Oxford Open Economics, 2022, **1**, 1–16
https://doi.org/10.1093/ooec/odab002
Advance access publication date 17 March 2022
Research Article

Cyclical signals from the labor market

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

[1]Chair for Empirical International Economics, University of Göttingen, 37073 Göttingen, Germany
[2]Department of Economics, University of Warwick, Coventry, CV4 7AL, United Kingdom
[3]School of Economics, University of Sydney, Sydney, NSW 2006, Australia
[4]Department of Econometrics and Business Statistics, Monash University, Melbourne, VIC 3145, Australia
*Correspondence address. School of Economics, Level 5, Social Sciences Building, University of Sydney, Sydney, NSW 2006, Australia. Tel. +612 9351 3368;
E-mail: james.morley@sydney.edu.au

Abstract

We consider which labor market variables are the most informative for estimating and nowcasting the US 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

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 & Roberts, 2011, Morley & Wong, 2020, Barigozzi & 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.

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 US 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 a rapid reduction in the unemployment rate, falling back down to under 7% by the end of the year. This swift recovery was in stark contrast to previous US recessions, including the Great Recession in 2007–2009, which featured a much more persistent increase 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 US 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 US labor force participation rate, highlighting the main concern expressed in the quotation 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

Received: October 13, 2021. **Revised:** November 22, 2021. **Accepted:** November 29, 2021

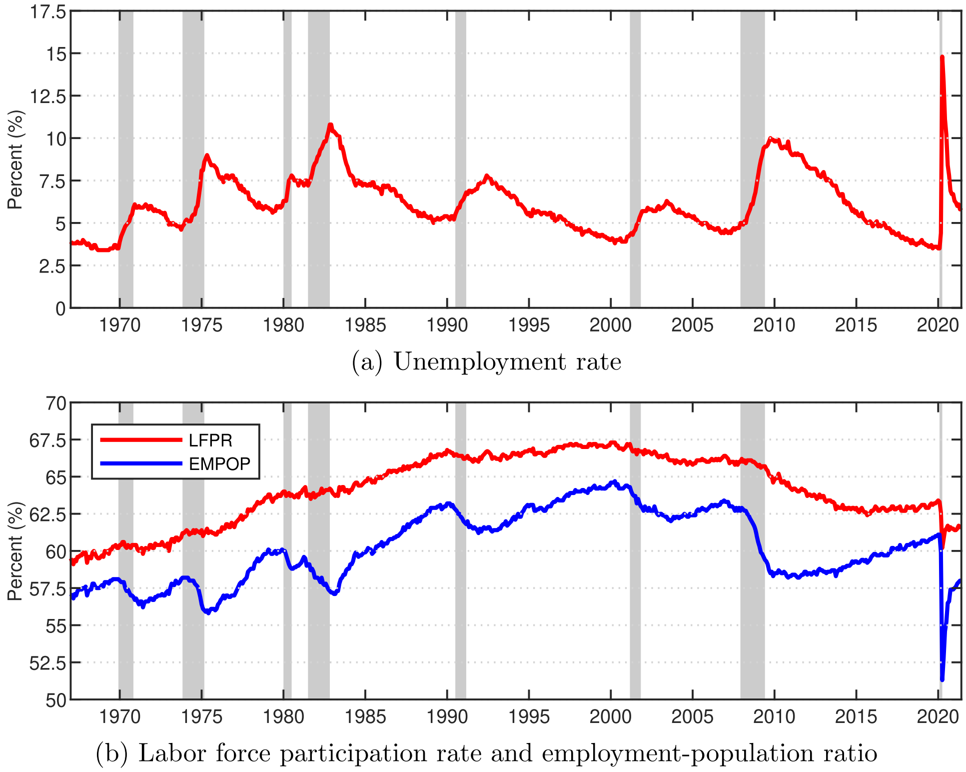

Figure 1. US labor market indicators and recessions.
Notes: Sample period is January 1967 to June 2021. LFPR stands for labor force participation rate and EMPOP stands for employment-to-population ratio. Shaded bars correspond to NBER-dated recessions.

for work (US Bureau of Labor Statistics, 2015), it could understate the true extent of the labor market damage. This possibility is also apparent in the blue line in the bottom panel of Figure 1, which shows a decline in the US 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 e.g. by van Zandweghe (2012), Erceg & 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 & Vierke (2017) and Faberman *et al.* (2020), using measures that account for both the extensive margin (labor force participation) and intensive margin of labor supply, find that US 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 a multivariate trend-cycle decomposition, 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.* (2022) place a lot of weight on information from the unemployment rate and updated estimates for the model imply a positive US 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 vector autoregressive (VAR) model used in Berger *et al.* (2022) (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 & 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 & 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 finding 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 & 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: The next section describes the data and methods employed in our analysis. The following section reports our empirical results for an application to US data, including subsections on informational decompositions, comparison with other estimates of the output gap, a number of robustness checks, and consideration of implications for the output gap since the onset of the Covid-19 pandemic. The last section concludes.

Data and methods

To estimate the output gap, we consider US quarterly log real Gross Domestic Product (GDP) as the target variable for multivariate trend-cycle decomposition. Our multivariate information set includes the following monthly indicators shown by Berger *et al.* (2022) 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 Consumer Price Index (CPI) inflation rate, industrial production (IP) growth, and growth in housing starts.

Morley & Wong (2020) show that the unemployment rate is a particularly important informational variable for multivariate trend-cycle decomposition of US 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 & 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 & 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 & Wong (2020) and also by the variables included in the Kansas City Fed's Labor Market Conditions Indicators (Hakkio & 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 Bureau of Labor Statistics (BLS): U-1 (15 weeks and over) unemployment rate and U-2 (job losers) unemployment rate (note: 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). 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, 'employment rate in hours' is constructed according to Berger & Vierke (2017). Following Berger *et al.* (2022), 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 same potential distortions as the overall unemployment rate.

The data series have a variety of sources including the Bureau of Economic Analysis (BEA) and BLS and were obtained from the Federal Reserve Economic Database (FRED) for the sample period of January 1967 to June 2021 (1967Q1–2021Q2 for real GDP). (Note: Some additional values of monthly indicators for July, August, and September 2021 were obtained for the nowcasting exercise in a subsection below on the output gap since the onset of the Covid-19 pandemic. Also, one of the employment series (Employed, Usually Work Full Time) that is considered in a model with a full set of labor market variables is only available from January 1968, so earlier values were backcast based on the initial observation. However, model selection results that drop this variable from our benchmark model are robust to starting the sample period in January 1968.) Following Berger *et al.* (2022), multivariate trend-cycle decomposition is based on the BN decomposition. Because the BN decomposition calculation presented below requires inversion of the companion matrix for a 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 & 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 Augmented Dickey–Fuller (ADF) test cannot reject a unit root (note: tests are performed on quarterly versions of the variables and significance is determined at the 5% level). For comparability with the Berger *et al.* (2022) results, the variables adopted from their specification are transformed exactly like they are in that paper. 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 matters of model specification, estimation, and implementation of trend-cycle decomposition and nowcasting closely follow Berger *et al.* (2022). Specifically, letting $m_{j,t-1+v}$ be the j^{th} variable observed at monthly frequency in quarter t, where $v \in \{1/3, 2/3, 1\}$ corresponds to the month within the quarter, we stack the k monthly indicators as

$$\boldsymbol{m}_{t-v} = \begin{bmatrix} \tilde{m}_{1,t-1+v} \\ \tilde{m}_{2,t-1+v} \\ \vdots \\ \tilde{m}_{k,t-1+v} \end{bmatrix},$$

where $\tilde{m}_{j,t-1+v} \equiv m_{j,t-1+v} - \mu_j$, and μ_j is the mean of the j^{th} monthly indicator. Denoting $\Delta\tilde{y}_t \equiv \Delta y_t - \mu_{\Delta y}$, where $\mu_{\Delta y}$ is the mean of real GDP growth, we then stack all of the demeaned variables observed at monthly frequency within the quarter, along with demeaned real GDP growth, which is observed at quarterly frequency, as follows:

$$\mathbf{Y}_t = \begin{bmatrix} \boldsymbol{m}_{t-1+1/3} \\ \boldsymbol{m}_{t-1+2/3} \\ \boldsymbol{m}_t \\ \Delta\tilde{y}_t \end{bmatrix}.$$

Letting $n = 3k + 1$, the $n \times 1$ vector process $\mathbf{Y}_t$ is assumed to have a VAR(p) structure at the quarterly frequency:

$$\mathbf{Y}_t = \boldsymbol{\Phi}_1\mathbf{Y}_{t-1} + \boldsymbol{\Phi}_2\mathbf{Y}_{t-2} + \cdots + \boldsymbol{\Phi}_p\mathbf{Y}_{t-p} + \boldsymbol{e}_t, \quad \boldsymbol{e}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}). \tag{1}$$

Parameters are estimated using Bayesian methods with a Minnesota-type shrinkage prior, where the shrinkage hyperparameter is set based on minimizing the one-step-ahead root mean squared forecast error for output growth, as in Morley & Wong (2020) and Berger *et al.* (2022). Then, following Morley (2002) and Morley & Wong (2020), the VAR(p) model can be cast into companion form:

$$\mathbf{X}_t = \mathbf{F}\mathbf{X}_{t-1} + \boldsymbol{H}\boldsymbol{e}_t, \tag{2}$$

where $\mathbf{X}_t$ is an $np \times 1$ vector of stationary and demeaned variables, $\boldsymbol{F}$ is the $np \times np$ companion matrix that includes the $\boldsymbol{\Phi}$ matrices from the VAR, $\boldsymbol{H}$ is an $np \times n$ matrix that maps the VAR forecast errors to the companion form and $\boldsymbol{e}_t$ is the $n \times 1$ vector of the forecast errors also given in equation 1. Let $\mathbf{s}'_{np,r}$ be a $np \times 1$ selector vector that consists of 1 as its r^{th} row and zeros otherwise. As demeaned real GDP growth is included as the n^{th} element of the vector $\mathbf{X}_t$ in equation (2), the BN cycle of y_t can be calculated following Morley (2002) as

$$c_t = -\mathbf{s}'_{np,n}\mathbf{F}(\mathbf{I} - \mathbf{F})^{-1}\mathbf{X}_t. \tag{3}$$

Following Berger *et al.* (2022), the nowcast for the output gap is then calculated as

$$c_{T+1|T+\omega} = -\mathbf{s}'_{np,n}\mathbf{F}(\mathbf{I} - \mathbf{F})^{-1}\left[\mathbf{F}\mathbf{X}_T + \boldsymbol{H}\boldsymbol{e}_{T+1|T+\omega}\right], \tag{4}$$

where, letting $\omega \in (0, 1)$ correspond to the fraction of the interval of time in which all data for a quarter are released that a particular i^{th}-ordered monthly indicator becomes available, $\boldsymbol{e}_{T+1|T+\omega} = \mathbf{B}\mathbf{z}_{T+1|T+\omega}$, with $\mathbf{z}_{T+1|T+\omega} = \left[\mathbf{z}^i_{T+1} \;\; \mathbf{0}\right]'$, $\mathbf{z}^i_{T+1} = \mathbf{B_i}^{-1}\boldsymbol{\varepsilon}^i_{T+1}$, and $\mathbf{B_i}$ corresponding to the first $i \times i$ elements of the lower-triangular Cholesky factor $\boldsymbol{B}$ of the forecast error variance–covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}'$. This approach is based on Waggoner & Zha (1999) and it should be noted that the ordering has no structural interpretation but only reflects the order in which data are released.

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, as in Berger *et al.* (2022). 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.

Empirical results

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. 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 of the full set of labor market variables, as in the BMW model.

The two estimates are reasonably similar for most of the sample period and both align with the National Bureau of Economic Research (NBER) reference cycle, implying the information captured by the alternative

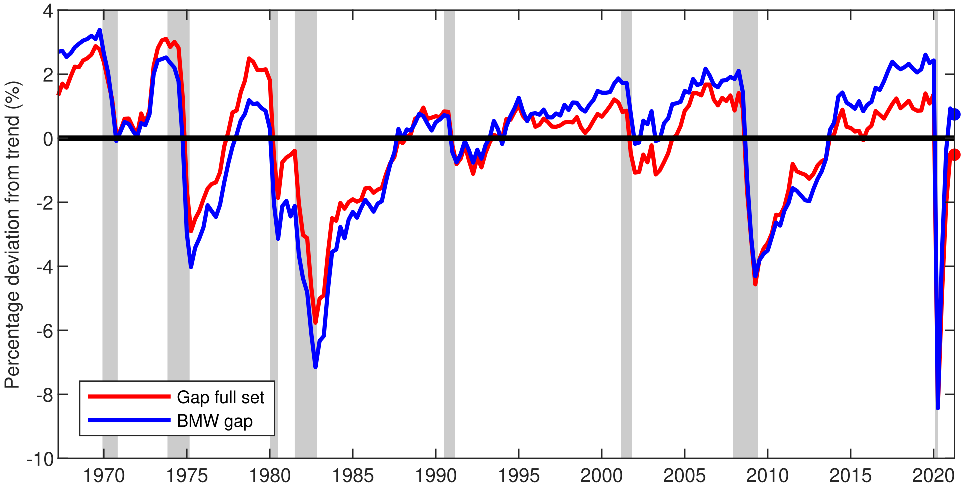

Figure 2. Estimated output gap with and without full set of labor market variables.
Notes: Sample period is 1967Q1 to 2021Q2. BMW refers to the Berger *et al.* (2022) model. Shaded bars correspond to NBER-dated recessions.

labor market variables is qualitatively similar to that captured by the unemployment rate. Meanwhile, the fact that they are not identical suggests that the unemployment rate does not summarize all of the relevant cyclical information in the labor market.

We modify the Morley & Wong (2020) informational decomposition and variable selection for a mixed-frequency setting with monthly indicators and a quarterly target variable by summing the contributions of forecast errors for the first, second, and third months in a quarter of each monthly indicator in the mixed-frequency VAR. Based on equations (2) and (3), the contribution of the j^{th} monthly indicator in a mixed-frequency setting is

$$c_{j,t} = -\sum_{i=1}^{3}\sum_{l=0}^{t-1} \mathbf{s}'_{np,n}\mathbf{F}^{l+1}(\mathbf{I}-\mathbf{F})^{-1}\mathbf{H}\mathbf{s}_{n,j+(i-1)k}\mathbf{s}'_{n,j+(i-1)k}\boldsymbol{e}_t, \quad (5)$$

where, as noted in the previous section, k is equal to the number of monthly indicators included in the model and $n = 3k + 1$ is equal to the total number of variables in the mixed-frequency VAR such that output growth is the n^{th} variable.

Following Morley & Wong (2020), the standard deviation of $c_{j,t}$ is used to quantify the j^{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 (note: average hourly earnings is included in the VAR in second differences, but for brevity is simply referred to as 'average hourly earnings'). Because these two variables are the only ones whose contributions are larger than the smallest contribution among the Berger *et al.* (2022) 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. (Note: 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 & 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 remaining 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.)

Our benchmark model with only the the U-2 unemployment rate and average hourly earnings as labor market variables produces 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. (Note: The only noticeable 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 the robustness subsection below.) The similarity is confirmed in Figure 4, which plots the output gap estimates before and after variable selection. The similarity provides a notable contrast to the comparatively larger differences between the estimates in Figure 2, suggesting that the U-2 unemployment rate and average hourly earnings are better than the overall unemployment rate at summarizing all of the relevant cyclical information in the labor market. 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.

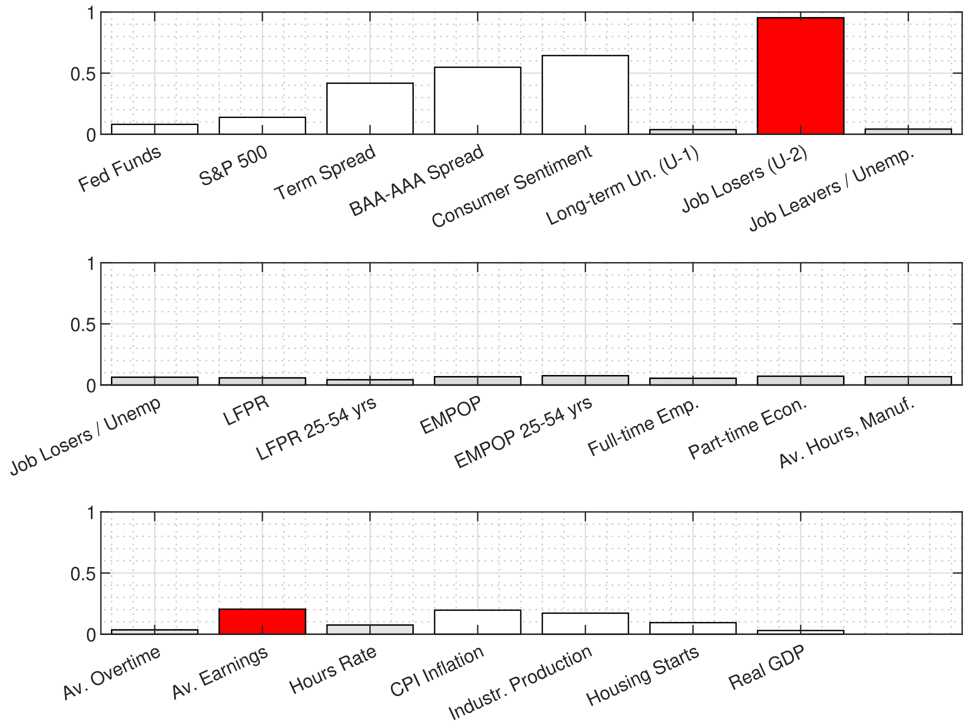

Figure 3. Informational decomposition with full set of labor market variables.
Notes: Standard deviations of the forecast-error contributions for each variable in the information set are reported. White bars correspond to non labor market variables also considered in Berger *et al.* (2022) and grey bars 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 employment-to-population ratio.

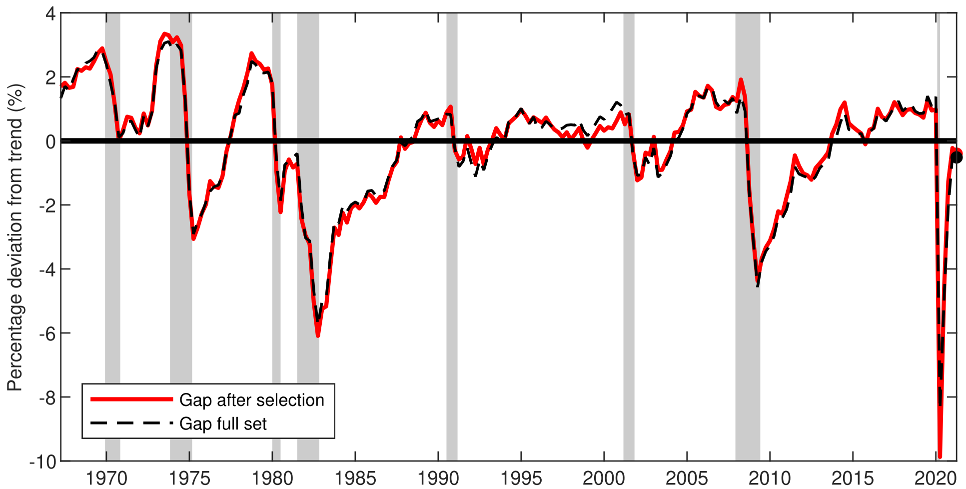

Figure 4. Comparison of output gap estimates before and after variable selection.
Notes: Sample period is 1967Q1 to 2021Q2. Shaded bars correspond to NBER-dated recessions.

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 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 captured by a linear VAR. (Note: Nonlinear VARs of output and the labor market have been considered in the literature; e.g. see Altissimo & Violante (2001). However, we leave analysis of such nonlinearities to future research.) At the same time,

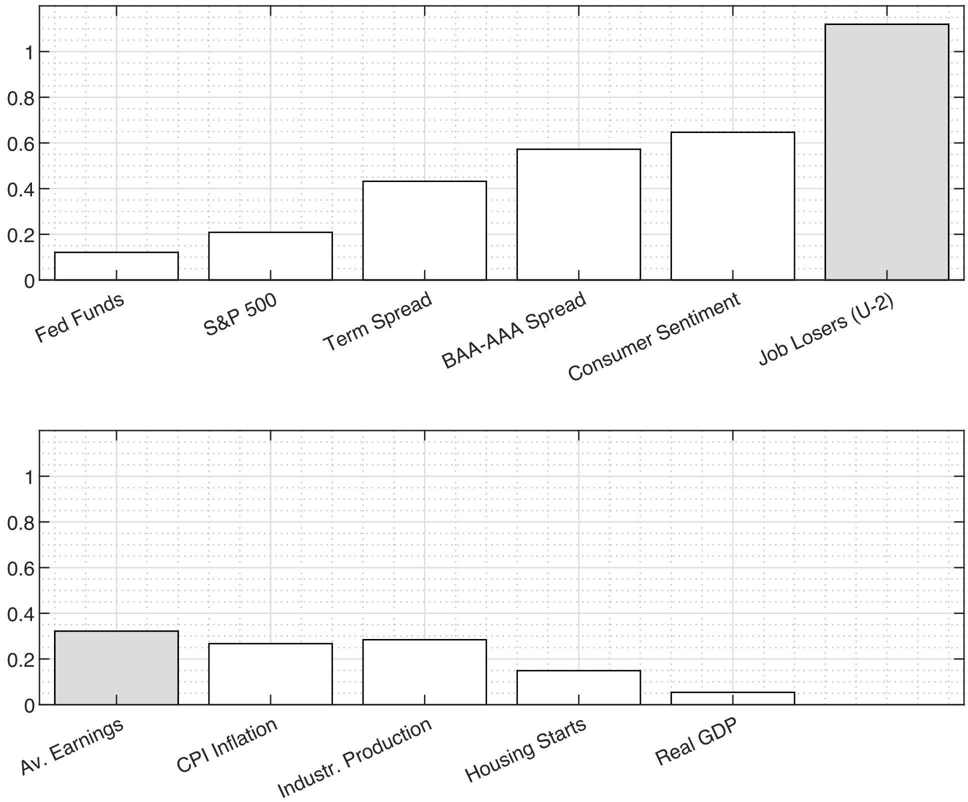

Figure 5. Informational decomposition with selected labor market variables.
Notes: Standard deviations of the forecast-error contributions for each variable in the information set are reported. White bars correspond to the non labor market variables also considered in Berger *et al.* (2022) and grey bars correspond to the selected labor market variables.

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 changes in labor force participation.

As can be seen in Figure 6, there are a lot of similar cyclical movements in the overall unemployment rate and the U-2 unemployment rate, but their difference is highly persistent. In particular, 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. (Note: The *P*-value for an ADF test allowing for a constant mean under the alternative and selecting lags based on AIC is 0.209. Also, the more powerful Elliott–Rothenberg–Stock (Elliott *et al.*, 1996) ADF-GLS test again using AIC is not significant at the 10% level. Results are robust to using BIC for lag selection. Meanwhile, a KPSS test can reject stationarity at the 5% level, supporting the unit root test results.) 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. (Note: 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), 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 more significant rejection of the unit root for the U-2 unemployment rate (ERS test is significant at 1% level versus 5% level for the overall 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

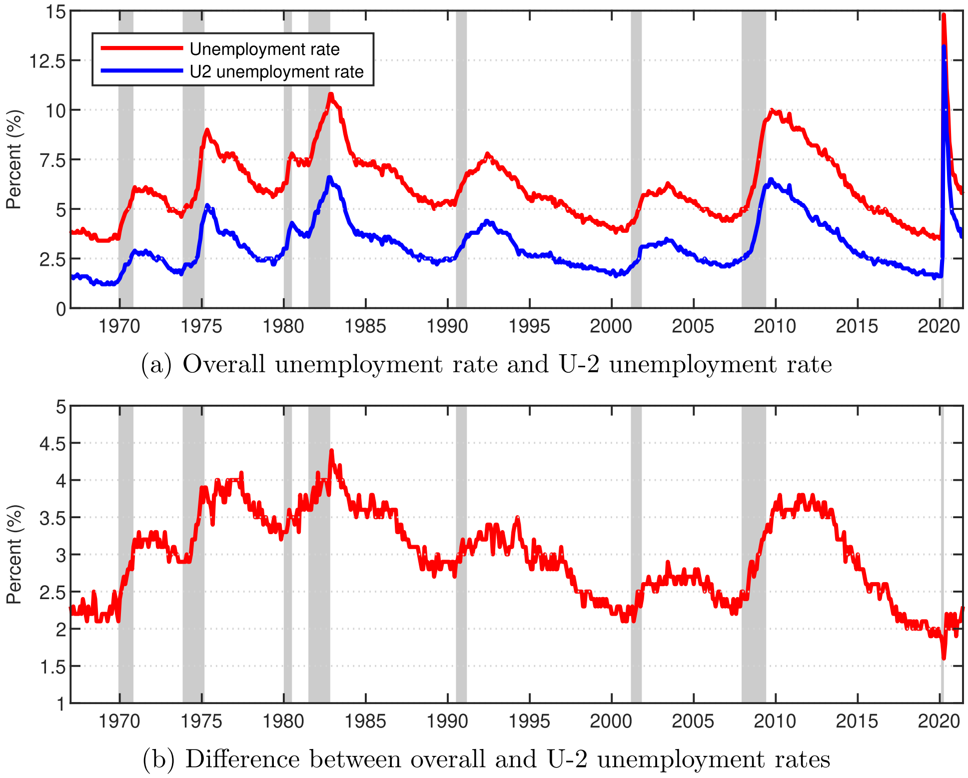

Figure 6. Different measures of unemployment.
Notes: Sample period is January 1967 to June 2021. Shaded bars correspond to NBER-dated recessions.

unemployment rate. Even if the difference between the unemployment rate and the U-2 unemployment rate is actually I(0) and the unit root test results for their difference reflect weak power against a highly persistent alternative, the higher persistence of the overall unemployment rate means that the estimated output gap will inherit more persistent movements than typically thought of as related to business cycle horizons.) Related, the higher average level of the difference between the two unemployment 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.

In terms of average hourly earnings, the informational contribution as a monthly indictor is less than for the U-2 unemployment rate, but is greater than for 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 reflecting 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 & 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.

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

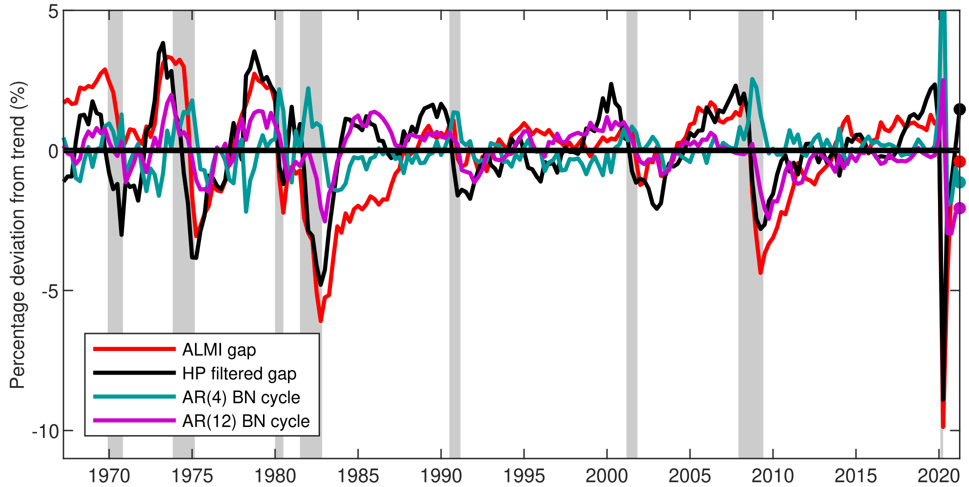

Figure 7. Comparison of the ALMI output gap with estimates from other methods.
Notes: Sample period is 1967Q1 to 2021Q2. ALMI refers to the benchmark 'alternative labor market indicators' model. HP refers to the Hodrick-Prescott filter with $\lambda = 1600$. AR refers to a univariate autoregressive model with the stated number of lags. BN refers to the Beveridge-Nelson decomposition. Shaded bars correspond to NBER-dated recessions.

with the comparison of univariate and multivariate BN decompositions in Evans & 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. (Note: Aastveit & Trovik (2014) use a 54-variable factor model to nowcast and forecast a common factors component of US 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 much 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.* (2022).)

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 CBO estimate generally exhibits more negative values than the ALMI and BMW estimates since 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.

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 $(\hat{c}_t, ln(GDP_{t+4}/GDP_t))$ and corr $(\hat{c}_t, ln(CPI_{t+4}/CPI_t))$, where $\hat{c}_t$ corresponds to the output gap estimate and time t corresponds to a quarterly frequency. ALMI refers to the benchmark 'alternative labor market indicators' model. BMW refers to the Berger *et al.* (2022) model. CBO refers to the Congressional Budget Office.

In order to assess the relevance of output gap estimates for policy, Morley & Wong (2020) proposed 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. (Note: Furthermore, as highlighted in Berger *et al.* (2022), 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.) For inflation, the ALMI estimate performs similarly to the CBO estimate, and much better than the BMW estimate.

Somewhat related, Tables 2 and 3 compare the in-sample and pseudo out-of-sample performance of the ALMI and BMW models in forecasting and nowcasting output growth. For each model, the mean absolute error (MAE) is reported for forecast/nowcast horizons from 6 quarters ahead (following Camba-Mendez & Rodriguez–Palenzuela, 2003) to within one month (1/3 quarter). For

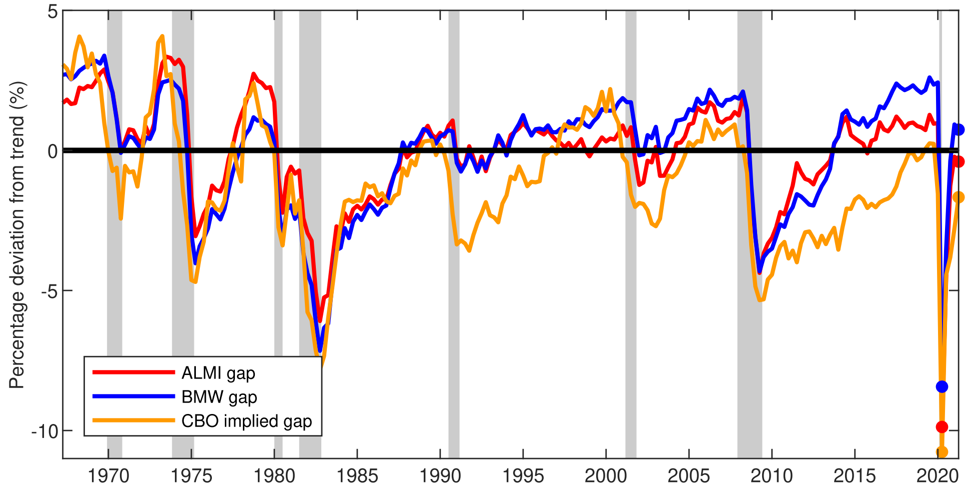

Figure 8. Comparison of the ALMI output gap with BMW and CBO estimates.
Notes: Sample period is 1967Q1 to 2021Q2. ALMI refers to the benchmark 'alternative labor market indicators' model. BMW refers to the Berger *et al.* (2022) model. CBO refers tot he Congressional Budget Office. Shaded bars correspond to NBER-dated recessions.

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

Horizon	Mean absolute error			DM *P*-value
	ALMI model	BMW model	Ratio (ALMI/BMW)	
6 quarters	0.541	0.544	0.996	0.590
5 quarters	0.533	0.534	0.997	0.704
4 quarters	0.519	0.518	1.002	0.763
3 quarters	0.513	0.515	0.998	0.757
2 quarters	0.477	0.483	0.988	0.167
1 quarter				
No monthly variables	0.418	0.427	0.980	0.193
Before labor indicators	0.404	0.417	0.970	0.041
After labor indicators	0.394	0.408	0.964	0.040
All monthly variables	0.369	0.377	0.978	0.084
2/3 quarter				
Before labor indicators	0.350	0.356	0.983	0.191
After labor indicators	0.348	0.353	0.988	0.459
All monthly variables	0.332	0.346	0.960	0.037
1/3 quarter				
Before labor indicators	0.329	0.341	0.964	0.042
After labor indicators	0.315	0.340	0.926	0.003
All monthly variables	0.306	0.334	0.916	0.001

Notes: This table reports the mean absolute error (MAE) for forecasts and nowcasts of output growth from the benchmark 'alternative labor market indicators' (ALMI) and Berger *et al.* (2022) (BMW) models for forecast/nowcast horizons from 6 quarters ahead to within one month (1/3 quarter), 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 ratio of the two MAEs and *P*-values for Diebold–Mariano (DM) tests of equal predictive accuracy between the two models based on a lin-lin loss function are also reported.

each horizon, the ratio of the MAEs is also reported, along with *P*-values from Diebold–Mariano tests for equal predictive accuracy between the two models based on a 'lin-lin' loss function, as also considered when evaluating nowcasts in Berger *et al.* (2022) (Note: For the Diebold–Mariano tests, we consider the modified approach proposed in Harvey *et al.* (1997) that corrects the scale of the test statistic for multi-period horizons and uses a long-run variance estimator with a Bartlett kernel that sets autocovariances at the horizon and higher to zero. Results are robust to allowing for non-zero higher-order autocovariances when estimating the long-run variance and to consideration of root mean squared error under a quadratic loss function.) For in-sample fit, the ALMI model is significantly more accurate than the BMW model at most nowcasting horizons, 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-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,

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

Now/forecast horizon	Mean absolute error			DM p-value
	ALMI model	BMW model	Ratio (ALMI/BMW)	
6 quarters	0.447	0.441	1.015	0.419
5 quarters	0.455	0.448	1.016	0.397
4 quarters	0.454	0.445	1.022	0.240
3 quarters	0.460	0.447	1.029	0.075
2 quarters	0.460	0.447	1.028	0.023
1 quarter				
No monthly variables	0.419	0.408	1.027	0.136
Before labor indicators	0.413	0.412	1.002	0.930
After labor indicators	0.403	0.405	0.994	0.806
All monthly variables	0.393	0.393	1.001	0.948
2/3 quarter				
Before labor indicators	0.390	0.389	1.001	0.968
After labor indicators	0.390	0.380	1.024	0.349
All monthly variables	0.374	0.373	1.002	0.948
1/3 quarter				
Before labor indicators	0.386	0.387	0.998	0.938
After labor indicators	0.365	0.386	0.945	0.157
All monthly variables	0.364	0.384	0.948	0.197

Notes: This table reports the mean absolute errors (MAE) for forecasts and nowcasts of output growth from the benchmark 'alternative labor market indicators' (ALMI) and Berger *et al.* (2022) (BMW) models for forecast/nowcast horizons from 6 quarters ahead to within one month (1/3 quarter), with parameters re-estimated at each point in time, and the forecast/nowcast 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/nowcast evaluation is performed over the remainder of the sample. The shrinkage parameter is set at the optimized value for the whole sample as Morley & 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 ratio of the two MAEs and P-values for Diebold–Mariano (DM) tests of equal predictive accuracy between the two models based on a lin-lin loss function are also reported.

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 quarter				
No monthly variables	0.303	-	0.280	-
Before labor indicators	0.236	0.000	0.242	0.000
After labor indicators	0.209	0.011	0.191	0.000
All monthly variables	0.207	0.557	0.191	0.976
2/3 quarter				
Before labor indicators	0.117	0.000	0.139	0.000
After labor indicators	0.100	0.000	0.101	0.000
All monthly variables	0.094	0.038	0.094	0.009
1/3 quarter				
Before labor indicators.	0.050	0.000	0.066	0.000
After labor indicators	0.041	0.000	0.046	0.000
All monthly variables	0.014	0.000	0.022	0.000

Notes: This table reports the mean absolute error (MAE) for nowcasts of the output gap from the the benchmark 'alternative labor market indicators' (ALMI) and Berger *et al.* (2022) (BMW) models for nowcast horizons from one quarter ahead to within one month (1/3 quarter). 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. Diebold–Mariano (DM) P-values for a test of equal predictive accuracy given the previous information set (i.e. the respective row above) and the current information set are also reported.

the results at least seem to indicate that, as a forecasting model, the ALMI model is comparable to the BMW model.

Table 4 compares the nowcasting performance for the output gap between the two models. The table reports MAE for nowcasts from the beginning of a quarter to within one month, 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 equal predictive accuracy between the previous nowcast (in the row above) and the current nowcast. Results are similar for both models, with ALMI errors slightly larger at the beginning of a quarter, but smaller by the end of the quarter. (Note: In this case, the MAEs are not strictly comparable given different final estimates of the output gap for the two models. Thus, we do not report a ratio of MAEs.) 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

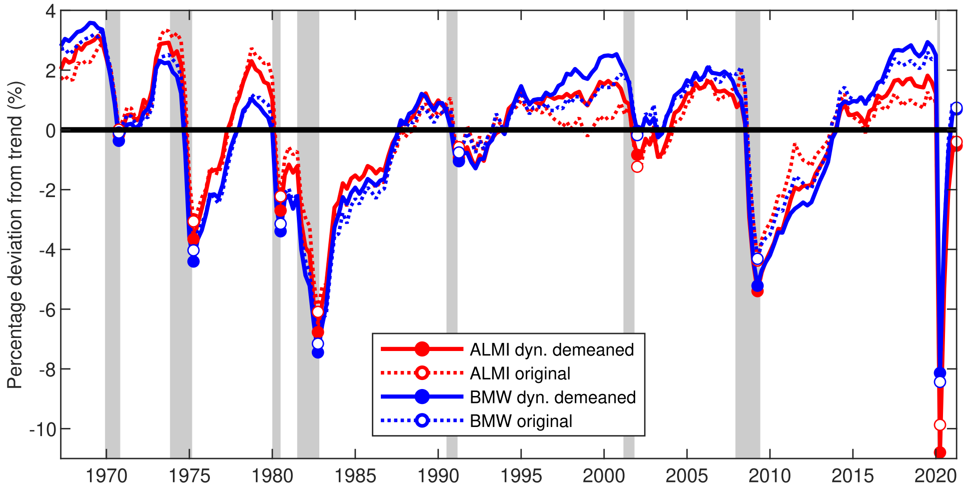

Figure 9. Dynamic demeaning and ALMI vs. BMW estimates of the output gap.
Notes: Sample period is 1967Q1 to 2021Q2. ALMI refers to the 'alterantive labor market indicators' model. BMW refers to the Berger *et al.* (2022) model. Shaded bars correspond to NBER-dated recessions.

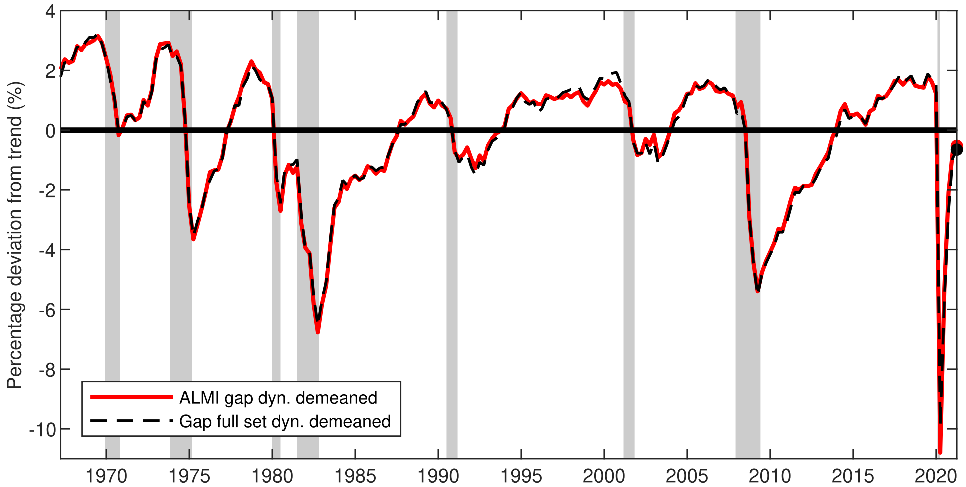

Figure 10. Dynamic demeaning and comparison of output gap estimates before and after variable selection.
Notes: Sample period is 1967Q1 to 2021Q2. ALMI refers to the 'alternative labor market indicators' model. Shaded bars correspond to NBER-dated recessions.

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.

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 & Wada (2009) critique, Berger *et al.* (2022) 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 10-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.

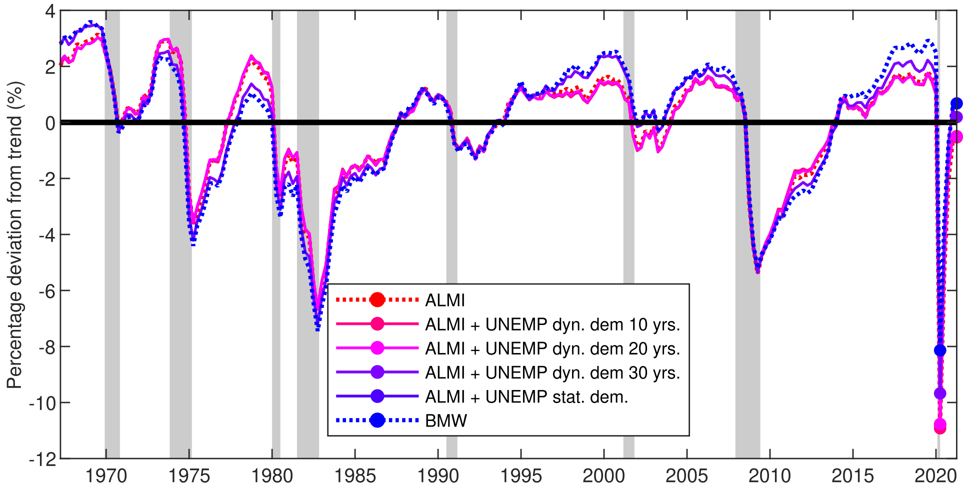

Figure 11. The output gap and dynamic demeaning of the unemployment rate.
Notes: This figure plots the estimated output gap for the benchmark 'alternative labor market indicators' (ALMI) model, four extensions of the ALMI model where the overall unemployment rate (UNEMP) is also included assuming a static (constant) mean and dynamically demeaned with windows of 10, 20, and 30 years, respectively, and the Berger *et al.* (2022) (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-dated 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 the unemployment rate 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 staticallydemeaned 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.

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 benchmark 'alternative labor market indicators' (ALMI) and Berger *et al.* (2022) (BMW) nowcasts avaialble 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.

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 in this subsection are obtained using all available monthly data up to and including September 2021 and quarterly real GDP until 2021Q2 (updated nowcasts are available at https://outputgapnow.com). 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,

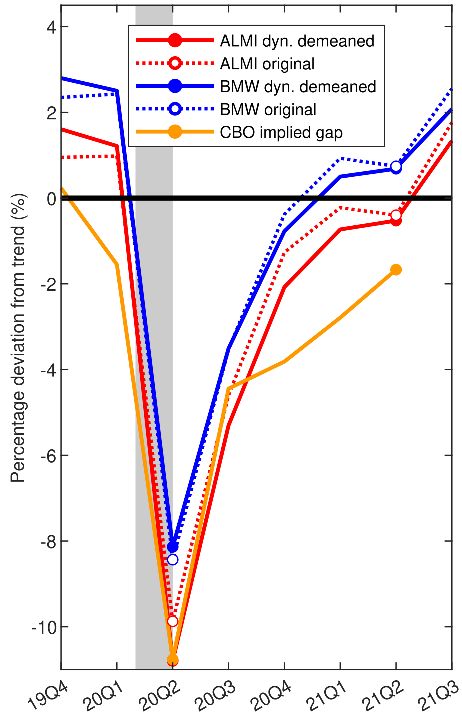

Figure 12. Estimated output gaps around the Covid pandemic.
Notes: The values for 2021Q3 are nowcasts constucted from the monthly data shown in Table 5. ALMI refers to the 'alternative labor market indicators' model. BMW refers to the Berger et al. (2021) model. CBO refers to the Congressional Budget Office. The shaded bar corresponds to the NBER-dated recession.

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 is consistent with the findings in the earlier subsection 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 US economy.

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 similar to Morley & 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 US 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 around 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.* (2022), with 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.

Acknowledgments

We thank the Editor, Ana Galvão, and two anonymous referees for helpful comments and suggestions. Nowcasts based on the methods and variables considered in this paper are automatically updated daily at https://outputgapnow.com and we thank Cristian Corrado for help with the webpage. This research was supported by Australian Research Council grants DP190100202 and DE200100693. The usual disclaimers apply.

References

Aastveit, K. A., and Trovik, T. (2014) 'Estimating the Output Gap in Real Time: A Factor Model Approach', *The Quarterly Review of Economics and Finance*, **54**: 180–93.

Altissimo, F., and Violante, G. L. (2001) 'The Non-Linear Dynamics of Output and Unemployment in the U.S.', *Journal of Applied Econometrics*, **16**: 461–86.

Barigozzi, M., and Luciani, M. (2022) 'Measuring the Output Gap Using Large Datasets', *The Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_01119.

Berger, T., and Vierke, H. (2017) 'Estimating the Natural Rate of Hours', *Macroeconomic Dynamics*, **21**: 1426–53.

Berger, T., Everaert, G., and Vierke, H. (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., Morley, J., and Wong, B. (2022) 'Nowcasting the Output Gap', *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2020.08.011.

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 Rodriguez-Palenzuela, D. (2003) 'Assessment Criteria for Output Gap Estimates', *Economic Modelling*, **20**: 529–62.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020) *Labor Markets During the COVID-19 Crisis: A Preliminary View*. NBER Working Paper No. 27017.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996) 'Efficient Tests for an Autoregressive Unit Root', *Econometrica*, **64**: 813–36.

Erceg, C. J., and Levin, A. T. (2014) 'Labor Force Participation and Monetary Policy in the Wake of the Great Recession', *Journal of Money, Credit and Banking*, **46**: 3–49.

Evans, G., and Reichlin, L. (1994) 'Information, Forecasts, and Measurement of the Business Cycle', *Journal of Monetary Economics*, **33**: 233–54.

Faberman, R. J., Mueller, A. I., Şahin, A., and Topa, G. (2020) 'The Shadow Margins of Labor Market Slack', *Journal of Money, Credit and Banking*, **52**: 355–91.

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**: 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 Gertler, M. (1999) 'Inflation Dynamics: A Structural Econometric Analysis', *Journal of Monetary Economics*, **44**: 195–222.

Hakkio, C., and Willis, J. (2014) 'Kansas City Fed's Labor Market Conditions Indicators (LMCI)', *Macro Bulletin*, 1–3.

Harvey, D., Leybourne, S., and Newbold, P. (1997) 'Testing the Equality of Prediction Mean Squared Errors', *International Journal of Forecasting*, **13**: 281–91.

Kamber, G., Morley, J., and Wong, B. (2018) 'Intuitive and Reliable Estimates of the Output Gap From a Beveridge–Nelson Filter', *The Review of Economics and Statistics*, **100**: 550–66.

Morley, J. C. (2002) 'A State–Space Approach to Calculating the Beveridge–Nelson Decomposition', *Economics Letters*, **75**: 123–7.

Morley, J., and Wong, B. (2020) 'Estimating and Accounting for the Output Gap With Large Bayesian Vector Autoregressions', *Journal of Applied Econometrics*, **35**: 1–18.

Perron, P., and Wada, T. (2009) 'Let's Take a Break: Trends and Cycles in US Real GDP', *Journal of Monetary Economics*, **56**: 749–65.

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**: 147–59.

US 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].

Waggoner, D. F., and Zha, T. (1999) 'Conditional Forecasts in Dynamic Multivariate Models', *Review of Economics and Statistics*, **81**: 639–51.

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: Available at https://www.federalreserve.gov/newsevents/speech/yellen20140822a.htm [Accessed: 27/05/2021]. Wyoming.

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

Appendix: Data and transformations

The table below reports the details of the data considered in our empirical analysis. The data were obtained from 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 ADF 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 A1. 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 US 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 · AWHMAN /40)	LNS12300060, AWHMAN	Δ	