BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM

DIVISION OF RESEARCH AND STATISTICS

Date: March 15, 2018

To: Andrew Figura

From: Travis Berge[1]

Subject: Measuring the uncertainty of judgmental estimates of unobserved variables

---

Knowing the output gap and natural rate of unemployment in real-time is central to good policy. However, these objects are unobserved and can only be estimated, even in hindsight. Model-based estimates of the output gap and natural rate are imprecise, so that output and unemployment rate gaps that fall comfortably within a model's confidence intervals can carry starkly different policy implications. The Board staff takes a different approach, producing judgmental estimates of these concepts without reporting its uncertainty about them.

This memo attempts to answer a critical question: How precise are staff estimates of the output gap, natural rate of unemployment, and unemployment rate gap? To do so, I develop a new, model-free method for calculating the uncertainty around estimates of unobserved variables. The uncertainty measure echoes what is presented for staff projections of real GDP growth, the unemployment rate, and price inflation in the Risks & Uncertainty section of Tealbook that derive from forecast errors of previous staff projections.[2] Of course, forecast errors cannot be calculated for unobserved variables because their "true" values are never observed. Instead, I use the magnitude of staff revisions to proxy the uncertainty around these variables. Revisions contain several sources of uncertainty: that the staff does not know the true model of the economy, possibly time-varying parameter values, and that the data itself is uncertain. However, revision uncertainty does not fully capture filtering uncertainty, the idea that a variety of paths of the unobserved variables may be consistent with a particular model of the economy.

Staff revisions to its estimates of the output gap, natural rate, and unemployment rate gap can be sizable, so that there is notable average revision uncertainty. For example, the 70 percent revision interval for the staff's estimate of the output gap in 2018Q1 ranges from 1.0 to 3.0 percent. Natural rate estimates are

[1]Niko Paulson provided superb research assistance.

[2]See also the November 25, 2014 Research & Statistics memo "Characterizing uncertainty around the staff forecast in the Tealbook and the FOMC forecasts reported in the Summary of Economic Projections (SEP)," written by Jeremy Nalewaik, Brad Strum and Adam Scherling.

also uncertain: based on previous revisions, the 70 percent interval around the staff's estimate of the natural rate in 2018Q1 ranges from 4.4 to 5.0 percent. Uncertainty only increases going forward: in 2020Q1, the 70 percent intervals for the staff estimate output gap and natural rate estimates are 0.6 to 5.1 percent and 4.1 to 5.5 percent, respectively.

1 Vintage Tealbook data

Calculating revisions requires a long history of real-time Tealbook estimates for the natural rate and output gap. An ongoing project in CMC is the development of a database that contains real-time vintage data for a large number of macroeconomic variables. The database also stores important forecasts, such as staff judgmental forecasts and CMC model-based forecasts.[3] Work on the database is ongoing, but it currently contains major elements of Tealbook projections since the March 1996 Tealbook. Staff forecasts of the natural rate of unemployment are available since May 1997; see table 1.

	Output gap	Natural rate	Unemployment rate
First Tealbook vintage	March 1996	May 1997	March 1996
Latest Tealbook vintage	January 2018	January 2018	January 2018
Number of vintages	176	167	176

Table 1: Tealbook vintages available for staff estimates of the output gap, natural rate of unemployment and unemployment rate.

To get a sense of the data, vintage Tealbook estimates of the output gap are shown in figure 1. Each gray line begins in 1990Q1 and represents a particular Tealbook estimate of the output gap. Most lines end 8-10 quarters following that Tealbook's publication date (recent Tealbook estimates all end in 2018Q1). The dark blue line is the output gap estimate from the January 2018 Tealbook. Since the first vintage shown in the figure is from March 1996, the vintage data cover roughly two complete business cycles.

Figure 2 focuses on the staff's judgmental assessment of the output gap, natural rate and unemployment rate gap in two specific quarters, 2005Q4 and 2008Q4. The quarters are arbitrary but chosen to illustrate of the kinds of revisions undertaken by the staff. The upper-left panel shows estimates of the output gap in 2005Q4 from each Tealbook produced between August 2003—the first Tealbook to include an output gap projection for 2005Q4—and January 2018. The pink shaded region delineates 2005Q4. Thus, dots to the left of the pink region are *forecasts*; dots within the pink region, of which there are two, are *nowcasts*;

[3]The primary source for Tealbook forecasts is archived backups of the Ruth database.

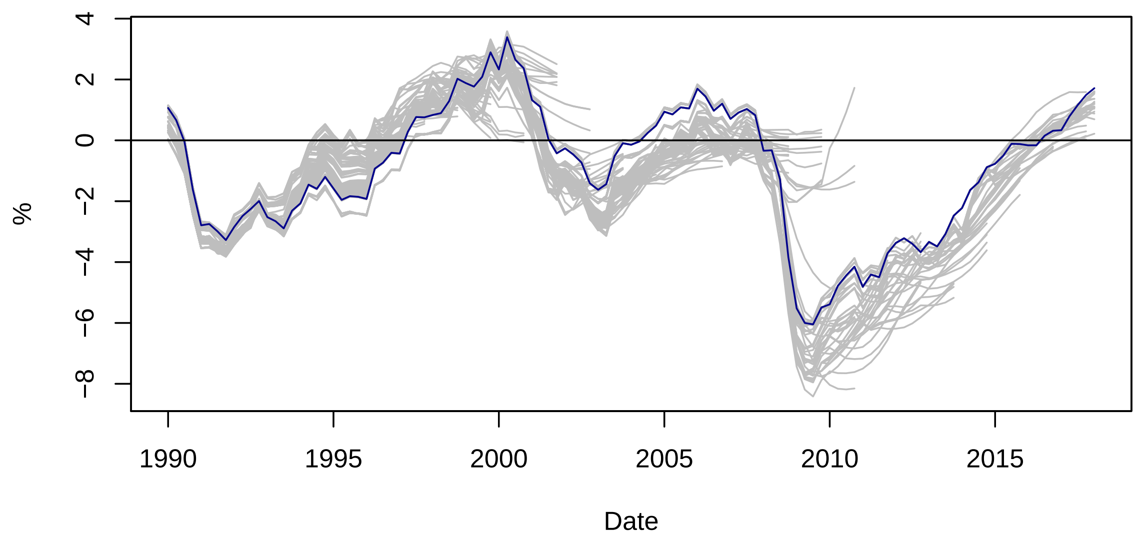

Figure 1: The output gap, 1990Q1–2018Q1. The blue line is the January 2018 estimate of the output gap. Each gray line is a Tealbook vintage of the path of the output gap (history and real-time forecast). First available Tealbook vintage is March 1996.

and dots to the right of the shaded area are *backcasts*, changes to the staff's beliefs about the value of the output gap after the fact.

Focusing first on the left column, throughout 2004 and 2005, the staff believed that the output gap at the end of 2005 would be slightly negative. The forecasts are somewhat erratic, reflecting staff revisions to its forecasts in real-time. After the fact, staff output gap estimates stabilize and eventually increase to about a percentage point. After the experience of the Great Recession and the subsequent recovery, the staff came to view the level of output in 2005Q4 as unsustainably high.

The right column contains staff estimates of the output gap, natural rate and unemployment rate gap in 2008Q4, near the trough of the most recent recession. While the staff did not forecast a large downturn, by the December 2007 Tealbook, the Tealbook projection called for a negative and falling output gap (not shown). The staff's view of the depth of the downturn and its view of the natural rate evolved throughout the recession and subsequent recovery, when large disinflation did not materialize and the unemployment rate fell.

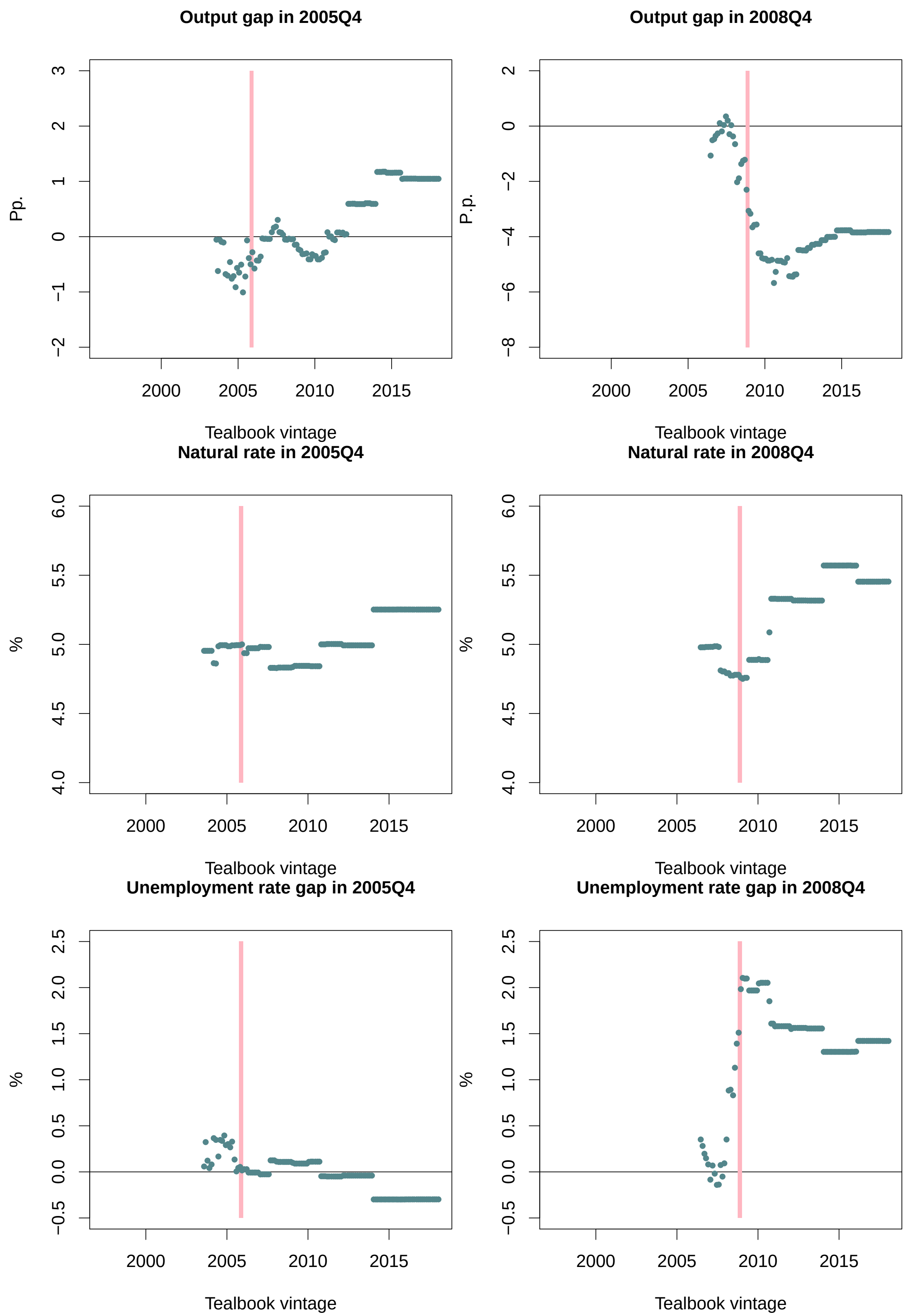

Figure 2: Estimates of output gap, natural rate, and unemployment rate gap across Tealbook vintages for two selected dates: 2005Q4 (left column), and 2008Q4 (right column). Pink shaded region denotes quarter being estimated.

2 Measuring the uncertainty around unobserved variables

Let $\hat{x}_{t+h,v}$ denote a forecast of a quarterly variable x in quarter $t+h$ from Tealbook v, $v \in t$. The condition $v \in t$ simply benchmarks the data, since Tealbook vintages are at a higher frequency than x.[4] For any forecast horizon h, we can calculate the revision from the Tealbook of vintage v until some alternative vintage, $\tilde{v}$:

$$\rho_{t+h,v} = x_{t+h,\tilde{v}} - x_{t+h,v}. \tag{1}$$

The sequence of revisions $\{\rho_{t+h,v}\}_{v=1...V}$ can be used to calculate an empirical distribution of staff revisions for each forecast horizon h. Since the staff backcasts the output gap and natural rate, h can run from large negative numbers to the number of quarters ahead in the medium-term projection. In the application below, I set $h \in \{-20, 10\}$ quarters.[5] As a baseline, I calculate staff revisions using the Tealbook estimate 40 vintages—about 5 years—after the original vintage; that is, $\tilde{v} = v + 40$. The distribution therefore measures the region into which the staff may update its estimates after five years, as implied by the staff's average behavior. Since I consider forecasts up to 10 quarters into the future, the amount of time between the quarter being forecast and the alternative vintage is at minimum about 2.5 years.

Figure 3 shows the sequence of output gap, natural rate, and unemployment rate gap estimates for the nowcast horizon $h = 0$. The blue dots are the staff's real-time, current quarter estimates, while the orange dots denote the staff estimate of the same period but from 40 Tealbook vintages later. Since the data ends with the January 2018 vintage, the final quarter shown in the figure is the first quarter of 2013. A couple of historical periods are of particular interest.

- *The late 1990s.* In the late 1990s, the staff believed that the output gap was positive in real time. As time progressed, the unemployment rate fell to low levels but price inflation did not accelerate. Measured productivity also picked up throughout this period. The combination of a low unemployment rate, low inflation, and high productivity eventually led the staff to lower its estimate of the natural rate. The lower natural rate reduced the staff's estimate of slack in the economy in this period.

- *Recovery from the Great Recession.* It is clear the staff understood the dire economic situation in

[4]For example, if v were the January 2018 Tealbook, then $\hat{x}_{t+0,Jan2018}$ would denote the January Tealbook estimate of x in 2018Q1.

[5]I choose 10 as the maximum forecast horizon due to a paucity of data when h is larger than 10. When $h = 8$ there are 90-100 observations, by $h = 10$ there are only about 30 observations, and fewer thereafter.

real-time at the trough of the last recession. Tealbooks produced in 2009 and 2010 estimated that output was 8 percent below its potential level. After that, however, the lack of a large disinflation, an apparent outward shift in the Beveridge curve, and the fact that the unemployment rate fell despite tepid output growth caused the staff to conclude that resource utilization was somewhat higher than it had first estimated. Eventually, the staff revised natural rate estimate higher.

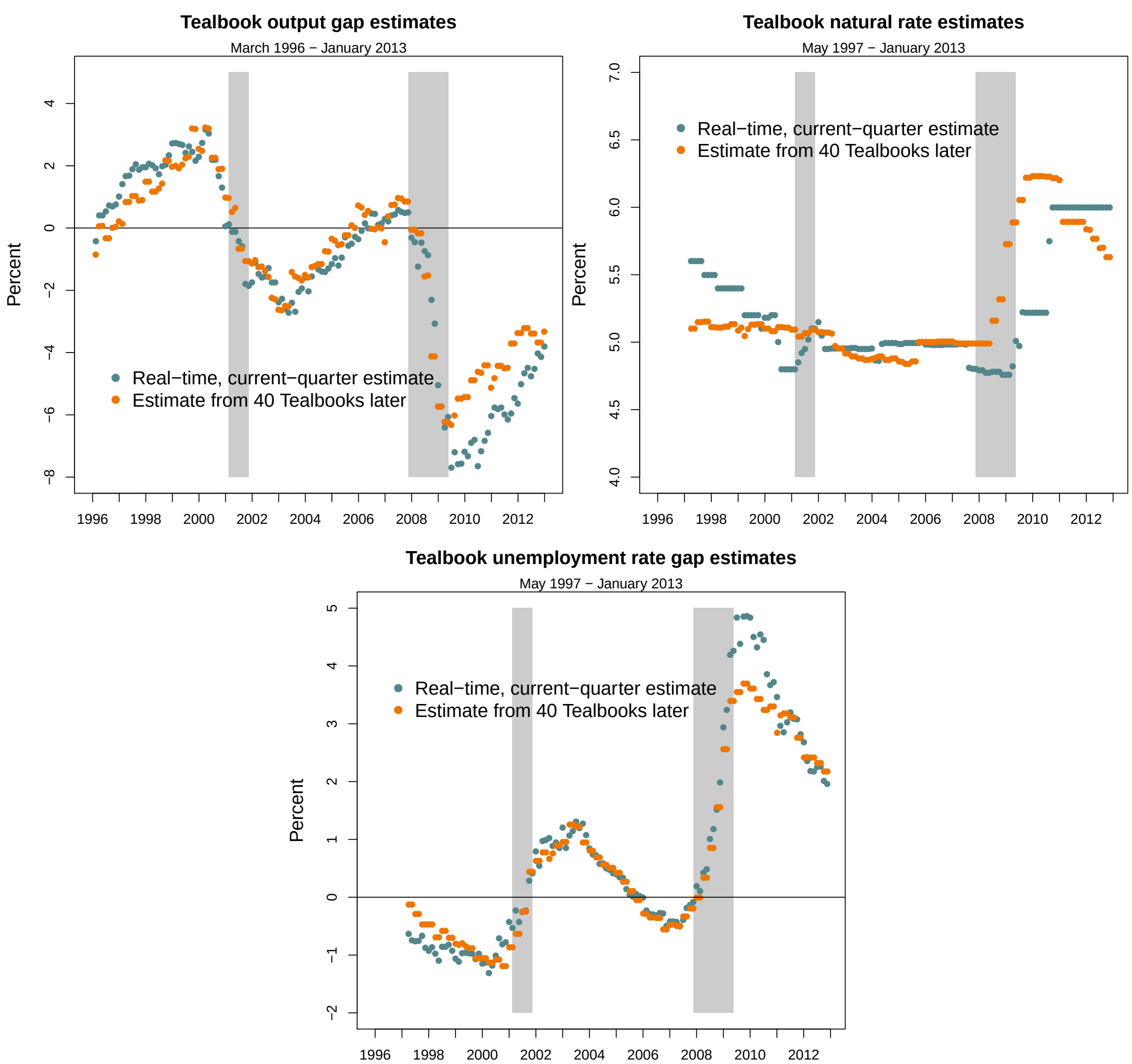

Figure 3: Output gap, natural rate, and unemployment rate gap estimates for forecast horizon $h = 0$. Shaded regions denote NBER-defined recessions.

Figure 4 presents the uncertainty around the January 2018 Tealbook estimates of the output gap, natural rate, and unemployment rate gap for forecast horizons $h = \{-20, 10\}$. The black line in each panel is the January 2018 Tealbook estimate. The figure shows 70 and 90 percent confidence intervals, calculated for each forecast horizon considered, and shown in blue and green, respectively.

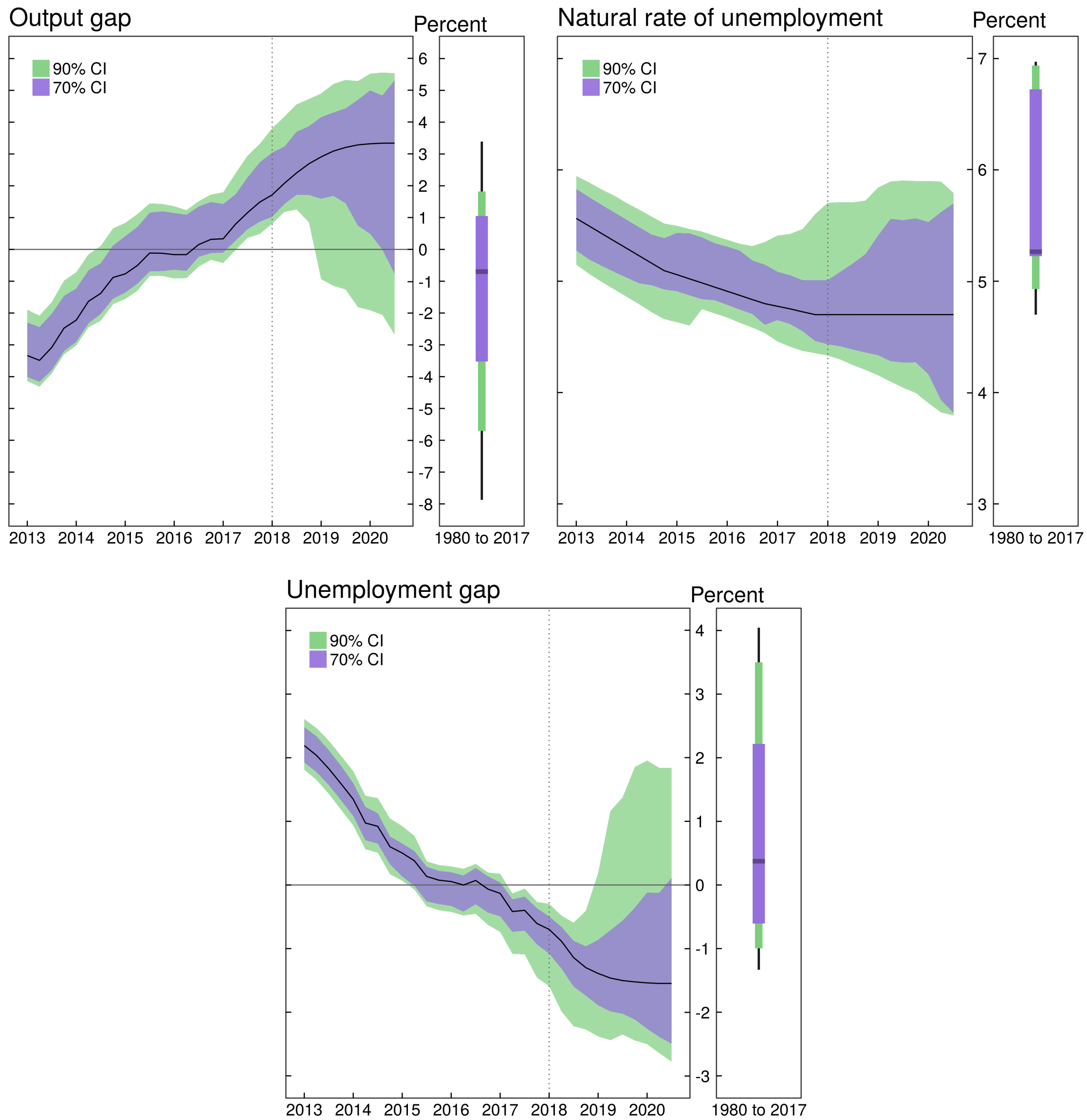

Figure 4: January 2018 Tealbook estimates of the output gap, natural rate, and unemployment rate gap. Blue and green shaded regions indicate 70 and 90 percent revision intervals, calculated using the estimate from 40 Tealbooks later. Vertical dotted line at 2018Q1. Forecast horizons $h \in \{-20, 10\}$ quarters. See text for details.

There is considerable uncertainty around staff estimates of the output gap, natural rate, and unemployment rate gap. The 70 percent revision interval around the staff's output gap estimate in 2018Q1, the vertical dotted line, ranges from 1 to 3 percent. The 70 percent revision interval around the natural rate estimate ranges from 4.4 to 5.0 percent, which, alongside the staff's unemployment rate projection, implies an unemployment rate gap between -1.1 to -0.5 pp. Since the staff revisits its estimates of these variables, sometimes well after the fact, considerable uncertainty remains around the staff's historical estimates.

Unsurprisingly, uncertainty increases as the projection moves into the future. However, the uncertainty going ahead is also somewhat asymmetric. Driven by the fact that the staff has made a few large downward revisions to its output gap estimate—typically during recessions—the asymmetry for the output gap is to the downside.[6] The asymmetry reflects the fact a recession is the largest downside risk to the medium-term projection. The risk of recession is also seen in the uncertainty about the projections of the natural rate and the unemployment rate gap.

Finally, the bands to the far right of each panel present the empirical distribution of the January 2018 Tealbook estimate of that variable. For example, the black bar in the upper-left panel shows that the range of the January 2018 output gap estimate between 1980 and 2017 is -7.9 to 3.4 percent.[7] The median value during this period is indicated by the flat line, -0.7, and the blue and green regions denote the 70 and 90 percent distributions. Comparing these distributions to the January Tealbook projection, it becomes clear that the staff medium-term forecast is for a robust economy: the output gap is forecast to be as high as it has been since 1980, while the natural rate and unemployment rate gaps are also near the end of their historical range.

3 Robustness checks

One concern about the approach laid out above is whether the uncertainty bands are sensitive to the five year period used to calculate revisions. To understand the impact of a five year window on the uncertainty measures, I compute two alternative sets of revision uncertainty. In the first, I shorten the amount of time between revisions to 24 Tealbook vintages, or three years. In the second, I compare the real-time estimates to the staff's most recent Tealbook.

Figure 5 shows the four sets of vintage data used to calculate revision uncertainty.[8] Naturally, the staff's estimates change across vintages. The effect of the different definitions on the revision bands is shown in table 2, which presents percentiles of revisions to current-quarter estimates for all three definitions of revision (i.e. $h = 0$).

[6] The largest negative revisions to the output gap occurred in 2007-2008. If I exclude these two years when calculatating the distribution of revisions, the bands remain asymmetric, although the asymmetry is less pronounced.

[7] The minimum value of the staff's output gap estimate is from 1982Q4.

[8] I can calculate more recent forecast revisions under these alternative definitions. However, I keep the sample size the same and limit forecast revisions to the sample period that was previously used.

The revision bands are qualitatively similar across definitions, especially for the 70 percent interval. The 70 percent interval for the output gap has a width of about 2 percentage points, while the width of the 70 percent confidence interval of the natural rate is about 1/2 percentage point. The width of the unemployment rate gap is a touch wider, reflecting uncertainty about the unemployment rate itself. For the output gap, the uncertainty bands do shift somewhat depending on the revision definition. The 70 percent interval for the output gap moves from $\{-0.9, 1.0\}$ under $v + 24$ to $\{-0.5, 1.7\}$ when the January 2018 Tealbook is used to define the revision. The shift reflects the fact that the staff has revised its output gap path higher, on average, since 2012.[9] The 90 percent revision intervals—especially the 5^{th} percentile—are somewhat sensitive to the particular definition used, reflecting a paucity of data available in the tails of the revision distributions.

The most notable difference for the natural rate bands is that the 70 percent interval becomes a bit wider moving from $v + 24$ to $v + 40$. When the January 2018 Tealbook is used, the 85^{th} percentile of revisions moves quite a bit higher, reflecting changes to the staff's estimates of the natural rate in the 2000s made following the 2009-2009 recession; see figure 5.

	Output gap				
Revision definition ($\tilde{v}$)	5%	15%	50%	85%	95%
v+24	-1.4	-0.9	0.1	1.0	2.1
v+40	-0.9	-0.7	0.3	1.3	2.1
January 2018 Tealbook	-1.3	-0.5	0.5	1.7	2.1
	Natural rate of unemployment				
Vintage definition ($\tilde{v}$)	5%	15%	50%	85%	95%
v+24	-0.5	-0.2	0.0	0.2	0.8
v+40	-0.4	-0.3	0.0	0.3	1.0
January 2018 Tealbook	-0.3	-0.2	0.1	0.5	0.7
	Unemployment rate gap				
Vintage definition ($\tilde{v}$)	5%	15%	50%	85%	95%
v+24	-0.7	-0.3	0.0	0.3	0.5
v+40	-0.9	-0.4	0.0	0.2	0.4
January 2018 Tealbook	-0.6	-0.5	-0.2	0.2	0.3

Table 2: Percentile estimates of revisions to the output gap, natural rate, and unemployment rate using different definitions: 24 Tealbooks, 40 Tealbooks, and the January 2018 Tealbook. All estimates for current-quarter forecast ($h = 0$). For the output gap, the vintage history is March 1996–December 2012; for the natural rate and unemployment rate gap the vintage history is May 1997–December 2012. See text for details.

[9] This fact is reflected in figure 1, where the January 2018 estimate of the output gap is close to the upper-envelope of all output gap estimates for large parts of the sample.

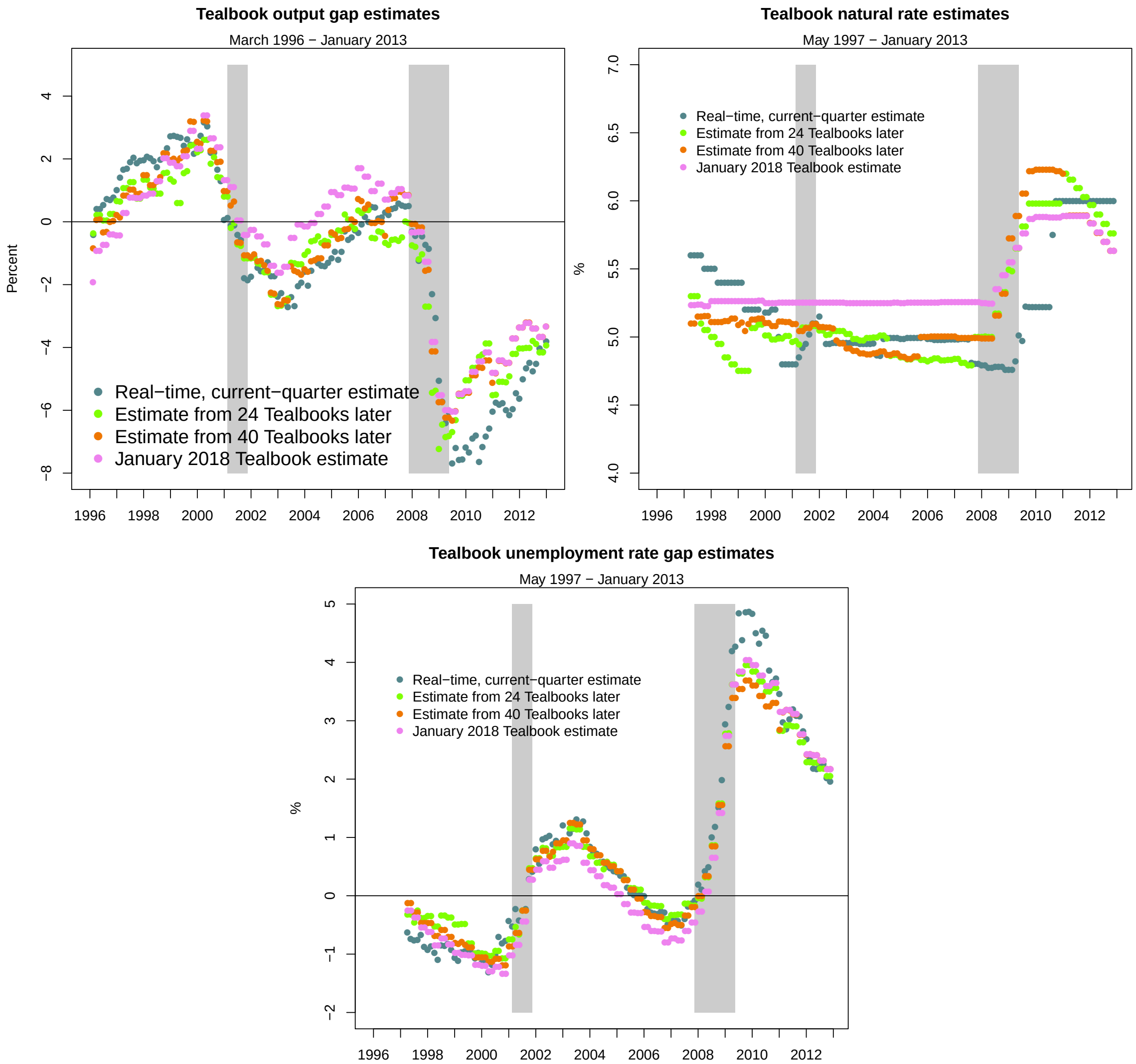

Figure 5: Output gap, natural rate, and unemployment rate gap estimates for forecast horizon $h = 0$. Shaded regions denote NBER-defined recessions.

4 Comparison to other estimates of uncertainty

Conceptually, there are at least four important sources of uncertainty for estimates of the output gap and natural rate:[10]

1. *Model uncertainty.* The true model of the economy is not known.

2. *Parameter uncertainty.* Conditional on a particular model, the true parameters of the model are not known.

[10] Stephanie Aaronson provided the basis for this taxonomy. Note that the list is not a precise mathematical condition, but instead a useful heuristic.

3. *Data uncertainty.* The economic variables that we observe are measured with error, and are subject to revision.

4. *Filtering uncertainty.* There are many paths of the unobserved variables consistent with a particular model, estimated parameters, and the observed data. Filtering uncertainty can be substantial, especially in a system with several unobserved variables, since changes in the path of one unobserved variable can be offset by changes to a different unobserved variable. For example, a lower output gap is consistent with the data if one also moves the natural rate of unemployment lower. As a result, filtering uncertainty can be significant, even *ex-post*. A component of filtering uncertainty is related to the end-point problem, the notion that trend-cycle decompositions are less precise at the end of a sample. The uncertainty associated with the end-point problem is reduced as more data is received, since additional data can rule out particularly unlikely paths of the unobservables.

It is difficult, perhaps impossible, to compute a precise decomposition of the contribution of each source of uncertainty to the total uncertainty around the unobserved states. Since the staff judgmental forecast is not conditioned on a particular model, and since the revisions occur in real-time, the revisions to the staff judgmental forecast capture model, parameter, and data uncertainty. However, the staff provides only modal forecasts, and does not report a measure of its uncertainty about its estimates. For this reason, staff revisions capture only the portion of filtering uncertainty associated with the end-point problem. In contrast, an econometric model can provide estimates of the filtering uncertainty as an output from the Kalman filter, but would not contain model or data uncertainty.

However, we can compare uncertainty estimates to get a sense of their relative importance. Table 3 compares the baseline measure of staff uncertainty developed above to two model-based measures computed using the CMC SS-PF model.[11] The lines labeled 'staff revisions' correspond to the uncertainty bands presented in figure 4. The first model-based measure, labeled 'SS-PF model,' is the output of the SS-PF model filter. These bands capture filtering and parameter uncertainty, but since they are from a particular model estimated with a particular data vintage, they do not measure model or data uncertainty.[12] The final measure of uncertainty, denoted 'SS-PF revisions,' mirrors staff revision uncertainty, using SS-PF estimates

[11] The SS-PF model is a descendant of Fleischman & Roberts (2011).

[12] The SS-PF model is estimated using Bayesian methods. One advantage of Bayesian estimation of state-space models is that the uncertainty bands around state estimates contain parameter uncertainty. State-space models estimated by maximum likelihood would capture only filtering uncertainty.

instead. They are calculated by comparing SS-PF filtered state estimates to its smoothed estimates, where the smoothed estimates are calculated with five additional years of data (also known as a *fixed lag smoother*).[13] The CMC database contains real-time data used by the SS-PF model only since 2005. Therefore the SS-PF revisions line is based on the current vintage of data and is pseudo real-time. The model is not re-estimated each quarter. Therefore, this line contains only the filtering uncertainty associated with the end-point problem, and does not contain model, parameter or data uncertainty.

The table shows the estimated gaps and natural rate in two periods, 2018Q1 and 2019Q1, as estimated at the time of the January 2018 Tealbook. Since they contain different sources of uncertainty, the bands differ across the three calculations. The tightest bands—those using revisions to SS-PF estimates—capture only end-point uncertainty. The staff revision bands are larger than the SS-PF revisions, since they include more uncertainty sources (model, parameter, data, and some filtering uncertainty). The two revision-based bands are not terribly different, suggesting that model, parameter, and data uncertainty are not the primary sources of uncertainty.[14] The largest bands come from the SS-PF model, which includes parameter and filtering uncertainty.

Given that parameter uncertainty appears to be a relatively small contribution to total uncertainty, filtering uncertainty is likely the largest source of imprecision in the estimates. Revision-based uncertainty measures incompletely capture filtering uncertainty. The staff forecast is of the 'most likely' path for the unobservables, and does not consider less likely alternative paths. Thus, while the revision uncertainty of the staff is intuitive and well-defined, it, if anything, understates the imprecision of staff estimates of these unobserved variables.

5 Discussion

In this memo, I developed a new measure of the uncertainty that surrounds judgmental estimates of unobserved economic variables. The measure proxies uncertainty with the size of the revisions to those estimates. I find that Tealbook estimates of the output gap, natural rate and unemployment rate gap revise considerably. Revision uncertainty is an intuitive and relevant measure of the uncertainty present in Tealbook estimates of

[13] Specifically, let $\hat{x}_{t|\tau}$ denote the posterior median estimate of the smoothed state vector, x in quarter t, given data through quarter τ. The lines labeled 'SS-PF revisions' presents uncertainty bands calculated from the empirical distribution of $\hat{x}_{t|t+20} - \hat{x}_{t|t}$.

[14] Of course, the comparison isn't a complete one. The SS-PF framework is by design quite similar to the staff, so the comparison likely understates model uncertainty. In addition, the staff observes a wide variety of data whereas the model is limited to a handful of the most important indicators.

				Output gap		
	Period	5%	15%	Estimate	85%	95%
Staff revisions	2018Q1	0.6	0.8	1.7	2.8	3.6
SS-PF model	2018Q1	-0.6	0.3	1.6	2.8	3.5
SS-PF model revisions	2018Q1	0.6	1.0	1.6	2.5	3.0
Staff revisions	2019Q4	-1.0	1.6	2.9	4.1	4.9
SS-PF model	2019Q4	-1.8	-0.8	1.4	3.3	4.7
SS-PF model revisions	2019Q4	-1.5	0.3	1.4	2.3	2.8
				Natural rate of unemployment		
		5%	15%	Estimate	85%	95%
Staff revisions	2018Q1	4.3	4.4	4.7	5.0	5.7
SS-PF model	2018Q1	3.5	3.9	4.7	5.4	5.9
SS-PF model revisions	2018Q1	4.1	4.3	4.7	5.4	5.6
Staff	2019Q4	4.2	4.3	4.7	5.4	5.8
SS-PF model	2019Q4	3.3	3.8	4.7	5.5	6.0
SS-PF model revisions	2019Q4	4.0	4.4	4.7	5.4	5.7
				Unemployment rate gap		
		5%	15%	Estimate	85%	95%
Staff	2018Q1	-1.6	-1.1	-0.8	-0.5	-0.3
SS-PF model	2018Q1	-1.9	-1.4	-0.7	0.0	0.5
SS-PF model revisions	2018Q1	-1.5	-1.2	-0.7	-0.4	-0.1
Staff	2019Q4	-2.4	-1.9	-0.8	-0.8	0.2
SS-PF model	2019Q4	-2.2	-1.7	-0.9	-0.0	0.4
SS-PF model revisions	2019Q4	-1.7	-1.4	-0.9	-0.4	0.0

Table 3: Estimates of output gap, natural rate, and unemployment rate gap in 2018Q1 and 2019Q1, staff and CMC supply-side model with production function. January 2018 Tealbook vintage estimates. See text for details.

unobserved variables, since it describes the region into which the staff is likely to revise its estimates in the future.

Finally, the revision uncertainty presented above is based on the staff's average behavior over the past two business cycles, and does not capture any time-variation in the uncertainty around these estimates. Figure 6 suggests that this time variation may be important. The figure shows the distribution of estimates of the natural rate of unemployment from the Survey of Professional Forecasters. (The SPF has asked its respondents for their estimate of the natural rate of unemployment in its third quarter survey in each year since 1996.) SPF responses suggest that uncertainty around the natural rate is time-varying. Since 1996, the spread between the 15th and 85th percentiles has been much larger during and immediately following recessions than during expansions. Extending the current framework to capture this time-variation is an interesting future project.

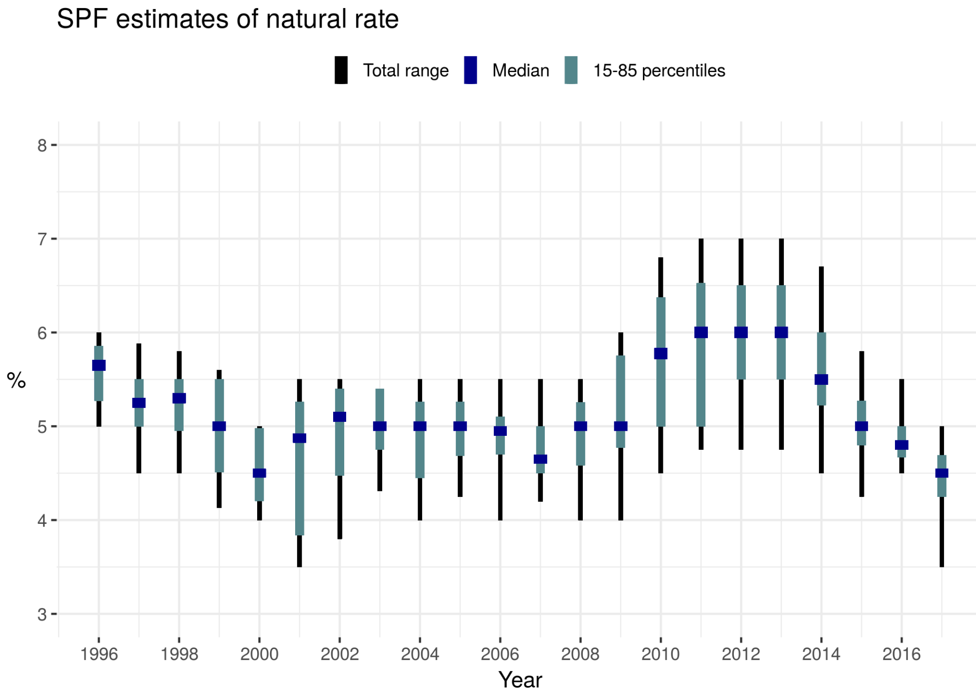

Figure 6: Estimates of the natural rate of unemployment from the Survey of Professional Forecasters. The dark blue square is the median estimate. The light blue region denotes 15^{th} and 85^{th} percentiles and black line is the range of estimates. Estimates provided in the third quarter survey each year; on average 40 respondents per survey.

References

Fleischman, C. A. & Roberts, J. M. (2011), From many series, one cycle: Improved estimates of the business cycle from a multivariate unobserved components model, Finance and Economics Discussion Series 2011-46, Board of Governors of the Federal Reserve System,, Washington.