Office of Economic Policy
Working Paper 2021-02
January 2021

The Expectations Gap

Alexander D. Schibuola and Andrew B. Martinez

The Expectations Gap

Alexander D. Schibuola* and Andrew B. Martinez†

January 14, 2021

Abstract

The output gap hinges on estimates of potential GDP, which can vary widely and are ultimately unobservable. Beckworth (2020) circumvents this using historical surveys of professional forecasters' nominal GDP projections to construct an "expectations gap," which shifts the benchmark for economic performance to what GDP was expected to be. However, this approach has some limitations, which we address as follows: (1) we assess the sensitivity of the expectations gap to alternative surveys of professional forecasters and different forecast horizons; (2) we use forecaster disagreement to capture the uncertainty surrounding the expectations gap; (3) we estimate both the nominal and real expectations gaps and compare them against other conventional output gap measures; (4) we produce real-time forward-looking estimates of the expectations gap and assess their performance near business cycle turning points. Though the expectations gap is closely correlated with existing output gap measures, we find that several historical episodes suggest that it contains additional useful information about the state of the economy.

Keywords: Business Cycles; Financial Stability; Forecast Accuracy; GDP Forecasts

JEL classifications: C18, E32, E37

**Office of Macroeconomic Analysis, U.S. Department of the Treasury. The views expressed here are those of the authors and do not necessarily represent the views of the Treasury Department or the U.S. Government.* *Thanks to David Beckworth, Stephen Miran, and Tara Sinclair for helpful comments and discussions.* *Contact: Alexander.Schibuola@treasury.gov.*

†Office of Macroeconomic Analysis, U.S. Department of the Treasury; H.O. Stekler Research Program on Forecasting, The George Washington University; Climate Econometrics at Nuffield College. Contact: Andrew.Martinez@treasury.gov.

Navigating by the stars can sound straightforward. Guiding policy by the stars in practice, however, has been quite challenging of late because our best assessments of the location of the stars have been changing significantly.

Federal Reserve Chair Jerome Powell, August 18, 2018

1 Introduction

The output gap – the percentage difference between actual GDP and potential GDP – is used to forecast inflation and inform economic policy.[1] However, muted inflationary pressures and the ascendance of financial stability concerns, along with the unobservability, sizeable revisions, and variety of potential GDP estimates, have reduced the reliability of output gaps as a measure of economic slack.[2] Potential GDP can be estimated using production functions,[3] structural and statistical approaches,[4] or assorted variations of the two.[5] Accordingly, different models produce very different estimates. For example, between 2012 and 2017, the Congressional Budget Office (CBO) reported a negative output gap, while the Laubach and Williams (2003) methodology consistently reported a positive output gap.

More recently, Beckworth (2020) calculates economic slack as the gap between nominal and "neutral" GDP. He estimates neutral GDP by averaging five years of professional forecasts for a particular quarter's nominal GDP.[6] Since professional forecasts are measured by the central tendency from surveys, they do not depend on any single model and reflect what actual expectations were at the time. This shifts the emphasis from "navigating by the stars" to navigating by what agents *expect*.

However, this "expectations gap" has several limitations. First, its sensitivity to alternative forecaster surveys and forecast horizons is unclear. Second, it is difficult to assess the strength of its signal about the economy since there are no measures of uncertainty. Third, estimation of the expectations gap in nominal terms complicates a direct comparison with conventional output gap measures, which are typically formulated in real terms. Fourth, it does not fully leverage the availability of real-time data to assess the performance near business cycle turning points.

We make the following contributions: First, we assess the sensitivity of the expectations gap to different surveys of professional forecasters and alternative forecast horizons. Second, we estimate the uncertainty surrounding the expectations gap by incorporating individual forecaster disagreement. Third, we evaluate how the nominal and real expectations gaps compare against the nominal and real output gaps produced by the CBO, Quast and Wolters (2020), and the

[1] E.g., the Taylor rule, see Taylor (1993), the output gap has also been considered for fiscal policy rules; see Taylor (2000) and Kumhof and Laxton (2013).
[2] See Powell (2020): "Before the Great Moderation, expansions typically ended in overheating and rising inflation. Since then … a series of historically long expansions had been more likely to end with episodes of financial instability…" Federal Reserve Chair Jerome Powell, August 27, 2020.
[3] See Shackleton (2018).
[4] See Fleischman and Roberts (2011) and Berge (2020).
[5] See Ball and Mankiw (2002), Laubach and Williams (2003) and Williams et. al. (2020), Coibion et. al. (2018).
[6] This is a refinement of one of the approaches evaluated in Coibion et. al. (2018).

nominal and real expectations gaps derived from the Federal Reserve Board's Greenbook forecasts. Fourth, we construct real-time and forward-looking estimates of the expectations gap and evaluate their performance, especially around business cycle turning points.

Our findings are as follows. First, the use of alternative professional forecaster surveys does not substantively alter expectations gap estimates. The lack of sensitivity to any particular survey suggests it is a robust economic indicator. However, the expectations gap measures are sensitive to the selection of forecast weighting and forecast horizon. Second, our analysis of forecaster disagreement shows that there is considerable uncertainty surrounding the expectations gap, which is consistent with traditional measures of the output gap; see Berge (2020). Despite the uncertainty, the expectations gap unambiguously signals economic overheating or underperformance in four distinct episodes. Third, we find meaningful differences between the nominal and real expectations gaps. This sharply contrasts with the CBO's nominal and real output gaps that are constructed to be identical. For example, during the mid-2000s when home values were judged to be overvalued, the real expectations gap and the CBO output gap indicate little-to-no economic overheating, while the nominal expectations gap does. Fourth, our real-time analysis of the expectations gap shows that significant revisions occur – mostly due to comprehensive BEA updates. However, these revisions are comparable in size to those made to CBO's output gap. Lastly, the forward-looking estimates outperform simple benchmarks and occasionally provide a leading indicator of business cycle turning points.

The next section describes the methodology and intuition of neutral GDP and the expectations gap. Section 3 estimates the expectations gap and explores the sensitivity to alternative surveys, forecast horizons, and forecaster disagreement. Section 4 compares the nominal and real expectations gaps with other output gap estimates. Section 5 derives real-time, forward-looking expectations gaps and evaluates their performance around business cycle turning points. Section 6 concludes.

2 Background and Methodology

This section explains the underlying intuition and estimation of neutral GDP. While conventional output gap measures are the difference between actual and *potential GDP*, the expectations gap measures the difference between actual and *neutral GDP*. Beckworth (2020) defines neutral GDP as "the public's expected growth path of nominal income." If actual GDP equates to neutral GDP, then, on average, households' and firms' expectations are validated by outcomes. Diagram 1 illustrates the key differences between the formulation of conventional output gaps and the expectations gap.

A key rationale of the expectations gap is that deviations of income from expectations that informed past adoption of debt obligations alters debt burdens, which can impinge on current and future spending decisions. Thus, while the conventional output gap measured in inflation-adjusted terms is geared towards predicting inflation vis-à-vis a Phillips curve, the expectations gap in nominal terms is more focused on risks to financial stability. Greater-than-expected income growth can lead to greater debt accumulation and leave agents more susceptible to negative shocks when lower-than-expected income episodes arise.

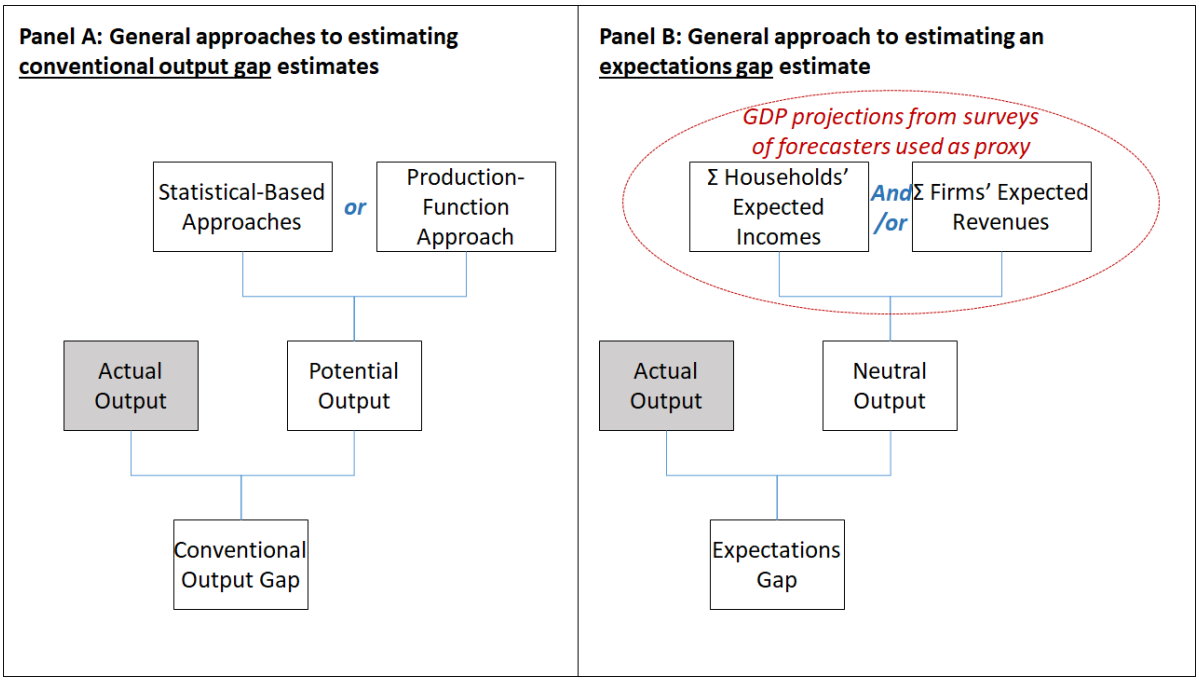


Diagram 1: Schematic Illustrating Different Components of Gap Measures

2.1 Intuition and Simplifying Assumptions

A household's expected income stream influences its spending decisions.[7] If, for example, a household expects an annual income of $60,000 indefinitely and finances the purchase of a home through a 30-year mortgage amortized by fixed payments of $2,000 per month, then 40 percent of the household's monthly income is tied up by this fixed debt obligation. If an economy-wide shock leads the household's realized income to fall 10 percent below expectations, then its monthly mortgage payments will encumber 44.4 percent of its income. This forces the household to reduce current spending, savings, or both, until the debt is paid off or restructured.[8] The deviation of actual outcomes from past expectations can influence both current and future spending via sticky nominal debt obligations. Analogously, a firm's allocation of resources can be affected when actual revenues deviate from expectations. The incorporation of these dynamics into macroeconomic models suggests sticky debt burdens play a more important role than sticky prices; e.g. see Sheedy (2014) and DiCecio and Bullard (2019).

Aggregated household income expectations or aggregated business revenue expectations represent the "neutral" level of income in the economy wherein expectations, in the aggregate, are consistent such that past spending and borrowing decisions prove to have been ideal. Since aggregate private agents' individual income projections do not exist, Beckworth (2020) uses the GDP projections from surveys of professional forecasters as a proxy to estimate neutral income.[9]

[7] E.g. see Romer (1990) for an argument of how collapsing future income expectations led to declines in purchases of durable goods during the Great Depression.
[8] This may especially be the case if the income loss is expected to be permanent.
[9] This equates aggregate income, spending, and output with GDP, which implies that changes in business inventories are irrelevant. This should be satisfied when (1) the expected contribution to GDP growth from business inventories

Though some households and firms may fare better or worse than expected, if actual GDP remains near the neutral level, these differences cancel out. In contrast, if actual GDP falls short of neutral GDP, this signifies that households' and firms' expectations were, *on average*, overly optimistic.

2.2 Description of Survey Forecasts

The Survey of Professional Forecasts (SPF) is a quarterly survey of about 40 mostly academic and business forecasters. It was first published in Q4-1968 as the ASA/NBER Economic Outlook Survey. Management of the survey was transferred to the Federal Reserve Bank of Philadelphia in 1990. In total, 443 unique forecasters have participated in the SPF since 1968.

The SPF is published near the middle of the second month of every quarter. Forecasts are reported for fixed-horizons, extending four-quarters ahead. In 1990, the SPF began including a forecast of the average annual real GDP growth rate over the next 10 years and the average annual CPI inflation rate over the next 10 years.[10] While the latter is updated with every quarter, the longer-run real GDP forecast is updated only in the first SPF survey of every year.

The Blue Chip Economic Indicators (BCEI) is a monthly survey of about 50 mostly business forecasters maintained by Wolters Kluwer that began in 1976. More than 115 unique forecasters have participated in the survey since 1980.

The survey is published on the tenth business day of every month with quarterly fixed-event forecasts that shrink from eight- to four-quarters ahead as a given year progresses (e.g., the monthly short-run forecast surveys for 2020 have a fixed endpoint of Q4-2021). Longer-run forecasts are updated twice per year in the March and October surveys. Initially only the 2-6 and 6-11 year average forecasts of variables were available. Since October 1982, the surveys have included annual growth forecasts for two-to-six years ahead and a five-year average forecast for 6-11 years ahead. Prior to 1990, only the average – "consensus forecast" – of variables was reported, however, since then, the average forecasts of the top 10 and bottom 10 forecasters have also been reported.

2.3 Estimating Neutral GDP and the Expectations Gap

Following Beckworth (2020), the neutral level of GDP at time t is measured as the average of the forecasts of time period t's GDP submitted over the past 20 quarters, including the nowcast:

$$\overline{Y}_{t,v} = \frac{1}{20}\sum_{h=0}^{19} Y_{t|t-h,v}, \qquad (1)$$

is minimal over longer forecast horizons and (2) the average historical contribution of changes in business inventories to GDP growth is 0.0 percent.

[10] The SPF and BCEI surveys prior to 1992 reported forecasts for GNP rather than GDP. We refer to GNP as GDP where applicable for simplicity.

where $Y_{t|t-h,v}$ is the aggregate forecast of GDP at time t and data vintage v based on expectations at time $t-h$. The aggregate forecast represents the central tendency of professional forecasters' expectations for the level of GDP at time t. Beckworth (2020) uses the SPF-reported median while we focus instead on the BCEI-reported consensus (i.e. average) forecast.[11]

This formulation does not impose restrictions on how forecasts are generated, i.e. it is not dependent on any one method or model. Quast and Wolters (2020) use a modified-Hamilton filter to estimate potential GDP. It is a special case in which $Y_{t|t-h,v}$ represents the forecasts from an autoregressive model with four-quarter lags. This provides a statistical justification for equation (1). It also implies that neutral and potential GDP estimates would be aligned if all forecasters used this model.[12] However, as the expected path of actual GDP can deviate from potential GDP, especially in the short-run, then measures of neutral GDP will differ from estimates of potential GDP.

Another important difference between the approach by Quast and Wolters (2020) and the expectations gap is the choice of forecast horizons. Neutral GDP estimates are centered on 10-quarters-ahead forecast with horizons between zero to 20-quarters ahead. This is considerably wider than the Quast and Wolters (2020) potential GDP estimates that are centered on the eight-quarters-ahead forecast with horizons ranging from four- to 12-quarters ahead. However, Quast and Wolters (2020) note that centering estimates around the 10-quarters-ahead forecast effectively allows for the inclusion of longer financial and credit cycles into the standard business cycle; e.g., see Beaudry et. al. (2020). This is consistent with the intuition underpinning the calculation of neutral GDP. The inclusion of longer horizons captures the rigidities stemming from fixed nominal debt obligations, which are assumed to be "sticky" for five years.

In practice, the GDP-level forecast for a particular horizon is not always available, and when it is, it is associated with a particular data vintage. Following Beckworth (2020), we construct the forecast in levels based on the underlying GDP growth rate forecasts so that

$$Y_{t|t-h,v} = \prod_{j=0}^{h} \left(1 + g_{t-h+j|t-h-1}\right) Y_{t-h-1,v}, \qquad (2)$$

where $g_{t-h+j|t-h}$ is the expected quarterly GDP growth rate for period t that is formulated at time $t-h$, and $Y_{t-h-1,v}$ is the initial level of GDP.

It is possible to obtain the quarterly-growth rate forecasts directly from forecaster surveys for up to four-quarters ahead for the SPF and between four- to eight-quarters ahead for the BCEI. To extend the SPF forecasts out 5-year ahead beyond the fixed horizon of four quarters, Beckworth

[11] Although the BCEI forecasts are released monthly, for simplicity we do not change the notation. Since three months of BCEI forecasts are released in a given quarter, they can be converted from monthly to quarterly by averaging the monthly forecasts.

[12] Quast and Wolters (2020) focus on the log of the level of GDP whereas here we emphasize the level of GDP so that this special case requires that the exponential of the AR(4) model forecasts of the log level are a good approximation of forecasts of the level of GDP. Depending on how forecasters generate their expectations, the estimate of the neutral level of GDP could also align with other estimates of potential GDP.

(2020) uses the latest 10-year real GDP growth rate forecast as a proxy for growth rates beyond four-quarters ahead. For the BCEI, we use the latest long-run annual growth rate forecasts to forecast GDP out to five-years ahead.

While the neutral level of GDP can be constructed in either nominal or real terms, we focus on the nominal measure following Beckworth (2020). While short-run nominal GDP growth and level forecasts are available from both SPF and BCEI, the SPF only produces real long-run growth forecasts. Therefore, Beckworth (2020) constructs a proxy long-run nominal GDP growth forecast based on the long-run real GDP growth forecast and the CPI inflation forecast. Beckworth (2020) also adjusts the CPI forecast to conform to the GDP deflator according to the historical wedge between short-run forecasts of CPI and the GDP deflator.

The choice of the initial value of output, i.e. $Y_{t-h,v}$, is important for ensuring that a consistent GDP-level is applied over time. For example, Beckworth (2020) focuses primarily on the latest vintage of GDP data, $v = T$. However, since real-time estimates are subject to substantial revisions over time, they may exhibit additional volatility beyond the underlying business cycle fluctuations; see Orphanides and Norden (2002). We explore this in Section 5.

Lastly, the expectations gap is constructed as the percentage difference between the observed level of GDP at time t and the estimate for neutral GDP at time t

$$GAP_{t,v} = \left(\frac{Y_{t,v} - \bar{Y}_{t,v}}{\bar{Y}_{t,v}}\right) * 100, \qquad (3)$$

where the choice of data vintage – e.g., first estimate or latest estimate of GDP – can affect the expectations gap estimates; see Sinclair and Stekler (2013).

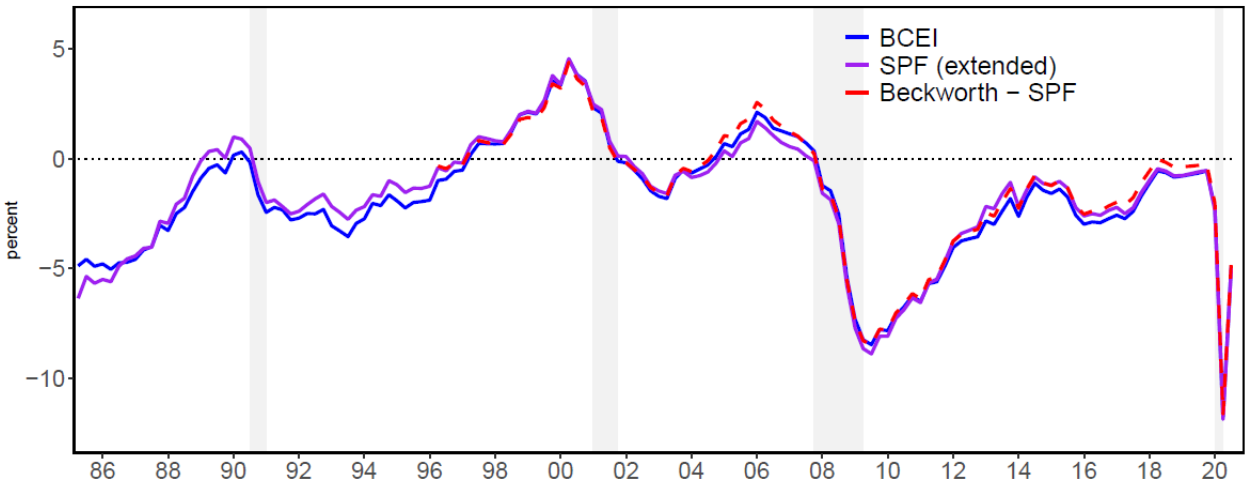
3 The Sensitivity of the Expectations Gap

This section evaluates the sensitivity of the nominal expectations gap along three dimensions: First, the sensitivity to the choice of forecaster survey; second, varying the forecast horizons; and third, accounting for forecaster disagreement. Ultimately, regardless of whether the SPF or BCEI survey is used, the expectations gap estimates exhibit little difference. In contrast, expectations gap estimates are sensitive to the selected forecast horizons and the weights assigned to forecasts, particularly at longer horizons. Forecaster disagreement quantifies the range of uncertainty surrounding the expectations gap and highlights the four distinct episodes in which the expectations gap gave an unambiguous signal about the state of the economy.

3.1 Alternative Forecaster Surveys

We start by calculating the expectations gap from Q1-1985 to Q3-2020 using the BCEI and SPF forecasts. To extend the latter back to Q1-1985, when long-run forecasts were not available we use the last published growth rate forecast for the four quarters beyond the current quarter as a proxy for the long-run growth rate up to five-years ahead (this is referred to as "SPF (extended)" in Figure 1).

Except for minor deviations around 1991-97, Figure 1 illustrates that the BCEI and SPF nominal expectations gaps are very similar, which implies that they capture features of the economic cycle rather than features specific to the surveys. Given the minor differences between the SPF and BCEI measures, the remaining analysis in this paper focuses on the BCEI measure.



Sources: Authors' calculations based on data and forecasts from BEA, SPF, and BCEI and Beckworth (2020)

Figure 1: Measures of the Nominal Expectations Gap (Q1-1985 – Q3-2020)

Figure 1 also illustrates that using the one-year-ahead growth rate forecasts as a proxy for longer-horizon SPF forecasts is tenable since there are no large discrepancies from the BCEI-derived measure between 1985 and 1990. Thus, it is possible to extend the SPF-derived expectations gap measure back to the early 1970s using this approach.

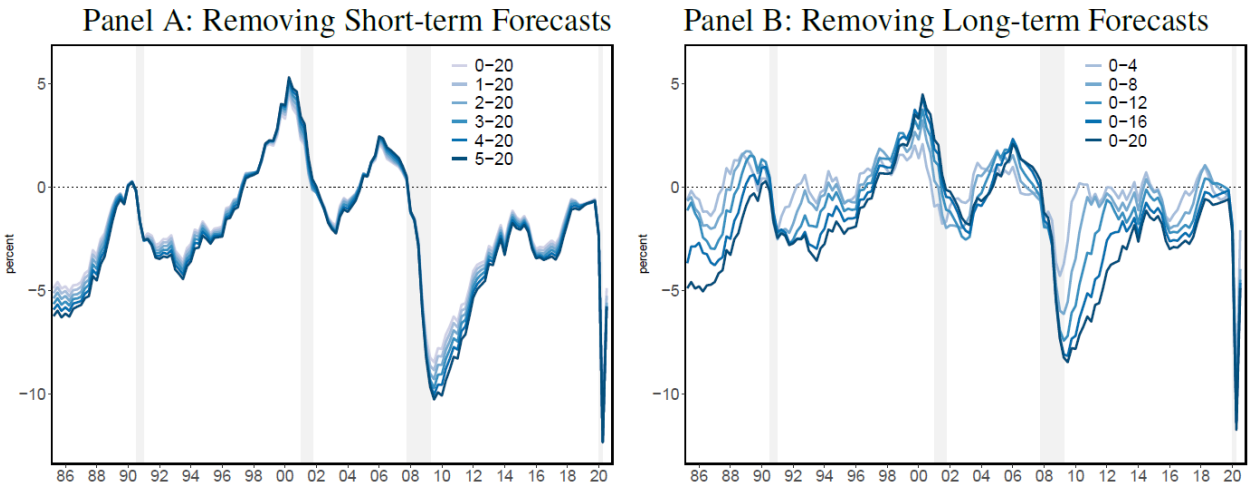
3.2 Alternative Forecast Horizons

Isiklar and Lahiri (2007) show that professional forecasters can capture both the trend and cycle in economic growth up to one year ahead. Therefore, including these short-term forecasts within the measure of neutral GDP may produce a more muted expectations gap. By removing the short-term forecasts from the estimates of neutral GDP, we can observe the consequent impact on the expectations gap.

In Panel A of Figure 2 lines go from lighter to darker as more short-term GDP-level forecasts are excluded from the estimates of neutral. The removal of short-term forecast horizons from the calculation of neutral GDP puts greater weight on the more distant forecasts of GDP made for a particular quarter. Differences tend to arise around peaks and troughs in the expectations gap series. This is especially pronounced following the trough in the expectations gap in 1985-88 and in 2009-12 where removing the short-term forecasts exacerbates the gap.

Conversely, removing forecasts made in the more distant past from a particular quarter's neutral GDP estimates signifies that debts are less sticky. At the limit, were neutral GDP derived solely from the nowcast, it would imply that debt obligations are fully flexible and are adjusted from quarter-to-quarter so as to leave the distribution of debt burdens unchanged. Panel B of Figure 2

shows the consequences of removing the more distant forecasts and depicts a faster convergence of the expectations gap to zero.



Sources: Authors' calculations based on data from BEA and forecasts from BCEI

Figure 2: Sensitivity of the Nominal Expectations Gap to Alternative Horizons

Figure 2 illustrates that the expectations gap – and the amount of economic slack it implies – depends on which level forecasts are included. Averaging five years of forecasts to estimate a quarter's level of neutral GDP is a plausible starting point if the average debt contract is assumed to be refinanced within five years from the contract date. However, if leverage is driven at times by an increase in shorter-term debt contracts, then the five-year average may artificially exacerbate the cyclicality of the expectations gap. This is especially pertinent if long-term debt obligations are less likely to be contracted during periods of high and volatile inflation.

In addition to the optimal range of forecast horizons, the equal weighting scheme for each forecast horizon merits consideration, as this assumes debts are contracted continuously and uniformly over time. Debts are accrued at different rates across the business cycle such that more lending takes place during an expansion than a recession. To better align neutral GDP estimates with actual debt burdens, then forecasts made during a recession should receive less weight than those made during an expansion. Future research should endogenize the optimal range and weighting of forecasts to vary over time and in response to economic variables, such as when debt burdens are accrued and the debt maturity structure.

Another concern is that many of the forecasts used to estimate the neutral level of GDP are derived from less frequently updated longer-term growth rate projections. Thus, the actual forecast path of GDP over longer-horizons may be more variable than the intermittently updated longer-run forecasts permit. However, this issue may be mitigated since the longer-run forecast GDP level path is heavily influenced by the short-term forecasts that are updated with every survey release. This is demonstrated by the SPF (extended) measure in Figure 1, which is constructed by extrapolating the one-year ahead growth rate forecast out over five years, very closely tracking the measures based on the less frequently updated long-term forecasts.

3.3 Measuring Uncertainty through Forecaster Disagreement

The expectations gap uses measures of central tendency to construct the neutral level of GDP, which ignores disagreements between individual forecasters that can be persistent and linked to the business cycle; see Patton and Timmermann (2010) and Bürgi and Sinclair (2020). A large dispersion around the central tendency indicates that aggregate expectations about the neutral level may be more uncertain. To quantify this uncertainty, we estimate the forecaster disagreement around the expectations gap.

There are two different ways to measure of forecaster disagreement: disagreement between individual forecaster measures of the neutral level and disagreement between the underlying forecasts of GDP. Adapting equation (1) to allow for individual forecasters allows us to focus on the first source:

$$\overline{Y}_{i,t,v} = \frac{1}{20}\sum_{h=0}^{19} Y_{i,t|t-h,v}, \qquad (4)$$

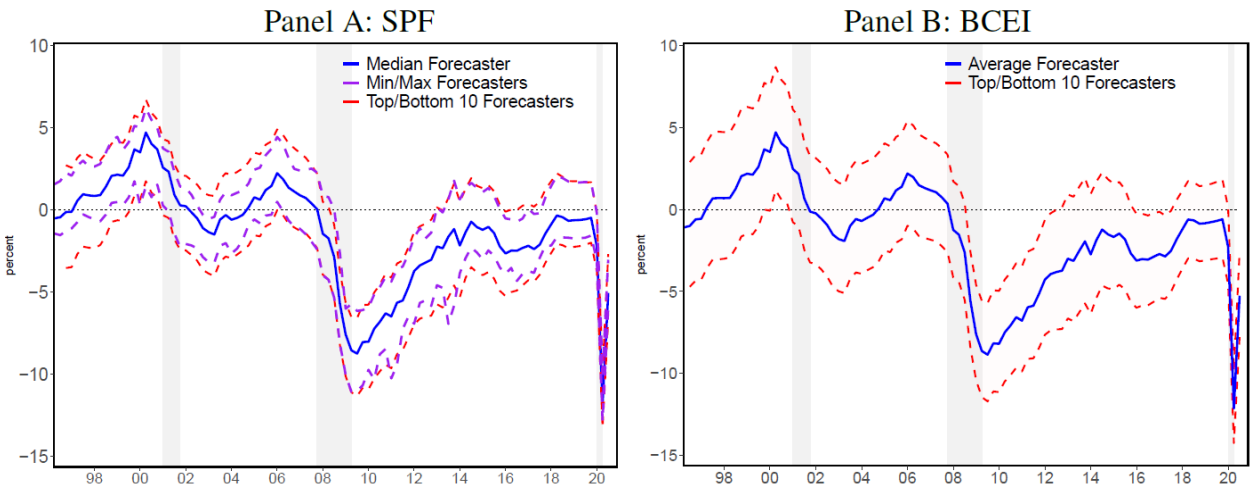
where the index i represents an individual forecaster. This illustrates that a measure of the neutral level for each individual forecaster can be constructed at each point in time. However, to do so, forecasters must have generated at least one forecast for time period t per quarter for 20 consecutive quarters (i.e., a forecaster is in the sample continuously for five years). In practice we relax this restriction by allowing forecasters to miss up to three consecutive surveys and interpolate between the missing forecasts to ensure retention of a sufficiently large sample over time.

While the BCEI does not release individual forecasts publicly, the SPF does. In total, 79 unique SPF forecasters satisfy the requirements (out of the full sample of 186 since 1990). We construct an expectations gap for each of these forecasters and then calculate the minimum and maximum individual expectations gaps at each point in time. This captures the full range of disagreement between forecasters and measures the uncertainty around aggregate expectations at each point in time; see Lahiri and Sheng (2010).

Panel A of Figure 3 plots this range (in purple) along with the median derived from the individual measures. The median of the individual expectation gaps is very similar to the expectations gap constructed from the median forecaster. This suggests that focusing on the subset of 186 forecasters does not bias our analysis. The dispersion between individual expectation gaps is fairly stable over time, ranging approximately three to four percentage points with an increase during the dotcom boom in 2000 and a prolonged increase during the housing boom and bust in 2005-10. Only in recent years has it gradually declined.

An alternative approach for measuring forecaster disagreement is to consider the forecasts from the top or bottom 10 forecasters at each point in time. This is beneficial because it does not rely on individual expectations gaps with ranges that could be biased by the entry and exit of individual forecasters over time. Conversely, it is not necessarily representative of any individual forecaster's views since the top / bottom forecasters vary across variables, horizons, and time. It

also does not capture the full range of forecaster dispersion, but each horizon corresponds analogously to a plus / minus one standard deviation in forecaster disagreement.[13]



Sources: Authors' calculations based on data from BEA amd forecasts from SPF and BCEI

Figure 3: Incorporating Forecaster Disagreement

Panel B of Figure 3 plots the top / bottom 10 forecaster measures of the expectations gap for the BCEI while Panel A plots it for the SPF (in red). The dispersion of the BCEI forecasters is larger than the SPF (7 percentage points vs. 4 percentage points), indicating that individual expectations in the BCEI are more diffuse than those in the SPF.

Both measures of forecaster disagreement illustrate that there are four episodes during which the economy clearly performed significantly above or below most individual forecasters' expectations. The first episode corresponds with the dot-com boom when the economy grew faster than expected; the second episode was during and after the 2008-09 recession when the economy grew much slower than expected; the third episode was during the so-called "invisible recession"[14] of 2015-16; and the fourth episode occurred during the COVID-19 pandemic induced recession in 2020. Each of these episodes corresponds to important turning points in the economy and so support the interpretation of the expectations gap as an indicator of business cycle turning points.

4 Comparing Nominal and Real Gaps

While this paper focuses on the nominal expectations gap, this section applies the same methodology from Section 3 to produce a real expectations gap from real GDP data and BCEI forecasts. Despite a high degree of correlation, the nominal expectations gaps appears to suggest additional useful information about the state of the economy.

[13] Reifenschneider and Tulip (2019) use a similar approach to calculate forecast uncertainty.
[14] See Neil Irwin in 2018: https://www.nytimes.com/2018/09/29/upshot/mini-recession-2016-little-known-big-impact.html, last accessed 17 September 2020.

We also compare the BCEI-based expectations gap with three alternative gap measures: the real-time output gaps from the Federal Reserve Board's Greenbook[15] staff forecasts and the nominal and real expectations gap estimates that we derive from them; the CBO's output gap; and the output gap generated using the modified Hamilton filter of Quast and Wolters (2020). There are similarities between measures, however, allowing for nominal fluctuations produces a more holistic measure of the economic cycle.

Panel A of Figure 4 plots the nominal and real expectations gaps derived from the BCEI surveys. From the mid-to-late 1980s the real gap was positive while the nominal gap was negative, which is consistent with the disinflation of that period. The differences between the real and nominal gaps was broadly stable throughout the 1990s even as both measures increased. Since the 2001 recession, the nominal and real measures have followed broadly similar patterns except for two episodes. First, during the mid-2000s, when home prices and construction were rapidly rising, the nominal gap indicates overheating whereas the real gap does not. The second episode occurs in 2015-16, when the nominal gap drops more sharply than the real gap, thus better capturing the "invisible recession" that was associated with declines in energy, agriculture, and manufacturing activity. These episodes suggest the nominal gap is more informative for business cycle analyses.

A comparison of the nominal and real expectations gap estimates with conventional output gap measures is complicated since they are derived from different sources. To provide a more "apples-to-apples" comparison of the expectations and output gaps, we construct the nominal and real expectations gaps from the Federal Reserve Board staff's Greenbook forecasts and compare these to their output gap forecasts.[16]

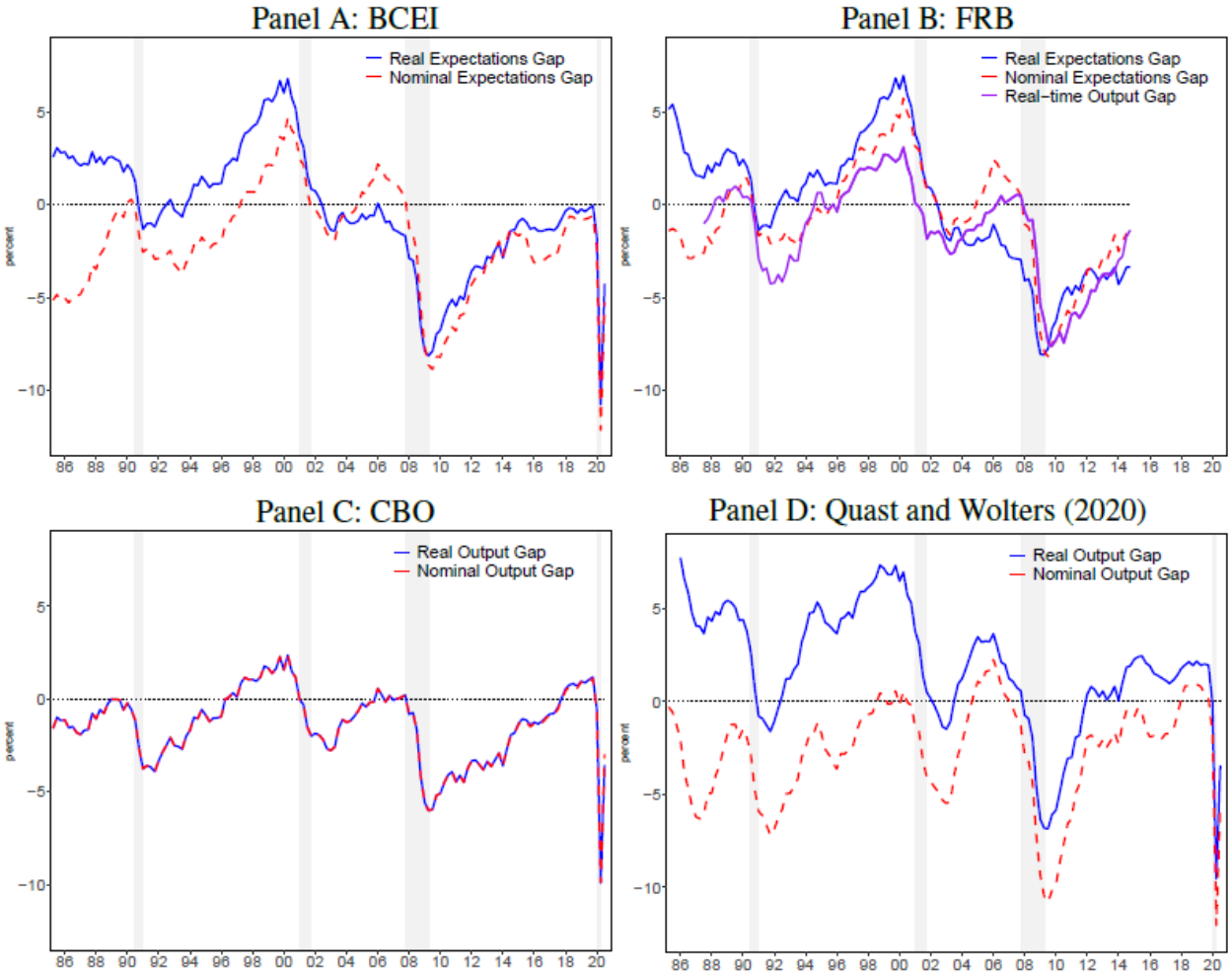
Panel B of Figure 4 plots all three measures up through 2015, which corresponds to the last publicly accessible Greenbook. Moreover, the Greenbook output gap was only released beginning in 1987. The Greenbook expectations gap estimates are similar to the BCEI expectation gaps. Notably, the nominal expectations gap is more closely correlated to the real-time output gap (0.92) than the real expectations gap (0.81). The Greenbook output gap shows a larger decline following the 1990-91 recession but suggests less overheating during the late 1990s and more quickly turns negative in 2000. It exhibits much less overheating in the mid-2000s than the nominal expectations gap but tracks both expectations gaps closely after 2008. Overall, the expectations gap for an individual forecaster can be thought of as a proxy for the implicit output gap used to construct the forecasts.

To explore how sensitive measures of the output gap are to being conveyed in nominal or real terms, we look at two alternative approaches for constructing the output gap. The first, shown in Panel C, is the latest measure by CBO that uses a production-function approach to estimate potential GDP and assumes the real and nominal measures of the output gap are identical. The measures do not indicate any overheating associated with the housing price boom in the mid-to-late 2000s nor do they capture the slowdown in 2015-16.

[15] Since mid-2010 the Greenbook has been replaced by the Tealbook.

[16] While the Greenbook does not explicitly contain long-term forecasts, we extrapolate long-term forecasts out five-years ahead by using the last quarterly growth rate forecast (typically between four- to eight-quarters ahead) as a proxy for the long-run growth rate.

Another statistical approach is the real-time measure by Quast and Wolters (2020) that employs a modified Hamilton filter to estimate potential GDP, which we replicate using both real and nominal GDP. The resulting output gap estimates are plotted in Panel D of Figure 4. Unlike the CBO, the nominal and real measures have large and persistent differences. While the real output gap remains positive with only occasional dips into negative territory around recessions, the nominal output is almost always negative except for at the peak of a cycle. Both versions suggest the economy was overheating during the mid-2000s but diverge in 2015-16 with the real output gap rising and the nominal output gap falling sharply.



Sources: Authors' calculations using data and forecasts from BEA, CBO, the Real-Time Data Set for Macroeconomists, FRB and BCEI

Figure 4: Real vs. Nominal Gaps

Our analysis suggests that the real and nominal gaps can convey different information. While the CBO assumes away these differences, the episodes in the mid-2000s and 2015-16 suggest that including nominal fluctuations provides a clearer picture of overheating and underperformance associated with episodes of broader price changes. Conversely, both the real expectations gap and conventional output gaps missed or downplayed these economic developments. Thus, the nominal expectations gap is more informative for analyses of business cycles.

However, there is a notable oddity in both the nominal and real expectations gap around the 2001 recession: these do not turn negative until the end of that recession. A positive expectations gap during a recession seems contradictory. It implies that actual GDP is still greater than expected, on average. A potential explanation is that more debts were accrued during the dotcom boom leading up to the 2001 recession.[17] In more distant forecasts prior to the dotcom boom are excluded then the gap turns negative more quickly. This interpretation is corroborated by Panel B of Figure 2, which shows that the expectations gap turns negative earlier near the 2001 recession when more distant forecasts are excluded from the estimation of neutral GDP and the expectations gap.

5 Real-Time and Forward-Looking Estimates

This section shows (1) how the expectations gap changes between first, second, third, and the latest estimates of GDP, (2) how a forward-looking expectations gap measure can be derived, and (3) how forward-looking gap estimates perform near business cycle turning points. The real-time exercise illustrates the measures that would have been observable at given moments of time. This section focuses exclusively on nominal measures. Therefore, any mention of the expectations gap, GDP, or neutral GDP refers to the nominal measures thereof.

5.1 Real-Time Estimates

As outlined in Section 2, forecasted GDP growth rates are combined with actual GDP levels to generate forecasts for the level of GDP. The average of these forecasts over a 5-year horizon yields a neutral GDP estimate, and the percentage difference between actual GDP and neutral yields the expectations gap. However, the BEA revises GDP data over time and this can lead to sizeable changes in the estimate expectations gap.

Estimating a real-time expectations gap involves taking the percentage difference between the real-time estimate of GDP and the real-time estimate of neutral GDP. The first aspect is simple as historical vintages of GDP estimates are readily available; see Croushore and Stark (2001). For example, for Q1-1997 the first, second, and third estimates of that quarter's GDP were released in April, May, and June 1997.

The second aspect, estimation of real-time neutral GDP, is more intensive. Each monthly vintage of GDP data must be combined with the 60 BCEI GDP growth rate forecasts that are needed to estimate neutral GDP in Q1-1997. For example, the first estimate of neutral GDP Q1-1997 uses the GDP series that became available in April 1997 when the first estimate of Q1-1997's GDP was released. The percentage differences between the first, second, and third estimates of GDP

[17] Initially, we hypothesized the comprehensive updates by the BEA over time may have biased the neutral GDP estimates downward, thus overstating how large the expectations gap was in 2001. (This would arise if actual GDP estimates were revised up from the addition of new, higher-growth spending categories while the real-time BCEI growth rate projections would be unchanged and would tend to underestimate growth.) However, as our real-time analysis in the next section shows, the expectations gap estimates derived from the latest GDP estimates are smaller than the real-time gap estimates.

are then taken against the first, second, and third estimates of neutral GDP that would have been available at each point in time.

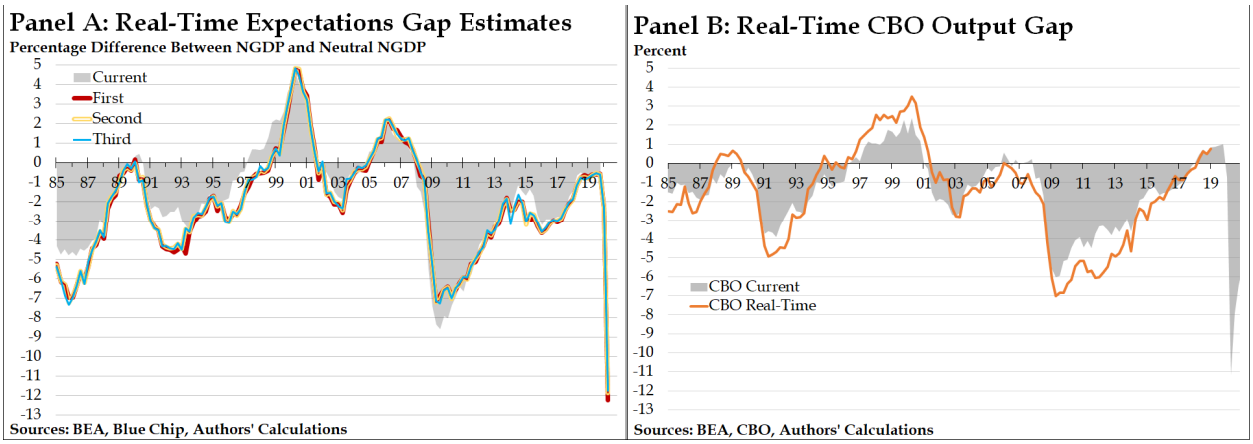


Figure 5: Real-time Estimates of the Expectations Gap

Panel A of Figure 5 shows the first-, second-, and third-estimates of the expectations gap that would have appeared when a given quarter's GDP numbers became available. For example, the Q1-1997 first, second, and third estimates (released in April, May, and June of 1997) indicated an expectations gap of about -2.0 percent. Since then, comprehensive updates made to the GDP estimates as of December 2020 indicate that the actual expectations gap (titled "Current" in Figure 5) for Q1-1997 GDP was only about -0.5 percent. Differences between the first, second, and third estimates of the expectations gap are very small, while comprehensive updates do lead to non-trivial revisions, especially in the late 1980s,1990s, and 2014-2018 when the latest expectations gap estimates tend to have a smaller absolute value than the real-time estimates. Exceptions occurred in 1997-99 and 2008-11 when the latest estimate was revised up or down respectively from its real-time counterpart.[18]

Panel B of Figure 5 computes a real-time CBO output gap series that we generated by combining as reported GDP data with CBO's estimates of potential GDP that were published near the charted quarters. For example, the real-time output gap around late 2008 to early 2009 is derived from BEA data from June 2009 and CBO's August 2009 Update to the Budget and Economic Outlook. The methodology used in Figure 5 combines the CBO's various outlook publications (generally two per year) with the vintage of BEA data that was used to make their forecasts. The results show that CBO's revisions to the output gap are just as large or even larger at times than those seen in the expectations gap.

[18] The BEA's comprehensive benchmark revision in 1999 to include spending categories on software, intellectual property products and research and development is likely behind many of the larger revisions in the late 1980s and 1990s as these categories were growing faster than average over this period so that their inclusion would imply a more positive (less negative) expectations gap.

5.2 Estimating a Forward-Looking Expectations Gap

Here we illustrate how a forward-looking expectations gap can be derived from incomplete survey data, using Q4-2021 as an example. As outlined in Section 2, calculating the expectations gap requires calculating the percentage difference between actual GDP and neutral GDP. The Q4-2021 neutral GDP estimate would be the average of forecasts for GDP in that quarter that were submitted between January 2016 and December 2021. If survey data is only available through September 2020 (Q3-2020), with 15 quarterly forecasts in hand and 5 to be submitted, there are two ways to derive a forward-looking neutral GDP estimate.

One approach is to average all forecasts that have been published to date:

$$\overline{Y}_{t+j|t,v} = \frac{1}{20-j}\sum_{h=0}^{19-j} Y_{t+j|t-h,v} \qquad (5)$$

which gives them all equal weight. An alternative approach gives the most recently published forecast additional weight by assuming that it will be repeated for the (5) forecasts that have yet to be submitted for the quarter of interest. Averaging the forecasts available-to-date and the last-published forecast repeated for the *j* missing quarters in between the present and the future quarter of interest yields:

$$\overline{Y}_{t+j|t,v} = \frac{1}{20}\sum_{h=0}^{19-j} \left(Y_{t+j|t-h,v} + jY_{t+j|t,v}\right) \qquad (6)$$

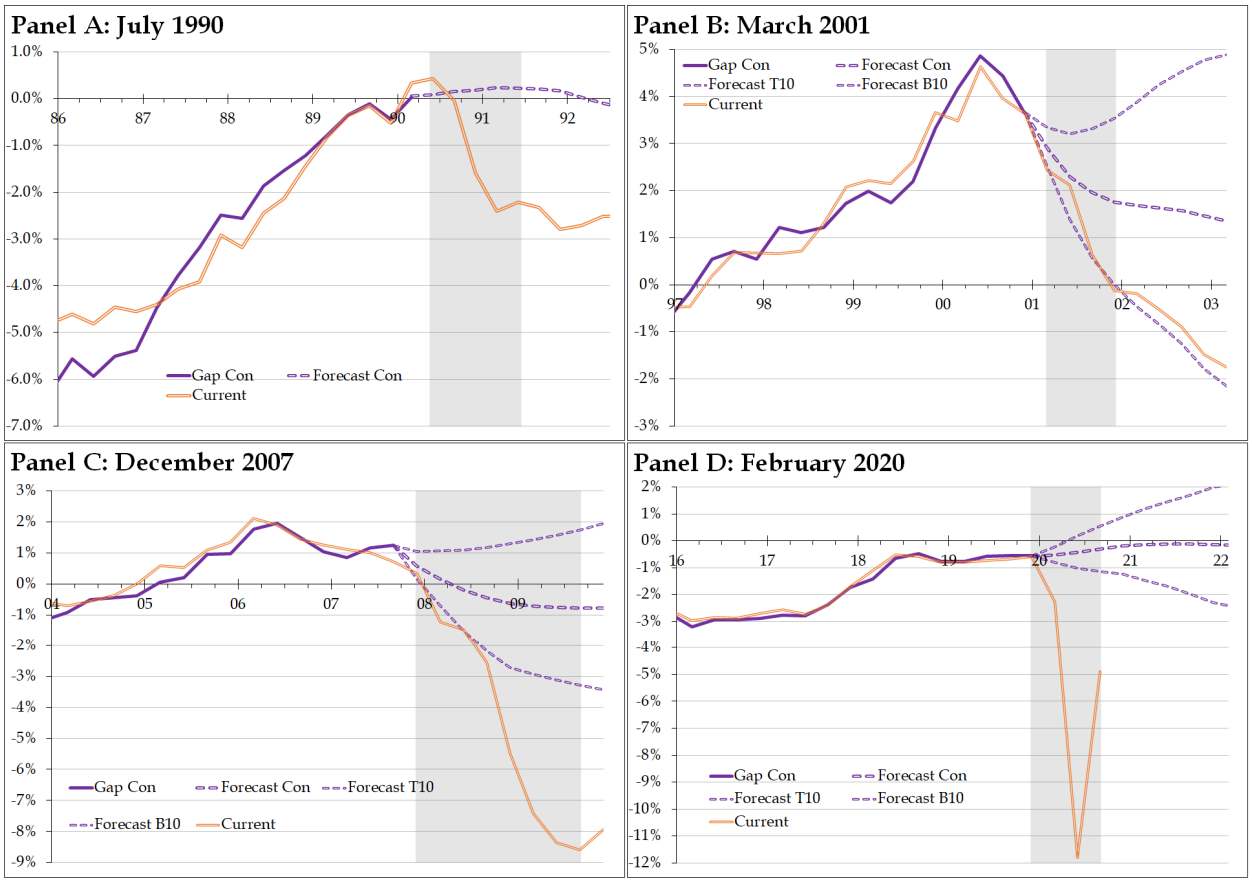
This latter approach places greater weight on the most recent forecast. It assumes that the future forecasts of GDP are more likely to be represented by what is the latest forecast than by the forecasts made in any period prior. We focus on this latter approach.

Taking the percentage differences between the latest forecast path of nominal GDP and a series of forward-looking neutral GDP estimates derived according to equation (5) or (6) yields a forward-looking nominal expectations gap series. The consensus forward-looking expectations gap series converges to zero at the five-year ahead horizon regardless of whether equation (5) or (6) is used.[19] Since this drives the movements in all forward-looking expectations gap estimates at longer-horizons, we focus on the forward-looking estimates up to two-years ahead.

Figure 6 illustrates how the BCEI-derived forward-looking expectations gap estimates would have appeared with the BCEI surveys released for July 1990 (panel A), March 2001 (panel B),

[19] For example, supposing the September 2020 BCEI survey was the latest, the consensus forecasts that Q3-2025 GDP will be $26,410.5 billion and is the only forecast available for Q3-2025. Assuming, as equation (6) does, that the forecast of $26,410.5 billion will be repeated 19 times for the surveys that will be published between October 2020 to September 2025, the calculation becomes 20*($26,410.5 billion)/20, an expectations gap of 0.0 percent. There is some underlying economic intuition for this assumption, in that sticky debt obligations that enable potential inefficiencies ex ante are expected to be resolved after 5 years.

December 2007 (panel C), and February 2020 (panel D) – the months a business cycle peak was eventually declared by the NBER for the last four recessions.



Panels show the precentage difference between nominal GDP and neutral nominal GDP
Sources: BEA, Blue Chip, Authors' Calculations

Figure 6: Forecasts of the Nominal Expectations Gap

The forward-looking expectations gap ranges are the percentage difference between the top 10 (bottom 10) nominal GDP projections and a forward-looking neutral nominal GDP estimate that is derived *only* from the consensus nominal GDP forecasts using equation (6). Thus, Figure 6, shows the most optimistic and most pessimistic forward-looking paths for the expectations gap. It ignores forecaster disagreement about neutral GDP estimates that was discussed in Section 3.3 and focuses only on the disagreement around the latest survey's forecast path for GDP.

We can evaluate the performance of the forward-looking expectations gap by computing the root mean square prediction errors (RMSE) across horizons. We compare the performance of the forward looking measures against a simple random walk that extends the last observed estimate of the expectations gap forward into the future and is a simplified version of the autoregressive prediction proposed in Beckworth (2020).

Table 1 shows that the real-time BCEI-based forward-looking measures consistently have smaller RMSEs than the random walk out through four-quarters ahead. This is unsurprising as

the forward-looking measures at short-horizons contain most of the same information as the actual estimates of the expectations gap do. The forecasts are particularly useful through 2-quarters ahead with RMSEs less than 1% but become less so at longer horizons where the 4-quarter-ahead forecasts have RMSEs of almost 1.5%. The performance at shorter horizons is less accurate when using the latest estimates of the expectations gap, which include large revisions in the underlying data.

Table 1. Expectations Gap Forecast Performance (1985 Q1 – 2020 Q2)

	1st		2nd		3rd		Latest	
Horizon	**BCEI**	**RW**	**BCEI**	**RW**	**BCEI**	**RW**	**BCEI**	**RW**
0	0.18	0.65	0.13	0.70	0.15	0.73	0.97	1.25
1	0.43	1.08	0.48	1.12	0.52	1.14	1.06	1.56
2	0.72	1.49	0.76	1.53	0.79	1.54	1.20	1.89
3	1.04	1.85	1.08	1.89	1.09	1.90	1.39	2.21
4	1.37	2.20	1.39	2.23	1.40	2.23	1.62	2.50

Focusing on periods that were later defined to be recessions by the NBER (Table 2), the forward-looking measure continues to outperform the random walk, though performance deteriorates substantially at longer horizons. Now, only nowcast and one-quarter-ahead forecasts have RMSEs that are less than 1%, while the 4-quarter-ahead forecasts have RMSEs of about 3.7%.

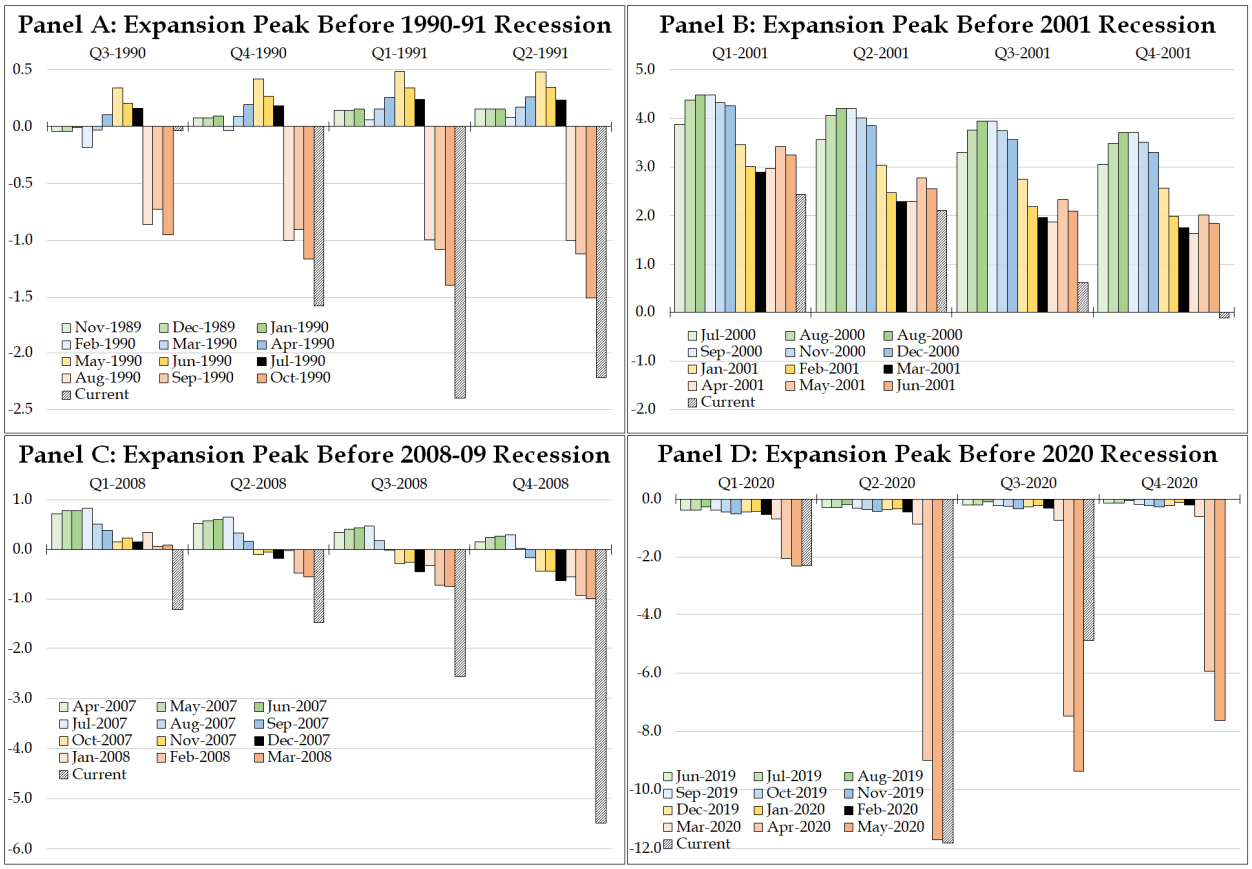
Table 2. Expectations Gap Forecast Performance During NBER Recessions

	1st		2nd		3rd		Latest	
Horizon	**BCEI**	**RW**	**BCEI**	**RW**	**BCEI**	**RW**	**BCEI**	**RW**
0	0.19	2.79	0.14	2.74	0.15	2.70	1.09	3.17
1	0.57	3.65	0.63	3.59	0.62	3.56	1.29	4.06
2	2.63	4.06	2.58	4.01	2.55	3.98	2.89	4.51
3	3.45	4.26	3.39	4.22	3.35	4.19	3.73	4.72
4	3.78	4.42	3.73	4.37	3.69	4.33	4.11	4.90

5.3 Forward-Looking Estimates Near Business Cycle Turning Points

Figures 7 and 8 show the evolution of the real-time estimates of the forward-looking expectations gap over four quarters near business cycle turning points. Within a particular quarter, the bars represent a consecutive monthly sequence of forward-looking expectations gap estimates for that quarter. The darkest-colored bar is the forward-looking expectations gap as it would have appeared in the month that NBER eventually declared to be the peak of the expansion. The bars that precede it are the forward-looking estimates from the BCEI surveys released the prior eight months. The three bars that follow it are from the three subsequent months' BCEI surveys. The right-most bar for each quarter represents the expectations gap as it is currently estimated.

For example, in Panel A of Figure 7, shows the quarters around the 1990-91 recession. July 1990 was the month that was ultimately declared to be the peak by the NBER. The sequence of forward-looking expectations gap for Q3-1990 is slightly negative for the BCEI surveys released between November 1989 and March 1990. After becoming positive between the April 1990 and July 1990 surveys, it turns sharply negative for the surveys submitted between August 1990 and October 1990.



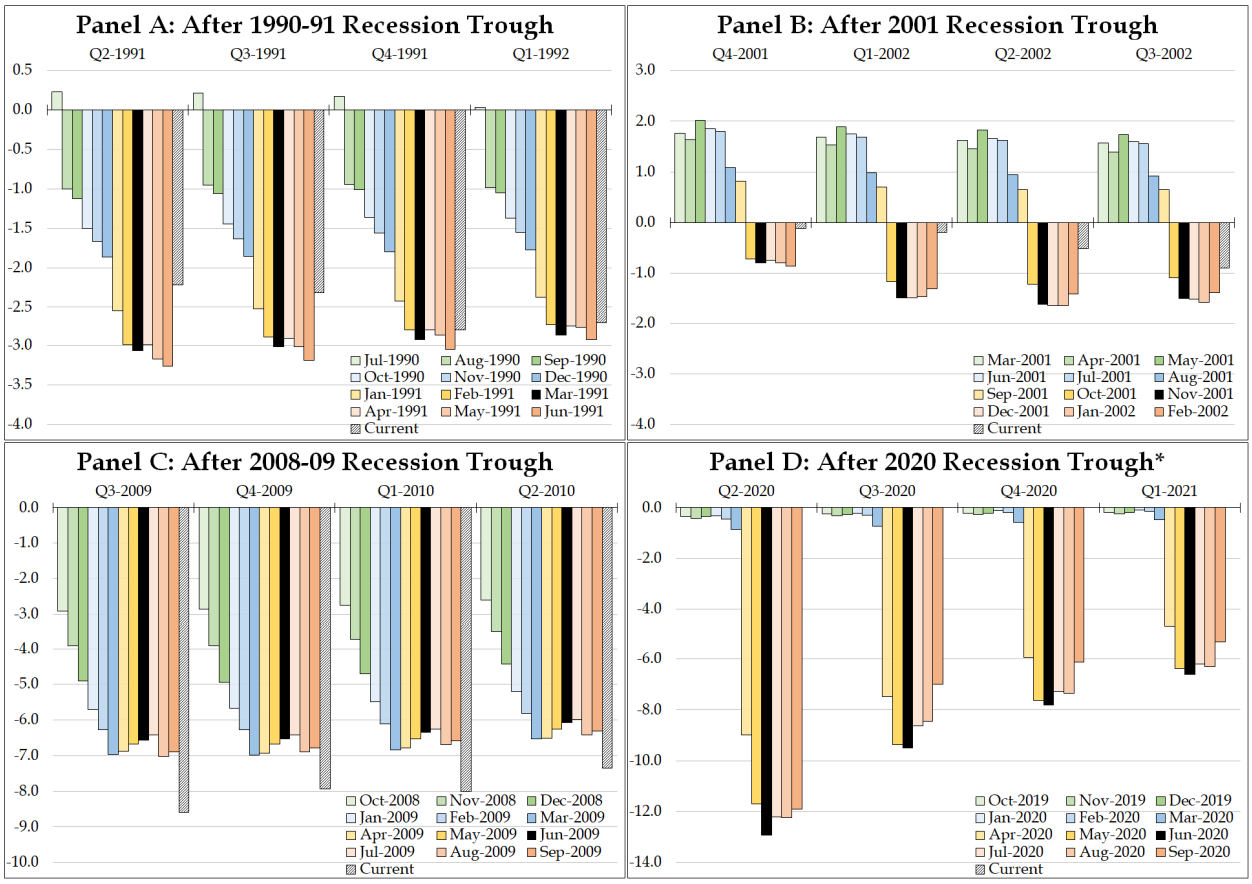
Sources: Bureau of Economic Analysis, Blue Chip, Authors' Calculations

Figure 7: Real-time Forward-Looking Expectations Gap Estimates Near Expansion Peaks

Before the 1990-91, 2001, and 2008-09 recessions the forward-looking gap estimates attain a high two- to five-months ahead of an expansions' peak month. This does not apply to the 2020 recession, which underscores the sudden and unexpected nature of the COVID-19 pandemic. The 2001 recession is peculiar in that the expectations gap remains positive. Thus, a negative expectations gap is not a sufficient condition to signify a recession. Instead, a recession may be portended by consistent downward revisions to a given quarter's forward-looking expectations gap. This is discussed further in Section 6.

Figure 8 shows the evolution of the real-time estimates of the forward-looking expectations gap near the trough of recessions. In these panels, the darkest bars represent the expectations gap estimates from the BCEI-survey released during the month the NBER eventually declared to be the trough of a recession (March 1991, November 2001, June 2009; for the 2020 recession, we assume that the NBER will declare a trough to have occurred in June 2020).

A pattern of forward-looking expectations gap estimates near a recession trough is even less discernable relative to those near an expansion peak. In Panel A, around the 1990-91 recession, the forward-looking expectations gap for the four quarters charted cease being revised down by March 1991, which corresponds to the recession trough.



Sources: Bureau of Economic Analysis, Blue Chip, Authors' Calculations
***Assuming a recession trough will be declared to have occurred in June 2020**

Figure 8: Real-time Forward-Looking Expectations Gap Estimates Near Troughs

In Panel B, the September 11, 2001 terrorist attacks have a sharp impact on forward-looking gap estimates. While the September 2001 survey (usually released the tenth day of the given month, i.e., in this case just before the attacks took place) signifies a positive forward-looking expectations gap, the subsequent measure based on the October survey drops sharply into negative territory. However, downward revisions to the forward-looking expectations gap ceased with the November survey, which is the month the NBER eventually declared to be a trough.

In Panel C for the 2008-09 recession, the forward-looking expectations gap improved somewhat following the March 2009 survey until the June survey, which coincides with the month the NBER declared to be the trough.

For the 2020 recession (Panel D), the forward-looking expectations gap estimates reached the lowest point in June. Since the forward-looking gap did not improve before and up to the June

2020 survey, it failed to serve as a leading indicator of the trough, assuming the NBER eventually declares June as the trough.

Thus, for the 1990-91, 2001, and 2020 recessions, the real-time forward-looking expectations gap did not predict an eventual trough. However, in the 2008-09 recession, a slight improvement in the forward-looking expectations gap did appear two months before the eventual trough.

Overall, persistent month-to-month differences between real-time forward-looking expectations gap estimates seems to predict business cycle turning points, but not always. This is because recessions and recoveries can often occur abruptly for reasons that were unforeseen by professional forecasters at the time. Therefore, a negative forward-looking expectations gap may be sufficient but is not necessary to signify a recession as illustrated by the 2001 recession.

6 Conclusion

In this paper we assess the expectations gap proposed in Beckworth (2020). The latter measure of economic slack is derived from professional forecasters' historical expectations about the trajectory of the economy in contrast to conventional output gap estimates that are derived from model-specific measures of potential GDP. To overcome some of the limitations of Beckworth (2020) we develop and evaluate several alternative measures of the expectations gap using alternative surveys, different forecast horizons, individual forecasters' estimates, nominal and real measures, and real-time and forward-looking estimates.

Although the expectations gap is not particularly sensitive to the choice of forecaster survey it is sensitive to the range of forecast horizons over which neutral GDP is calculated. Moreover, there is considerable disagreement between individual forecasters about the future trajectory of the economy. Despite this, we identify four distinct episodes in which the entire range of individual forecasters' expectations gaps provides an unambiguous signal about the state of the economy.

There are meaningful and informative differences between the nominal and real expectations gaps, which contrasts with conventional output gap measures. This is exemplified by the nominal expectations gap detecting overheating during the run-up in housing prices during the mid-2000s that preceded the 2008-09 recession. Also, in 2015-16 the nominal expectations gap detects the so-called invisible recession while the real expectations gap and conventional output gap measures do not. Furthermore, although there have been considerable historical revisions to the expectations gap due to comprehensive updates in GDP, these revisions are in line with those experienced by conventional measures.

Forward-looking measures of the expectations gap is generally informative up to two-quarters ahead and are always more informative than assuming there will be no change. Although forward-looking measures of the expectations gap do not always signal an oncoming expansion peak or recession trough, the direction of updates to the forward-looking expectations gap occasionally precedes a business cycle turning point.

Overall, the expectations gap is a useful measure of economic slack. Consequently, it may be useful to shift the emphasis from "navigating by the stars" to navigating by what agents *expect*.

References

Ball, L. and Mankiw, N. G. (2002). The NAIRU in Theory and Practice. *Journal of Economic Perspectives*, 16(4):115–136.

Beckworth, D. (2020). The Stance of Monetary Policy: The NGDP Gap. Mercatus Policy Brief, The Mercatus Center at George Mason University, Arlington, VA.

Beaudry, P. and Galizia, D. and Portier, F. (2020). Putting the Cycle Bank into Business Cycle Analysis. *American Economic Review,* 110(1): 1-47.

Berge, T. J. (2020). Time-varying Uncertainty of the Federal Reserve's Output Gap Estimate. Finance and Economics Discussion Series 2020-012, Board of Governors of the Federal Reserve System, Washington D.C.

Coibion, O. and Gorodnichenko, Y. and Ulate, M. (2018). The Cyclical Sensitivity in Estimates of Potential Output. Brookings Papers on Economic Activity, Fall 2018. Washington D.C.

Croushore, D. and Stark, T. (2001). A Real-time Data Set for Macroeconomists. *Journal of Econometrics*, 105(1):111–130.

Bullard, J. and DiCecio, R. (2019). Optimal Monetary Policy for the Masses. Federal Reserve Bank of St. Louis, Working Paper 2019-009C.

Bürgi, Constantin and Sinclair, Tara M. (2020). What does forecaster disagreement tell us about the state of the economy? *Applied Economics Letters.*

Fleischman, C. A. and 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 2011-046, Board of Governors of the Federal Reserve System, Washington D.C.

Isiklar, G. and Lahiri, K. (2007). How Far Ahead Can We Forecast? Evidence from cross-country surveys. *International Journal of Forecasting*, 23(2): 167-187.

Kumhof, M. and Laxton, D. (2013). Simple Fiscal Policy Rules for Small Open Economies. *Journal of International Economics*, 91(1):113–127.

Lahiri, K. and Sheng, X. (2020). Measuring Forecast Uncertainty by Disagreement: The Missing Link. *Journal of Applied Econometrics*, 25(4): 514-538.

Laubach, T. and Williams, J. C. (2003). Measuring the Natural Rate of Interest. *Review of Economics and Statistics*, 85(4):1063–1070.

Orphanides, A. and Norden, S. v. (2002). The Unreliability of Output-gap Estimates in Real Time. *Review of Economics and Statistics*, 84(4):569–583.

Patton, A. J. and Timmermann, A. (2010). Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion. *Journal of Monetary Economics*, 57(7):803–820.

Powell, J. H. (2018). Monetary Policy in a Changing Economy. Remarks by Chair Jerome H. Powell at the Changing Market Structure and Implications for Monetary Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming. Board of Governors of the Federal Reserve System, Washington, DC.

Powell, J. H. (2020). New Economic Challenges and the Fed's Monetary Policy Review. Remarks by Chair Jerome H. Powell at the Changing Market Structure and Implications for Monetary Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming. Board of Governors of the Federal Reserve System, Washington, DC.

Quast, J. and Wolters, M. (2020). Reliable Real-Time Output Gap Estimates Based on a Modified Hamilton Filter. *Journal of Business & Economic Statistics*, forthcoming.

Reifenschneider, D. and Tulip, P. (2019). Gauging the Uncertainty of the Economic Outlook Using Historical Forecasting Errors: The Federal Reserve's Approach. *International Journal of Forecasting*, 35(4): 1564-1582.

Romer, C. (1990). The Great Crash and the Onset of the Great Depression. *The Quarterly Journal of Economics,* 105(3): 597-624.

Shackleton, R. (2018). Estimating and Projecting Potential Output Using CBO's Forecasting Growth Model. Congressional Budget Office.

Sheedy, K. D. (2014). Debt and Incomplete Financial Markets: A Case for Nominal GDP Targeting. *Brookings Papers on Economic Activity*, Spring, 301-361.

Sinclair, T. M. and Stekler, H. O. (2013). Examining the quality of early GDP component estimates. *International Journal of Forecasting*, 29(4):736–750.

Taylor, J. B. (1993). Discretion versus Policy Rules in Practice. In *Carnegie-Rochester conference series on public policy*, volume 39, pages 195–214. Elsevier.

Taylor, J. B. (2000). Reassessing Discretionary Fiscal Policy. *Journal of Economic Perspectives*, 14(3):21–36.

Williams, P. and Abdih, Y. and Kopp, E. (2020). Reading the Stars. IMF Working Paper WP/20/136. International Monetary Fund. Washington, D.C.