDPO data (matrix is constructed on Blue Book data)
5. GVA growth ~ GDP growth (as in ONS monthly GDP)
6. Nominal GDPO data in Blue Books doesn't match GDPO data in the corresponding input-output table prior to 2017

Input-output model: not real-time feasible

$$\boldsymbol{y_T^e} \equiv \boldsymbol{y_T^o}\, \boldsymbol{R_T}\, \boldsymbol{C_T} + \boldsymbol{p_T}\, M_T + \boldsymbol{T_T}$$

taken from I-O tables

$\boldsymbol{p_T}$ (1x6) = vector whose elements, which sum to 1, represent the share of total imports (whether consumed directly or indirectly) finding their final use in the corresponding final expenditure aggregate to $\boldsymbol{y^e}_T$;
M_T = total imports

Several reasons why this identity can't be used to give us perfect nowcasts:

1. Nominal vs. real
2. Annual vs. quarterly
3. Matrices only available with several years' lag
4. Measurement error in early vintages of GDPO data (matrix is constructed on Blue Book data)
5. GVA growth ~ GDP growth (as in ONS monthly GDP)
6. Nominal GDPO data in Blue Books doesn't match GDPO data in the corresponding input-output table prior to 2017

Input-output: feasible model

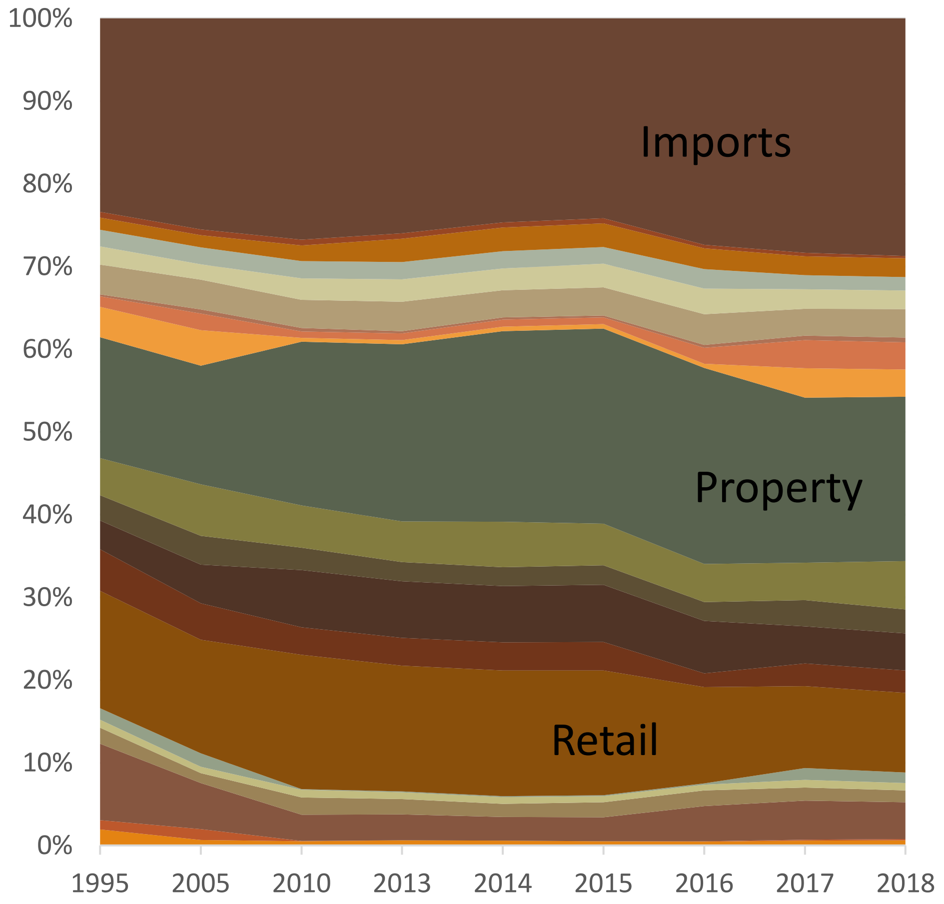
$$\widehat{\widehat{\boldsymbol{y}}}^{\boldsymbol{e}}{}_{\boldsymbol{t}} = \widehat{\widehat{\boldsymbol{y}}}^{\boldsymbol{o}}{}_{\boldsymbol{t}}\, \boldsymbol{R_S}\, \boldsymbol{K_S} + \boldsymbol{\pi_S}\, \widehat{M}_t$$

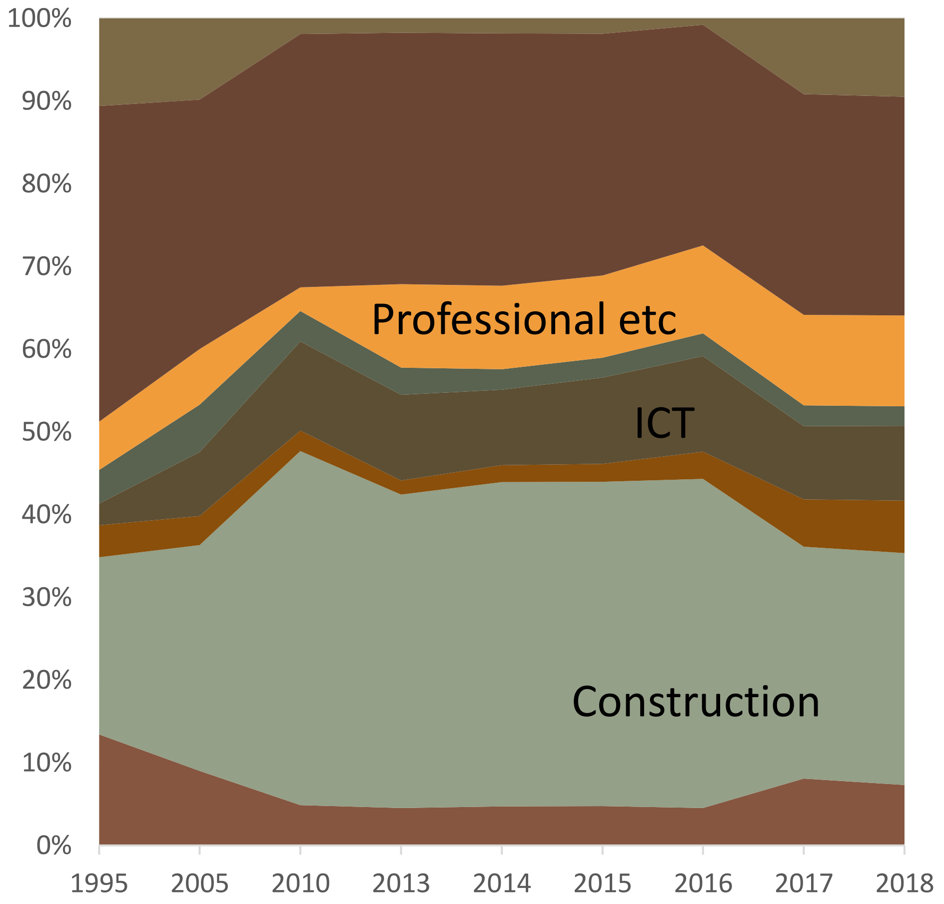
$\widehat{\widehat{\boldsymbol{y}}}^{\boldsymbol{e}}{}_{\boldsymbol{t}}$ (1x6) = estimated quarterly growth rates of elements of $\boldsymbol{y^e}_{\boldsymbol{t}}$;
$\widehat{\widehat{\boldsymbol{y}}}^{\boldsymbol{o}}{}_{\boldsymbol{t}}$ (1x6) = estimated GVA growth in each domestic industry in quarter t;
$\boldsymbol{R_S}$ (20x20) = matrix whose diagonal elements are the ratios between GVA by industry and GVA by product in year *S*, which is always at least four years earlier than the year containing quarter t, owing to the time taken to produce the input-output tables;
$\boldsymbol{K_S}$ (20x6) = matrix whose element c_{ij} corresponds to the share of aggregate (column) *j* expended on products in industry (row) *i* for year *S* – for example: the expenditure in pounds sterling of consumption on agriculture products in 2018 – divided by the total GVA for expenditure aggregate (column) *j*;
$\boldsymbol{\pi_S}$ (1x6) = vector, each of whose elements represents the share made up by imports in the corresponding expenditure aggregate in year *S*;
$\widehat{M}_t$ = estimated imports in quarter t

Data 2010Q1-2019Q4, ONS 2022Q4 first estimate vintage.

How stable are (domestic) input-output relationships?

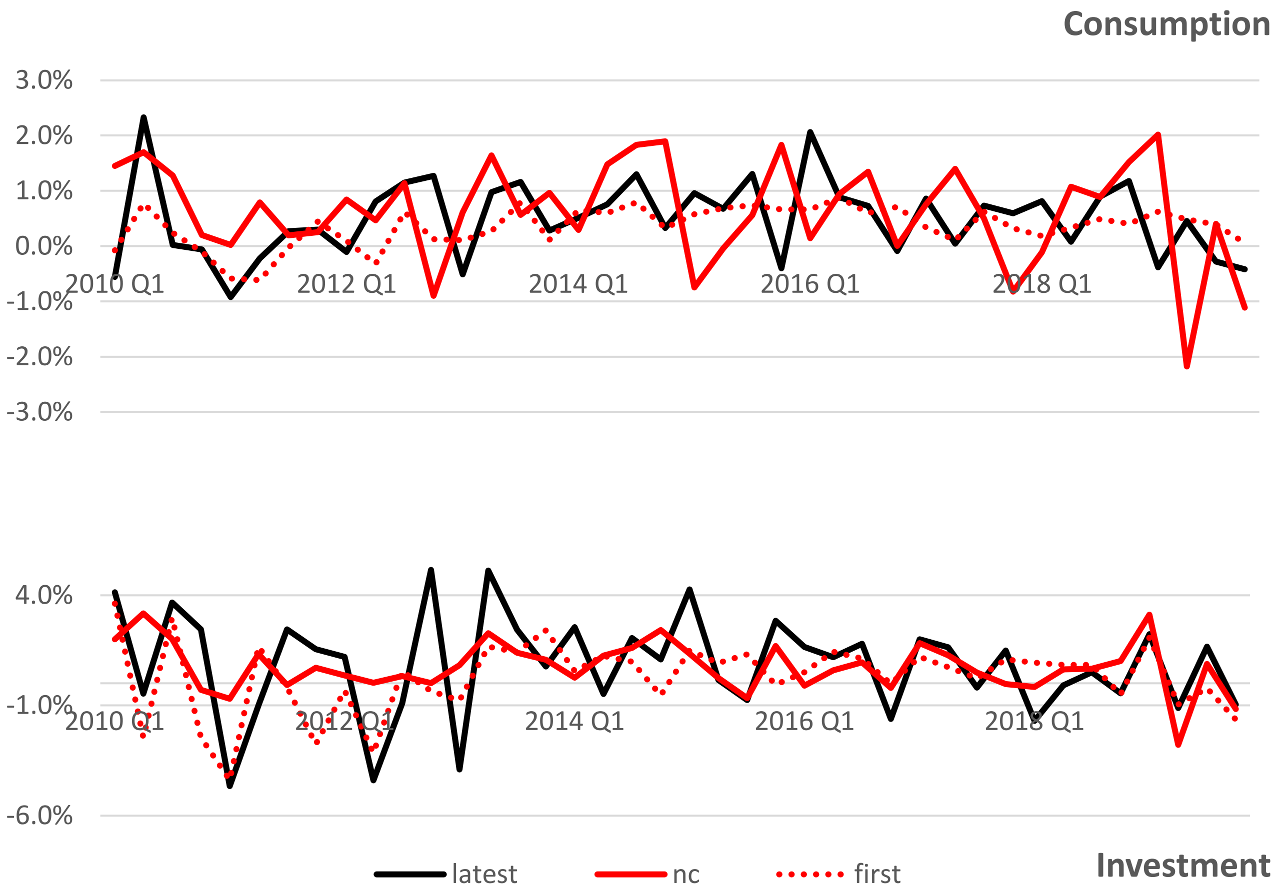
C





GFCF

Investment nowcasts are more accurate than consumption



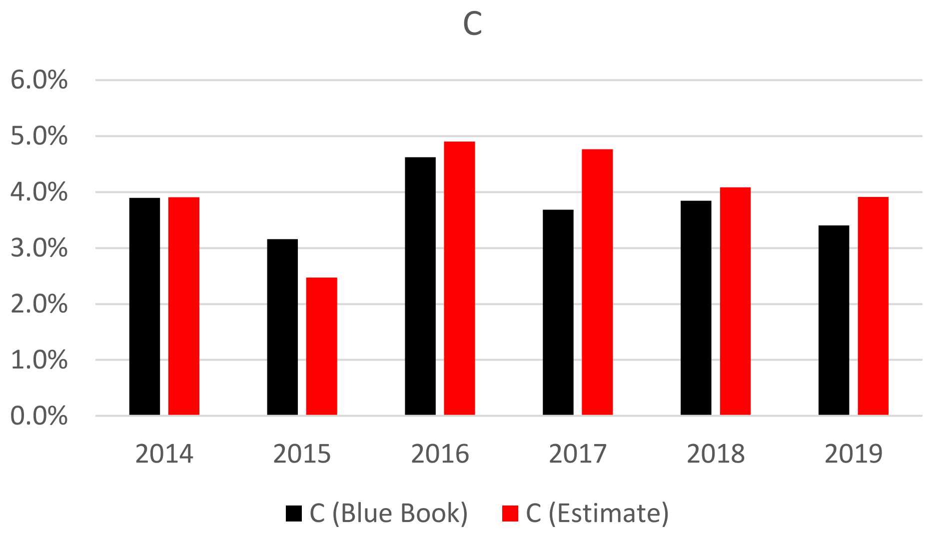
	Consumption		Investment	
	Nowcast	First estimate	Nowcast	First estimate
RMSFE	1.14%	0.61%	2.09%	2.08%
Correlation coefficient	0.03	0.54	0.41	0.39
Mean error (bias)	0.13%	-0.14%	-0.09%	-0.56%

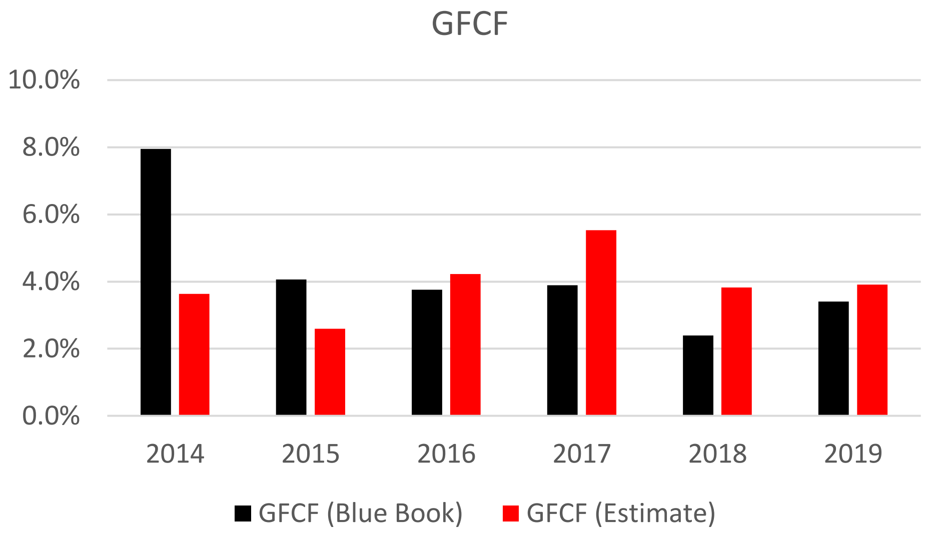
Using real-time GDP(output) data doesn't seem to hinder nowcasts

- Potential source of error from revisions to output data: if early vintages of GDP(O) sectoral growth (and imports) are inaccurate, this will add to forecast error.
- Limited ability to test this, because we only have a real-time database at SIC industry level going back to 2018: not only a short period but covering quarters which have not been revised much at all.
- Indicative comparison between nowcasts for 2018-2019 created using real-time data and those using the latest vintage.
  - RMSFE for consumption 1.70% (real-time) versus 1.45% (latest vintage)
  - RMSFE for GFCF 1.39% (real-time) versus 1.00% (latest vintage)

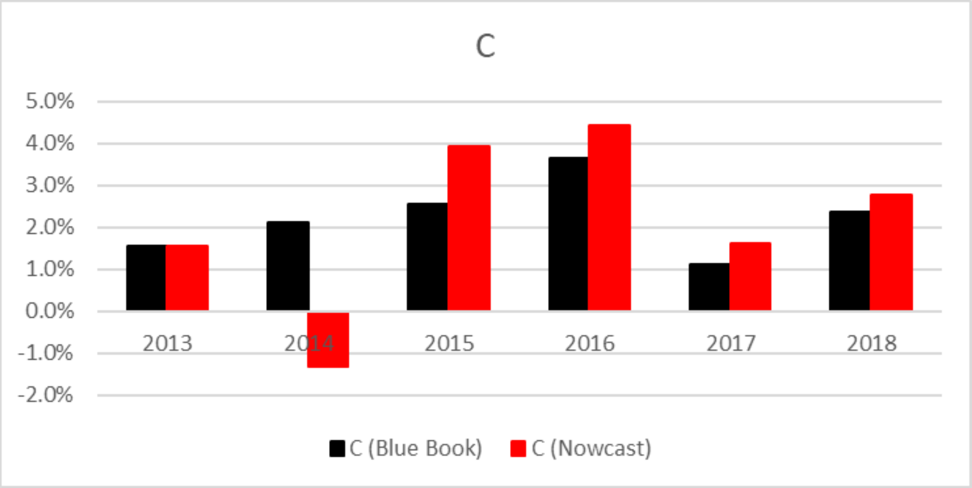
How big a problem are lagged matrices?

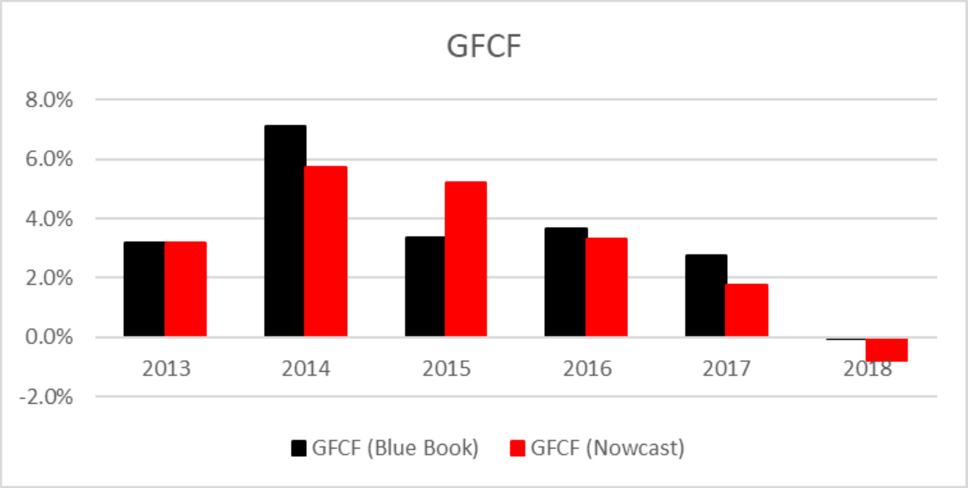
	Consumption		GFCF	
	Lagged matrices	In-year matrices	Lagged matrices	In-year matrices
RMSE	1.01%	0.98%	1.66%	1.67%
Mean error (bias)	0.10%	0.05%	-0.07%	-0.05%





Nominal v real: the role of relative prices in forecast errors





Conclusions and future work

Early suggestions that both real-time data and input-output tables can contribute to improving early estimates.

Input-output tables approach, several future paths to consider:

- Working in (cumulative) levels rather than quarterly changes
- Using sector price indices to work in nominal terms
- Modelling inventories growth
- More detailed data on (goods) imports
- Improving allocation of secondary production

Conclusions and future work

- Constraining GDP to the level implied by sector growth indices
  - Using contemporaneous trade (and government spending) data to constrain estimation
    - Preliminary results from (non-feasible) approach using supply-use tables; assuming for each low-level SIC) product that each nominal demand component is a fixed share of total domestic demand

	Households	Non-profit institutions serving households	Central government	Local government	Total	Gross fixed capital formation
	2019					
Actual	2.63%	5.84%	7.93%	4.49%	3.65%	4.46%
Predicted	3.39%	4.86%	5.80%	6.15%	3.94%	3.76%
	2020					
Actual	-12.48%	-7.58%	14.29%	5.79%	-6.84%	-9.29%
Predicted	-11.17%	-2.91%	9.70%	0.80%	-6.86%	-9.63%

2019

2020