Measuring GDP Growth Data Uncertainty

Ana Beatriz Galvão and James Mitchell

University of Warwick
Funding by ESCOE is gratefully acknowledged.

November 2018

Macroeconomic Uncertainty

- Many measures of uncertainty rise during recessions: stock market volatility, macroeconomic forecasting uncertainty, professional forecasters disagreement and economic policy uncertainty (Bachman et al , 2013; Jurado et al, 2015; Baker et al, 2016; Rossi et al, 2016; surveyed by Bloom (2014)).

Data Uncertainty

- Economic statistics uncertainty has two components (Manski, 2014). *Transitory statistical uncertainty*, early data releases that are revised as new information arrives; and *permanent statistical uncertainty* from data incompleteness or the inadequacy of data collection which does not diminish over time.

Data Uncertainty

- Economic statistics uncertainty has two components (Manski, 2014). *Transitory statistical uncertainty*, early data releases that are revised as new information arrives; and *permanent statistical uncertainty* from data incompleteness or the inadequacy of data collection which does not diminish over time.
- For GDP growth, we are mainly concerned about the transitory uncertainty.

Data Uncertainty

- Economic statistics uncertainty has two components (Manski, 2014). *Transitory statistical uncertainty*, early data releases that are revised as new information arrives; and *permanent statistical uncertainty* from data incompleteness or the inadequacy of data collection which does not diminish over time.
- For GDP growth, we are mainly concerned about the transitory uncertainty.
- We expect this *epistemic* uncertainty, that is uncertainty due to lack of knowledge about current and past data, to diminish as additional data are collected and both corrections and improvements to these data are made by the statistics office.

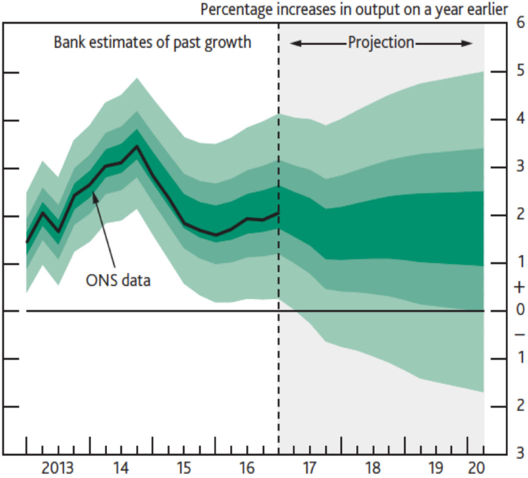
Data Uncertainty

- Economic statistics uncertainty has two components (Manski, 2014). *Transitory statistical uncertainty*, early data releases that are revised as new information arrives; and *permanent statistical uncertainty* from data incompleteness or the inadequacy of data collection which does not diminish over time.
- For GDP growth, we are mainly concerned about the transitory uncertainty.
- We expect this *epistemic* uncertainty, that is uncertainty due to lack of knowledge about current and past data, to diminish as additional data are collected and both corrections and improvements to these data are made by the statistics office.
- We show that GDP data uncertainty may affect policy decisions by adding a layer of uncertainty on the measurement of the current state of the economy.

Communication of UK GDP Growth Data Revision Uncertainty

- While the Office for National Statistics emphasise the uncertainty of early GDP data releases by indicating that their data will be revised, it is the Bank of England that provide quantitative estimates of GDP data uncertainty, as perceived by their Monetary Policy Committee.

Communication of UK GDP Growth Data Revision Uncertainty: Inflation Report Fan Chart June 2017.



Communication of UK GDP Growth Data Revision Uncertainty

- The fan charts shows "how the MPC's best collective judgement of the most likely path for the mature estimate of GDP growth, and the uncertainty around it, both over the past and into the future."
- "To the left of the first vertical dashed line, the centre of the darkest band of the fan chart gives the Committee's best collective judgement of the most likely path for GDP growth once the revisions process is complete. The estimate is based on an analysis of business surveys and the past pattern of official data revisions."

In this paper,

- We compare MPC's expected uncertainty estimates with realised *ex post* measures to

In this paper,

- We compare MPC's expected uncertainty estimates with realised *ex post* measures to
  - provide the first evaluation of the Bank of England's predictive densities for revised GDP growth values;

In this paper,

- We compare MPC's expected uncertainty estimates with realised *ex post* measures to
  - provide the first evaluation of the Bank of England's predictive densities for revised GDP growth values;
  - construct a new measure of data uncertainty, that captures only the unforecastable component to data revisions.

In this paper,

- We compare MPC's expected uncertainty estimates with realised *ex post* measures to
  - provide the first evaluation of the Bank of England's predictive densities for revised GDP growth values;
  - construct a new measure of data uncertainty, that captures only the unforecastable component to data revisions.
- We propose a generic, loss-function based approach to measuring uncertainty; and show how density forecast calibration tests can then be constructed from this.

In this paper,

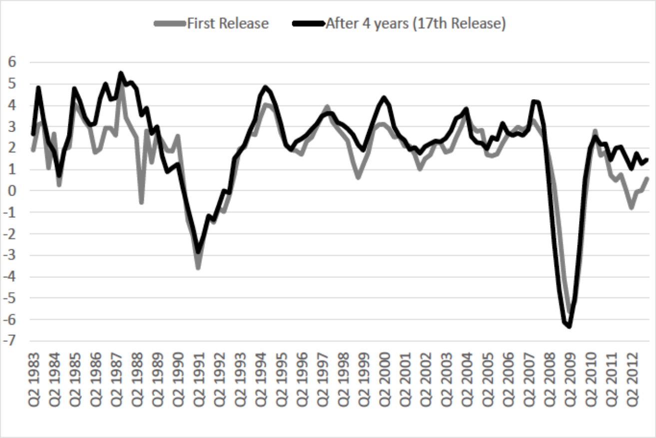
- We compare MPC's expected uncertainty estimates with realised *ex post* measures to
  - provide the first evaluation of the Bank of England's predictive densities for revised GDP growth values;
  - construct a new measure of data uncertainty, that captures only the unforecastable component to data revisions.
- We propose a generic, loss-function based approach to measuring uncertainty; and show how density forecast calibration tests can then be constructed from this.
- We first show that UK GDP revisions matter, emphasising that they are time-varying and contain 'news'. Then we provide evidence that our measure of GDP data uncertainty increased at the onset of the 2018-2019 recession and is correlated with measures of macroeconomic uncertainty.

UK Y-on-Y Growth Data Revisions Characteristics

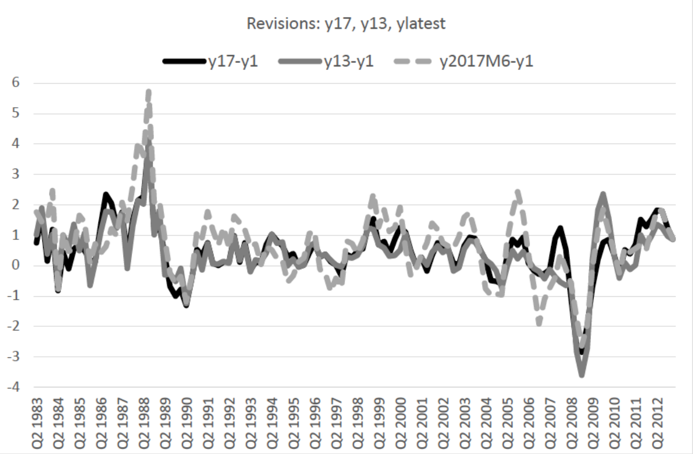
- GDP growth values published after 3 years, y_t^{t+13}, are a popular measure of 'mature' data in the US data revision literature.
- But UK GDP growth revisions between y_t^{t+13} and y_t^{t+17} are mainly news (not correlated with the earlier release) and imply a non-zero mean revision suggest that y_t^{t+17} might be a better measure of 'mature' data for the UK.
- The revisions $y_t^{t+17} - y_t^{t+1}$ have a 0.4 sample mean (1993Q2-2013Q1), which is statistically different from zero, and are uncorrelated with y_t^{t+1}. Mean absolute revision is 0.7.

UK Y-on-Y Growth Data Revisions

Display of y_t^{t+1} and y_t^{t+17}.



UK Y-on-Y Growth Data Revisions



A Time-varying model for data revisions I

- The model describes "mature" data, y_t^{t+l} with respect to an earlier estimate y_t^{t+b} $(l > b, b \geq 1)$ as:

$$y_t^{t+l} = y_t^{t+b} + \mu_t + u_t,$$

where μ_t is a time-varying local mean revision. The disturbance, u_t, characterises the measurement error assumed mean zero, so that $Var(u_t)$ measures the degree of measurement error in the initial release.

- The time-varying local mean, μ_t, follows a random walk process implying that the average revision moves slowly over time:

$$\begin{aligned} \mu_t &= \mu_{t-1} + e^{.5(h_0 + w_h \tilde{h}_t)} \zeta_{b,t} \\ \tilde{h}_t &= \tilde{h}_{t-1} + \zeta_{h,t}; \\ & \zeta_{b,t}, \zeta_{h,t} \text{ are both } iidN(0,1). \end{aligned}$$

A Time-varying model for data revisions II

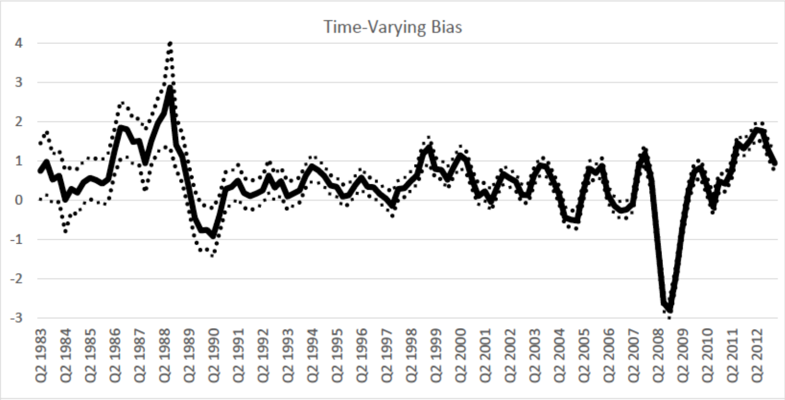
- And the measurement error also has a time-varying volatility:

$$(y_t^{t+l} - y_t^{t+b}) = rev_t^{(l-b)} = \mu_t + e^{.5(g_0 + w_g \tilde{g}_t)} \zeta_{u,t}$$

- We use a Bayes factor approach to check whether we need both stochastic volatility process. The results support the model above for $l = 17$ and $b = 1, 4, 8, 12$.

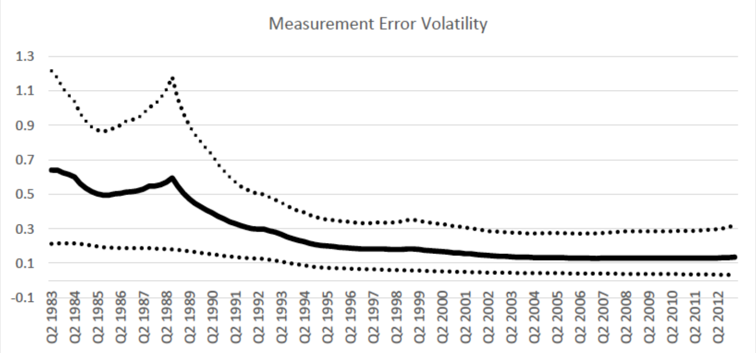
Time-Varying Local Mean Revision

Estimates for $y_t^{t+17} - y_t^{t+1}$.



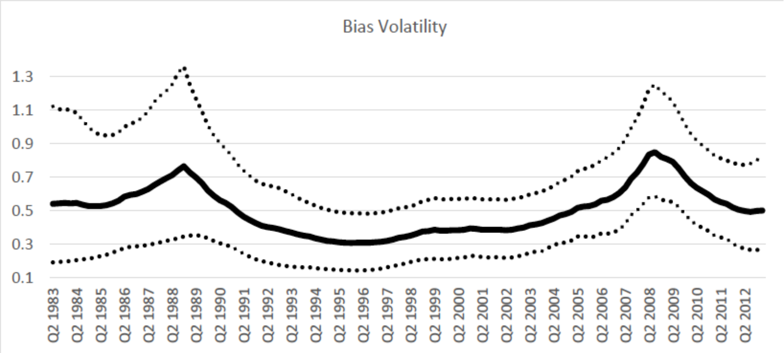
Measurement Error Volatility

Estimates for $y_t^{t+17} - y_t^{t+1}$.



Local Mean Volatility

Estimates for $y_t^{t+17} - y_t^{t+1}$



Measuring Uncertainty

- We are interested in the unforecastable component of data uncertainty.
- Uncertainty is the difference between the *ex post* (or realised) and the *ex ante* (or expected) values of the chosen loss function (or scoring rule):

$$Unc_t = L(f_t, y_t) - E_{f_{t-h}}\left[L(f_t, Y_t)\right],$$

  generalising Clements (2014), Jurado et al (2015) and Rossi et al (2016).
- $E_{f_{t-h}}\left[L(f_t, Y_t)\right]$ are in expectations computed by (honest and loss-minimising) agents who assume (*ex ante*) that their forecast is as good as it can be.
- Unc_t is a (realised) shock to confidence, the definition of ambiguity in Ilut and Schneider (2014), with confidence measured, in our framework, by $E_{f_t}\left[L(f_t, Y_t)\right]$ i.e. the expected "risk" of the forecast.

Test for Calibration given a Loss Function

Given loss function $L(f_t, y_t)$, correct unconditional average calibration of the forecast f_t with respect to the realisation y_t, is defined as when

$$H_0^U : E\left(Unc_t\right) = 0.$$

If we view Unc_t as capturing Knightian uncertainty, under H_0^U while there may be data risk there is no data uncertainty - as users of the forecast, f_t, are correctly capturing all aspects of Y_t that are relevant given $L(f_t, Y_t)$.

Uncertainty Measures for Specific Loss (score) functions: I

1. Mean squared error loss:

$$Unc_t^{MSE} = (y_t - \hat{y}_t)^2 - \hat{\sigma}_t^2$$

2. Interval loss:

$$Unc_t^{Int} = x_t(\alpha) - \alpha$$

where $x_t(\alpha) = 1(y_t \in J(\alpha)); J(\alpha) = [lower_t, upper_t]$.

3. Logarithm score (-1) for Gaussian densities:

$$Unc_t^{\log S} = 0.5(z_t^2 - 1)$$

as $E_{f_{t-h}}\left[L^{\log S}(f_t, Y_t)\right] = 0.5(1 + \log 2\pi) + \log(\hat{\sigma}_t)$ and $z_t = \left(\frac{(y_t - \hat{y}_t)}{\hat{\sigma}_t}\right)$.

Uncertainty Measures for Specific Loss (score) functions: II

4. CRPS for Gaussian densities:

$$Unc_t^{CRPS} = ((y_t - \hat{y}_t))\left(2\Phi\left(z_t\right) - 1\right) + 2\hat{\sigma}_t \phi\left(z_t\right) - 2(\hat{\sigma}_t / \sqrt{\pi})$$

as $E_{f_t}(L^{CRPS}(f_t, y_t)) = (\hat{\sigma}_t / \sqrt{\pi})$.

Bank of England's Probabilistic Backcasts

- Let $\hat{y}^t_{t-B}, ..., \hat{y}^t_{t-2}, \hat{y}^t_{t-1}$ denote the MPC's point estimate of growth ending in quarter $t-1, ..., t-B$, as announced by the MPC in quarter t; and let $\hat{\sigma}^t_{t-b}, ..., \hat{\sigma}^t_{t-2}, \hat{\sigma}^t_{t-1}$ denote the corresponding set of standard deviation estimates.

Bank of England's Probabilistic Backcasts

- Let $\hat{y}^t_{t-B}, ..., \hat{y}^t_{t-2}, \hat{y}^t_{t-1}$ denote the MPC's point estimate of growth ending in quarter $t-1, ..., t-B$, as announced by the MPC in quarter t; and let $\hat{\sigma}^t_{t-b}, ..., \hat{\sigma}^t_{t-2}, \hat{\sigma}^t_{t-1}$ denote the corresponding set of standard deviation estimates.
- Because estimates are computed after ONS first release, the predicted revision is $\hat{y}^t_{t-1} - y^t_{t-1}$.

Bank of England's Probabilistic Backcasts

- Let $\hat{y}^t_{t-B}, ..., \hat{y}^t_{t-2}, \hat{y}^t_{t-1}$ denote the MPC's point estimate of growth ending in quarter $t-1, ..., t-B$, as announced by the MPC in quarter t; and let $\hat{\sigma}^t_{t-b}, ..., \hat{\sigma}^t_{t-2}, \hat{\sigma}^t_{t-1}$ denote the corresponding set of standard deviation estimates.
- Because estimates are computed after ONS first release, the predicted revision is $\hat{y}^t_{t-1} - y^t_{t-1}$.
- In general the predicted revision is $\hat{y}^{t+b}_t - y^{t+b}_t$, where y^{t+b}_t is ONS earlier b^{th} estimate and $\hat{y}^{t+b}_t$ is a MPC's prediction for mature GDP values that uses information up to $t+b$.

Bank of England's Probabilistic Backcasts

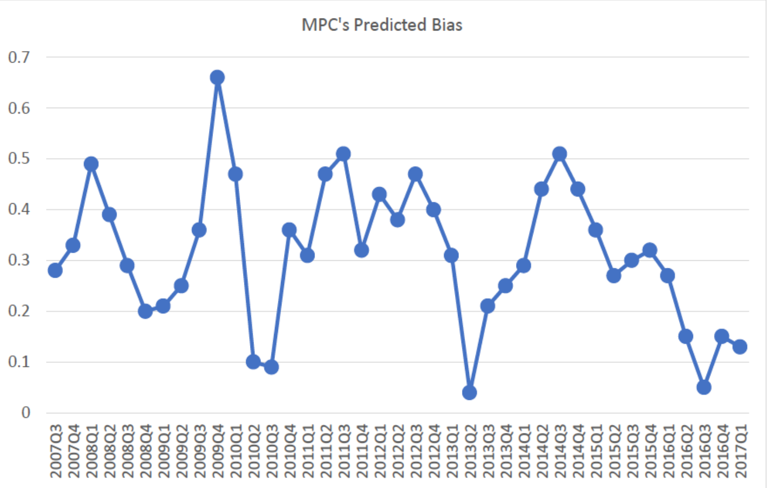
- Let $\hat{y}^t_{t-B}, ..., \hat{y}^t_{t-2}, \hat{y}^t_{t-1}$ denote the MPC's point estimate of growth ending in quarter $t-1, ..., t-B$, as announced by the MPC in quarter t; and let $\hat{\sigma}^t_{t-b}, ..., \hat{\sigma}^t_{t-2}, \hat{\sigma}^t_{t-1}$ denote the corresponding set of standard deviation estimates.
- Because estimates are computed after ONS first release, the predicted revision is $\hat{y}^t_{t-1} - y^t_{t-1}$.
- In general the predicted revision is $\hat{y}^{t+b}_t - y^{t+b}_t$, where y^{t+b}_t is ONS earlier b^{th} estimate and $\hat{y}^{t+b}_t$ is a MPC's prediction for mature GDP values that uses information up to $t+b$.
- "Mature" values are observed (released by the ONS) at $t+l$: the ONS GDP estimation error is $y^{t+l}_t - y^{t-b}_t$ and the MPC's backcast error is $y^{t+l}_t - \hat{y}^{t+b}_t$.

Bank of England's Probabilistic Backcasts

- Let $\hat{y}^t_{t-B}, ..., \hat{y}^t_{t-2}, \hat{y}^t_{t-1}$ denote the MPC's point estimate of growth ending in quarter $t-1, ..., t-B$, as announced by the MPC in quarter t; and let $\hat{\sigma}^t_{t-b}, ..., \hat{\sigma}^t_{t-2}, \hat{\sigma}^t_{t-1}$ denote the corresponding set of standard deviation estimates.
- Because estimates are computed after ONS first release, the predicted revision is $\hat{y}^t_{t-1} - y^t_{t-1}$.
- In general the predicted revision is $\hat{y}^{t+b}_t - y^{t+b}_t$, where y^{t+b}_t is ONS earlier b^{th} estimate and $\hat{y}^{t+b}_t$ is a MPC's prediction for mature GDP values that uses information up to $t+b$.
- "Mature" values are observed (released by the ONS) at $t+l$: the ONS GDP estimation error is $y^{t+l}_t - y^{t-b}_t$ and the MPC's backcast error is $y^{t+l}_t - \hat{y}^{t+b}_t$.
- We set $l = 17$ so we evaluate backcasts for $b = 1, ..., 16$ for $t = 2007Q3 - 2013Q1$.

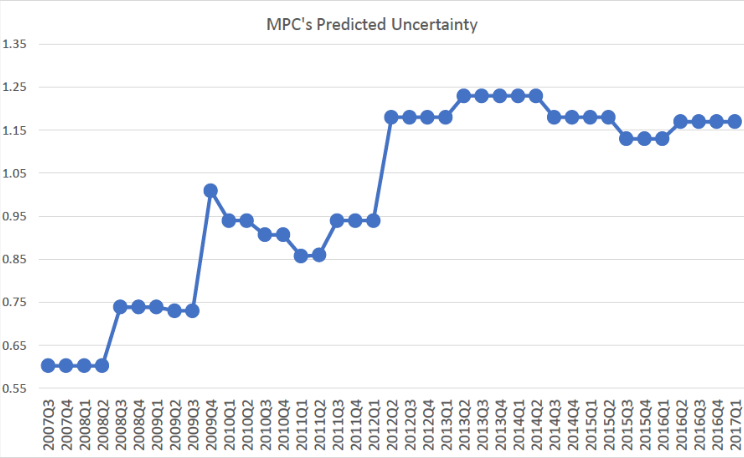
MPC's Predicted Revision with b=1

Computed as $\hat{y}_t^{t+1} - y_t^{t+1}$

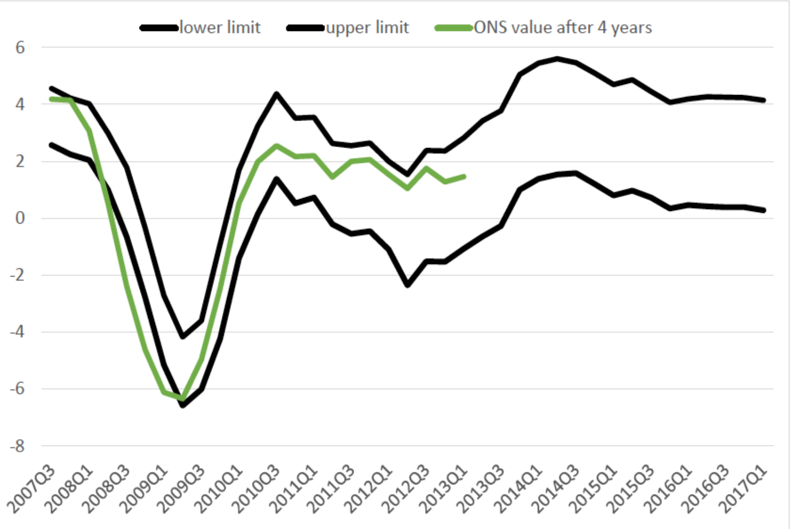


MPC's Predicted Standard Error (Uncertainty) with b=1

$$se(y_t^{t+l} - \hat{y}_t^{t+1})$$



MPC's Backcasts 90% Predicted Interval with b=1



Benchmark Backcasting Model I

- Because time-series models of stationary data converge to the unconditional mean and variance at long horizons, we compute predictions using the recursively updated (in real-time) unconditional mean and standard deviation of the revisions.
- The probabilistic backcasts for the observation t, using information up to $t+b$, which includes a time series of ONS revisions between the l^{th} data release and the b^{th} release for observations up to $t-l+1$ is

$$N(\hat{y}_t^{t+b,unc}, \hat{\sigma}_t^{2,t+b,unc})$$

Benchmark Backcasting Model II

where

$$
\begin{aligned}
\hat{y}_t^{t+b,unc} &= y_t^{t+b} + \hat{\mu}_t^{t+b,unc} \text{ for } t = T-(b-1)+1, ..., T+P \\
\hat{\mu}_t^{t+b,unc} &= \frac{1}{t-l+1} \sum_{\tau=1}^{\tau=t-l+1} rev_\tau^{(l-b)} \\
\hat{\sigma}_t^{t+b,unc} &= \sqrt{\frac{1}{t-l+1} \sum_{\tau=1}^{\tau=t-l+1} \left(rev_\tau^{(l-b)} - \hat{\mu}_t^{t+b} \right)^2} \\
rev_\tau^{(l-b)} &= y_\tau^{\tau+l} - y_\tau^{\tau+b}.
\end{aligned}
$$

Accuracy of GDP Estimates (or Backcasts)

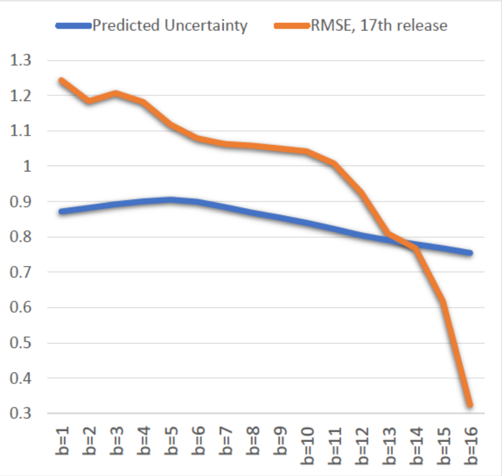
Note that ONS average revision with $b = 1$ is 0.33 over the period.

	Bias	RMSE	RMSE Ratio to ONS est.
MPC, b=1	-0.046	1.243	0.931
Uncond., b=1	-0.223	1.335	1.000
MPC, b=4	-0.024	1.138	0.951
Uncond., b=4	-0.034	1.173	0.980
MPC, b=8	-0.070	0.953	0.973
Uncond., b=8	-0.067	0.971	0.991
MPC, b=12	0.092	0.770	0.948
Uncond, b=12	0.139	0.806	0.992
MPC, b=16	-0.025	0.283	0.997
Uncond, b=16	-0.005	0.283	0.997

Accuracy of Revisions Predictive (or Backcasts) Densities

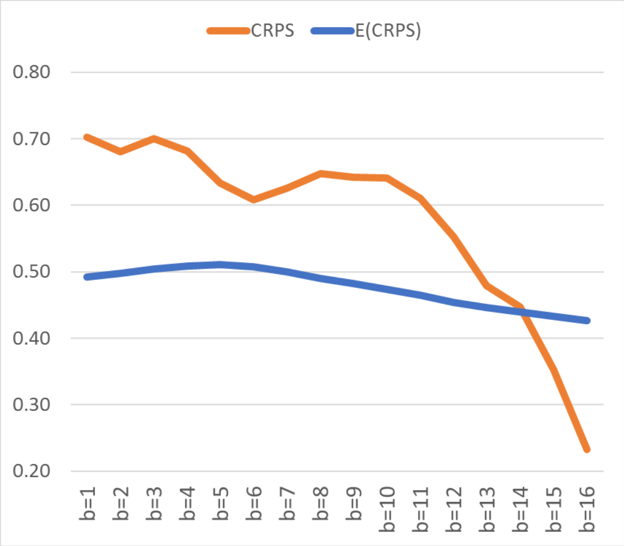
	Logscore	CRPS	90%Cov.	75%Cov.	50%Cov.
MPC, b=1	2.070	0.702	83%	65%	43%
Uncond., b=1	2.118	0.735	74%	61%	43%
MPC, b=4	1.898	0.655	87%	61%	39%
Uncond., b=4	2.631	0.687	62%	38%	23%
MPC, b=8	1.513	0.559	74%	65%	35%
Uncond., b=8	2.055	0.595	67%	37%	20%
MPC, b=12	1.215	0.520	74%	74%	39%
Uncond, b=12	2.533	0.483	62%	44%	29%
MPC, b=16	0.571	0.196	96%	91%	91%
Uncond, b=16	1.166	0.110	89%	89%	84%

Expected (ex ante) and Realised (ex post) Uncertainty: MSE

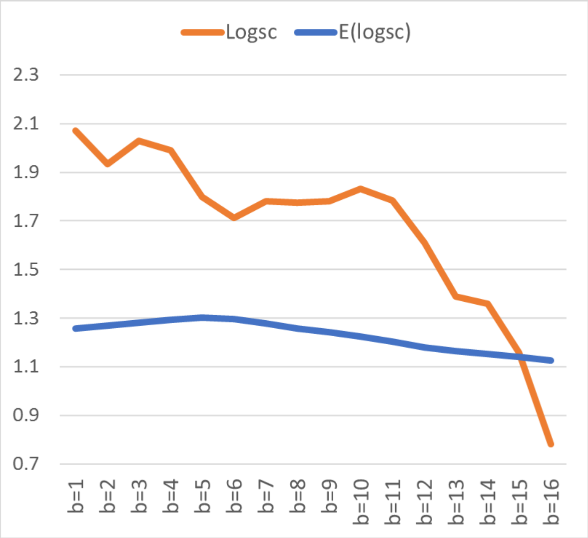


As in Clements (2014).

Expected (ex ante) and Realised (ex post) Uncertainty: CRPS



Expected (ex ante) and Realised (ex post) Uncertainty: Logscore

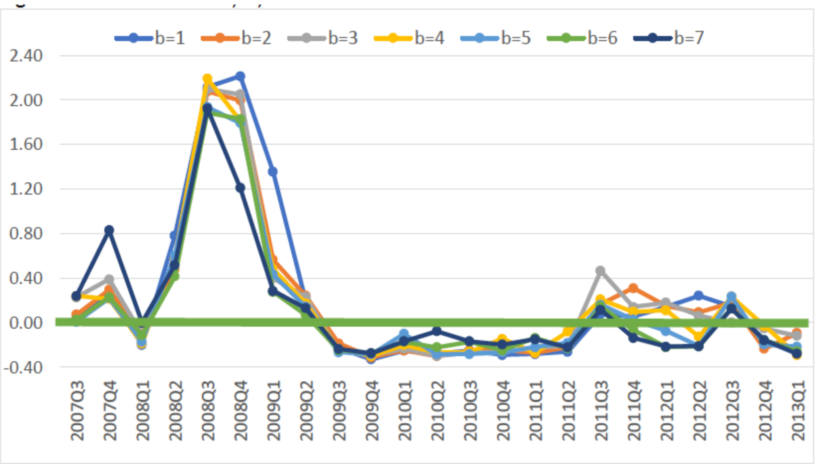


Test for calibration of MPC's backcasts (for a given loss function) for

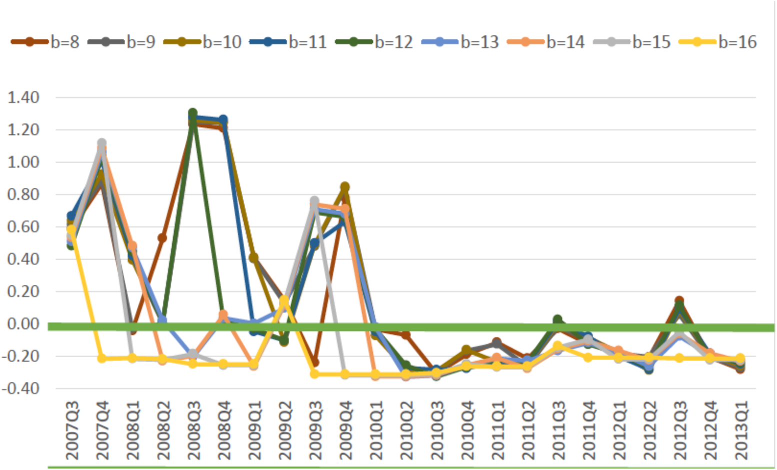
$t = 2007Q3 - 2013Q1.$

	MSE	CRPS	Logscore	Cov, 90%, pv	Cov, 75%, pv	Cov, 50%, Pv
b=1	1.34	1.38	1.61	0.28	0.30	0.53
b=2	1.22	1.36	1.47	0.28	0.14	0.53
b=3	1.29	1.45	1.53	0.10	0.14	0.14
b=4	1.19	1.31	1.40	0.64	0.14	0.30
b=5	0.92	0.98	1.20	0.64	0.30	0.83
b=6	0.75	0.82	1.11	0.64	0.71	0.83
b=7	0.86	1.12	1.36	0.28	0.90	0.30
b=8	1.19	1.57	**1.90**	**0.03**	0.30	0.14
b=9	1.20	1.57	**1.89**	**0.03**	0.30	0.14
b=10	1.23	1.60	**2.00**	**0.01**	0.30	0.30
b=11	1.08	1.38	**1.80**	**0.01**	0.56	0.30
b=12	0.74	1.04	1.43	**0.03**	0.90	0.30
b=13	0.03	0.41	0.95	0.10	0.71	0.83
b=14	-0.19	0.09	0.79	0.10	0.71	0.30
b=15	-1.20	-1.02	0.07	0.64	0.38	**0.02**
b=16	**-5.49**	**-4.82**	**-2.90**	0.31	**0.04**	**0.00**

CRPS deviations for b=1,...,7

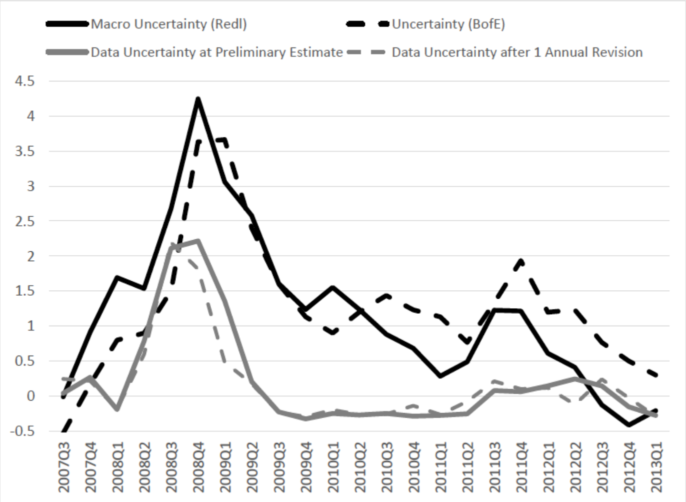


CRPS deviations for b=8,...,16



UK GDP Growth Data Uncertainty and other Macro Uncertainty Measures

Max correlation is at the time of first release and approx. 0.8.



Conclusions I

- We propose a measure of data uncertainty. The measure is calculated for UK GDP growth data; and is based on identifying the component of future data revisions that the Bank of England is unable to predict correctly after observing earlier ONS growth estimates.
- We find that UK data uncertainty rises at the onset of recessions; and is positively correlated with measures of UK macroeconomic uncertainty, such as the measures computed by Redl (2017) and Bank of England (2016).
- Data uncertainty might be interpreted as an additional souce and layer of uncertainty relative to the more traditional macroeconomic uncertainty measures discussed in Bloom (2014).

Conclusions II

- We find that the MPC's point estimates of historical GDP growth are more accurate measures of revised ONS data than the equivalently timed estimates from the ONS themselves.
- We find that the MPC's probabilistic backcasts for GDP growth are, on average, well-calibrated and perform well relative to a benchmark model; but the MPC do appear to have over-estimated (ex ante) data uncertainty for observations in the 2010-13 period. Data revisions to mature ONS data ($b > 8$) are harder to predict, because of the unknown impact of future benchmark revisions.

Conclusions III

- We commend the Bank of England for communicating the (*ex ante*) predictable component of data revisions in their published fan charts;
- and we recommend the ONS reconsiders if and how they communicate the uncertainty associated with their early GDP estimates.
- We do believe, however, that the Bank of England would improve communication further if they stated explicitly what data vintage they seek to forecast.