Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Cyclical Signals from the Labor Market

CAMA Working Paper 91/2021
October 2021

Tino Berger
University of Göttingen
Centre for Applied Macroeconomic Analysis, ANU

Paul David Boll
University of Warwick

James Morley
University of Sydney
Centre for Applied Macroeconomic Analysis, ANU

Benjamin Wong
Monash University
Centre for Applied Macroeconomic Analysis, ANU

Abstract

We consider which labor market variables are the most informative for estimating and now-casting the U.S. output gap using a multivariate trend-cycle decomposition. Although the unemployment rate clearly contains important cyclical information, it also appears to reflect more persistent movements related to labor force participation that could distort inferences about the output gap. Instead, we show that the alternative U-2 unemployment rate (job losers as a percentage of the labor force) provides a more purely cyclical indicator of labor market conditions. To a lesser extent, but consistent with a link of the output gap to real labor costs in a New Keynesian setting, we also find that average hourly earnings are informative about the output gap.

Keywords

Nowcasting, output gap, Covid-19, U-2 unemployment rate, average hourly earnings

JEL Classification

C53, E24, E32

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

Cyclical Signals from the Labor Market*†

Tino Berger[1,5], Paul David Boll[2], James Morley[3,5], and Benjamin Wong[4,5]

[1]University of Göttingen

[2]University of Warwick

[3]University of Sydney

[4]Monash University, Australia

[5]CAMA

October 12, 2021

Abstract

We consider which labor market variables are the most informative for estimating and nowcasting the U.S. output gap using a multivariate trend-cycle decomposition. Although the unemployment rate clearly contains important cyclical information, it also appears to reflect more persistent movements related to labor force participation that could distort inferences about the output gap. Instead, we show that the alternative U-2 unemployment rate (job losers as a percentage of the labor force) provides a more purely cyclical indicator of labor market conditions. To a lesser extent, but consistent with a link of the output gap to real labor costs in a New Keynesian setting, we also find that average hourly earnings are informative about the output gap.

Keywords: Nowcasting; output gap; Covid-19; U-2 unemployment rate; average hourly earnings

JEL Classification: C53; E24; E32

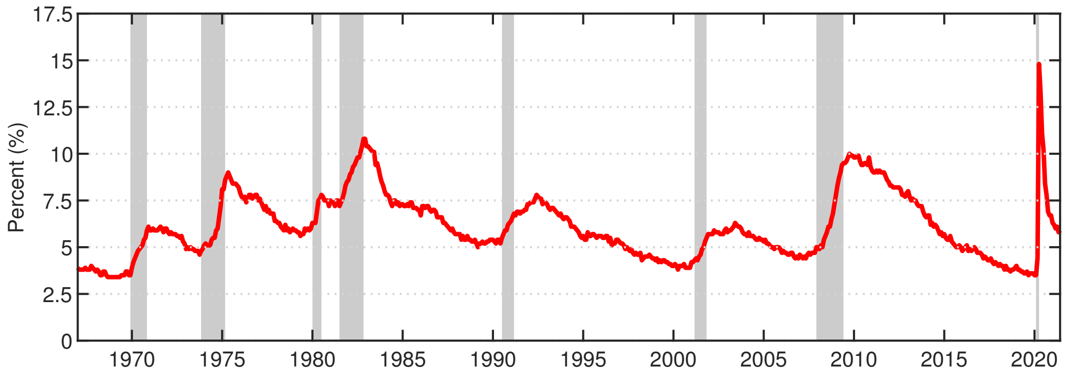
*Contact: tino.berger@wiwi.uni-goettingen.de, david.boll@warwick.ac.uk, james.morley@sydney.edu.au , benjamin.wong@monash.edu

†This research was supported by ARC DP190100202 and ARC DE200100693. The usual disclaimers apply.

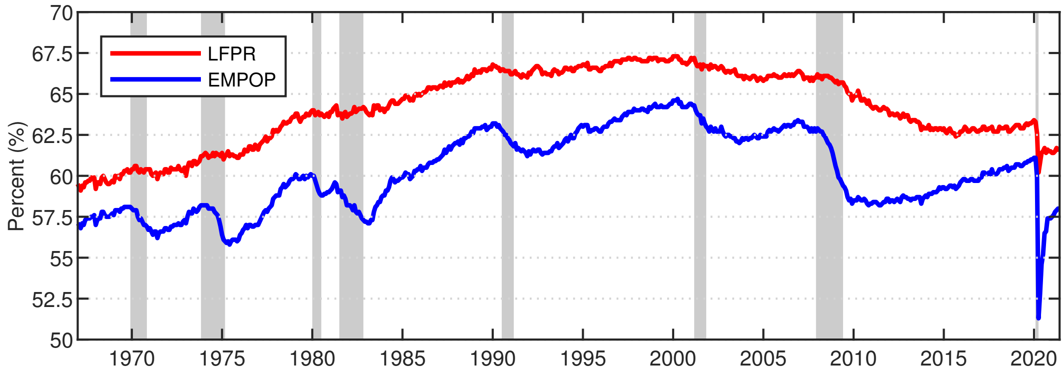
1 Introduction

Applications of multivariate methods of trend-cycle decomposition often find the unemployment rate to be one of the most important variables in informing estimates of the output gap, even when many other variables are considered (e.g., Fleischman and Roberts 2011, Morley and Wong 2020, Barigozzi and Luciani 2022). However, since the Great Recession and particularly with the Covid-19 pandemic, there has been some concern among policymakers about the reliability of the unemployment rate in capturing labor market conditions.

Figure 1: U.S. labor market indicators and recessions



(a) Unemployment rate



(b) Labor force participation rate and employment-population ratio

Notes: Sample period is January 1967 to June 2021. Source is BLS and data were obtained from FRED. LFPR stands for labor force participation rate and EMPOP stands for employment-to-population ratio. Shaded bars correspond to NBER recessions.

The onset of the Covid-19 pandemic and the associated restrictions on economic activity starting in March 2020 led to a sharp rise in the U.S. unemployment rate from 3.5% in February to 14.8% in April, the largest increase in the postwar period. However, as the top panel in Figure 1 shows, the spike in April was immediately followed by rapid reductions in the unemployment rate, falling back to under 7% by the end of the year. This swift recovery was in stark contrast to previous U.S. recessions, including the Great Recession in 2007–2009, which featured much more persistent increases in unemployment and a very slow recovery afterwards. While this different behavior of the unemployment rate is likely due to the unusual nature of the pandemic recession that led to temporary business closures followed by relatively quick re-openings once lockdowns were lifted (Cajner et al. 2020; Powell 2021), there is also a possibility that the unemployment rate understated the impact of the pandemic on the labor market. This view was expressed by Jerome Powell, Chair of the U.S. Federal Reserve Board of Governors, in February of 2021:

> "After rising to 14.8 percent in April of last year, the published unemployment rate has fallen relatively swiftly, reaching 6.3 percent in January. But published unemployment rates during Covid have dramatically understated the deterioration in the labor market. Most importantly, the pandemic has led to the largest 12-month decline in labor force participation since at least 1948. [...] In addition, the Bureau of Labor Statistics reports that many unemployed individuals have been misclassified as employed. Correcting this misclassification and counting those who have left the labor force since last February as unemployed would boost the unemployment rate to close to 10 percent in January." (Powell 2021)

The red line in the bottom panel of Figure 1 shows the evolution of the U.S. labor force participation rate, highlighting the main concern expressed in the quote above. The onset of the pandemic coincided with a rapid decline in labor force participation, which only partially rebounded and stabilized around 2 percentage points below pre-pandemic levels. This, too, was very different to previous recessions, which had much smaller, if any, immediate impact on the labor force participation rate. To the extent that this decline captures "discouraged workers", meaning workers who would prefer to be working but have given up searching for work, it represents a loss in employment due to the pandemic (Coibion et al. 2020). Because the unemployment rate only captures those who are not employed and actively searching for work (U.S. Bureau of Labor Statistics 2015), it could understate the true extent of labor market damage. This possibility is also apparent in the blue line in Figure 1b, which shows a decline in the U.S. employment-to-population ratio by 3.7 percentage points between February 2020 and January 2021. Attributing this loss of employment to unemployment rather than to a decline in labor force participation would lead to an implied unemployment rate of over 9% in January 2021.

Related concerns about the unemployment rate as an indicator of labor underutilization were already raised following the Great Recession (e.g., Yellen 2014). In addition to a distortion through cyclical movements in labor force participation, as documented, for example, by van Zandweghe (2012), Erceg and Levin (2014), Fujita (2014), the unemployment rate does not account for the intensive margin of labor supply. This ignores reductions in hours worked which typically occur during recessions, again understating the true extent of labor underutilization. In the United States, both the Great Recession and the Covid-19 pandemic led the number of workers working "part-time for economic reasons" to more than double. Berger and Vierke (2017) and Faberman et al. (2020), using measures which account for both the extensive margin (labor force participation) and intensive margin of labor supply, find that U.S. labor underutilization was significantly higher after the Great Recession than the unemployment rate suggested. Faberman et al. (2020), in particular, argue that the unemployment rate has increasingly understated labor market slack since the Great Recession.

Given the importance of the unemployment rate in estimating the output gap via multivariate trend-cycle decompositions, these concerns regarding its ability to sufficiently capture labor market conditions suggest that output gap estimates could be improved by considering alternative labor market variables. If the unemployment rate has mixed informational value with respect to labor utilization, especially since the Great Recession, estimates of output gaps that rely heavily on the unemployment rate may not accurately reflect the business cycle. Notably, output gap nowcasts based on the model in Berger et al. (2021) place a lot of weight on information from the unemployment rate and updated estimates for the model imply a positive U.S. output gap in the first quarter of 2021, which is in contrast to other measures of the output gap, such as the production-function-based Congressional Budget Office (CBO) estimate.

In this paper, we extend the mixed-frequency Bayesian VAR model used in Berger et al. (2021) (the 'BMW model' hereafter) to consider a number of alternative labor market variables and apply the multivariate Beveridge-Nelson (BN) decomposition with the variable selection procedure proposed in Morley and Wong (2020). Our aim is to determine which labor market variables are most informative about the output gap and how they affect output gap estimates, particularly in recent times. A key aspect of the Morley and Wong (2020) implementation of the multivariate BN decomposition is that the calculation of trend and cycle requires the variables in the forecasting model to be stationary. The unemployment rate generally tests as being stationary, but it also appears to have some very persistent movements beyond business cycle horizons. These can lead to persistent movements in the estimated output gap that do not reflect the business cycle. By contrast, our preferred alternative labor market variable of the U-2 (job losers as a percentage of the labor force) unemployment rate is clearly stationary and avoids implying movements in the output gap that persist beyond business cycle horizons. This result makes sense as the U-2 unemployment rate should not reflect changes in the long-term rate of unemployment due to changes in labor force participation, but rather will capture changes in the unemployment rate for cyclical reasons such as the onset of a recession. We also find that, to a lesser extent, average hourly earnings are informative about the output gap, consistent with the link of the output gap to real labor costs in a New Keynesian setting (Galí and Gertler 1999). Interestingly, other labor market variables, such as the labor force participation rate, the employment-to-population ratio, and measures of work hours, do not appear to be informative about the output gap once accounting for the U-2 unemployment rate and average hourly earnings.

The rest of this paper is organized as follows: Section 2 describes the data and briefly discusses estimation of the mixed-frequency Bayesian VAR used for multivariate trend-cycle decomposition. Section 3 reports our empirical results for an application to U.S. data, including a number of robustness checks and consideration of implications for the output gap since the onset of the Covid-19 pandemic. Section 4 concludes.

2 Data and methods

To estimate the output gap, we consider U.S. quarterly log real GDP as the target variable for multivariate trend-cycle decomposition. Our multivariate information set includes the following monthly indicators shown by Berger et al. (2021) to be useful for within-quarter nowcasting of the output gap: the federal funds rate in first differences, the 10-year-minus-1-year term spread for Treasuries, the BAA-minus-AAA corporate bond credit risk spread, S&P 500 stock market returns, the University of Michigan consumer sentiment index, the unemployment rate, the CPI inflation rate, industrial production (IP) growth, and growth in housing starts.

Morley and Wong (2020) show that the unemployment rate is a particularly important informational variable for multivariate trend-cycle decomposition of U.S. real GDP. Estimates of the output gap are generally robust to the inclusion of 8, 23, and 138 variables in the VARs used in Morley and Wong (2020) to conduct the BN decomposition when the unemployment rate is included in the set of variables, but they are highly sensitive to the removal of the unemployment rate from even the 23-variable model (see Figure 4 in Morley and Wong 2020). Given this importance of the unemployment rate in estimating the output gap, we substitute a set of related labor market variables for the unemployment rate in the BMW model in order to better understand the underlying source of cyclical signals from the unemployment rate and the labor market more generally. Our choice of additional monthly indicators is motivated by the labor market variables used in the 138-variable VAR from Morley and Wong (2020) and also by the variables included in the Kansas City Fed's Labor Market Conditions Indicators (Hakkio and Willis 2014). This includes the labor force participation rate, the employment-to-population ratio, average work hours, average hourly earnings, and two of the six "alternative measures of labor underutilization" from the BLS: U-1 (15 weeks and over) unemployment rate and U-2 (job losers) unemployment rate.[1] The "prime-age" (25–54 years) labor force participation rate and employment-to-population ratio measures are also included, as they should be less susceptible to distortions from demographic factors (Powell 2021). Additionally, the "employment rate in hours" is constructed according to Berger and Vierke (2017). Following Berger et al. (2021), total nonfarm payroll employment growth and initial claims for unemployment insurance are not included because they both exhibit extreme values in March and April 2020 that are out of proportion to the broader economic developments and likely due to legislative changes that altered the measurement of the variables, at least temporarily. Other measures related to the unemployment rate, such as more disaggregated versions by age or length of unemployment, are excluded, since they are most likely subject to the

[1] U-3 corresponds to the overall unemployment rate, while the other three measures, U-4, U-5, and U-6, are only available from 1994.

same potential distortions as the overall unemployment rate.

The data series have a variety of sources including the BEA and BLS and were obtained from FRED for the sample period of January 1967 to June 2021 (1967Q1-2021Q2 for real GDP).[2] Following Berger et al. (2021), multivariate trend-cycle decomposition is based on the BN decomposition. Because the BN decomposition calculation presented in the next section requires inversion of the companion matrix for the VAR, all of the labor market variables need to be tested for nonstationarity and suitably transformed prior to inclusion in the model. First, natural logarithms are taken for all variables describing levels rather than percentages or rates. Second, following Morley and Wong (2020), the data are differenced if either a Chow test rejects a change in mean between the two halves of the sample or an ADF test cannot reject a unit root.[3] For comparability with the Berger et al. (2021) results, the variables adopted from their specification are transformed exactly like they are in Berger et al. (2021). This includes the unemployment rate, which tests as stationary in levels, but the possible nonstationarity of which we will discuss in the next section. Full details of the data, including transformations, are provided in the appendix.

Most other matters of model specification, estimation, and implementation of trend-cycle decomposition and nowcasting are as in Berger et al. (2021) and the reader is referred to that paper for full details. The VAR lag order is set to $p = 4$ in quarterly terms (i.e., 12 lags in monthly terms) and the estimated output gap is converted from log differences to percentage deviations. Parameter estimation is based on the pre-Covid sample period of 1967 to 2019 only, although inferences about the output gap are reasonably robust to updating parameter estimates to the full sample period even though there are some outlier observations during the Covid-19 pandemic. In terms of the timing of nowcasts, we note that all of the alternative labor market variables are released at or around the same time as the unemployment rate.

[2] Some additional values of monthly indicators for July, August, and September 2021 were obtained for the nowcasting exercise in Section 3.4.

[3] Tests are performed on quarterly versions of the variables and significance is determined at the 5% level.

3 Empirical results

3.1 Informational decompositions

In order to determine the relevant labor market variables, a mixed-frequency Bayesian VAR is first estimated with the non labor market variables from the BMW model and the full set of alternative labor market variables discussed in the previous section. Specifically, following Morley and Wong (2020), we cast an n-variable VAR(p) model into companion form:

$$\boldsymbol{X}_t = \boldsymbol{F}\boldsymbol{X}_{t-1} + \boldsymbol{H}\boldsymbol{e}_t, \tag{1}$$

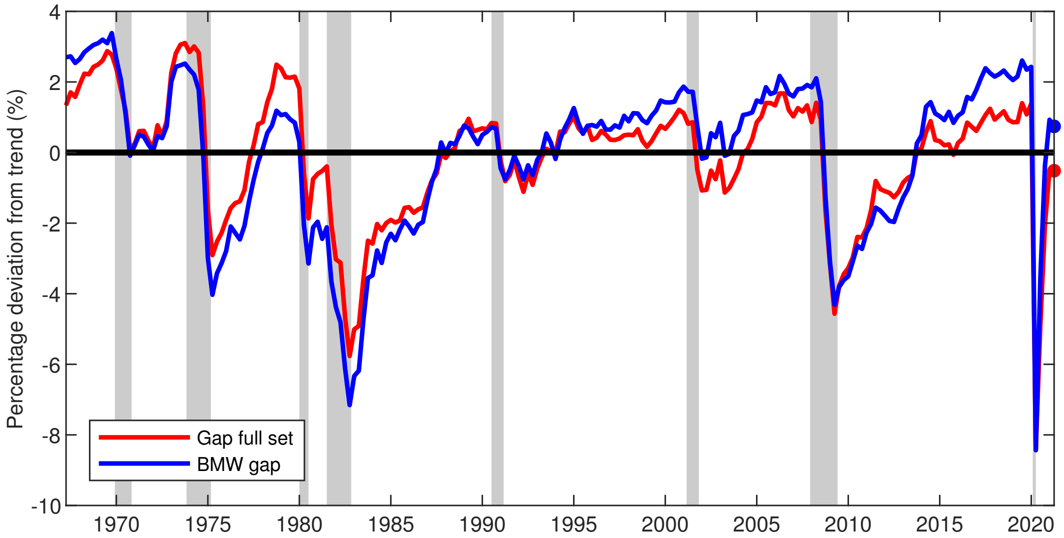
where $\boldsymbol{X}_t$ is a vector of stationary and demeaned variables, $\boldsymbol{F}$ is the companion matrix, $\boldsymbol{H}$ is a matrix that maps the forecast errors to the companion form, and $\boldsymbol{e}_t$ is a vector of forecast errors. Let $\mathbf{s_{np,r}}$ be a $np \times 1$ selector vector which consists of 1 as its r^{th} row and zeros otherwise. If real GDP growth Δy_t is included as the n^{th} element of the vector $\boldsymbol{X}_t$ in equation (1), the BN cycle of y_t can be calculated following Morley (2002) as

$$c_t = -\mathbf{s}'_{\mathbf{np,n}}\boldsymbol{F}(\boldsymbol{I} - \boldsymbol{F})^{-\mathbf{1}}\boldsymbol{X}_{\mathbf{t}}. \tag{2}$$

The red line in Figure 2 plots the estimated output gap when considering the full set of labor market variables. For comparison, the blue line plots the estimated output gap when considering the unemployment rate instead, as in the BMW model. The two estimates are reasonably similar for most of the sample period and both align with NBER recession dates, implying the information captured by the alternative labor market variables is qualitatively similar to that captured by the unemployment rate.

To apply the Morley and Wong (2020) informational decomposition and variable selection to a mixed-frequency model with monthly indicators, it is necessary to sum the contributions of forecast errors for the first, second, and third months in a quarter of each monthly indicator in

Figure 2: Estimated output gap with and without full set of labor market variables



Notes: Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER recessions.

the mixed-frequency VAR. Thus, the contribution of the k^{th} monthly indicator is

$$c_{k,t} = -\sum_{i=1}^{3}\sum_{j=0}^{t-1} \mathbf{s}'_{np,l}\boldsymbol{F}^{j+1}(\mathbf{I}-\boldsymbol{F})^{-1}\boldsymbol{H}\mathbf{s}_{n,k_i}\mathbf{s}'_{n,k_i}\boldsymbol{e}_t, \tag{3}$$

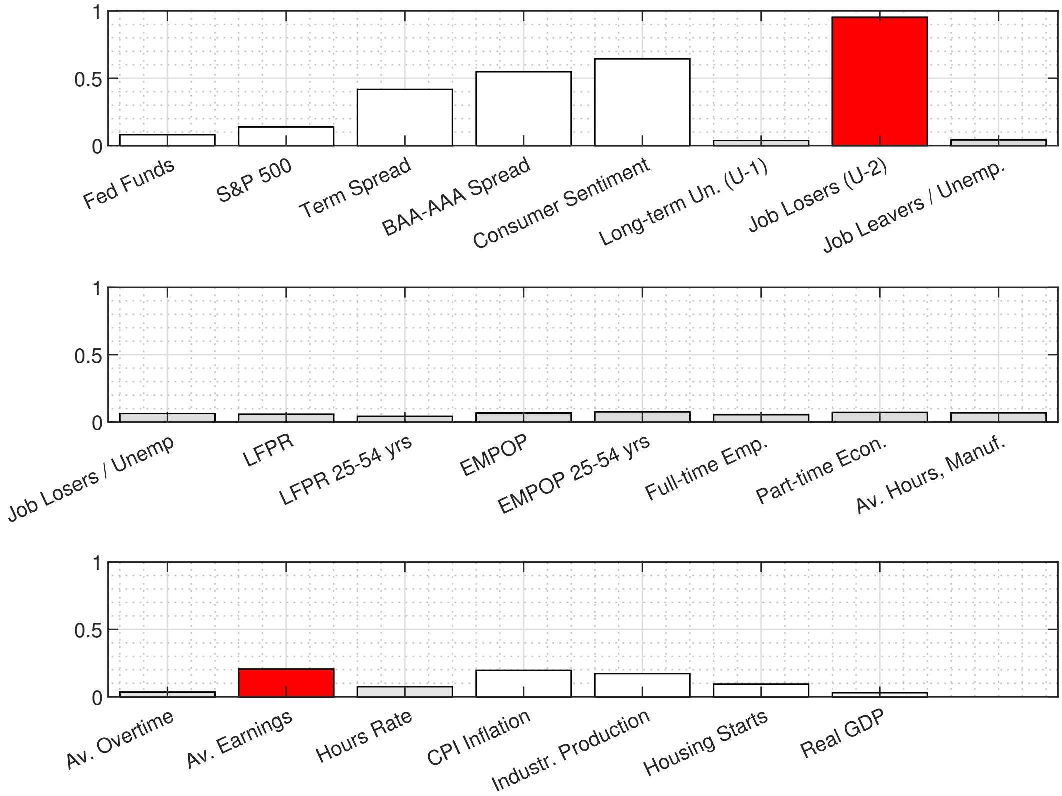
where $k_i = k + (i-1)m$, with m equal to the number of monthly indicators included in the model and $n = 3m + 1$ is equal to the total number of variables in the mixed-frequency VAR such that output growth is the n^{th} variable.

Following Morley and Wong (2020), the standard deviation of $c_{k,t}$ is used to quantify the k^{th} variable's informational contribution. The results for the full set of variables can be found in Figure 3. As can be seen, the contributions are relatively small for all of the labor market variables except for the U-2 unemployment rate and, to a lesser extent, average hourly earnings.[4] Because these two variables are the only ones whose contributions are larger than the smallest contribution among the Berger et al. (2021) monthly indicators (i.e., the federal funds rate), we include them but not the other labor market variables in our benchmark model for the remainder of our empirical analysis.[5]

[4]Average hourly earnings is included in the VAR in second differences, but for brevity is simply referred to as 'average hourly earnings'.

[5]It is possible that a high degree of multicollinearity between the labor market variables could result in informationally-relevant variables not to being included when eliminating variables simply based on the informational decomposition for the model with the full set of variables. Therefore, following Morley and Wong (2020), we repeat the variable selection by sequentially dropping the variable with the smallest contribution and re-estimating the model. However, we find identical results, with the U-2 unemployment rate and average hourly earnings re-

Figure 3: Informational decomposition with full set of labor market variables



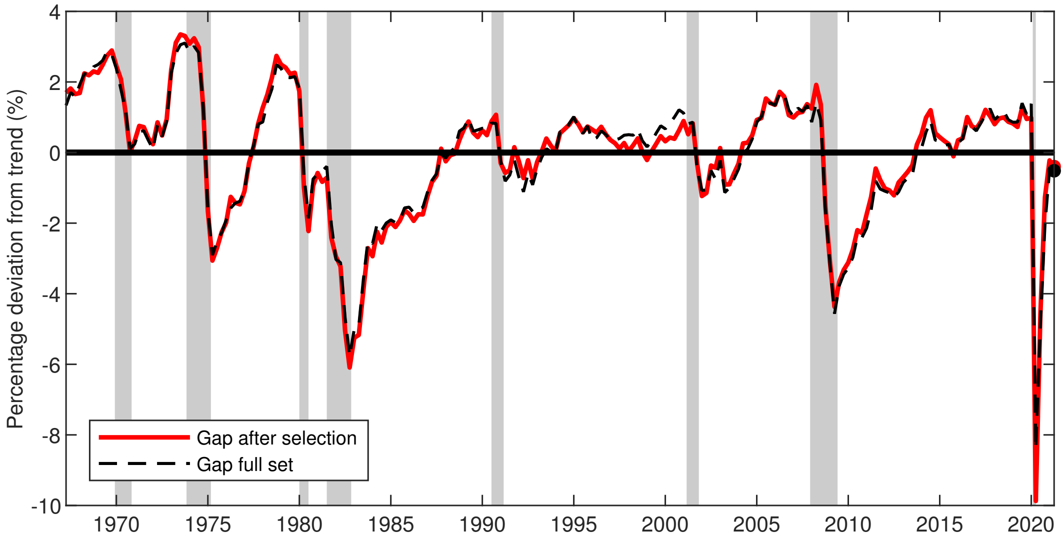
Notes: Standard deviations of the forecast-error contributions for each variable in the information set are reported. White bars correspond to the non labor market variables also considered in BMW and grey bars correspond to labor market variables, with the two labor market variables with the most substantial informational contributions highlighted in red. LFPR stands for labor force participation rate and EMPOP stands for to employment-to-population ratio.

Our benchmark model with only the the U-2 unemployment rate and average hourly earnings as labor market variables results in an output gap estimate that has a correlation of 0.993 with the gap estimated using the full set of labor market variables, suggesting that essentially no relevant information is lost by ignoring the other labor market variables when estimating the output gap.[6] The similarity is confirmed in Figure 4, which plots the output gap estimates before and after variable selection. Figure 5 reports the informational decomposition for the benchmark model and confirms the substantial cyclical information content in the U-2 unemployment rate and average hourly earnings.

maining as the labor market variables with the highest shares, while none of the other labor market variables exhibit any substantial contribution at any stage of this iterative procedure.

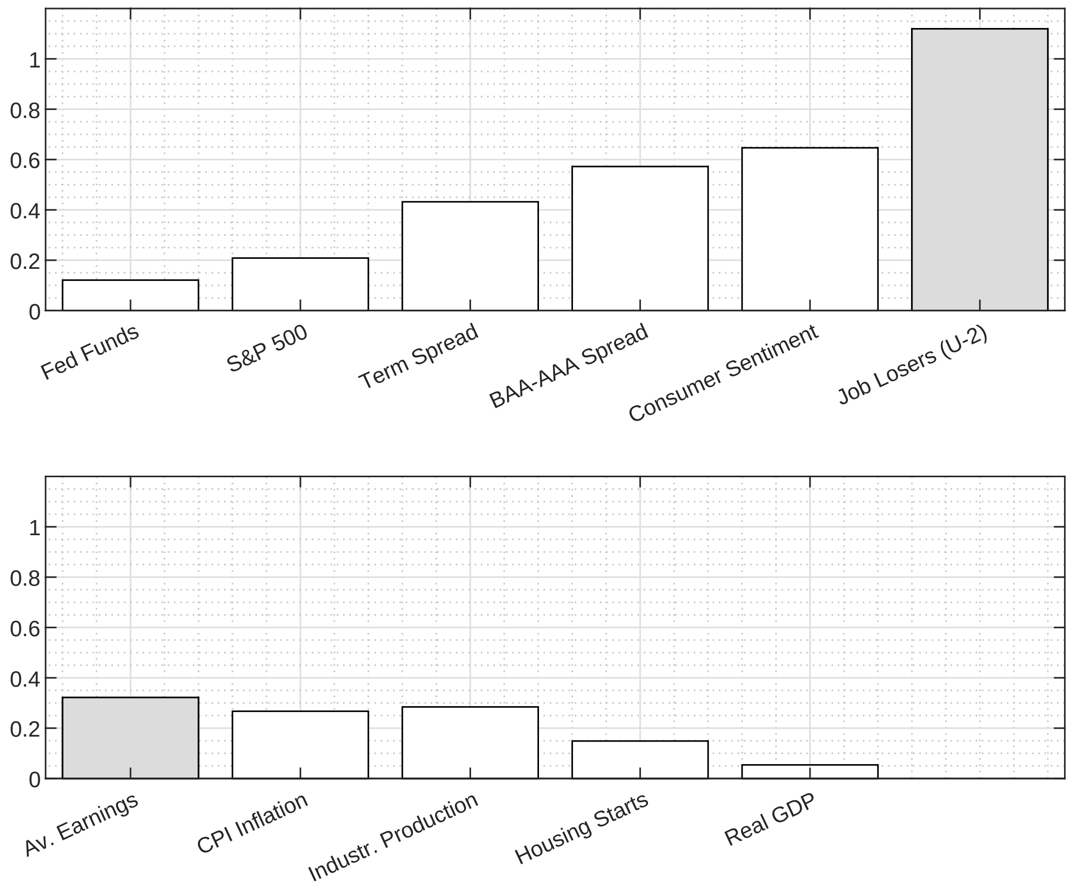
[6]The only substantial difference between the gaps estimated from the full model and the more parsimonious one occurs in 2020Q2, where the latter is roughly 1.4 percentage points more negative than the former. This disparity is reduced when allowing for a structural break in long-run output growth; see Section 3.3.

Figure 4: Comparison of output gap estimates before and after variable selection



Notes: Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER recessions.

Figure 5: Informational decomposition with key labor market variables



Notes: Standard deviations of the forecast-error contributions for each variable in the information set are reported. White bars correspond to the non labor market variables also considered in BMW and grey bars correspond to the selected labor market variables.

It is possibly surprising that other labor market variables, such as the labor force participation rate, the employment-to-population ratio, or the measures of work hours, do not seem to be informative about the output gap beyond any common information captured by the U-2 unemployment rate and average hourly earnings. However, some of these variables, like the labor force participation rate, show little immediate correlation with business cycle fluctuations except for during the Covid-19 recession (see Figure 1), making them less obviously suitable for inclusion in a forecasting model to capture cyclical variation in output. Other variables like the employment-to-population ratio and the measures of work hours, while correlated with the business cycle, also exhibit trends driven by exogenous factors such as demographics, implying that they do not have a strong linear relationship with output growth that can be simply captured by a linear VAR.[7] At the same time, it is highly plausible that the U-2 unemployment rate is a valid and arguably more consistent measure of labor market conditions than the overall unemployment rate, at least around recessions. By only capturing job losers rather than also including unemployed new entrants or re-entrants into the labor force, U-2 may be more robust to more persistent movements related to labor force participation.

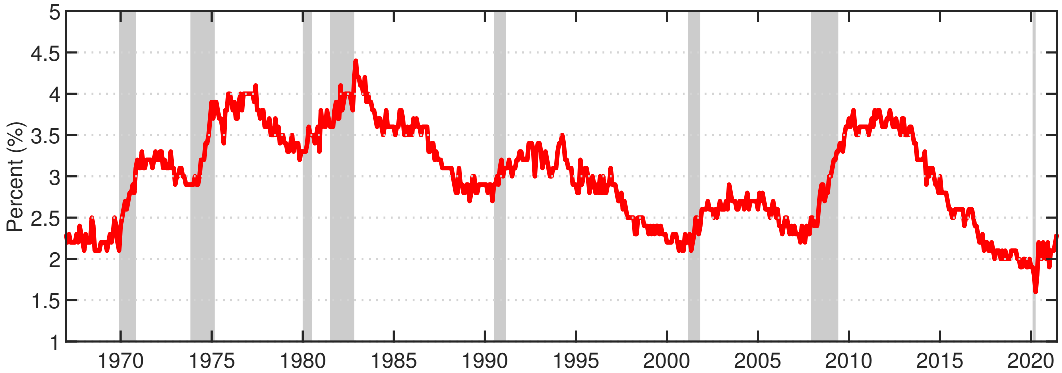
As can be seen in Figure 6, the difference between the overall unemployment rate and the U-2 unemployment rate is highly persistent. The difference appears to increase with recessions and stays elevated for years afterwards until gradually decreasing as expansions become considerably more mature. This is consistent with basic search and matching models that imply the steady-state unemployment rate is inversely related to labor market tightness, at least assuming a greater sensitivity of the job finding rate for new entrants to vacancies than for job losers captured in the U-2 unemployment rate. The idea of slow moving changes in the natural rate of unemployment related to search and matching frictions is also supported by a unit root test for the difference between the unemployment rate and the U-2 unemployment rate, which cannot reject the presence of a unit root. Thus, there may actually be a stochastic trend in the unemployment rate that is obscured when testing for a unit root by large cyclical movements in addition to the smaller persistent changes.[8] Related, the higher average level of the difference between the two unem-

[7]Nonlinear VARs of output and the labor market have been considered; e.g., see Altissimo and Violante (2001). However, we leave analysis of such nonlinearities to future research.

[8]This result is somewhat the opposite of cointegration where two persistent series test as I(1), but a linear combination tests as I(0). In this case, the unemployment rate and the U-2 unemployment rate both test as I(0),

ployment measures in the first half of the sample than the second half directly suggests that the unemployment rate could be overstating labor utilization in the second half of the sample relative to the first half. For example, the peak unemployment rate in the early 1980s recession (10.8%) was nontrivially higher than the peak rate in the Great Recession (10.0%), whereas the peak U-2 unemployment rates were much more similar across the two recessions (6.6% and 6.5%), which is also more consistent with the relative declines in real GDP across the two recessions.

Figure 6: Difference between the overall unemployment rate and the U-2 unemployment rate



Notes: Sample period is January 1967 to June 2021. Shaded bars correspond to NBER recessions.

In terms of average hourly earnings, the informational contribution as a monthly indictor is less than the U-2 unemployment rate, but is greater than the federal funds rate, stock returns, CPI inflation, IP growth, and growth of housing starts. It is also a highly plausible measure of labor market conditions in terms of measuring wage pressures resulting from relative changes in labor supply and demand. Furthermore, there is a direct link of the output gap to real labor costs in a New Keynesian setting (Galí and Gertler 1999), with the joint inclusion of average hourly earnings and CPI inflation in our model, thus capturing information related to changes in real labor costs.

but a linear combination (in this case, the difference between the two series) tests as I(1). It is well-known that unit root tests can have severe size distortions if the variance of the stochastic trend shocks is small relative to the transitory component (Schwert 1989). This could explain why the unemployment rate tests as I(0) even if it has a stochastic trend with small trend shocks. Our attribution of the stochastic trend to the overall unemployment rate rather than the U-2 unemployment rate is motivated by a much more significant rejection of the unit root for the U-2 unemployment rate and the theoretical reasoning that the overall unemployment rate should be more susceptible to persistent movements related to search and matching frictions than the U-2 unemployment rate.

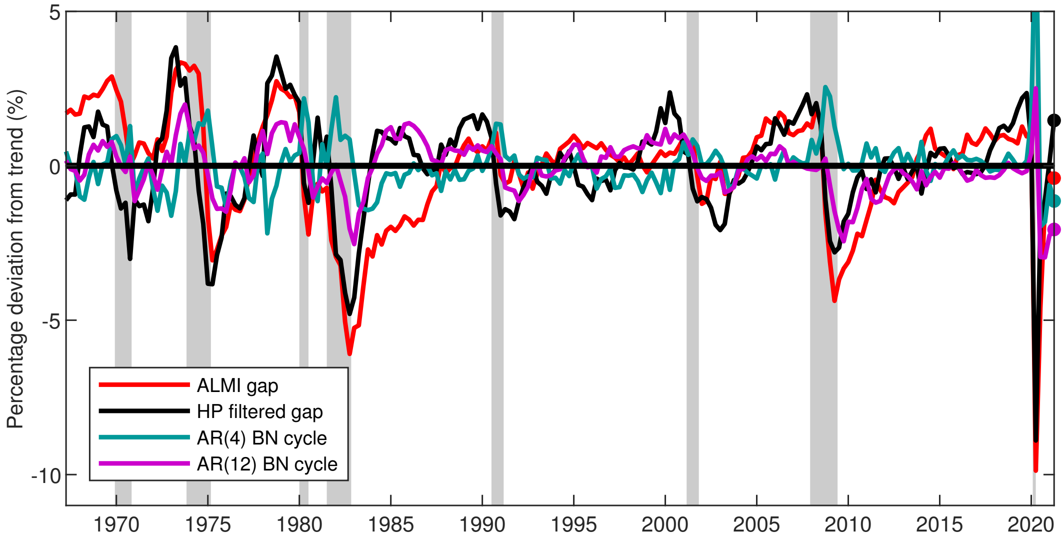
3.2 Comparison with other output gap estimates

Figure 7 compares the estimated output gap from our benchmark 'alternative labor market indicators' (ALMI) model (i.e., the model with U-2 unemployment rate and average hourly earnings as the selected labor market variables) with univariate estimates based on the Hodrick-Prescott (HP) filter (with $\lambda = 1600$, as is standard for quarterly data) and the BN decomposition for AR(4) and AR(12) models of output growth (with parameters also estimated using the pre-Covid sample from 1967-2019). The AR(4) model produces an estimated output gap which, compared to the ALMI and HP filter estimates, is far smaller in amplitude and usually of opposite sign, including often being positive during recessions. The estimate based on the AR(12) model is more intuitive in terms of sign, but is also relatively small in amplitude. These results suggest that the AR(4) model fails to capture negative autocorrelation at longer lags, leading it to interpret large negative innovations in recessions as trend rather than cyclical movements, whereas the AR(12) model seems to capture at least some negative autocorrelation, although it still attributes most of the variance in output to trend movements. Including multivariate information in the ALMI model implies much more predictability in output growth, increasing the amplitude of the estimated output gap in accordance with the comparison of univariate and multivariate BN decompositions in Evans and Reichlin (1994). The HP estimate is similar in amplitude to the ALMI estimate and usually positively correlated with it, although there are substantial differences. Arguably, the ALMI estimate is much more plausible when there are differences, such as with the HP output gap during the Great Recession only being slightly larger in magnitude than during the 2001 recession. Furthermore, the HP filter is unreliable at the end of the sample, rendering it less suitable for consideration of recent developments during the Covid-19 pandemic.[9]

Figure 8 compares our ALMI output gap with the estimate based on the BMW model, which uses the unemployment rate as the only labor market indicator, and the output gap implied by the production-function-based CBO estimate of potential output as a reference point. All three estimated output gaps are of similar amplitude and mostly the same sign, although the

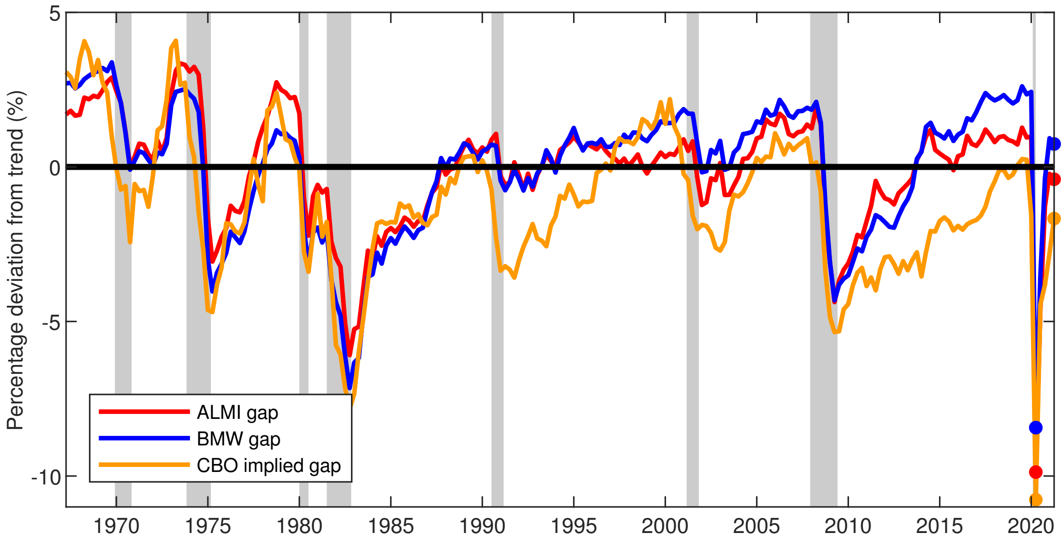
[9] Aastveit and Trovik (2014) use a 54-variable factor model to nowcast and forecast a common factors component of U.S. real GDP and then apply the HP filter to the common factors component in order to estimate the output gap. They find that augmenting common factor estimates with forecasts yields relatively more reliable real-time estimates of the output gap. However, there are still considerably larger revisions in the output gap estimates than with the direct nowcasts of the output gap based on the BN decomposition considered in Berger et al. (2021).

Figure 7: Comparison of the ALMI output gap with estimates from other methods



Notes: Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER recessions.

Figure 8: Comparison of the ALMI output gap with BMW and CBO estimates



Notes: Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER recessions.

CBO estimate exhibits more negative values than the ALMI and BMW estimates from the 1990s. Comparing the ALMI output gap with the BMW output gap, it is noticeable that the former is consistently higher than the latter in the first half of the sample and lower in the second half of the sample. Moreover, the BMW estimate is closer to the CBO estimate in the recessions of the 1970s and 1980s, but the ALMI estimate is closer in the early 2000s recession and during the Covid-19 pandemic.

In order to assess the relevance of output gap estimates for policy, Morley and Wong (2020) propose looking at correlations with one-year-ahead output growth and one-year-ahead inflation, respectively, under the premise that an accurate estimate should have a negative correlation with future output growth and a positive correlation with future inflation. Table 1 reports these correlations for the ALMI, BMW, and CBO estimates, respectively. For output growth, the ALMI estimate performs better than the BMW estimate, and both perform better than the CBO estimate.[10] For inflation, the ALMI estimate performs similarly to the CBO estimate, and much better than the BMW estimate.

Table 1: Correlations with future output growth and inflation

	ALMI gap	BMW gap	CBO gap
Output growth	−0.46	−0.41	−0.32
Inflation	0.25	0.01	0.30

Note: This table reports corr $(\hat{c}_t, \ln(\mathrm{GDP}_{t+4}/\mathrm{GDP}_t))$ and corr $(\hat{c}_t, \ln(\mathrm{CPI}_{t+4}/\mathrm{CPI}_t))$, where $\hat{c}_t$ corresponds to the output gap estimate and time t corresponds to a quarterly frequency.

Somewhat related, Tables 2 and 3 compare the in-sample and pseudo out-of-sample performance of the ALMI and BMW models in nowcasting and forecasting output growth. For each model, the RMSE is reported for nowcasts 1/3, 2/3 and one quarter ahead, and for forecasts one to six quarters ahead (following Camba-Mendez and Rodriguez-Palenzuela 2003). For each horizon, p-values from Diebold-Mariano tests for equality of forecast accuracy between the two models are reported. For in-sample fit, the ALMI model is significantly more accurate than the BMW model at most nowcasting horizons and for one-quarter-ahead forecasts, while it is never significantly less accurate at a 5% significance level. For pseudo out-of-sample performance, there is less of a difference, with only the two- and three-quarter-ahead forecasts for the BMW model significantly more accurate. The importance of these results should not be exaggerated, however, as the ALMI model includes one more variable, giving it an advantage for in-sample-fit, while the shrinkage parameter for each model is specifically set to optimize its one-step-ahead out-of-sample forecast performance. However, the results at least seem to indicate that, as a forecasting model, the ALMI model is comparable to the BMW model.

[10] Furthermore, as highlighted in Berger et al. (2021), the CBO estimate, like the HP filter estimate, is revised considerably over time and its final-vintage values appear to benefit from a look-ahead bias in forecasting future output growth. By contrast, the BMW output gap is much more reliable in real time, with the real-time BMW estimate actually predicting final-vintage CBO and HP filter values better than the respective real-time CBO and HP filter values.

Table 2: In-sample output growth nowcast and forecast fit

Now/forecast horizon	Root mean square error (RMSE)			DM p-value
	ALMI model	BMW model	Difference	
1/3 quarter				
all monthly variables	0.402	0.434	−0.032	0.010
after labor indicators	0.410	0.440	−0.030	0.013
before labor indicators	0.425	0.441	−0.016	0.019
2/3 quarter				
all monthly variables	0.432	0.450	−0.018	0.016
after labor indicators	0.451	0.461	−0.011	0.087
before labor indicators	0.454	0.463	−0.010	0.067
1 quarter				
all monthly variables	0.470	0.481	−0.011	0.054
after labor indicators	0.542	0.560	−0.018	0.035
before labor indicators	0.550	0.566	−0.016	0.016
no monthly variables	0.564	0.581	−0.016	0.023
2 quarters	0.659	0.661	−0.003	0.533
3 quarters	0.720	0.718	0.002	0.589
4 quarters	0.735	0.731	0.004	0.356
5 quarters	0.748	0.748	0.001	0.859
6 quarters	0.759	0.761	−0.003	0.594

Notes: This table reports the root mean squared forecast errors for nowcasts and forecasts of output growth from the ALMI and BMW models for nowcast/forecast horizons between 1/3 and 6 quarters, with parameters estimated over the pre-Covid sample 1967Q1–2019Q4. Nowcast errors are given at three points in time within a month: before the respective labor market variables are included, immediately after they are included, and when all monthly indicators are included. The third column reports the difference between the two RMSEs. The last column reports p-values for Diebold-Mariano (DM) tests of equal forecast accuracy between the two models.

Table 3: Pseudo-out-of-sample output growth nowcast and forecast performance

Now/forecast horizon	Root mean square error (RMSE)			DM p-value
	ALMI model	BMW model	Difference	
1/3 quarter				
all monthly variables	0.493	0.502	−0.010	0.590
after labor indicators	0.483	0.501	−0.019	0.282
before labor indicators	0.497	0.503	−0.006	0.593
2/3 quarter				
all monthly variables	0.481	0.491	−0.010	0.373
after labor indicators	0.489	0.483	0.006	0.592
before labor indicators	0.491	0.486	0.005	0.555
1 quarter				
all monthly variables	0.492	0.488	0.004	0.618
after labor indicators	0.536	0.536	0.000	0.1000
before labor indicators	0.548	0.545	0.003	0.639
no monthly variables	0.540	0.533	0.008	0.265
2 quarters	0.616	0.603	0.013	0.038
3 quarters	0.637	0.618	0.019	0.025
4 quarters	0.630	0.612	0.018	0.080
5 quarters	0.627	0.614	0.013	0.241
6 quarters	0.618	0.607	0.011	0.299

Notes: This table reports the root mean squared forecast errors for nowcasts and forecasts of output growth from the ALMI and BMW models for now/forecast horizons between 1/3 and 6 quarters, where parameters are re-estimated at each point in time, and the now/forecasts are constructed before the data for the respective period are included in the estimation. The first 80 quarters are used as the training sample and the forecast evaluation is performed over the remainder of the sample. The shrinkage parameter is set at the optimized value for the whole sample because Morley and Wong (2020) do not find evidence for time-varying shrinkage. Nowcast errors are given at three points in time within a month: before the respective labor market variables are included, immediately after they are included, and when all monthly indicators are included. The third column reports the difference between the two RMSEs. The last column reports p-values for Diebold-Mariano (DM) tests of equal forecast accuracy between the two models.

Table 4 compares the nowcasting performance for the output gap between the two models. Following Berger et al. (2021), we consider mean absolute error (MAE) for evaluating the accuracy of the nowcasts. The table reports MAE for nowcasts 1/3, 2/3 and one quarter ahead, and in each case differentiating between the nowcasts before the respective labor market variables are included, after their inclusion, and after all monthly variables are included. Diebold-Mariano p-values are also reported for tests of equality of forecast accuracy between the previous nowcast (in the row below) and the current nowcast. Results are similar for both models, with ALMI errors slightly higher at the beginning of a quarter, and BMW errors slightly higher later in a quarter. The reduction in MAE through the inclusion of the respective labor market variables is a bit higher for the BMW model, indicating that it places more relative weight on the unemployment rate than the ALMI model places on the U-2 unemployment rate and average hourly earnings compared to the non labor market variables. However, the contributions of the labor market variables to the output gap nowcast are significant at a 1% significance level for both models.

Table 4: Nowcasting performance for the output gap for the ALMI and BMW models

Nowcast horizon	ALMI model		BMW model	
	MAE	DM p-value	MAE	DM p-value
1/3 quarter				
all monthly variables	0.015	0.000	0.022	0.000
after labor indicators	0.038	0.000	0.042	0.000
before labor indicators.	0.048	0.000	0.064	0.000
2/3 quarter				
all monthly variables	0.089	0.004	0.089	0.006
after labor indicators	0.095	0.000	0.096	0.000
before labor indicators	0.113	0.000	0.135	0.000
1 quarter				
all monthly variables	0.200	0.075	0.185	0.271
after labor indicators	0.202	0.008	0.185	0.000
before labor indicators	0.231	0.000	0.237	0.000
no monthly variables	0.298	-	0.276	-

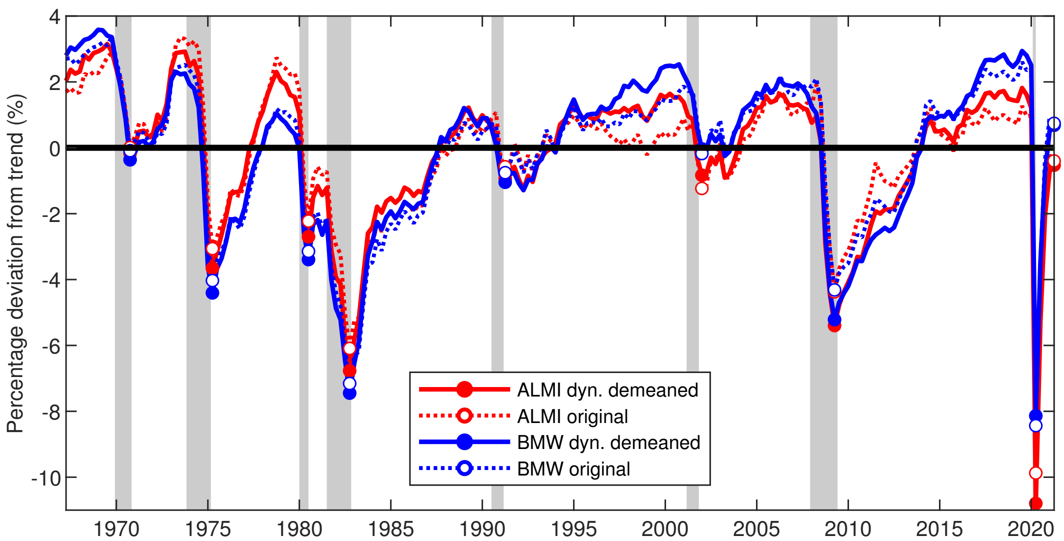
Notes: This table reports the mean absolute errors for nowcasts of the output gap from the ALMI (column 2) and BMW (column 4) models. Nowcast errors are given at three points in time within a month: before the respective labor market variables are included, immediately after they are included, and when all monthly indicators are included. Columns 3 and 5 report the Diebold-Mariano (DM) p-values for a test of equal forecast accuracy given the previous information set (the respective row below) and the current information set.

3.3 Robustness

Next, we assess the robustness of the estimated output gap to consideration of possible structural change in long-run output growth and the level of the unemployment rate.

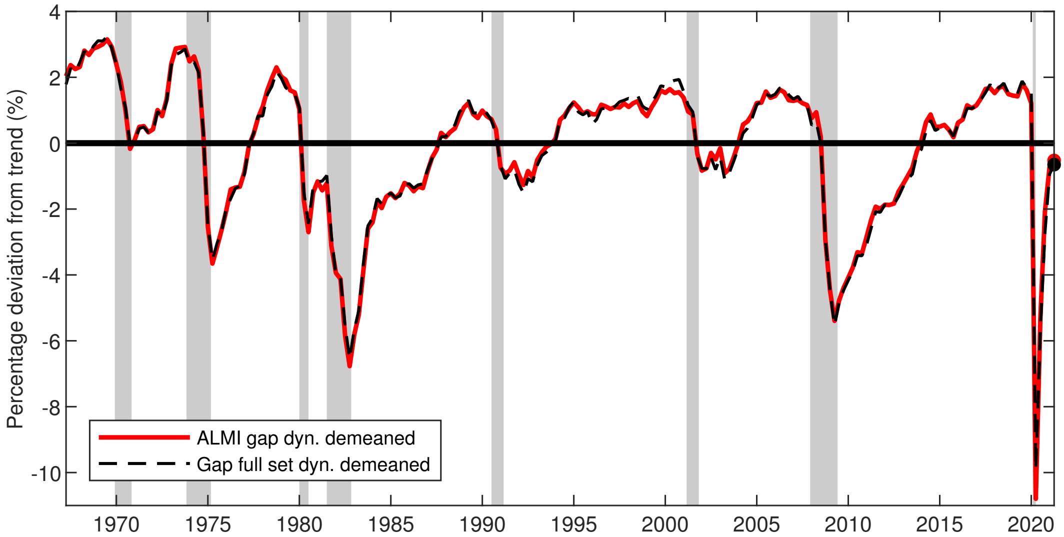
First, to address the Perron and Wada (2009) critique, Berger et al. (2021) examine the robustness of the estimated output gap when allowing for structural breaks in the unconditional mean of output growth $\mu_{\Delta y}$. Specifically, they use the dynamic-demeaning procedure proposed by Kamber et al. (2018), in which $\mu_{\Delta y}$ is estimated using a backward-looking ten-year rolling average of output growth rather than assuming a constant mean for the whole sample. This rolling-window approach is useful in a nowcasting setting by avoiding the problem of having to determine the exact number of structural breaks and their timing. Figure 9 displays the ALMI and BMW estimates with dynamically-demeaned output growth, as well as the constant-mean equivalents for comparison. As is apparent, the qualitative differences between the ALMI and BMW estimates are unchanged, with the ALMI estimate mostly higher in the first and lower in the second half of the sample. The differences between output gap estimates assuming a constant or dynamic mean are generally larger in the second half of the sample, with particularly large differences in the late 1990s, the Great Recession, and, for the ALMI model, the Covid-19 pandemic. Notably, as shown in Figure 10, dynamic demeaning leads to an even smaller difference between an estimate using the full set of labor market variables and the benchmark ALMI model, with the difference in 2020Q2 substantially reduced.

Figure 9: Dynamic demeaning and ALMI vs. BMW estimates of the output gap



Notes: Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER recessions.

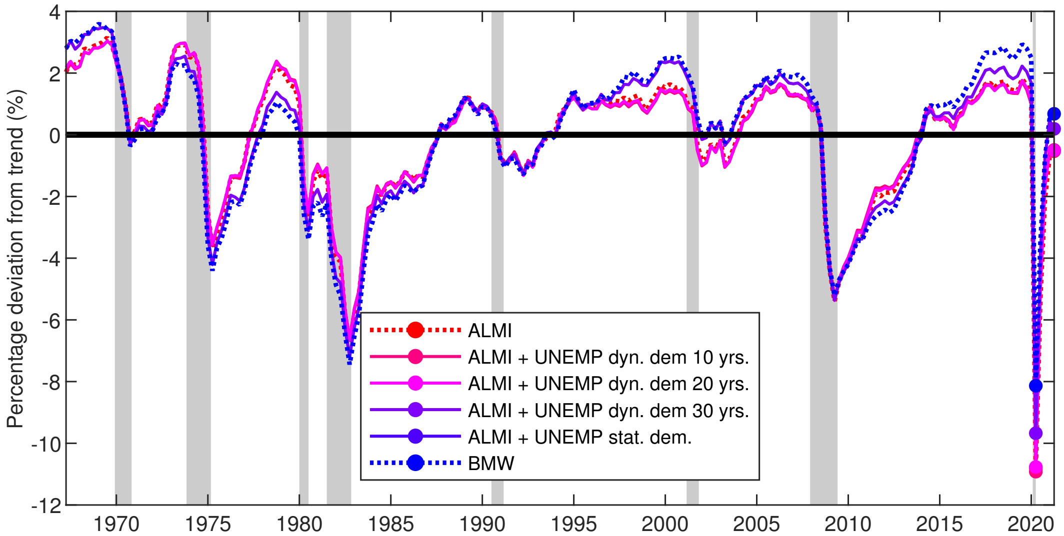
Figure 10: Dynamic demeaning and comparison of output gap estimates before and after variable selection



Notes: Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER recessions.

Second, we assess the robustness of the ALMI estimate to the inclusion of the overall unemployment rate in addition to the U-2 unemployment rate and average hourly earnings. Following the earlier discussion noting the persistence in the difference between the overall unemployment rate and the U-2 unemployment rate, we consider dynamic demeaning of the unemployment rate to capture persistent movements beyond business cycle horizons. Three different lengths (10, 20 and 30 years) are allowed for the rolling window used when dynamically demeaning in order to assess robustness to the window length and to account for the possibility that unemployment recovers more sluggishly from recessions than output (Berger et al. 2016), which would suggest the need for an even longer window to smooth over business cycle fluctuations. Figure 11 shows the results for these four specifications, as well as the benchmark ALMI and BMW estimates for comparison. Strikingly, the extended ALMI estimates with demeaning windows of 10 and 20 years are practically identical to the ALMI estimate without the overall unemployment rate, whereas the 30-year version lies mostly between the ALMI and BMW estimates, while the statically demeaned version is very close to the BMW estimate. Thus, the unemployment rate dominates the alternative indicators if its long-run level is assumed to be constant, but does not add much relevant information beyond the alternative indicators if its long-run level is allowed to vary slowly over time. Given that the difference in estimates reflects persistent movements in the unemployment rate beyond business cycle horizons, it would appear preferable to consider the U-2 unemployment rate or, similarly, the overall unemployment rate allowing for dynamic demeaning when trying to capture cyclical signals from the labor market.

Figure 11: The output gap and dynamic demeaning of the unemployment rate



Notes: This figure plots the estimated output gap for the benchmark ALMI model, four extensions of the ALMI model where unemployment is included assuming a constant mean and dynamically demeaned with windows of 10, 20 and 30 years, respectively, and the BMW model. In each case, output growth is also dynamically demeaned with a window of 10 years. Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER recessions.

3.4 Implications for the Covid-19 recession and recovery

Finally, we consider the implications of the ALMI model for the output gap since the onset of the Covid-19 pandemic. All nowcasting results presented are obtained using all available monthly data up to and including September 2021 and quarterly real GDP until 2021Q2. The monthly data are displayed in Table 5.

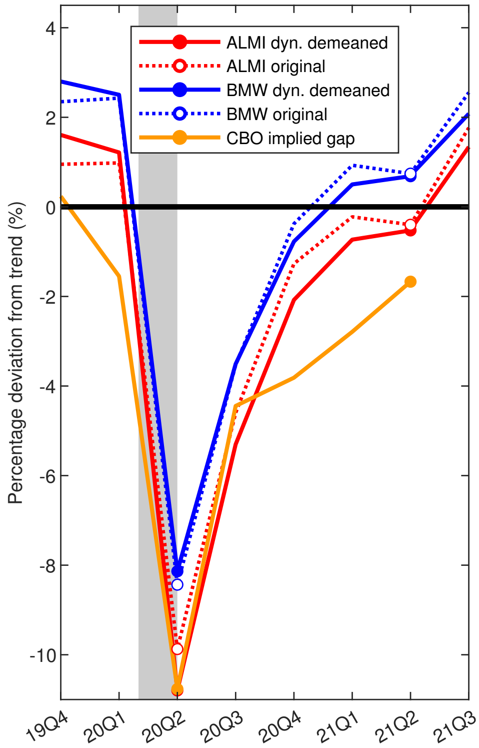
Figure 12 plots the evolution of the final output gap estimates from 2019Q4 to 2021Q2 for the ALMI and BMW models, each with and without dynamically demeaned output growth, and based on the CBO estimate of potential output. In addition, for the ALMI and BMW estimates, the nowcasts for the output gap in 2021Q3 are plotted. As can be seen, the estimates are broadly similar, but the BMW estimates are consistently higher than the ALMI and CBO estimates. This

Table 5: Monthly variables for 2021Q3

Variable	July	August	September
Non-labor indicators			
Federal funds rate (%)	0.08	0.09	0.08
S&P 500 returns (%)	5.90	0.64	0.06
Term spread (%)	1.24	1.21	1.30
BAA−AAA spread (%)	0.67	0.69	0.70
Consumer sentiment (index)	81.2	70.3	72.8
Monthly CPI Inflation (%)	0.47	0.27	
IP growth (%)	0.84	0.40	
Housing starts growth (%)	−6.22	3.93	
ALMI labor indicators			
U-2 job losers rate (%)	3.10	2.80	2.50
Av. earnings growth (%)	0.54	0.58	0.54
BMW labor indicator			
Unemployment rate (%)	5.40	5.20	4.80

Notes: This table reports all monthly data used in the ALMI and BMW nowcasts available in mid October 2021. For clarity, the federal funds rate is displayed non-differenced instead of first-differenced and average hourly earnings is displayed in growth rates rather than first differences of growth rates. Log changes are converted to percentage changes.

Figure 12: Estimated output gaps around the Covid pandemic



Notes: The values for 2021Q3 are nowcasts constructed from the monthly data shown in Table 5. The shaded bar corresponds to the NBER recession.

is consistent with the findings in Section 3.1 concerning the signals from the unemployment rate in the second half of the sample. The ALMI estimates are very similar to the CBO estimate at the trough of the Covid-19 recession, but higher before and after. Possibly surprisingly, the BMW model implies a positive output gap for 2021Q1 and Q2 at just under 1%, despite the unemployment rate, consumer sentiment, and other variables still being far from pre-pandemic levels. However, even the ALMI model nowcast in 2021Q3 is positive, which is consistent with recent heightened inflationary pressures in the U.S. economy.

4 Conclusion

Multivariate trend-cycle decomposition depends crucially on variable selection. The unemployment rate is often considered in multivariate trend-cycle decomposition to estimate the output gap and appears to be an informative variable for cyclical movements in real GDP. However, the unemployment rate also appears to exhibit persistent movements beyond business cycle horizons. So a natural question arises as to whether other labor market variables contain similar information about cyclical movements in output, without the more persistent movements that affect the unemployment rate. Recent concerns among policymakers about whether the unemployment rate fully captured labor market conditions in the Great Recession and the Covid pandemic further motivate our consideration of alternative labor market indicators. Variable selection based on informational decompositions as in Morley and Wong (2020) suggest that the U-2 unemployment rate, which captures unemployment due to lost jobs in particular, and, to a lesser extent, growth in average earnings are informative alternative labor market indicators for the U.S. output gap.

The output gap estimates resulting from these alternative labor market indicators are broadly similar to those obtained using the unemployment rate and other common methods of trend-cycle decomposition. However, in comparison to estimates obtained when using the unemployment rate, the output gap implied by the alternative labor market indicators generally imply a lower level of the output gap since the mid-1990s and especially with the Covid-19 pandemic. The multivariate forecasting model based on the alternative labor market indicators appears to be credible in terms of in-sample and out-of-sample performance compared to the original multivariate forecasting model using the unemployment rate in Berger et al. (2021), with the estimates robust to allowing for structural changes in long-run output growth.

The fact that the difference between the overall unemployment rate and the U-2 unemployment rate tests as having a unit root, as well as the similar results for a model that also includes the overall unemployment rate but considers dynamic demeaning to remove highly persistent movements in the unemployment rate over time, supports the idea that the U-2 unemployment rate provides clearer cyclical signals for estimating the output gap than the overall unemployment rate. Also, average hourly earnings appear to contain relevant information about the output gap, consistent with New Keynesian theory.

References

Aastveit, K. A. and T. Trovik (2014). Estimating the output gap in real time: A factor model approach. *The Quarterly Review of Economics and Finance 54*, 180–193.

Altissimo, F. and G. L. Violante (2001). The non-linear dynamics of output and unemployment in the U.S. *Journal of Applied Econometrics 16*(4), 461–486.

Barigozzi, M. and M. Luciani (2022). Measuring the output gap using large datasets. *The Review of Economics and Statistics, forthcoming.*

Berger, T., G. Everaert, and H. Vierke (2016). Testing for time variation in an unobserved components model for the U.S. economy. *Journal of Economic Dynamics and Control 69*, 179–208.

Berger, T., J. Morley, and B. Wong (2021). Nowcasting the output gap. *Journal of Econometrics, forthcoming.*

Berger, T. and H. Vierke (2017). Estimating the natural rate of hours. *Macroeconomic Dynamics 21*(6), 1426–1453.

Cajner, T., L. Crane, R. Decker, J. Grigsby, A. Hamins-Puertolas, E. Hurst, C. Kurz, and A. Yildirmaz (2020). The U.S. labor market during the beginning of the pandemic recession. NBER Working Paper No. 27159.

Camba-Mendez, G. and D. Rodriguez-Palenzuela (2003). Assessment criteria for output gap estimates. *Economic Modelling 20*(3), 529–562.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020). Labor markets during the COVID-19 crisis: A preliminary view. NBER Working Paper No. 27017.

Erceg, C. J. and A. T. Levin (2014). Labor force participation and monetary policy in the wake of the Great Recession. *Journal of Money, Credit and Banking 46*(S2), 3–49.

Evans, G. and L. Reichlin (1994). Information, forecasts, and measurement of the business cycle. *Journal of Monetary Economics 33*(2), 233–254.

Faberman, R. J., A. I. Mueller, A. Şahin, and G. Topa (2020). The shadow margins of labor market slack. *Journal of Money, Credit and Banking 52*(S2), 355–391.

Fleischman, C. A. and J. M. Roberts (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), 1–39.

Fujita, S. (2014). On the causes of declines in the labor force participation rate. Technical report, Federal Reserve Bank of Philadelphia, Research Rap Special Report.

Galí, J. and M. Gertler (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics 44*(2), 195–222.

Hakkio, C. and J. Willis (2014). Kansas City Fed's Labor Market Conditions Indicators (LMCI). *Macro Bulletin*, 1–2.

Kamber, G., J. Morley, and B. Wong (2018). Intuitive and reliable estimates of the output gap from a Beveridge-Nelson filter. *The Review of Economics and Statistics 100*(3), 550–566.

Morley, J. and B. Wong (2020). Estimating and accounting for the output gap with large Bayesian vector autoregressions. *Journal of Applied Econometrics 35*(1), 1–18.

Morley, J. C. (2002). A state–space approach to calculating the Beveridge–Nelson decomposition. *Economics Letters 75*(1), 123–127.

Perron, P. and T. Wada (2009). Let's take a break: Trends and cycles in US real GDP. *Journal of Monetary Economics 56*(6), 749–765.

Powell, J. H. (2021). Getting back to a strong labor market. Speech on February 10th, 2021, at the Economic Club of New York. Available at `https://www.federalreserve.gov/newsevents/speech/powell20210210a.htm` [Accessed: 02/04/2021].

Schwert, G. W. (1989). Tests for unit roots: A Monte Carlo investigation. *Journal of Business & Economics Statistics 7*(2), 147–159.

U.S. Bureau of Labor Statistics (2015). How the government measures unemployment. Online Resource. Available from `https://www.bls.gov/cps/cps_htgm.htm#unemployed` [Accessed 28/05/2021].

van Zandweghe, W. (2012). Interpreting the recent decline in labor force participation. *Economic Review, Federal Reserve Bank of Kansas City 97*(Q I), 5–34.

Yellen, J. L. (2014). Labor market dynamics and monetary policy. Speech on August 22nd, 2014 at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming. Available at `https://www.federalreserve.gov/newsevents/speech/yellen20140822a.htm` [Accessed: 27/05/2021].

A Appendix

A.1 Data and transformations

The table below reports the details of the data considered in our empirical analysis. The data were obtained from the Federal Reserve Economic Database (FRED) and the FRED mnemonic is provided. The "Adjust" column refers to any data transformations: 'ln' indicates natural logarithms have been taken and 'Δ^i' indicates the variable has been differenced i times. Differences are taken if a Chow test for a change in mean from the first half to the second half of the sample rejects at a 5% level and/or an augmented Dickey-Fuller test fails to reject a unit root at the 5% level. An 'x' in the 'BM' column indicates that a variable is included in our benchmark model.

Table 6: Details of the data and transformations Details of the data and transformations

Series	Mnemonic	Adjust	BM
Variables also considered in BMW			
Effective Federal Funds Rate	FEDFUNDS	Δ	x
S&P 500	SP500	ln, Δ	x
10 Year $-$ 1 Year Treasury Term Spread	DGS10, DGS1		x
Corporate BAA $-$ AAA Spread	BAAFFM, AAAFFM		x
University of Michigan: Consumer Sentiment	UMCSENT		x
Unemployment Rate	UNRATE		
Consumer Price Index for All Urban Consumers: All Items in U.S. City Average	CPIAUCSL	ln, Δ	x
Industrial Production: Total Index	INDPRO	ln, Δ	x
New Privately-Owned Housing Units Started: Total Units	HOUST	ln, Δ	x
Real Gross Domestic Product	GDPC1	ln, Δ	x
Additional labor market variables			
Unemployment Rate	UNRATE		
Percent of Civilian Labor Force Unemployed 15 Weeks and Over (U-1)	U1RATE	Δ	
Unemployment Rate - Job Losers (U-2)	U2RATE		x
Job Leavers as a Percent of Total Unemployed	LNS13023706	Δ	
Job Losers as a Percent of Total Unemployed	LNS13023622	Δ	

Labor Force Participation Rate	CIVPART	Δ	
Labor Force Participation Rate - 25-54 Yrs.	LNS11300060	Δ	
Employment-Population Ratio	EMRATIO	Δ	
Employment-Population Ratio - 25-54 Yrs.	LNS12300060	Δ	
Employed, Usually Work Full Time	LNS12500000	ln, Δ	
Employment Level - Part-Time for Economic Reasons, All Industries	LNS12032194	ln, Δ	
Average Weekly Hours of Production and Nonsupervisory Employees, Manufacturing	AWHMAN	ln, Δ	
Average Weekly Overtime Hours of Production and Nonsupervisory Employees, Manufacturing	AWOTMAN	ln, Δ	
Average Hourly Earnings of Production and Nonsupervisory Employees, Total Private	AHETPI	Δ^2	x
Employment Rate in Hours (LNS12300060 $\cdot$ AWHMAN /40)	LNS12300060, AWHMAN	Δ	