FEDERAL RESERVE BANK *of* ATLANTA **WORKING PAPER SERIES**

Nowcasting with Mixed Frequency Principal Components When Not All Months Are the Same

Patrick Higgins

Working Paper 2026-15
September 2026

Abstract: This paper develops algorithms for estimating a principal components (PC) model for monthly macroeconomic data mixed with either weekly or daily data. The algorithms' utilization of high frequency data allows it to account for differences in the temporal aggregation patterns among the monthly series that are simple aggregates or averages and those that have an alternative aggregation pattern such as data from the monthly employment report where the aggregation pattern is often series- and month-specific and depends on the timing of the twelfth day of the month. The algorithms are applied to three respective models with mixed frequency price related, labor market related, or economic activity related data, the latter of which is similar to the factor model used by the Atlanta Fed's GDPNow nowcasting model to forecast GDP source data. When aggregated appropriately, the mixed frequency activity model PC closely resembles both the Chicago Fed National Activity Index and GDPNow's dynamic factor. Moreover, the PCs from all three models can be well approximated using the subset of data that would remain available during a widespread government shutdown.

JEL classification: E37, C38, C53

Key words: nowcasting, forecasting, macroeconometric forecasting, principal components, factor models

https://doi.org/10.29338/wp2026-15

The views expressed here are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Patrick Higgins is in the Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309-4470, patrick.higgins@atl.frb.org.

I. Introduction

In many macroeconomic forecasting applications – including the GDPNow model for nowcasting/tracking quarterly real GDP growth I designed (see Higgins [2014]) based on the work of others – most of the data come in and are used are at the monthly or quarterly frequency. This is partly because the data from the government source agencies for this and other models using macroeconomic data come in this form. However, with the proliferation of alternative data sources used to estimate or shed light on major macroeconomic aggregates that we've seen in the past 10-15 years and especially in the early months of the COVID-19 pandemic, there are more and more sources of higher than monthly frequency data – often but not always from non-federal government sources – that can be used to supplement the data inputs used by GDPNow and other models using macroeconomic data. As the presentations at the Fall 2025 Advancing Economic Measurement NBER conference illustrate[1], much headway along these lines has already been made. While this is partly a luxury resulting from the increased prevalence of previously untapped sources of data, such work may become more necessary given reduced response rates for many of the surveys underlying federal government macro data, lower funding rates for those surveys, and the increased prevalence of government shutdowns when data releases from federal agencies are unavailable.

Although updated GDPNow forecasts continued to be released during the 2018-2019 and 2025 government shutdowns, those updates were both less frequent and less informative[2]. In particular, the releases in those shutdowns included very little coverage of the GDP source data the model uses in it's so-called bridge equations used to relate that data to (often) granular[3] GDP subcomponents. The additional data beyond what GDPNow currently uses generally does include much in the way of alternative source data that could replace unreleased GDP source data, but it does include additional non-government data that are often classified as "soft" and can be used to help forecast the missing source data until it's released.

[1]See `https://www.nber.org/conferences/advancing-economic-measurement-fall-2025`.

[2]In the 2018-19 shutdown, Department of Labor and industrial production data continued to be released, partly offseting the lack of Census Bureau and BEA releases. But, in both over 1-month shutdowns, the GDP nowcasts during the shutdown differed by no more than 0.1 percentage point from their last readings heading into the shutdown.

[3]By granular, we essentially mean a subcomponent within a subcomponent. As a (very) granular example, residential investment in manufactured homes are "bridged" using a nowcast padded aggregate of the estimated "real" value of mobile home shipments based on nominal mobile home shipments data from the Census Bureau deflated by a PPI for mobile homes.

As we show below, we can track the estimated factors used to forecast GDP spending subcomponent source data and – in the case of the extension of GDPNow described in Higgins [2024] – forecast GDP subcomponent price deflator source data.

"Alternative" or non-government data often is – or can be – available at a weekly or data frequency. Our work below, which uses daily price data to estimate a (mixed-frequency) price principal component (PC) and weekly activity and labor market data to estimate PCs for each of those two categories, exploits this advantage. As the work from Knotek and Zaman [2017] – and the daily inflation nowcast updates from the Cleveland Fed derived from it – show, high frequency oil and gasoline price data are useful for nowcasting consumer energy prices that account for much of the volatility of headline CPI and PCE inflation on a monthly or quarterly basis. There are less clear benefits to doing the same thing for activity and labor market data – since unemployment claims and other weekly data used in work such as the Dallas Fed's Weekly Economic Index – are not so closely tied to the major quarterly and/or monthly output and labor market benchmark metrics such as real GDP growth or the unemployment. But as expert users of the Current Population Survey (CPS) and establishment survey data from the monthly labor report are well aware, those statistics are based on the employment status in the calendar week or pay period containing the 12th day of the month. This effectively weights statistics from those reports more towards the first and middle thirds of the month than they do for the price and spending GDP source data that are usually calculated or reported as simple aggregates or averages. "Soft" data, such as the regional Fed analogues to the national ISM Manufacturing Surveys or the consumer attitudes surveys from The Conference Board and the University of Michigan are often collected and released before the end of the month giving rise to an intra-month timing mismatch with conventional monthly price and spending data. Our use of daily and weekly data will allow us to largely correct for this mismatch for what we call "ISM style" surveys that we use as a blanket term to encompass any temporal aggregation pattern – such as the ones used for the household and establishment survey based-data in the monthly employment report – that differs from simple monthly averaging or aggregation.

The next section of this manuscript discusses the related literature on factor model and/or principal component estimation as it relates to forecasting, both with and without higher than monthly frequency data. Section III introduces the mixed monthly-daily frequency principal components/factor model for prices and Section IV does the same for the monthly-weekly frequency PC/factor models that are separately used for both the labor market and for economic activity more generally. Section V gives a short overview of the data used in the models followed by a discussion of the model results in Section VI. A brief conclusion is provided in Section VII.

II. Related Literature

The author's primary interest in the factors/principal components constructed here are for use as a forecasting tool. Stock and Watson [1999] and the pair of 2002 papers Stock and Watson [2002a] and Stock and Watson [2002b] showed the theoretical motivation and the empirical evidence for why principal components of a fairly large number of macroeconomic aggregates could be useful as forecasting tools for economic activity and inflation. In part based on this work, the Chicago Fed began regularly updating an index of economic activity in the early 2000s as the first principal component of (now) 85 time series all related to, and if necessary transformed to be compatible with via differencing or log differencing, GDP growth. By design, most of these 85 series were included in the dynamic factor model module of the GDP nowcasting model – which itself is largely based on work from Doz et al. [2011] and Giannone et al. [2008] – called GDPNow that was introduced in Higgins [2014] and has been regularly updated by the Atlanta Fed ever since.

In GDPNow, the dynamic factor is used to forecast monthly GDP source data that has not yet been released via factor augmented autoregressions (FAARs) much like Stock and Watson used in their pioneering research. But the factor could also be used to nowcast GDP growth directly as it was in Giannone et al. [2008]. The framework for estimating the dynamic factor model in this paper was largely adapted by the New York Fed in their initial version of The New York Fed Staff Nowcast model (see Bok et al. [2018]) that, much like GDPNow, also nowcasts real GDP growth though it has since evolved to incorporate stochastic volatility to better handle outliers induced by the early months of the COVID pandemic (see Almuzara et al. [2023]). The Giannone et al. [2008] study showed that as a particular GDP forecasting cycle evolved, one of the releases associated with the biggest improvements in average nowcast accuracy was the Federal Reserve Bank of Philadelphia's Business Outlook Survey (of manufacturers) since that release was the first one of the month for the stylized data release calendar they considered. The purported month t observation of the Philly Fed survey is typically the third Thursday of month t, and therefore more than a week before the end of that month[4]. Thus the survey month and activity period for the Philly

[4]Both GDPNow and the New York Fed Staff Nowcast model include the Philly Fed manufacturing survey variable based on the response comparing the "Current month" vs. the "Previous month" for the following

Fed Survey respondents won't exactly match up. Setting aside this inconsistency – which we will be able to accommodate in large part with our mixed frequency approach – the Philly Fed manufacturing survey variables play a less important role in GDPNow than they could in the Giannone et al. [2008] paper because, by construction, GDPNow does not include any month t values in its dynamic factor model until the month t ISM Manufacturing Report is released (usually) on the first business day of month $t+1$. This choice was made to keep the Philly Fed survey, as well as the consumer attitudes variables from The Conference Board and the University of Michigan surveys whose month t values are also generally released before the end of month t, from playing an overly influential role in forecast changes. But it has the opposite effect of increasing the importance of the ISM Manufacturing variables when their importance in the Kalman gain is magnified in the last observation of the data sample for the Kalman filter and smoother – by removing the corresponding rows of the selection matrices as described in Aruoba et al. [2008] – for most of the variables in the data matrix whose time t values are not yet released. This magnified importance was dampened to some extent by the April 2018 change to GDPNow's factor model that accounted for error persistence in the observation equations of the factor model effective with the model's first nowcast of 2018:Q2 GDP growth[5], but only partially.

The introduction of higher frequency daily or weekly variables allows for greater flexibility of the mothly aggregation pattern that account for differences between "monthly-average" style timing and "ISM style" timing. Most of the time the exact intramonth aggregation pattern of monthly variables is likely not an issue of large importance, except perhaps at the end of the sample when only a subset of the variables are available. But this is not always the case, as illustrated by the behavior of employment variables in March 2020. Because establishment survey employment variables cover the pay period including the 12th day of the month, those variables put more weight on employment and hours worked in

question: "What is your evaluation of the level of general business activity?" GDPNow also includes a variable based on the response comparing "Six months from now" vs. the "Current month". The sample collection period is the 8-day period ending on the Monday three days before the Philly Fed Index is published, and is thus administered in the early-middle to middle part of the month. Of course, the "Current month" vs. the "Previous month" comparison could embed a nowcast of what average activity level the panelists anticipate for the remainder of the month. Even in that case though, that expectation wouldn't incorporate unanticipated but ultimately realized activity levels.

[5]See `https://www.atlantafed.org/-/media/Project/Atlanta/FRBA/Documents/cqer/researchcq/gdpnow/ModificationsToGDPNowModel.pdf`.

the beginning and middle of the month than they do at the end of it and therefore the March 2020 payroll employment and hours were stronger than they would have been had they been equally weighted across the entirity of March 2020[6]. In order to not understate productivity growth in 2020:Q1 because economic activity appeared to be much weaker in March 2020 than employment due to the timing mismatches in the source data, the BLS adjusted their normal procedure for estimating March 2020 employment and hours worked in their productivity database measures by using weekly unemployment insurance claims data to estimate the weekly pattern of employment and hours worked within the month[7].

Figure 1 plots the dynamic factor estimated by the standard published Atlanta Fed GDPNow model on August 7, 2026 against the first principal component (PC) of the (unbalanced) data that is used to estimate many of the parameters used by the model's Kalman filter and smoother. We can see the two series are very highly correlated, with correlation $\rho = 0.990$. In the plot, the two standardized series have been truncated three standard deviations above and below their means so that their April 2020 minimum values just under -20 and the May 2020 maximum values over 6 are outside the plot limits. But the correlation of the two series drops only modestly to $\rho = 0.978$ if the March 2020 through June 2020 are removed from the correlation calculation.

In the initial stages of this research project, we embedded mixed monthly and weekly frequency activity data into a multi-step principal component followed by Kalman filter/smoother framework that is very similar to the approach originated by Doz et al. [2011] and used by GDPNow as documented in Higgins [2014]. In this preliminary work, we found such an approach with mixed monthly and weekly data to be feasible, but computationally demanding both from a coding and a run-time perspective. A less computationally intensive approach integrating principal components into a factor model is developed in Bräuning and Koopman [2014]. Their approach, which they note and demonstrate is general enough to handle mixed frequency data, embeds one or more of the first PCs of the data into a smaller state space model designed to forecast a subset of the data series rather than the entire dataset used to estimate the PCs, thereby "reduc[ing] the observation and parameter dimensions

[6]In 2020 there were 208k initial unemployment insurance claims in the week ending March 7th, 273k in the week ending March 14th, 2.914 million in the week ending March 21st, 5.946 million in the week ending March 28th, and 6.137 million in the week ending April 4th.

[7]See Bureau of Labor Statistics [2020] and `https://www.bls.gov/bls/productivity-and-costs-adjusting-march-2020-employment-and-hours-to-account-for-effects-o` `htm`.

substantially, with minimal information loss". Because there are many monthly GDP source series that need to be forecast, this advantage may be less applicable to GDPNow and the extension of it to forecasting the price deflators of GDP and its subcomponents in Higgins [2024]. However, the very high correlation between GDPNow's first PC and dynamic factor evident in Figure 1 suggests we may be able to stop after the first PC step of the Doz et al. [2011] algorithm and, like the earlier Stock and Watson efforts, use the PC(s) to directly forecast GDP source data without having to implement the Kalman filter or smoother. The issue with this approach is that the first PC estimate in Figure 1 ends in May 2026 – the last month of the sample where no observations are missing – while the smoothed factor ends in July 2026. This difference is not problematic for GDPNow's implementation of the Doz et al. [2011] algorithm because the OLS regression estimates used to parameterize the Kalman filter/smoother matrices do not change much with the addition of two additional months of data. But this convenience of being able to drop the last one or few observations of most series in the sample won't be implementable since we want to use the PCs as forecast tools. Instead, we will have to "fill in" missing observations in the mixed frequency data at end of the sample with an approach similar to the EM algorithm methodology developed by Stock and Watson [2002b] that GDPNow uses to "fill in" missing observations at the beginning of the sample for those series which begin before 1967.

In terms of its empirical byproducts, the studies producing content most similar to this one are Aruoba et al. [2009], which introduces what is now known as the Philadelphia Fed's Aruoba-Diebold-Scotti (ADS) Business Conditions Index[8] and the pair of studies Lewis et al. [2020] and Lewis et al. [2021] which document an index of weekly activity whose descendant is now hosted as the Dallas Fed's Weekly Economic Index (WEI)[9]. The ADS index is a mixed frequency state space model estimated with weekly initial unemployment insurance claims, (quarterly) real GDP and four monthly indicators that have historically been used by the NBER to identify business cycle expansions and contractions. Because of the small number of variables, the ADS model can be estimated at the daily frequency using fairly standard MLE techniques. That approach would be less well suited to our environments where between 50 and 175 time series of activity, labor, or price variables are included in the estimation. Like our measures, the Dallas Fed's WEI incorporates multiple high frequency indicator variables

[8]The ADS series, and related documentation on methodological updates and other material is available at `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/ads`.

[9]See `https://www.dallasfed.org/research/wei`.

– in their case 10 weekly activity variables – and also like our activity and labor market factors, their measures are estimated with the first PC of the high frequency series. Because the Dallas Fed measure does not incorporate any lower frequency variables, there is no need to interpolate lower frequency variables for inclusion in the principal component calculation. As documented in Lewis et al. [2020], after each Thursday initial claims report the WEI has only 8 of their 10 weekly variables available for the most for the most recent full week[10]. They handle the 2 missing values by directly forecasting the WEI with a regression equation using a constant, two lagged values of the WEI, and the contemporaneous full week values of the 8 weekly variables that are available at the time of the Thursday update[11]. This nowcasting solution for the "ragged" or "jagged" edge of the end of the data sample where only a proper subset of the variables is available is not easily implementable in our activity or labor based settings since weeks do not evenly divide into months. Also, unlike the WEI related studies, we use seasonally adjusted high frequency data rather than 52 week changes or 364-day changes in the case of daily prices.

One possible advantage of the WEI is its use of primarily non-government data sources. A pre-cursor to the WEI, also referred to at the time as the Weekly Economic Index and constructed in a very similar manner and with a number of the same series, was constructed and released by the Council of Economic Advisors (CEA) in October 2013 using non-government data indicators to assess how economic activity was impacted by the government shutdown in that month that preempted a number of economic releases[12]. Since then, there have been even longer government shutdowns in December 2018-January 2019 and October 2025-November 2025 with even larger disruptions to government provided data. Only 3 of the 10 series in the current iteration of the WEI are from or derived from federal government agency data releases – federal withholding tax collections and both initial and continuing unemployment insurance claims – but the latter two of these could nearly be replicated during the 2025 government shutdown using state level unemployment insurance claims data released

[10]The unavailable values are continuing claims and the ASA Staffing Index from the American Staffing Association.

[11]The Lewis et al. [2020] study also notes an even earlier Tuesday update using the 3 of the 10 values that are available at the time of that update using a similarly structured regression projection with 2 lags of the WEI and 3 instead of 8 contemporaneous values that are available. But this update is not include in the regular posting schedule for the Dallas Fed WEI.

[12]James Stock, one of the originators of the Dallas Fed's WEI was on the CEA in October 2013.

by state agencies (see, for example, Gould and Fast [2025][13]). And though the majority of the monthly indicators used to construct the standard (principal component based) Kansas City Fed Labor Market Conditions Indicators[14] are sourced from the federal government, the KC Fed was able to construct September 2025 and October 2025 values of the index during the shutdown primarily using the remaining indicators sourced by non-government organizations (see Mustre-del Río et al. [2025a] and Mustre-del Río et al. [2025b]). The indices we construct here all utilize enough data from non federal-government sources to allow us to create "shutdown insulated" variants and compare them to the complete versions of the indices.

Although more narrowly focused on consumer spending rather than economic activity more broadly as the WEI and the analog constructed here are, the approach used in the pair of studies Brave et al. [2021] and Brave et al. [2024] is similar to ours and differs from that used in constructing the WEI both in its use of seasonally adjusted high frequency data and in mixing monthly with higher frequency data. In particular, the working paper Brave et al. [2021] documents how the Chicago Fed CARTS[15] program is able to forecast the Census Bureau measure of retail sales using higher frequency "weekly" series that are seasonally adjusted and aggregated to be consistent with 4-week months – where the last "week" of the month covers anywhere between 7 and 10 days – that can then be tied to monthly frequency variables within a state-space system using what are now standard cumulator aggregation methods from Harvey [1990] and Mariano and Murasawa [2003] among others. As the Brave et al. [2021] study details, the "standard" seasonal adjustment programs for weekly data that we utilize have to be tweaked in the CARTS environment due to the short (less than 10 year) length of time series they work with. The "weekly" data – which for most of the Chicago Fed series can be and are built up from daily data – to monthly data conversion approach the Chicago Fed uses could in principle also be used here. But, in part because of our inclusion of what we call "ISM-style" monthly variables, which unlike retail sales are not based on simple monthly aggregates, we use a less clean but more flexible weekly to monthly aggregation approach than the CARTS approach uses. Because we end our estimation step after (iterative) computation of high frequency principal components without using the additional steps of the Kalman filter and smoother, we need to embed a nowcasting step that

[13]Available at https://www.epi.org/blog/amid-the-shutdown-data-blackout-state-unemployment-i

[14]See Hakkio and Willis [2014].

[15]CARTS stands for Chicago Fed Advance Retail Trade Summary.

handles the staggered nature of the monthly and higher frequency data series. In the Chicago Fed's National Activity Index principal component based calculation, this has primarily been handeled via forecast padding of series with delayed release dates with an autoregressive projection rather than a projection based on the EM algorithm outlined in Stock and Watson [2002b][16]. As we describe below, we essentially handle the stagerring of data releases via a combination of autogressive and EM algorithm based forecast augmentation.

III. Daily factor model for prices

Before describing our approach for estimating principal components with mixed daily and monthly frequency price variables and using them to nowcast these prices, we provide a few approximate aggregation identities we will utilize.

Let $x^D_{i,m,d}$ denote the first difference of the log of an observed daily price variable indexed by i – such as WTI oil – on the dth business day of month m, with the notational convention that $x^D_{i,m,1}$ is the change in the logged price between the first business day of month m and the last business day of month $m-1$. The lowercase x is used to denote a daily observation. Moreover, the log of the daily values have been seasonally adjusted using the daily seasonal adjustment package dsa in R that is described in Ollech [2018] and maintained by Daniel Ollech. The 8 daily price variables used in this study are: 1) Brent and 2) WTI oil prices, the S&P Goldman Sachs indices of 3) overall, 4) agricultural, 5) industrial metals, and 6) livestock commodity commodity prices, and finally the Foundation for International Business and Economic Research (FIBER) indices of 7) overall and 8) core industrial materials prices. The price variables are all available since June 2, 1986, where we start the daily portion of our price sample. However, the Ollech [2018] daily seasonal adjustment routine requires a continuous interval of non-missing values even on weekends and holidays. To satisfy this mathematical requirement, we linearly interpolate the missing daily data and seasonally adjust the daily series that are each augmented with interpolated values. We then remove the weekend observations and common holidays where all of the daily variables have missing values such as December 25, 2025. However, we retain the holiday observations where only a proper subset of the daily variables have missing values such as the Martin Luther King Jr. holiday observed on January 19, 2026 when the Brent oil price was the only daily series with a non-missing value.

[16]See `https://www.chicagofed.org/-/media/publications/cfnai/background/cfnai-background-pdf.pdf`. For the single month of April 2020, the CFNAI used the EM algorithm based forecasts.

We denote the change in the log of the monthly average of this daily price from month $m-1$ to month m as $X^D_{i,m}$, with the use of a capital letter X meant to denote a monthly observation and the D superscript used to denote the highest observed frequency of the variable (so that it could be the price of oil, but not a CPI subcomponent for instance). Suppose there are $B(m)$ business days in month m and $B(m-1)$ business days in the prior month. Then, we can use the following approximation:

$$X^D_{i,m} \approx \sum_{d=2}^{B(m-1)} \frac{(d-1)x^D_{i,m-1,d}}{B(m-1)} + \sum_{d=1}^{B(m)} \frac{(B(m)+1-d)x^D_{i,m,d}}{B(m)} \tag{1}$$

In the case that $B(m-1) = B(m)$, the weights for the business days in the two months in sequential order are $\{\frac{1}{B(m)}, \frac{2}{B(m)}, ..., \frac{B(m)}{B(m)}, ..., \frac{2}{B(m)}, \frac{1}{B(m)}\}$. Of course $B(m) > 3$, but approximation 1 would work equally well with $B(m) = 3$ when it would be the same approximation of a quarterly (log) growth rate as the weighted sum of (log) monthly growth rate rates used in a number of state-space model applications such as the New York Fed Staff Nowcast Model (see Almuzara et al. [2023]).

For a monthly price variable indexed by j with a native frequency that is monthly such as a CPI subcomponent, we denote the difference of the monthly logged prices as $X^M_{j,m,d}$. Of course, with such a variable we don't observe the daily inflation rate $x^M_{j,m,d}$, but we treat this variable as latent and use the following aggregation identity similar to approximation (1):

$$X^M_{j,m} = \sum_{d=2}^{B(m-1)} \frac{(d-1)x^M_{j,m-1,d}}{B(m-1)} + \sum_{d=1}^{B(m)} \frac{(B(m)+1-d)x^M_{j,m,d}}{B(m)} \tag{2}$$

A possible shortcoming of this approach is that it does not account for the impact of transactions on weekends or holidays, which we know often take place. Since commodities markets are generally closed on these days, we acknowledge but ignore this issue and don't opt for an alternative such as interpolating a Friday to Monday price change to encompass Saturday and Sunday price changes[17].

Finally, we have price variables from the ISM manufacturing and services survey and both of their analogs from surveys separately administered by each of the Dallas, Kansas City, New York, Philadelphia, and Richmond Fed Banks[18]. The regional Fed surveys all close

[17]Such an interpolation strategy would be more palatable if Friday to Monday commodity price changes were substantially more volatile than they were for the last four days of the working week, but we found this not to be the case.

[18]Each of the manufacturing and services surveys from the 5 regional Feds has to price measures: one for prices paid and one for prices received. We use the latter in each case since we are interested in forecast

before the last business day of the month and hence, equation (1) clearly will not hold for the variables from these surveys. But we can easily generalize this equation as

$$X^I_{j,m} = \sum_{h=-2}^{0} \sum_{d=1}^{B(m+h)} a^I_{j,m-h,d} x^I_{j,m-h,d} \tag{3}$$

where the I superscript is meant to denote an ISM style index, and, as we discuss below, the $a^I_{j,m-h,d}$ coefficients are nonnegative, sum to 1.0 and, in many cases, with a significant number of them exactly equal to 0.

We illustrate our approach with the Dallas Fed's Texas Manufacturing Outlook Survey, which is generally released on the fourth Monday of the month if it's not a holiday and the day after that if it is. For Monday release dates, survey responses are collected over the ten day period ending on the Wednesday before the Monday release[19]. We assume that responses are submitted with equal response shares of the total for each of the business days within the ten day window except for the two end-dates which are assumed to have half the response shares as each of the other days inbetween[20]. Respondents are asked other than the normal seasonal change whether prices received for finished goods have increased, decreased, or are not changed in the current month relative to the prior month. For each of the collection dates, we construct the day weights so that respondents are effectively comparing the average price over the seven days ending on the response date with the average price over seven days ending on the same numerical day of the month in the prior month[21]. The weights for the non-business days within each collection date's comparison window are discarded and we repeat this weight assignment and cumulate them for all of the up to 8 business days within the collection window. Figure 2 plots the Texas Manufacturing Outlook Survey business day weights for each of the 12 months of 2020. We see that business days late in month $m-1$ or early in month m have a positive weight for the month m survey. But business days near the middle of month m will have a positive weight both for the month m and the month $m+1$ surveys.

consumer and other final goods and services prices. The two national ISM surveys only have a prices paid index both of which we also use.

[19]For the occasional Tuesday release, the collection period also has the same ten day period and timing.

[20]For periods without any holidays, the opening Monday and closing Wednesday of the survey collection period both account for $\frac{1}{14}$th of the response shares and the seven business days in-between those dates each account for $\frac{1}{7}$th of the response share.

[21]This implies that for each collection date, the day weights within the comparison window are proportional to the sequence $\frac{1}{7}, \frac{2}{7}, ..., \frac{6}{7}, \frac{7}{7}, ..., \frac{7}{7}, \frac{6}{7}, ..., \frac{2}{7}, \frac{1}{7}$.

In the Dallas Fed survey, the business day weights for the month two months before the survey month are all 0 – i.e. $a^I_{j,m-2,d} = 0$ – but for the New York and Richmond survey where the collection period starts near the beginning of the month, we will have $a^I_{j,m-2,d} > 0$ for $d = B(m-2)$ and a few of the business days before $d = B(m-2)$. Another noteworthy feature of Figure 2 is the peak value of the monthly weights, which can vary a bit from month to month. The height of these peaks is a function of the length of the survey collection period. The differences in peaks are even more pronounced for the Richmond Fed Manufacturing Survey which will have smaller peaks for surveys ending 5 weeks after the previous survey [e.g. January, April, July, and October 2020] than it will for surveys ending 4 weeks after the previous survey [e.g. the other 8 months in 2020]. This is similar to the fluctuating 4-week vs. 5-week interval between reference periods that the Bureau of Labor Statistics adjusts for when seasonally adjusting establishment survey data (see Cano et al. [1996]), though we do not make any adjustments beyond modifying the values of the $a^I_{j,m-h,d}$ coefficients such as adjusting the variance of the error term.

With the 8 daily price variables and $x^D_{i,m,d}$ and the 50 monthly price variables $X^M_{j,m}$ and $X^I_{j,m}$, many of which coincide those used to estimate the monthly price factor model used in Higgins [2024] to forecast the source data for the subcomponents of the GDP deflator, our algorithm for estimating the daily principal components of these 58 variables is the following:

Step 1: Estimate the first two principal components $\hat{f}^1_{m,d}$ and $\hat{f}^2_{m,d}$ of the 8 standardized daily series $z^D_{i,m,d}$ all of which span the entire daily sample.

Step 2: Forecast the estimated daily factors a predetermined number of months (likely somewhere between two and six) beyond the end of the daily data sample using a 60-lag bi-variate VAR with no constant term and the two preliminary daily factors. Then iteratively forecast each of the 8 standardized daily price series $z^D_{i,m,d}$ using a factor augmented autoregression (FAAR) out to that same end-date using only the contemporaneous value of the forecasted factor and the autoregressive lags of each variable. For each daily price series, the number of autoregressive lags is determined by the Akaike Information Criterion (AIC). The forecast extended standardized series are denoted $\hat{z}^D_{i,m,d}$.

Step 3: Aggregate the VAR extended forecasts of the daily factors to the monthly frequency using equation (2) thereby creating the monthly factors $\hat{F}^{M,1}_m$ and $\hat{F}^{M,2}_m$ from the aggregated daily factors. Use these estimated factors to forecast the 38 non-ISM style standardized price variables $Z^M_{j,m}$ out to the same date where the extended daily variables end using factor augmented autoregressions (FAAR). The maximum lag length k of the factors and the number of autoregressive lags p is again chosen by the AIC, with $0 \leq k \leq 3$ and $0 \leq p \leq 6$. $k = 0$ corresponds to only the contemporaneous factor values being used. Some of the variables start after the estimation sample begins in 1986 such as the Manheim Used Vehicle Value Index which begins in early 1997 and thus needs to be "completed" for the initial 10+ years of the sample where it is missing. Although the adapted Fernández [1981] interpolation procedure that we use and is described in the appendix can fill in the estimated initial values for the monthly price variables that begin after the 1986 starting period, it doesn't always do so in a plausible way, particularly for those variables with little persistence. Therefore we backcast $Z^M_{j,m}$ to the beginning of the sample period when necessary by estimating a "reverse-FAAR" with p leads of $Z^M_{j,m}$, where $0 \leq p \leq 6$, and the contemporaneous and first k leads of the factor where $0 \leq k \leq 3$. Once estimated, the initial values of $Z^M_{j,m}$ and the factor values at and before that time are iteratively used to backcast the series back to the beginning of the sample. The forecast and backcast extended monthly price variables $\hat{Z}^M_{j,m}$ are then interpolated to a daily frequency variable $\hat{z}^M_{j,m,d}$ using the daily factors $\hat{f}^1_{m,d}$ and $\hat{f}^2_{m,d}$ using the adapted Fernández [1981] interpolation procedure.

Step 4: For each of the 12-ISM style price variables, aggregate the forecast extended daily price factors $\hat{f}^1_{m,d}$ and $\hat{f}^2_{m,d}$ to the monthly frequency variables $\hat{F}^{I,1}_m$ and $\hat{F}^{I,2}_m$ consistent with it's temporal aggregation pattern using equation (3). Similar to **Step 3**, we forecast and backcast $Z^I_{j,m}$ using FAARs with the same permissible lag lengths for the factors and autoregressive lags as in that step. Again as in **Step 3**, once extended the price variable $\hat{Z}^I_{j,m}$ is interpolated to be the daily frequency price variable $\hat{z}^I_{j,m,d}$ using the adapted Fernández [1981] interpolation procedure.

Step 5: Collecting the 8-daily price variables $\hat{z}^D_{i,m,d}$ from **Step 2**, the 38 daily frequency non-ISM style interpolated monthly price variables $\hat{z}^M_{j,m,d}$, and the 12 daily frequency ISM style interpolated monthly price variables $\hat{z}^I_{j,m,d}$ into a 58 column matrix $\mathbf{Z}^D$, we aggregate the series into a 58 column matrix of monthly price variables $\mathbf{Z}^M$ by applying equation (2) to each of the columns of $\mathbf{Z}^D$. $\mathbf{Z}^M$ will contain closely approximated log differences of monthly averages, which would correspond to the same measurement units that would be used for a standard principal components calculation with the monthly aggregated daily and non-ISM style price variables. However, the original temporal aggregation of the "ISM style" price variables are all converted to the "monthly average" equation (2) style aggregation used for CPI style variables and thus will not be exactly consistent with their original values due to the change made in their temporal aggregation weights. The variables in $\mathbf{Z}^M$ are all standardized over the month ending in the last month of the sample where all the price variables are still available, and the first two principal components of this matrix of monthly variables over this period is calculated with its 58 $\times$ 2 matrix of associated factor weights $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2]$. Standardizing $\mathbf{F}^D = \mathbf{Z}^D\mathbf{V}$ produces a new two-column matrix of our daily nowcast extended principal component factors accounting for all of the 58 price variables, and not just the 8 original daily commodity price series.

Step 7: Repeat **Step 6** and **Steps 3-5**, in that order where **Step 6** is the following:

Step 6: Iteratively forecast each of the 8 standardized daily commodity price series using the extended standardized factors $\mathbf{F}^D$ using the same FAAR specifications as in **Step 2** except for removing the factor terms from the forecasting equations used to project the daily data[22].

The **Step 7** daily factors used in the repeatitions of **Steps 3-5** are the last ones calculated as $\mathbf{F}^D = \mathbf{Z}^D\mathbf{V}$.

A natural question is why **Step 5** aggregates the daily price variables, most of which are interpolated values, to the monthly frequency before computing the principal component factor weights? Because the 8 daily price variables are all commodity prices, in general they are more highly correlated with the second principal component "supply shock" factor than the first principal component "inflation trend" or "underlying inflation" factor[23]. Because of this, their (first 2) principal components will also be better interpolaters of monthly variables that are highly correlated with supply shocks (such as the import price for petroleum and petroleum products) than they will be for monthly variables that are highly correlated with underlying inflation (such as the Cleveland Fed and Dallas Fed measures of trimmed mean CPI and PCE inflation). Computing the principal components on the interpolated daily data would overweight the predominantly supply shock related variables relative to the original monthly factor model for the data available or entered in at that frequency. Since we are predominantly interested in forecasting monthly price variables, the monthly aggregation PCA calculation better suits our purposes.

[22]Because the daily data end after the coverage period of the monthly data, the factors would be pure forecasts in these FAARs. We use the presumption that commodity price changes are unpredictable, and the omission of the factors improves the convergence properties and speed of the algorithm.

[23]This will be seen in VI.

IV. The weekly factor models

We have two weekly based PCA/factor models that we estimate and assess in this treatment: 1.) An economic activity model that is an offshoot of the dynamic factor model used in the Atlanta Fed's GDPNow model (see Higgins [2014]) to forecast GDP source data, and 2.) A labor market PCA/factor model with only weekly and monthly labor market variables that bears some similarity to the labor market momentum measure of the Kansas City Fed's Labor Market Conditions Indicators (see Hakkio and Willis [2014]). While the approach outlined below is similar in many respects to the one used for estimating principal components for daily and monthly price measures and indicators, there are some important differences. The most obvious ones are inherent to using a weekly period rather than a daily one as the highest frequency used and in estimating only the first principal component in both the activity and labor market environments rather than the first two PCs estimated with price data[24]. A subtler difference is allowing for a break in the first half of 2020 with both the activity and labor market indicators that was not needed or used with the price data, because of the more significant impacts of the early months of the pandemic on output and employment than on prices.

Using a similar notation as was used in the prices section above, let $x^M_{i,m,w}$ denote a latent weekly observation of a stationary monthly economic activity or labor market variable $X^M_{i,m}$ that is unobserved, where i indexes the variable, m the month, and w the week with $w = 1$ corresponding to the week beginning on the first Sunday of the month, $w = 2$ corresponding to the week beginning on the second Sunday of the month and so on so that $w \leq 5$. We are ultimately interested in best approximating monthly aggregates $X^M_{i,m}$ with weekly aggregates based on the approximate aggregation

$$X^M_{i,m} \approx \sum_{t=1}^{T} \sum_{w=1}^{W(t)} a_{i,m,t,w} x^M_{i,t,w} \tag{4}$$

where $W(t)$ is the number of Sundays in month t and $a_{i,m,t,w}$ is the nonnegative weight that week w of month t has on the month m growth rate. We use the notational conventions that $x^M_{i,t,W(t)+1} = x^M_{i,t+1,1}$ and $a_{i,m,t,W(t)+1} = a_{i,m,t+1,1}$ since the Sunday after the wth Sunday

[24]We utilize two principal components with the price data – as was done in Higgins [2024] – because in the partial monthly price data through July 2026 the second principal component has a relatively high explanatory power – accounting for 11.6% of the variance in the EM completed standardized price data compared to 31.8% for the first PC and 5.9% for the third PC – and because, as we will see in VI, the second PC accounts for much of the variability in the daily price data when aggregated to the monthly frequency.

of month t may be either the $w+1$st Sunday of month t or the first Sunday of month $t+1$. Why do the coefficients $a_{i,m,t,w}$ have two month subscripts (m and t)? Because $x^M_{i,t,1}$ will usually have positive weights associated with it for both $X^M_{i,t}$ and $X^M_{i,t+1}$. This is analogous to a March monthly growth rate having a positive weight for a quarterly growth in both the first and second quarters of its year.

Our goal is to approximate the (standardized) monthly growth rate as the (standardized) weighted sum of weekly growth rates. For illustrative purposes, we use the difference of the log of initial weekly unemployment insurance claims since that variable is included in both the activity and the labor market factor models. In the notation of the previous section, this weekly growth rate in claims would be denoted $x^W_{i,t,w}$ since claims are observed weekly, but here we use $x^M_{i,t,w}$ since this aggregation would also work for variables only available at the monthly frequency such as industrial production growth.

Letting d denote the dth day of week w with Sundays corresponding to $d=1$, we approximate a weekly growth rate as the weighted sum of daily growth rates via

$$x^M_{i,t,w} = \sum_{h=w-1}^{w} \sum_{d=1}^{7} \alpha_{h-(w-1),d}\, y^M_{i,t,w,d} \tag{5}$$

where $y^M_{i,t,w,d}$ is a latent daily growth rate and $\alpha_{0,d}$ and $\alpha_{1,d}$ are weighting coefficients. In the labor model, we take these coefficients to be $\alpha_{0,d} = \frac{d-1}{7}$ and $\alpha_{1,d} = \frac{8-d}{7}$ using an analog to equation 1 and the presumption that persons can file an unemployment insurance claim on any day of the week[25]. In the activity model, the weekly variable is often a measure of economic activity which we assume occurs on weekdays. Thus, for the activity model we assume $x^M_{i,t,w}$ measures the log change in the aggregates for consecutive workweeks and is consistent with the coefficients $\alpha_{0,d} = \frac{d-2}{5}$ for $3 \le d \le 7$, $\alpha_{1,1} = 1$, $\alpha_{1,d} = \frac{7-d}{5}$ for $2 \le d \le 6$ and $\alpha_{0,1} = \alpha_{0,2} = \alpha_{1,7} = 0$.

Similar to equation 2, we use a daily to monthly aggregation conversion which in this instance is

$$X^M_{j,m} = \sum_{d=2}^{D(m-1)} \frac{(d-1)x^M_{j,m-1,d}}{D(m-1)} + \sum_{d=1}^{D(m)} \frac{(D(m)+1-d)x^M_{j,m,d}}{D(m)}, \tag{6}$$

where $D(t)$ is the number of calendar days in month t.

Given all of these preliminaries, our algorithm for populating the weekly weights $a_{i,m,t,w}$ in equation 4 associated with monthly average growth rate $X^M_{i,m}$, is to first initialize all of

[25]Weekly initial unemployment insurance claims data are aggregated over weeks ending on a Saturday.

them to 0. Then, going calendar day by calendar day from the first calendar day of the daily sample to the last, determine what day d of the week the day is, how many Sundays w there have been in that day's month so far up to and including the day, and what month m the day is in. After determining these, we recursively increase the following four weights via

$$\begin{aligned} a_{i,m,m,w} &= a_{i,m,m,w} + \frac{\alpha_{1,d}(D(m)+1-d)}{D(m)}, \\ a_{i,m,m,w+1} &= a_{i,m,m,w+1} + \frac{\alpha_{0,d}(D(m)+1-d)}{D(m)}, \\ a_{i,m,m+1,w} &= a_{i,m,m+1,w} + \frac{\alpha_{1,d}(d-1)}{D(m)}, \\ a_{i,m,m+1,w+1} &= a_{i,m,m+1,w+1} + \frac{\alpha_{0,d}(d-1)}{D(m)}. \end{aligned} \tag{7}$$

After iterating through all the days of the sample and recursively computing the sums, we insert the $a_{i,m,t,w}$ entries into a matrix $\mathbf{A}_i$ where each row corresponds to a particular month m and the columns run through all of the combinations of t and w. The ratio of the number of columns to the number of rows in the matrix $\mathbf{A}_i$ is very close to the number of weeks in a month, which is $\frac{365.25}{7*12} \approx 4.35$. We then divide each of the rows of $\mathbf{A}_i$ by the sum of their entries so that the monthly growth rates are approximated as a weighted average of the weekly growth rates. The mth row of $\mathbf{A}_i$ corresponds to the monthly growth rate on the left hand side of equation 4 while the entries in the mth row identify the coefficients on the right hand side of that equation. Since principal components operate on standardized data, we don't multiply $\mathbf{A}_i$ by $\frac{365.25}{7*12} \approx 4.35$ which we would need to do if we wanted to relate non-standardized weekly growth rates to non-standardized monthly growth rates.

This aggregation scheme with $\mathbf{A}_i$ derived as above works with monthly variables that are entered as log differences such as, say, industrial production, but not for "ISM style" variables which we also worked with Section III with prices. Just as in that section though, for ISM style variables we can simply replace equation 4 with

$$X^I_{j,m} = \sum_{h=-2}^{0} \sum_{d=1}^{D(m+h)} b^I_{j,m-h,d} x^I_{j,m-h,d} \tag{8}$$

where the above equation is analogous to equation 3 except the sum is over all calendar days of each month rather than just the business days and that the two months m and $m-1$ with non-negative weights in equation 4 now includes a third month $m-2$ with, possibly, non-negative weights. The coefficients $b^I_{j,m-h,d}$ then are put in place of $\frac{(D(m)+1-d)}{D(m)}$ and $\frac{(d-1)}{D(m)}$ in a sequence of recursively iterated equations analogous to the four equations in 7 with the addition of two additional equations corresponding to month $m-2$. For the ISM style variables, the coefficients in equation 4 are also collected in a matrix $\mathbf{A}_i$ specific to ISM-style variable i.

For the labor model, the employment sub-indices are used in place of the price sub-indices in both the national surveys and the 5 regional Fed surveys of both manufacturing and services and, when available[26], the average workweek subindex is used as well. In both the activity and labor models, "ISM style" weighting is also used for the variables from the Current Population Survey (CPS), the Current Employment Statistics (CES) establishment or payroll survey of employment, the University of Michigan's Survey of Consumers, and The Conference Board's Consumer Confidence Survey. The weighting for the payroll survey variables is industry specific and is based on the 2023 length of the pay periods for that industry as reported by the BLS[27]. Because payroll survey variables cover pay periods including the 12th day of the month, the CES employment variables for the "financial activities" industry, where 18.4 percent are paid monthly and 14.0 percent are paid weekly[28] is effectively weighted to be later in the month than the CES employment variables for the construction industry where 65.4 percent of workers are paid weekly and only 5.9 percent are paid monthly. Finally, for the labor model, we treat private job openings from the BLS Job Openings and Labor Turnover Survey (JOLTS) as an "ISM-style" variable because the survey measures the stock of openings on the last business day of the month rather than its monthly average[29].

As mentioned at the beginning of this section, the weekly activity and labor market models include COVID pandemic dummies. From the week ending March 14, 2020 to the week ending March 21, 2020 initial claims spiked from 273k to 2.914 million, an increase of 2.37 log points. This is more than 6 times larger than the previous record increase of 0.37 log points in the week ending July 25, 1992. The ISM manufacturing index and subindices were

[26]In the manufacturing surveys, this applies to all the regional banks except Richmond. In the services surveys, this applies to the Dallas and Philadelphia Fed surveys only.

[27]See `https://www.bls.gov/ces/publications/length-pay-period.htm`.

[28]The remaining workers are paid biweekly or semimonthly.

[29]See `https://www.bls.gov/jlt/jltover.htm`.

impacted much less dramatically relative to prior recessions[30]. This less dramatic impact is necessarily the case because ISM survey variables and many of those in the Regional Fed analogues are bounded below by 0.

Because of this extreme increase in volatility during the height of the pandemic, and it's differential impacts across variables, we implement the following regime specific standardization of log-differenced claims and the other native weekly activity and labor market variables in our factor models, where $z^W_{i,t,w} = x^W_{i,t,w} - \bar{x}^W_i$ and $z^{*W}_{i,t,w} = \frac{x^W_{i,t,w} - \bar{x}^W_i}{\sigma^{x^W_i}_{t,w}}$, where $\sigma^{x^w_i}_{t,w} = \sigma^{x^W_i}_1 = \sqrt{E((z^W_{i,t,w})^2)}$ for combinations of t and w between and including the weeks ending between January 4, 2020 and July 4, 2020, and $\sigma^{x^W_i}_{t,w} = \sigma^{x^W_i}_0$ for combinations of t and w that correspond to weeks ending either before January 4, 2020 or after July 4, 2020.

For the variables with a native weekly frequency, the equation relating the observed regime specific standardized values to the weekly factors

$$z^{*W}_{i,t,w} = \gamma_{i,0}(1 - D_{t,w})f_{t,w} + \gamma_{i,1}D_{t,w}f_{t,w} + \epsilon_{i,t,w} \tag{9}$$

where $D_{t,w}$ is a dummy variable that assumes the value 1 for each the 27 weeks ending between January 4, 2020 and July 4, 2020 and assumes a value of 0 for all of the remaining values in the sample. This specification allows for some variable specific flexibility in the degree in which the pandemic impacted the weekly activity variables in the first half of 2020.

To standardize the non-weekly monthly variables, we define the monthly, variable i specific dummy variables, $d_{i,m}$ that has a month m value that is the following weighted average of weekly dummy variables:

$$d_{i,m} = \sum_{t=1}^{T} \sum_{w=1}^{W(t)} a_{i,m,t,w} D_{t,w} \tag{10}$$

where $D_{t,w}$ is the weekly dummy variable defined above that is 0 apart from virtually all of the first half of 2020 when it is 1. Thus $d_{i,m}$ will be 0 for months with no positively weighted weeks inside the 27 week period ending on July 4, 2020.

Setting $\tilde{X}^M_{i,t} = X^M_{i,t} - \bar{X}^M_i$, where the mean $\bar{X}^M_i$ is computed over the entire monthly sample including the height of the pandemic, we calculate the two monthly standardization factors

[30]For example, the ISM manufacturing employment subindex bottomed out at 28.0 in April 2020 in that year's recession, identical to it's February 2009 trough of 2007-09 Great Recession and just above it's May 1982 trough of 27.8 in the 1981-82 recession.

$$\sigma_{i,1} = \sqrt{\frac{\sum_t d_{i,t}(\tilde{X}_{i,t}^M)^2}{\sum_t d_{i,t}}} \tag{11}$$

and

$$\sigma_{i,0} = \sqrt{\frac{\sum_t (1-d_{i,t})(\tilde{X}_{i,t}^M)^2}{\sum_t (1-d_{i,t})}} \tag{12}$$

where $d_{i,t}$ is the coeffcient weighted sum of dummy variables defined in equation (10). We standardize the monthly observations $X_{i,t}^M$ as

$$X_{i,t}^{*M} = \frac{X_{i,t}^M - \bar{X}_i^M}{\sigma_{i,t}} \tag{13}$$

where $\sigma_{i,t} = \sqrt{(1-d_{i,t})\sigma_{i,0}^2 + d_{i,t}\sigma_{i,1}^2}$

Our algorithm for estimating the mixed frequency factor model essentially combines the two step principal component analysis (PCA) + Kalman filter/smoother estimation approach with the Stock and Watson [2002b] EM algorithm estimation of the mixed frequency PCA implementation. More specifically, we use the following algorithm

Step 1: Estimate the first principal component $\hat{f}_{t,w}$ (PC) of the weekly series that span the entire sample that are denoted as $z_{i,t,w}^{*W}$ above and have been standardized with separately calculated standard deviations for both the weeks between and ending January 4, 2020 and July 4, 2020 and, separately, the union of the pre-January 4, 2020 and post-July 4, 2020 sample period. Then, forecast those factors to the end of the pre-determined sample+forecast period.

Step 2: Forecast the missing or yet to be released values of $z_{i,t,w}^{*W}$ using the forecast augmented factors and equation (9) .

Step 3: As in Stock and Watson [2002b], iterate between and repeat steps 1 and 2 until the first step weekly factors $\hat{f}_{t,w}$ estimated with the forecast augmented weekly series $\hat{z}_{i,t,w}^{*W}$ only converge.

Step 4: Recalling that that monthly to weekly conversion coefficients are collected into the matrix $\mathbf{A}_i$, with rows corresponding to months and columns corresponding to weeks, we collect the estimated and forecasted weekly factors into the vector $\hat{\mathbf{f}}^{\mathbf{W}}$ and create the monthly "factors" $\hat{\mathbf{f}}_{\mathbf{i}}^{\mathbf{J}} = \mathbf{A}_i\hat{\mathbf{f}}^{\mathbf{W}}$, where J is M if i is a standard monthly average or and J is I if i is an ISM-type variable. We estimate one or a pair of FAARs for $X_{i,t}^{*M}$ or $X_{i,t}^{*I}$, one with autoregressive lags and lags of the monthly "factors" in $\hat{\mathbf{f}}_{\mathbf{i}}^{\mathbf{M}}$, and one, if needed, with leads of those terms just as in **Step 3** of the algorithm describing estimation of the daily price factors in Section III. Those FAARs are both estimated over the non-COVID portion of the sample only and used to forecast and possibly backcast the series to the beginning and end of the monthly sample. As with the prices, the combined estimated forecasted and possibly backcasted data $\hat{X}_{i,t}^{*J}$ are then interpolated using the adapted Fernández [1981] interpolation procedure described in the appendix with the two interpolator interaction variables $D_{t,w}\hat{f}_{t,w}$ and $(1-D_{t,w})\hat{f}_{t,w}$ thereby creating the interpolated forecast and possibly backcast augmenteted weekly series $\hat{z}_{i,t,w}^{*J}$ for either a non-ISM or an ISM style variable.

Step 5: Standardizing and combining the series $\hat{z}_{i,t,w}^{*W}$ $\hat{z}_{i,t,w}^{*M}$, and $\hat{z}_{i,t,w}^{*I}$, compute the first PC $\hat{f}_{t,w}$ of both the native and interpolated weekly forecast augmented series.

Step 7: Repeat steps 6, 4, and 5, in that order until the factor PC estimates $\hat{f}_t$ converge, where

Step 6: Using the updated weekly first PC last computed in **Step 5**, forecast[31] and estimate any missing values for $z_{i,t,w}^{*W}$ just as in **Steps 1** and **2**.

V. Data overview

The data – for which we furnish additional details in Section V – are generally chosen with a particular end in mind. For the activity and price principal component models, the aim is to estimate factors that will be useful in forecasting a wide range of source data for the GDP subcomponents and there associated price deflators. For the activity factor model in the GDPNow model and the extension pursued it in Higgins [2024], the data indicators used were the GDP spending source data that needed to be forecasted and many of the 85 indicators used by the Chicago Fed to estimate the CFNAI because of their proven utility in forecasting activity related indicators in the earlier Stock and Watson. However, there were many (over 200) PPI and other monthly price deflators used in the Higgins [2024] model for forecasting just the non consumption and non-inventory GDP subcomponent price deflators. This is partly because one simple addition can be used to aggregate nominal spending levels for GDP subcomponent source data, but this strategy does not work as well for the corresponding price deflators. One could perhaps aggregate the price data first, but the subindices often have a short – less than 20-year – history. Thus, the price model deflators and other variables are chosen to span a wide range of spending types – including

[31] As with prices, we do not use factors to forecast the weekly data beyond the initial factor estimation in to improve convergence performance of the algorithm.

consumption, investment, imports and exports – while also not being redundant with other series in the collection. For example, we generally don't include PCE subcomponents for which there is a CPI equivalent or near-equivalent since the latter series is released first. A very similar strategy was used in the dynamic factor model underpinning the discontinued FRBNY Staff Underlying Inflation Gauge documented in Amstad et al. [2014]. But, as seen in Tables 1 and 1, consumption or related price indices were overrepresented relative to other types due to the 60 to 70 percent expenditure weight of PCE within nominal GDP. For the labor market PC, the indicators themselves are all chosen or transformed to be consistent with measuring momentum rather than slack or utilization. While not chosen to forecast a particular indicator, average hours worked and employment by industry in the payroll survey are heavily represented since those make up the labor market momentum aggregates of most interest (growth in employment and hours worked). The weekly ADP employment measure in the so-called National Employment Report Pulse[32] is not used because that series appears to be much smoother before June 2025 than after.

VI. Results

Tables 1 and 2 provide the parameter estimates and correlations for the indicator in the mixed daily-monthly inflation model, with the first two numerical columns showing the first and second principal component weights for the combined interpolated and non-interpolated daily data, where the weights in each column have been normalized to sum to 100. We can see the weights for the first PC are all modestly positive. The weights for the second PC are a mix of positive and negative values with the weights on commodity price related variables generally positive and the weights for the "underlying inflation" and "ISM-style" variables often negative. The daily commodity price variables load more heavily on the second PC than the first PC. The middle two numerical columns of Tables 1 and 2 show the correlation of the aggregated daily to monthly factors with the monthly variables with the aggregation type – monthly average or "ISM style" – dependent on the variable in question. The last 2 columns show the correlations of the daily factors with the (usually interpolated) daily data. For most variable and factor pairs, the variable/PC correlations with the daily data are similar in magnitude to the correlations of the monthly data.

Figure 3 shows three estimates for the first PC/factor – the smoothed factor for a variant of the monthly dynamic factor model used in Higgins [2024] (red solid line), the first PC

[32]See `https://www.adpresearch.com/main-street-macro/the-ner-pulse`.

for the baseline mixed daily/monthly price variable model (solid blue line), and the first PC for an alternative daily/monthly price variable model (dashed blue line) that doesn't use federal government data[33] – and Figure 4 does the same for the second PC/factor. We can see that the first PCs (monthly aggregated if necessary) at the monthly frequency are all highly correlated. The second PCs are also positively correlated, but not as strongly as the first PCs. The first PCs, or the factor that's related to it, all reach their maximum level in late 2021 to early 2022.

Figures 5 and 6 show the results of an experiment where we impose large 7 standard deviations shocks to each of the daily price variables in the price model on August 10, 2026 and leave the price levels unchanged for the remaining four days of the week. These shocks imply WTI and Brent oil prices spike up 18.5 and 16.4 log percentage points, respectively, on August 10th which is a bit more than one-third the size of their run-up over the entirity of March 2026. The remaining commodity prices increase between 2.5 and 9.0 log percentage points on August 10, 2026 in the experiment. The result of this experiment, as we can see in Figures 5 and 6, are that relative to the baseline unconditional forecast extensions, there are upward spikes in the first two aggregated principal components that are similar in their cumulative magnitudes to the increases seen in either March (for the second PC) or April (for the first).

Tables 3 through 5 provide the parameter estimates for the weekly labor model, with the first numerical columns showing the PC weights that have been normalized to sum to 100. Except in a few cases, where the weights are virtually 0, these weights are modest to moderate positive values. The weights for the (seasonally adjusted) 21 weekly variables – the ASA staffing and the national and 9 Census Division values for both initial and continuing unemployment insurance claims – are all positive and sum to just under 35%. The variables from either the national or regional Fed "ISM type" surveys account for 25% of the PC weights while the remaining variables account for just over 40% of them. The last two columns of the tables have the correlations of the PCs and the data at the monthly and weekly frequencies, respectively where the latter correlation calculation is done with interpolated and backcasted data where necessary[34]. At the monthly frequency, the correlation with the "ISM" type survey variables is higher than with the other monthly variables (0.69 vs. 0.28). This is perhaps not surprising because of the greater persistence and the lower-magnitude

[33]The regional Fed Bank series that use CPI or PCE subcomponents are also dropped.

[34]The monthly correlations only use published data.

high frequency fluctuations in the “ISM” diffusion index type survey labor variables, but it does suggest that a non-federal government data based version of the labor PC may closely resemble the standard version which we can see is in fact the case in Figure 7. In that figure, the red line denotes PC estimated with the variables and parameter estimates in Tables 3 through 5, while the green dashed line is estimated after dropping variables sourced from the establishment survey (Payroll) the Current Population Survey (CPS), or the JOLTS survey since these variables would not be available in a government shutdown that affected the Department of Labor[35]. We can see that these two series are highly correlated ($r = 0.944$). In the “non-federal” government series version of the model corresponding to the green dashed line in Figure 7, we have consolidated the 18 Census Division level series on initial and continuing unemployment insurance claims into 8 series on those claims by Census Region. This consolidation is made to approximately maintain the aggregate importance weights of the claims data relative to the other labor market variables.

The blue line in Figure 7 is the PC model estimated without the 18 weekly series on initial and continuing unemployment insurance claims by Census Division, though it includes the two national claims series in addition to the ASA staffing index. In spite of only including these three weekly series, the corresponding estimated PC is very highly correlated with the benchmark PC ($r = 0.979$). However, Figure 8 illustrates that the two series can react very differently to incoming weekly data. In the figure, the solid red and blue lines match the solid lines of the same color in Figure 7. The dashed lines are constructed by extending the weekly series through the remainder of the summer and some of the Fall. In particular, U.S. initial weekly unemployment insurance claims are assumed to steadily increase from 199k in the week ending August 1, 2026 to 400k in the week ending August 29, 2026 and remain at that elevated level through the rest of November. Continued unemployment insurance claims are assumed to remain at their prior week level in the week ending August 1, 2026 before increasing steadily by 40k each week through November. And the ASA Staffing index is assumed to shrink at a constant rate of 0.25 percentage points per week. Additionally, in the 21-weekly variable variant of the labor model, initial and continued claims for each of the nine Census Divisions are assumed to have the same proportional increases as their respective national counterparts over the same roughly 3-month period. In Figure 8, the dashed red line corresponds to the estimated labor market PC through October conditional

[35]We do not remove the unemployment claims data because that data can be closely estimated based on state level data that was largely reported by nearly all the states during the 2025 government shutdown.

on the assumed path for the 21 weekly variables while the dashed blue line represents the estimated PC conditional on the assumed path for the 3 national weekly variables when those are the only weekly series in the model. We can see the red line bottoms out at a level only modestly above its trough in the 2001 recession. In that recession, initial claims peaked at 517k in the last week of September 2001, which was twice as large as their year 2000 trough of 259k in April of that year so that the proportional increase was on par with our conditional scenario. The blue line in Figure 8 increases much less, illustrating that lowering the effective PC weight of the claims data in the 3-weekly variable model results in reducing its importance as a conditioning variable relative to the 21-weekly variable model.

Figures 7 and 8 show the trough of the (aggregated) monthly labor PC bottoming out at a lower level in the 2007-09 recession than in the pandemic recession. This phenomenon – due to our dummy variable treatment of pandemic dynamics and of standardizing the 2020:H1 data – is at odds with the April 2020 changes in payroll employment and the unemployment rate showing much more adverse labor market outcomes than the worst months of the 2007-09 recession in late 2008 and 2009. To remove this effect of how we handled the 2020 pandemic recession, we maintain the baseline PC weighting coefficients that are estimated in the benchmark model and recursively iterate between interpolating the labor market data and re-calculating the weekly PC with the non-COVID adjusted standardized data. The resulting PC – aggregated to the monthly frequency with standard monthly averaging for growth rate aggregation – after this calculation is plotted as the blue dashed line in Figure 9 alongside the standard model (21-weekly series) aggregated monthly PC represented by the red line in the same figure. The red and blue lines are both standardized to have mean 0 and standard deviation 1 over the entire 1967-2026 sample. The blue line bottoming out at a much lower level in the 2020 recession effectively means it has to bottom out at higher levels in the non-COVID recessions inorder to maintain the equal 1-unit standard deviations of both series. But when we remove the December 2019 - July 2020 observations of both series, the correlation is very high ($r = 0.996$). We can also see the "non-COVID" adjusted PC does in fact bottom out at a much lower level in the 2020 pandemic recession than it does in the 2007-09 Great Recession. However, we also see that there is an upward spike in February 2020. This upward spike is a consequence of the interpolated weekly factor having to best fit the huge monthly fluctuations in each of the monthly variables with various aggregation patterns in 2020:H1.

One possible criticism of the baseline PC is that it does not account for the impact of the slower growth rate of the potential labor force due beginning sometime in 2024, in large part, to the decline in net migration that has had the effect of reducing the "breakeven" rate in payroll employment growth needed to maintain a flat unemployment rate (see, for example, Butcher et al. [2026] and Cheremukhin et al. [2026]). To account for this, we divide the upward trending variables that measure counts of persons or aggregate hours worked by a smoothed measure of the effective potential labor force growth. More specifically, we used the smoothed CPS measures of population and labor force statistics from Coglianese et al. [2025] that were updated in July 2026 to form a Tornqvist index based labor force share weighted growth rate of the (seasonally adjusted) population using 16 age-sex cells to partition the age 16+ population and use the Christiano and Fitzgerald [1999] band-pass filter to remove frequencies higher than 24 months from the resulting series. The annualized (logarithmic) monthly growth rate of the smoothed demographically adjusted series – used to scale a subset of the labor market principal component source data – is the black solid line in Figure 10, which we can see is under 0.1 percentage points in July 2026. The red dashed line is the smoothed population growth rate when the population shares are not adjusted by labor force participation, and it is 0.5 percentage point in July 2026. The higher value is due to relatively high population growth for the oldest age 65+ groups that puts population growth above "potential" labor force growth. Figure 11 compares the labor force adjusted labor market PC (thicker solid blue line) with the standard PC (thinner solid red line). Both series are standardized to have mean 0, standard deviation 1 over the entire nearly 60-year sample and we can see that the adjusted series has a value close to 0 and close to its prepandemic-level at the end of the sample[36]. Without the adjustment, the labor market PC is moderately below both its average value of 0 and below its pre-pandemic level. But the overall impact of the demographic adjustment is fairly moderate.

Tables 6 through 10 provide the names and parameter estimates for the mixed weekly/monthly activity model. Nearly all the variables in the standard GDPNow activity factor model are also included in this version. In the standard GDPNow model, the variable selection was

---

[36]We end the forecast augmented sample in October 2026 because the weekly data through August will get positive weights in September for many series. The addition of 1 extra month is used to smooth out the interpolation. But the change from August 2026 to October 2026 are qualitatively minor.

based on including most of the 85 indicators used to estimate the Chicago Fed National Activity Index with an additional layer of monthly indicators used in GDPNow's bridge equations that relate granular GDP subcomponents to either monthly source or closely related data. Additionally, we include around 25 monthly indicators from the ISM manufacturing and services surveys – and their regional Fed analogues – as well as additional non-federal government sources to increase the estimation power in the event of a federal government shutdown.

The first numerical column of Tables 6 through 10 show the PC weights that have been normalized to sum to 100. The first 9 rows correspond to the (seasonally adjusted) weekly variables included at that frequency in the model. Cumulatively, they account for 6.6 percent of the PC variable weights. By source, four data releases account for just over 55 percent of the PC weight: the monthly labor report including both the CPS and payroll survey (26.8 percent), the Fed's industrial production report (12.5 percent), and the national ISM Manufacturing and Services surveys (8.8 percent and 7.3 percent, respectively). There are 47 of the 152 monthly indicators that aren't from or sourced from the federal government[37] and, together with the weekly indicators, they cumulatively account for 40.7 percent of the PC weights in the full activity model. Ten of the full model indicators – all of which are used in the GDPNow factor model and in of GDPNow's bridge equations – have (slight) negative weights, but together they only account for -0.4 percent of the cumulative PC weight. The average correlation of a monthly series' value with the full model activity PC aggregated to a monthly frequency with that series' aggregation weighting is 0.34 in the full model and 0.45 for the subset of 47 "non-federal government" indicators. The series with the highest

[37]Much of the initial estimates in the Fed's Industrial Production and Capacity Utilization report (IP) come from payroll employment survey so that IP releases in the Fall of 2025 were delayed until the end of the government shutdown. In addition to the series from that report, the variables not included in the "non federal government" version of the model come from the Advance Economic Indicators report (AdvEconInd), the Construction Spending Value Put-in-Place report (Construct), the BEA's Underlying detail NIPA tables (Detail), the Advance and Full International Trade Reports (ForeignTrade), the Monthly Labor Report (Labor), the M3-1 and M3-2 Manufacturing Reports from the Census Bureau (M31), the Census Bureau New Residential Sales report (NewHomes), the Personal Income and Outlays report (PersIncPCE), the Retail Sales and Inventories reports (Retail), the Census Bureau New Residential Construction report (Starts), and the Wholesale Trade report. We do however include the three unit sales variables for autos and light trucks (Autos) from the BEA because they can almost be exactly estimated using sales data from the Autodata Corporation that are generally released the afternoon before the BEA numbers are released in the subsequent morning.

monthly correlations with the aggregated factor are often concentrated in the ISM reports and their regional Fed analogs and other "soft-data" surveys, as well as the aggregate and industry level series in the payroll employment and industrial production reports.

The mixed frequency weekly+monthly first principal component of the economic activity related data – conceptually similar to both the Chicago Fed National Activity Index (CFNAI) and the first state PC computation used to estimate the dynamic activity factor in the GDPNow factor model – has the same COVID dummy adjustments used in the labor market setting. Because the CFNAI and GDPNow analogs do not have those adjustments, we need to "undo" the COVID adjustments by iteratively interpolating and applying the PC coefficients in the same way described above for the labor market PC. After that adjustment, we compare the three series over the common post-1980 sample period in Figure 12. We can see the three series are all highly correlated and have much lower troughs in the pandemic recession than the earlier ones. Nevertheless, there are some notable discrepancies between the mixed frequency PC and the other 2 indices during the height of the pandemic. In particular, the CFNAI and standard GDPNow first stage factor are more highly correlated with each other ($r = 0.985$) than either the former ($r = 0.888$) or latter ($r = 0.911$) is individually with the aggregated mixed frequency PC. Some of this is to be expected given how the mixed model allows for different monthly aggregation patterns unlike the other two measures. But the different treatment of 2020:H1 (dummy variable treatment vs. no treatment when deriving the PC weights) and the different sample start date for computing the PCs (1981 for the mixed frequency based measure due to availability limitations for the weekly data vs. 1967 for the other 2 PC measures) likely play a role as well.

Figure 13 shows our baseline measure of the mixed weekly-monthly economic activity (momentum) indicator that account for pandemic dynamics in 2020:H1 as the thick solid blue line while an analogous measure using only non-federal government data series – apart from unemployment claims and unit auto and light truck sales – is depicted with the thinner red line. The "private-only" measure uses much less monthly indicators than the full one. Nevertheless, the two series are highly correlated ($r = 0.915$). Among the largest discrepancies between the two measures are that the troughs of the full-variable model PC are lower in the 2007-09 and 2020 recessions than they are in the non-federal government indicator version of the model. This may be related to the higher share of diffusion index type variables – like the ISM variables and their regional Fed bank analogues – in the latter model, suggesting that one may be able to detect a recession nearly as well without without the usually "hard" data type federal government indicators, but it's more difficult to gauge it's severity.

To this point we have not given any focus to the daily or weekly factors at their highest frequency. Figure 14, which shows the weekly PC from the full variable activity model (blue line) in addition to its 4-week trailing moving average (red line), illustrates why this is the case. Though one can recognize the downturn in the weekly series in the 2007-09 recession and, less clearly, in the 1981-82 recession, the series is still quite noisy[38]. In particular, there are a couple of weeks at the end of 2008 and beginning of 2009 where the PC is positively valued unlike either its 4-week moving average or its monthly aggregate plotted in Figure 13. And even the 4-week moving average is somewhat noisy relative to the corresponding series in that plot. This noisiness suggests that forecasting monthly data in a FAAR specification with just the disaggregated weekly PCs would require a number of the lags of these terms that may be unyieldly.

VII. Conclusion

The work here has gone some distance in showing that one can augment a set of monthly indicators, used to estimate a monthly principal component or factor, with higher frequency daily or weekly data in order to estimate a higher frequency PC that maintains a close correspondence with the original monthly PC or factor when it is aggregated back up to the monthly level. Moreover, in the price and labor market environments we explored, we showed that propagation of extended shocked values of daily or weekly data to the estimated monthly PC could be of a quite plausible magnitude. In the labor market setting, this plausibility required having more than a few weekly variables implying that some judiciousness and finesse is needed when choosing the mix of lower and higher frequency variables that are included in the model. We did not explore how adding an additional observation of the national ISM variables and their regional Fed analogues – since those are released before most of the other variables – but that would be an obvious next step. Additional next steps would be assessing the forecast performance of monthly variables inside or outside the PC calculation that are of particular interest and, ultimately, embedding those forecasts inside the GDPNow model and the Higgins [2024] extension of the model to nowcast the price deflators of GDP and it's subcomponents. To be fruitful, this latter step would likely require

[38]Because of the dummy variable treatment of 2020:H1, the troughs in that period are significantly attenuated.

a more general handling of the forecast augmentation. For example, one would want to use daily oil prices to forecast monthly gasoline and petroleum import prices in addition to or rather than simply using the price factors. Since the monthly factors or PCs can be well approximated by a PC calculation with just the non-federal government indicators, another natural step would be to see how much mileage one could get by nowcasting GDP – or other variables of interest – with only “private” or non-shutdown effected variables.

We’ve also overlooked the computing time aspect of this exercise, which can take 20 minutes or so if one requires the PC(s) to converge close to a fixed point[39], which may make it difficult to continuously update on, say, a regular semiweekly basis. But the general approach of harmonising simple and “ISM” and ”labor report” style monthly variables with different temporal aggregation patterns with the introduction of higher frequency data. A similar approach to the one we’ve used could perhaps be implemented in, for example, a mixed frequency VAR to allow for sharper identification of monetary policy, oil, or other shocks of interest. We leave these efforts to future research.

References

Martin Almuzara, Katie Baker, Hannah O’Keeffe, and Argia Sbordone. The New York Fed staff nowcast 2.0. Staff nowcast technical paper, Federal Reserve Bank of New York, June 2023.

Marlene Amstad, Simon Potter, and Robert Rich. The FRBNY staff underlying inflation gauge: UIG. Staff Report 672, Federal Reserve Bank of New York, 2014.

S. B. Aruoba, F. X. Diebold, and C. Scotti. Real-time measurement of business conditions. Working Paper 08-19, Federal Reserve Bank of Philadelphia, 2008.

S. B. Aruoba, F. X. Diebold, and C. Scotti. Real-time measurement of business conditions. *Journal of Business and Economic Statistics*, 27(4):417–427, 2009.

Brandyn Bok, Daniele Caratelli, Domenico Giannone, Argia M. Sbordone, and Andrea Tambalotti. Macroeconomic nowcasting and forecasting with big data. *Annual Review of Economics*, 10(1):615–643, 2018.

Scott A. Brave, Michael Fogarty, Daniel Aaronson, Ezra Karger, and Spencer Krane. Tracking U.S. consumers in real time with a new weekly index of retail trade. Working Paper 2021-05, Federal Reserve Bank of Chicago, 2021.

[39] This is usually achieved after 60 to 100 iterations of the repeated steps in the algorithm.

Scott A. Brave, Ross Cole, Katherine Jolley, and Ezra Karger. Tracking holiday spending with CARTS 2.2 and a new dashboard. *Chicago Fed Insights*, December 2024.

Falk Bräuning and Siem Jan Koopman. Forecasting macroeconomic variables using collapsed dynamic factor analysis. *International Journal of Forecasting*, 30(3):572–584, 2014.

Bureau of Labor Statistics. Productivity and costs, first quarter 2020, revised. Bureau of Labor Statistics News Release, June 2020.

Kristin Butcher, Anne Fournier, Camilo Garcia-Jimeno, Aneesh Kudrimoti, and Alan Mathew. Reading the labor market when population is a moving target: A lesson from the post-pandemic immigration surge. *Chicago Fed Letter*, (525), August 2026.

Stephanie Cano, Patricia Getz, Jurgen Kropf, Stuart Scott, and George Stamas. Adjusting for a calendar effect in employment time series. Statistical Survey Paper 1, Bureau of Labor Statistics, 1996.

Anton Cheremukhin, Daniel Wilson, and Xiaoqing Zhou. Break-even employment declines as unauthorized immigration outflows continue. *Dallas Fed Economics*, March 2026.

Gregory C. Chow and An-loh Lin. Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53 (4):372–375, 1971.

Lawrence J. Christiano and Terry J. Fitzgerald. The band pass filter. NBER Working Paper 7257, National Bureau of Economic Research, 1999.

John Coglianese, Seth Murray, and Christopher J. Nekarda. Harmonized population and labor force statistics. Finance and Economics Discussion Series Working Paper 2025-057, Federal Reserve Board, 2025.

Frank T. Denton. Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, 66 (333):99–102, 1971.

Catherine Doz, Domenico Giannone, and Lucrezia Reichlin. A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164(1):188–205, 2011.

Roque B. Fernández. A methodological note on the estimation of time series. *The Review of Economics and Statistics*, 63(3):471–476, 1981.

Domenico Giannone, Lucrezia Reichlin, and David Small. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676, 2008.

Elise Gould and Joe Fast. Amid the shutdown data blackout, state unemployment insurance claims continue to shed light on the labor market. Economic Policy Institute Working Economics Blog Entry, October 2025.

Craig S. Hakkio and Jonathan L. Willis. Kansas city Fed's labor market conditions indicators (LMCI). *Federal Reserve Bank of Kansas City, The Macro Bulletin*, August 2014.

Andrew Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, 1990.

Patrick Higgins. GDPNow: A model for GDP nowcasting. Working Paper 2014-7, Federal Reserve Bank of Atlanta, 2014.

Patrick Higgins. The price is wrong. The deterioration of GDP nowcasting accuracy this decade and the role inflation played in it. Unpublished manuscript presented at August 2024 Joint Statistical Meetings, 2024.

Edward S. Knotek and Saeed Zaman. Nowcasting U.S. headline and core inflation. *Journal of Money, Credit and Banking*, 49(5):931–968, 2017.

Daniel J. Lewis, Karel Mertens, James H. Stock, and Mihir Trivedi. Measuring real activity using a weekly economic index. Staff Report 920, Federal Reserve Bank of New York, 2020.

Daniel J. Lewis, Karel Mertens, James H. Stock, and Mihir Trivedi. High-frequency data and a weekly economic index during the pandemic. *AEA Papers and Proceedings*, 111: 326–330, 2021.

R. S. Mariano and Y. Murasawa. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4):427–443, 2003.

José Mustre-del Río, Johnson Oliyide, and Emily Pollard. An alternative version of the KC Fed LMCI suggests little change in the labor market in September. *Federal Reserve Bank of Kansas City Economic Bulletin*, October 2025a.

José Mustre-del Río, Johnson Oliyide, and Emily Pollard. An alternative version of the KC Fed LMCI suggests the level of activity was little changed but momentum decelerated sharply in October. *Federal Reserve Bank of Kansas City Economic Bulletin*, November 2025b.

Daniel Ollech. Seasonal adjustment of daily time series. Discussion Paper 41/2018, Deutsche Bundesbank, 2018.

James H. Stock and Mark W. Watson. Forecasting inflation. *Journal of Monetary Economics*, 44(2):293–335, 1999.

James H. Stock and Mark W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002a.

James H. Stock and Mark W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162, 2002b.

Appendix A. Adapted Fernandez (1981) interpolation

Suppose we have a lower frequency time series X_l spanning L periods that we want to interpolate with one or more higher frequency factors that we collect in the vector or matrix $\mathbf{F}$. When interpolating a macroeconomic time series with the standard Fernández [1981] approach or the related approaches by Denton [1971] and Chow and Lin [1971], the lower frequency indexed by l is annual or quarterly while the higher frequency series spanning H periods and indexed by h is typcally quarterly ($H = 4L$) or monthly ($H = 3L$ when l is quarterly and $H = 12L$ when l is annual). However, in our application, l is monthly, X_l is usually a monthly growth rate or an ISM style series, and h is either daily or weekly. The high frequency interpolated series x_h is specified as a linear function of the factors plus an error term

$$x_h = F_h\beta + \epsilon_h \tag{14}$$

while the lower frequency series is a linear function of the higher frequency interpolated series:

$$X_l = \sum_{h=1}^{H} a_{l,h}x_h \tag{15}$$

.

If l were quarterly and h monthly, then the non-zero coefficients in equation 15 would satisfy $\mathbf{a_l} = [a_{l,h}, a_{l,h+1}, a_{l,h+2}, a_{l,h+3}, a_{l,h+4}] = [\frac{1}{3}, \frac{2}{3}, \frac{3}{3}, \frac{2}{3}, \frac{1}{3}]$. However, in our application, $\mathbf{a_l}$ will depend on factors such as the number of business days the month and preceding month have or the timing of weeks wholly are partly in or adjacent to the month in question. We rewrite equation 14 in matrix form as

$$\mathbf{x} = \mathbf{F}\beta + \epsilon \tag{16}$$

and equation 15 in matrix form as

$$\mathbf{X} = \mathbf{A}'\mathbf{x} \tag{17}$$

where $\mathbf{A} = [\mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_{L-1}}, \mathbf{a_L}]$.

As in the original Fernández [1981] implementation, the adapted algorithm minimizes the sum of squared differences in the consecutive error terms in equation 14, which in matrix form is

$$\min_{\mathbf{x},\beta}(\mathbf{x} - \mathbf{F}\beta)'\mathbf{D}'\mathbf{D}(\mathbf{x} - \mathbf{F}\beta) \tag{18}$$

where

$$\mathbf{D} = \mathbf{I}_H + \begin{bmatrix} \mathbf{0}_{1\times(H-1)} & 0 \\ -\mathbf{I}_{H-1} & \mathbf{0}_{(H-1)\times 1} \end{bmatrix} \tag{19}$$

.

The inverse of $\mathbf{D}'\mathbf{D}$, used below is

$$(\mathbf{D}'\mathbf{D})^{-1} = [d_{i,j} = min(i,j)] \tag{20}$$

Because it is often the case that $H \approx 10,000$ when interpolating a monthly price series to the business day frequency, it is much more efficient to use equation 20 directly rather than compute $(\mathbf{D}'\mathbf{D})^{-1}$ directly via brute force. As in Fernández [1981], the high frequency series is solved for via the equations

$$\beta = (\mathbf{F}'\mathbf{A}(\mathbf{A}'(\mathbf{D}'\mathbf{D})^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{F})^{-1}(\mathbf{F}'\mathbf{A}(\mathbf{A}'(\mathbf{D}'\mathbf{D})^{-1}\mathbf{A})^{-1})\mathbf{X} \tag{21}$$

and

$$\mathbf{x} = \mathbf{F}\beta + (\mathbf{D}'\mathbf{D})^{-1}(\mathbf{A}'(\mathbf{D}'\mathbf{D})^{-1}\mathbf{A})^{-1}(\mathbf{X} - \mathbf{A}'\mathbf{F}\beta) \tag{22}$$

.

Table 1. Variable names and summary statistics (Part 1)

Description	Transform	Start	PCA1 Wt	PCA2 Wt	Corr M1	Corr M2	Corr D1	Corr D2
GSCI (Daily)	DLog	198610	1.0	14.3	0.19	0.53	0.06	0.73
Brent Oil (Daily)	DLog	198610	0.9	12.9	0.16	0.47	0.05	0.67
WTI Oil (Daily)	DLog	198610	0.9	12.8	0.16	0.46	0.05	0.63
Agric (Daily)	DLog	198610	0.5	6.8	0.09	0.27	0.03	0.34
IndusMat (Daily)	DLog	198610	0.5	10.6	0.09	0.43	-0.03	0.42
Livestock (Daily)	DLog	198610	0.5	4.3	0.13	0.15	0.11	0.16
FIBER-All (Daily)	DLog	198610	0.8	15.1	0.14	0.63	-0.09	0.63
FIBERCore (Daily)	DLog	198610	0.7	13.7	0.11	0.59	-0.09	0.50
CPI Food Home	DLog	198610	1.5	-3.3	0.42	-0.27	0.34	-0.54
CPI Food Away	DLog	198610	2.3	-5.8	0.63	-0.50	0.59	-0.41
CPI Alcohol	DLog	198610	0.8	-4.0	0.21	-0.27	0.21	-0.70
CPI Rent	DLog	198610	1.6	-8.7	0.43	-0.61	0.28	-0.90
CPI Oth Lodge	DLog	198610	0.9	2.3	0.25	0.06	0.21	0.36
CPI OER	DLog	198610	1.8	-7.8	0.49	-0.60	0.34	-0.87
CPI Fuel+Utilities	DLog	198610	1.4	4.3	0.37	0.15	0.30	0.51
CPI HH Furn+Oper	DLog	198610	1.9	-3.2	0.52	-0.31	0.40	-0.68
CPI Apparel	DLog	198610	1.3	2.3	0.34	0.03	0.30	0.27
CPI Priv. Trans.	DLog	198610	1.5	12.3	0.33	0.48	0.08	0.92
CPI Vehicles	DLog	199302	1.4	2.5	0.36	0.12	0.37	0.05
CPI Pub. Trans.	DLog	198902	1.2	4.9	0.31	0.20	0.30	0.25
CPI MedC Commod	DLog	198610	0.3	-2.7	0.07	-0.20	0.09	-0.56
CPI MedC Svcs	DLog	198610	0.5	-1.2	0.12	-0.11	0.13	-0.27
CPI Recreat.	DLog	199302	1.4	-2.7	0.33	-0.24	0.32	-0.64
CPI Educ Tuition+	DLog	198610	0.1	-1.2	0.02	-0.09	0.06	-0.50
CPI Commun.	DLog	199802	0.3	0.5	0.08	0.01	0.08	0.10
CPI Oth Gds+Svcs	DLog	198610	0.6	-2.1	0.14	-0.18	0.17	-0.37
Manheim UVI	DLog	199702	0.4	8.0	0.10	0.36	0.00	0.85
NYFed GSCPI	DLog	199709	2.5	0.1	0.68	-0.13	0.66	-0.12
FRB Dallas Trim	DLog	198610	2.9	-4.7	0.76	-0.51	0.68	-0.62
FRB Clev Med PCE	DLog	198610	2.7	-6.2	0.72	-0.57	0.57	-0.76

Table 2. Variable names and summary statistics (Part 2)

Description	Transform	Start	PC1 Wt	PC2 Wt	Corr M1	Corr M2	Corr D1	Corr D2
FRB Clev Med Cpi	DLog	198610	2.7	-7.4	0.72	-0.64	0.57	-0.79
FRB Clev Trim Cpi	DLog	198610	3.2	-3.3	0.85	-0.44	0.79	-0.51
Atl Sticky Ex Shel	DLog	198610	2.4	-2.2	0.63	-0.32	0.57	-0.57
Atl Core Flex	DLog	198610	1.6	3.1	0.42	0.10	0.44	0.06
PPI Core Fin Gds	DLog	198610	1.5	-2.9	0.41	-0.24	0.31	-0.74
PPI Cap Equip	DLog	198610	2.4	-4.0	0.65	-0.37	0.45	-0.76
PPI Core Interm Mat	DLog	198610	2.1	10.3	0.55	0.43	0.32	0.78
PPI Core Crude Mat	DLog	198610	0.7	11.5	0.16	0.52	0.03	0.86
NFIB % Raise P	DLog	198610	3.1	-5.3	0.86	-0.52	0.80	-0.10
Imp PI Petrol	DLog	198610	1.1	13.5	0.24	0.52	0.04	0.92
Imp PI Nonpet	DLog	198610	1.7	8.7	0.44	0.36	0.24	0.77
Exp PI Ag	DLog	198610	0.8	7.7	0.18	0.31	0.00	0.89
Exp PI NonAg	DLog	198610	1.8	12.1	0.45	0.48	0.14	0.91
Pr Home Constr	DLog	198610	1.2	0.5	0.33	-0.05	0.28	0.17
PrPCENonMarket	DLog	198702	2.6	7.9	0.66	0.17	0.38	0.74
CLogic SFRI	DLog	200402	2.5	-2.0	0.64	-0.19	0.67	-0.15
Dall Mfg Pr	Lev	200406	2.7	0.2	0.85	-0.16	0.75	-0.12
KC Mfg Pr	Lev	200107	3.0	-1.3	0.89	-0.29	0.81	-0.04
NY Mfg Pr	Lev	200107	3.1	-3.2	0.86	-0.32	0.84	-0.03
Phil Mfg Pr	Lev	198610	2.7	0.3	0.76	-0.14	0.74	-0.06
Rich Mfg Pr	Lev	199706	3.1	-4.2	0.83	-0.43	0.82	-0.37
ISM Mfg Pr	Lev	198610	2.1	6.5	0.55	0.25	0.53	0.19
Dall Svcs Pr	Lev	200701	3.0	-2.2	0.82	-0.30	0.80	-0.02
KC Svcs Pr	Lev	201401	3.2	-1.6	0.83	-0.20	0.86	-0.30
NY Svcs Pr	Lev	200409	2.9	-4.2	0.78	-0.37	0.79	-0.24
Phil Svcs Pr	Lev	201103	2.5	-6.2	0.61	-0.40	0.63	-0.61
Rich Svcs Pr	Lev	199311	2.3	-7.0	0.66	-0.57	0.64	-0.26
ISM Svcs Pr	Lev	199707	3.1	0.8	0.83	-0.09	0.82	0.04

Note: Diff denotes first difference, DLog denotes first differenced log value, and Lev denotes no transformation. "PC1 Wt" and "PC2 Wt" denote first and second principal component weights after they have been standardized to sum to 100. "Corr M1" and "Corr M2" denotes correlation of transformed monthly series with the first and second price principal components (PC), respectively, after latter have been aggregated to the monthly frequency with the same aggregation pattern as the data series. "Corr D1" and "Corr D2" denote correlation of daily interpolated (when applicable) and possibly backcasted values of price series with the daily first and second PCs.

Table 3. Labor variable names and summary statistics (Part 1 of 3)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
ASA Staff. Index (W)	DLog	ASA	200701	1.97	0.53	0.29
Init Claims US (W)	DLog	Claims	196702	1.21	0.38	0.49
Init Claims CD1 (W)	DLog	Claims	198801	1.13	0.23	0.23
Init Claims CD2 (W)	DLog	Claims	198801	1.56	0.29	0.37
Init Claims CD3 (W)	DLog	Claims	198801	1.33	0.22	0.32
Init Claims CD4 (W)	DLog	Claims	198801	1.38	0.24	0.29
Init Claims CD5 (W)	DLog	Claims	198801	1.51	0.34	0.27
Init Claims CD6 (W)	DLog	Claims	198801	1.30	0.22	0.29
Init Claims CD7 (W)	DLog	Claims	198801	0.98	0.24	0.15
Init Claims CD8 (W)	DLog	Claims	198801	1.62	0.39	0.31
Init Claims CD9 (W)	DLog	Claims	198801	1.38	0.32	0.22
Cont Claims US (W)	DLog	Claims	196702	1.90	0.66	0.62
Cont Claims CD1 (W)	DLog	Claims	198801	1.84	0.55	0.46
Cont Claims CD2 (W)	DLog	Claims	198801	1.84	0.50	0.40
Cont Claims CD3 (W)	DLog	Claims	198801	1.85	0.55	0.46
Cont Claims CD4 (W)	DLog	Claims	198801	2.00	0.57	0.53
Cont Claims CD5 (W)	DLog	Claims	198801	2.15	0.64	0.47
Cont Claims CD6 (W)	DLog	Claims	198801	2.00	0.57	0.47
Cont Claims CD7 (W)	DLog	Claims	198801	1.72	0.55	0.35
Cont Claims CD8 (W)	DLog	Claims	198801	2.15	0.69	0.47
Cont Claims CD9 (W)	DLog	Claims	198801	1.92	0.55	0.39
ADP Emp	DLog	ADP	201002	1.65	0.53	0.51
ChallengGrayChris	Level	Challenger	198903	1.36	0.56	0.59
CB Plenty Jobs	Diff	Conf Board	197707	0.89	0.35	0.51
CB Jobs Hard Get	Diff	Conf Board	197707	0.97	0.36	0.60
U6	Diff	CPS	199402	1.81	0.62	0.78
Unemp % Quiters	Diff	CPS	196702	0.64	0.21	0.51
Unemp % Losers	Diff	CPS	196702	1.11	0.39	0.72
LFPR 25-54	Diff	CPS	196702	0.32	0.10	0.35
UR 25-54	Diff	CPS	196702	1.53	0.56	0.65
LFPR 16-24	Diff	CPS	196702	0.38	0.13	0.39
UR 16-24	Diff	CPS	196702	0.97	0.34	0.61

Table 4. Labor variable names and summary statistics (Part 2 of 3)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
Dall Mfg Emp	Level	Dallas Mfg	200406	1.05	0.77	0.52
Dall Mfg Avg Hrs	Level	Dallas Mfg	200406	1.55	0.77	0.81
Dall Svcs Emp	Level	Dallas Svcs	200701	0.67	0.67	0.43
Dall Mfg Avg Hrs	Level	Dallas Svcs	200701	1.04	0.76	0.63
ISM Mfg Emp	Level	ISM Mfg	196702	1.78	0.71	0.53
ISM Svcs Emp	Level	ISM Svcs	199707	1.40	0.69	0.54
JOLTS Quits	DLog	JOLTS	200101	0.85	0.26	0.07
JOLTS Priv Open	DLog	JOLTS	200101	1.34	0.29	0.80
JOLTS Priv Hires	DLog	JOLTS	200101	0.52	0.22	0.35
KC Mfg Emp	Level	KC Mfg	200107	1.30	0.77	0.59
KC Mfg Avg Hrs	Level	KC Mfg	200107	1.53	0.67	0.73
KC Svcs Emp	Level	KC Svcs	201401	1.99	0.62	0.88
NFIB Cant Fill	Diff	NFIB	198602	0.54	0.16	0.56
NFIB Few Qual App	Diff	NFIB	199305	0.45	0.12	0.47
NFIB Few Qual App	Diff	NFIB	198602	0.25	0.08	0.12
NFIB Avg Chg Emp	Diff	NFIB	198602	0.29	0.08	0.27
NY Mfg Emp	Level	NY Mfg	200107	1.53	0.75	0.63
NY Mfg Avg Hrs	Level	NY Mfg	200107	1.14	0.67	0.47
Rich Svcs Emp	Level	Rich Svcs	200409	0.79	0.62	0.37
Emp MiningLogg	DLog	Payroll	196702	0.41	0.17	0.30
Emp Constr	DLog	Payroll	196702	1.51	0.57	0.86
Emp Dur Manuf	DLog	Payroll	196702	1.89	0.73	0.83
Emp NDur Manuf	DLog	Payroll	196702	1.83	0.67	0.80
Emp Whole Trade	DLog	Payroll	196702	1.85	0.73	0.83
Emp Ret Trade	DLog	Payroll	196702	1.50	0.55	0.84
Emp Trans Ware	DLog	Payroll	197202	1.48	0.54	0.81
Emp Utilities	DLog	Payroll	196702	0.20	0.04	0.13
Emp Informat	DLog	Payroll	196702	0.49	0.17	0.59
Emp Fin Act	DLog	Payroll	196702	1.31	0.49	0.66
Emp PBS	DLog	Payroll	196702	1.92	0.73	0.91
Emp Educ+HCS	DLog	Payroll	196702	0.66	0.21	0.75

Table 5. Labor variable names and summary statistics (Part 3 of 3)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
Emp LeisHosp	DLog	Payroll	196702	1.29	0.46	0.73
Emp OthSvcs	DLog	Payroll	196702	1.19	0.43	0.80
Emp FedExCencus	DLog	Payroll	196702	0.01	0.01	0.16
Emp State	DLog	Payroll	196702	−0.01	−0.03	−0.11
Emp Local	DLog	Payroll	196702	0.26	0.07	0.36
AvHrs MiningLogg	DLog	Payroll	196702	0.42	0.11	0.63
AvHrs Constr	DLog	Payroll	196702	0.43	0.13	0.76
AvHrs Dur Manuf	DLog	Payroll	196702	0.73	0.19	0.77
AvHrs NDur Manuf	DLog	Payroll	196702	0.56	0.11	0.72
AvHrs Whole Trade	DLog	Payroll	197202	0.51	0.15	0.68
AvHrs Ret Trade	DLog	Payroll	197202	0.18	0.08	0.55
AvHrs Trans Ware	DLog	Payroll	197202	0.36	0.10	0.64
AvHrs Utilities	DLog	Payroll	197202	0.01	−0.02	−0.22
AvHrs Informat	DLog	Payroll	196702	0.28	0.07	0.53
AvHrs Fin Act	DLog	Payroll	196702	0.11	0.03	0.16
AvHrs PBS	DLog	Payroll	196702	0.28	0.08	0.59
AvHrs Educ+HCS	DLog	Payroll	196702	0.30	0.07	0.62
AvHrs LeisHosp	DLog	Payroll	196702	0.38	0.10	0.66
AvHrs OthSvcs	DLog	Payroll	196702	0.30	0.09	0.68
Payroll Diffusion	Level	Payroll	197701	2.03	0.81	0.87
Phil Mfg Emp	Level	Philly Mfg	196805	1.85	0.76	0.58
Phil Mfg Avg Hrs	Level	Philly Mfg	196805	1.91	0.75	0.78
Phil Svcs Emp	Level	Philly Svcs	201103	1.71	0.27	0.83
Rich Mfg Emp	Level	Richmond Mfg	199311	1.83	0.77	0.76
Rich Svcs Emp	Level	Richmond Svcs	199311	1.65	0.67	0.76

Note: Diff denotes first difference, DLog denotes first differenced log value, and Lev denotes no transformation. (W) denotes series that are weekly. "PC Wt" denotes first principal component weights after they have been standardized to sum to 100. "Monthly Corr" denotes correlation of the transformed monthly series with the labor (PC) after the latter has been approximately converted to monthly with the same aggregation pattern as the paired series. "Weekly Corr" denotes correlation of weekly interpolated (when applicable) and possibly backcasted values of series with the weekly labor PC.

Table 6. Activity variable names and summary statistics (Part 1 of 5)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
LumberProduction (W)	DLog	WestWoods	198102	0.22	0.09	0.42
LumberOrders (W)	DLog	WestWoods	198102	0.11	0.04	0.31
LumberShipments (W)	DLog	WestWoods	198102	0.21	0.08	0.37
CoalProduction (W)	DLog	EIA	198401	0.54	0.12	0.48
RailTraffic (W)	DLog	AAR	201101	1.54	0.20	0.71
JobPostings (W)	DLog	Indeed	202002	1.40	0.78	0.50
Init Claims US (W)	DLog	Claims	198102	0.74	0.39	0.34
Cont Claims US (W)	DLog	Claims	198102	1.08	0.61	0.57
RawSteelProd (W)	DLog	AISI	198201	0.76	0.38	0.46
ADP Priv Emp	DLog	ADP	201002	0.67	0.45	0.64
Retail I/S Ratio*	DLog	AdvEconInd	198102	0.19	0.06	0.79
LightTruck sale	DLog	Auto	198102	0.21	0.08	0.56
Auto sales	DLog	Auto	198102	0.29	0.11	0.75
Dom Auto sales	DLog	Auto	198102	0.30	0.11	0.74
Software Invest	DLog	BEA NIPA	200907	-0.21	-0.07	-0.21
Cass Freight SA	DLog	Cass	199002	0.47	0.21	0.71
CB ConsConf	Diff	CBoard	198102	0.37	0.21	0.63
CB ConsConf Pres	Diff	CBoard	198102	0.68	0.41	0.66
CB ConsConf Exp	Diff	CBoard	198102	0.10	0.06	0.53
ChallengGrayChrisSA	Diff	Challenger	198904	0.12	0.04	0.14
Priv Cons Nonres	DLog	Construct	200707	0.37	0.16	0.42
NewHousConst VPIP	DLog	Construct	198102	0.84	0.46	0.59
Pub Constr VPIP	DLog	Construct	198102	-0.01	-0.02	0.29
FedG Constr VPIP	DLog	Construct	198102	-0.10	-0.05	-0.20
StLoc Constr VPIP	DLog	Construct	198102	0.02	-0.01	0.35
UMich ConsSent	Diff	CSent	198102	0.10	0.03	-0.05
UM C Sent Exp	Diff	CSent	198102	0.06	0.01	0.04
UM C Sent Curr	Diff	CSent	198102	0.14	0.06	-0.17
Dall Mfg BusAct	Level	DallasMan	200406	0.73	0.78	0.29
Dall Svcs BusAct	Level	DallasSvc	200701	1.22	0.81	0.83
Autos Cons Sh*	Diff	Detail	198102	-0.06	-0.01	0.09
HeavyTruck sale	DLog	Detail	198102	0.45	0.23	0.72
Trucks Bus Sh	Diff	Detail	198702	0.25	0.11	0.41

Table 7. Activity variable names and summary statistics (Part 2 of 5)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
Trucks Bus Sh	Diff	Detail	198702	0.25	0.11	0.41
Exist 1Fam H Sale	DLog	ExistHomes	198102	0.18	0.06	0.66
valExHomeSales	DLog	ExistHomes	199601	0.21	0.08	0.37
NondefAirShp	DLog	ForeignTrade	199002	0.02	0.00	-0.42
ExportsComput	DLog	ForeignTrade	199101	0.20	0.08	0.38
CoreCapGoodsExp	DLog	ForeignTrade	200904	0.55	0.16	0.76
G ExpExCompCoreCap	DLog	ForeignTrade	199907	0.67	0.28	0.81
GoodsExports	DLog	ForeignTrade	199907	0.69	0.29	0.83
SvcExports	DLog	ForeignTrade	200907	0.70	0.21	0.81
ImportsComput	DLog	ForeignTrade	199101	0.23	0.09	0.53
CoreCapGoodsImp	DLog	ForeignTrade	200904	0.90	0.25	0.77
Imports Svcs BOP	DLog	ForeignTrade	200907	0.81	0.27	0.77
Imports Goods BOP	DLog	ForeignTrade	200004	0.62	0.26	0.78
G ImpExCompCoreCap	DLog	ForeignTrade	200004	0.61	0.26	0.77
IP Man	DLog	IP	198102	1.23	0.60	0.92
IP Tot	DLog	IP	198102	1.19	0.59	0.92
CapU Man	DLog	IP	198102	1.20	0.59	0.92
IP Fin Prod+Nonind Supp	DLog	IP	198102	1.13	0.55	0.92
IP Nonind Supp	DLog	IP	198102	1.24	0.64	0.93
IP Mat	DLog	IP	198102	1.06	0.53	0.91
IP Fin Prod	DLog	IP	198102	0.97	0.46	0.90
IP Bus Equip	DLog	IP	198102	0.99	0.49	0.91
IP Cons Goods	DLog	IP	198102	0.71	0.31	0.87
IP Dur Cons Goods	DLog	IP	198102	0.72	0.33	0.87
IP NDur Cons Goods	DLog	IP	198102	0.39	0.15	0.75
CapU Total	DLog	IP	198102	1.19	0.59	0.92
IP Oil Drill	DLog	IP	198102	0.49	0.29	0.24
ISM Mfg	Level	ISMMan	198102	1.34	0.79	0.49
ISM Mfg Prod	Level	ISMMan	198102	1.30	0.76	0.50
ISM Mfg Emp	Level	ISMMan	198102	1.10	0.65	0.40
ISM Mfg Norder	Level	ISMMan	198102	1.28	0.76	0.42
ISM Mfg Invent	Level	ISMMan	198102	0.83	0.51	0.36
ISM Mfg SupDel	Level	ISMMan	198102	0.94	0.59	0.42
ISM Mfg Exports	Level	ISMMan	198801	0.86	0.50	0.36
ISM Mfg Imports	Level	ISMMan	198910	1.17	0.68	0.70

Table 8. Activity variable names and summary statistics (Part 3 of 5)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
ISM Svcs	Level	ISMSvc	199707	1.28	0.86	0.81
ISM Svcs Bus Act	Level	ISMSvc	199707	1.28	0.82	0.86
ISM Svcs NewOrd	Level	ISMSvc	199707	1.37	0.82	0.90
ISM Svcs Emp	Level	ISMSvc	199707	1.12	0.71	0.65
ISM Supp Del	Level	ISMSvc	199707	0.38	0.52	0.03
ISM Svs Exp SA	Level	ISMSvc	199707	1.00	0.55	0.76
ISM Svs Imp SA	Level	ISMSvc	199707	0.82	0.50	0.57
KC Mfg Compos	Level	KCMan	200107	0.59	0.74	0.38
KC Svcs Compos	Level	KCSvc	201401	1.48	0.58	0.96
Emp Priv Ind	DLog	Labor	198102	1.44	0.83	0.95
Emp Goods	DLog	Labor	198102	1.44	0.85	0.96
Emp Nonfarm	DLog	Labor	198102	1.41	0.81	0.95
Emp Manuf	DLog	Labor	198102	1.38	0.81	0.88
Emp Dur Manuf	DLog	Labor	198102	1.34	0.79	0.86
Emp Priv Svcs	DLog	Labor	198102	1.24	0.71	0.91
Emp Ret Trade	DLog	Labor	198102	1.01	0.54	0.83
Emp Whole Trade	DLog	Labor	198102	1.21	0.72	0.65
Emp NDur Manuf	DLog	Labor	198102	1.11	0.64	0.74
Emp Constr	DLog	Labor	198102	1.14	0.66	0.92
CPS Emp Nonag	DLog	Labor	198102	0.85	0.47	0.80
CPS 16+	DLog	Labor	198102	0.84	0.48	0.81
UR 25-54 Male	Diff	Labor	198102	0.91	0.50	0.67
Emp Fin Act	DLog	Labor	198102	0.89	0.50	0.65
Avg Hrs Prod Manuf	Diff	Labor	198102	0.53	0.24	0.86
Avg OT Prod Manuf	Diff	Labor	198102	0.47	0.21	0.70
Emp Mining	DLog	Labor	198102	0.34	0.21	0.53
Unemp Duration	Diff	Labor	198102	0.24	0.16	0.22
Emp PBS	DLog	Labor	198102	1.31	0.76	0.90
Emp Educ+HCS	DLog	Labor	198102	0.25	0.11	0.67
Emp LeisHosp	DLog	Labor	198102	0.68	0.39	0.43
Emp OthSvcs	DLog	Labor	198102	0.75	0.41	0.75
Avg Hrs Prod Tot	DLog	Labor	198102	0.49	0.23	0.87
URJobLoserLayoff*	Diff	Labor	198102	0.53	0.28	0.67
Unemp Rate*	Diff	Labor	198102	1.01	0.56	0.81
LFPR	DLog	Labor	198102	0.21	0.12	0.52
Emp CompSysDes	DLog	Labor	199002	0.67	0.43	0.50
Emp Constr ResBuil	DLog	Labor	198502	1.08	0.59	0.87
Emp Res Remod Prid	DLog	Labor	199002	0.76	0.38	0.83
Fed Govt Emp	DLog	Labor	198102	0.05	-0.01	0.15
Emp StateLocal	DLog	Labor	198102	0.20	0.07	0.53
Emp DOD	DLog	Labor	198102	-0.07	-0.05	-0.13
Emp ProfSciTech	DLog	Labor	198302	1.10	0.66	0.81

Table 9. Activity variable names and summary statistics (Part 4 of 5)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
DGoods NewOrd	DLog	M31	200904	0.13	0.04	0.48
Mfg Shipments	DLog	M31	198102	0.67	0.32	0.83
Mfg I/S Ratio*	DLog	M31	198102	0.58	0.25	0.85
CapGoodsShip	DLog	M31	198102	0.69	0.37	0.82
CoreCapGoodsShip	DLog	M31	200904	0.91	0.26	0.76
ComputerShip	DLog	M31	199101	0.20	0.08	0.57
DefensShipments	DLog	M31	198102	-0.04	-0.04	0.61
Mob Home U Ship	Diff	MobileHomes	198102	0.29	0.12	0.71
Mob Home Val Ship	DLog	MobileHomes	198102	0.38	0.17	0.73
Fed Out NatDed SA	DLog	MTS	198202	0.00	0.01	0.26
Fed Out Tot SA	DLog	MTS	198102	0.00	0.02	0.53
NAHB HMI	Level	NAHB	198501	0.87	0.58	0.21
NewHomeMthSupp	DLog	NewHomes	198102	0.06	0.07	-0.63
NewHomeForSale	Diff	NewHomes	198102	0.76	0.49	-0.05
Exist 1Fam H Sold	Diff	NewHomes	198102	0.10	0.03	0.71
valNewHomeSales	DLog	NewHomes	199601	0.18	0.07	0.52
NFIB Sm BusOpt	Level	NFIB	198102	0.91	0.58	0.37
NatlRestaurAss Perf	Level	NRA	200206	1.21	0.72	0.80
NY Mfg BusCond	Level	NYMan	200107	1.12	0.66	0.88
NY Svcs BusAct	Level	NYSvc	200409	0.67	0.63	0.15
Real PI Ex Transf	DLog	PersIncPCE	198102	0.57	0.31	0.76
Real DPI	DLog	PersIncPCE	198102	0.20	0.10	0.84
Real PCE	DLog	PersIncPCE	198102	0.57	0.27	0.85
Real PCE DGoods	DLog	PersIncPCE	198102	0.36	0.15	0.82
Real PCE Svcs	DLog	PersIncPCE	198102	0.54	0.28	0.70
Real PCE NDGoods	DLog	PersIncPCE	198102	0.39	0.18	0.78
WageSal Disburs Gov	DLog	PersIncPCE	198102	-0.19	-0.16	0.17

Table 10. Activity variable names and summary statistics (Part 5 of 5)

Description	Trans-form	Source	Start	PC Wt	Monthly Corr	Weekly Corr
Philly Mfg Future	Level	PhillyMfg	198102	0.32	0.17	0.27
Philly Mfg Curr	Level	PhillyMfg	198102	1.30	0.78	0.49
Phil Svcs Activ	Level	PhillySvc	201103	0.69	0.55	0.47
JohnsonRebook	DLog	Redbook	199802	0.26	0.10	0.46
CoreRetailSales	DLog	Retail	198102	0.58	0.28	0.85
Ret+FoodSvc Sale	DLog	Retail	198102	0.45	0.20	0.85
Ret Sales Res Eq	DLog	Retail	199801	0.45	0.18	0.84
Ret Sales BuildSupp	DLog	Retail	200905	0.39	0.14	0.86
Rich Mfg Compos	Level	RichmndMan	199311	1.18	0.70	0.66
Rich Svcs Rev	Level	RichmndSvc	199311	0.92	0.63	0.55
SP500	DLog	SP500	198102	0.21	0.12	0.35
SP500 DivYld	Diff	SP500	198102	0.12	0.06	0.48
H Starts	Diff	Starts	198102	0.25	0.10	0.76
H Permits	Diff	Starts	198102	0.27	0.11	0.72
H Starts W	Diff	Starts	198102	0.09	0.04	0.11
H Starts S	Diff	Starts	198102	0.20	0.08	0.65
H Starts MW	Diff	Starts	198102	0.16	0.05	0.76
H Starts NE	Diff	Starts	198102	0.11	0.05	0.44
H Starts 1 Unit	Diff	Starts	198102	0.24	0.09	0.74
Val Starts 1 Unit	DLog	Starts	198102	0.28	0.11	0.73
H Starts MF	Diff	Starts	198102	0.16	0.07	0.60
Hous Complet	DLog	Starts	198102	0.20	0.11	0.07
Clev WARN Fact	Diff	WARN	199608	0.27	0.10	0.61
Merch W Sales	DLog	Wholesale	198102	0.60	0.28	0.84
Whol I/S Ratio*	DLog	Wholesale	198102	0.51	0.22	0.84

Note: Diff denotes first difference, DLog denotes first differenced log value, and Lev denotes no transformation. (W) denotes series that are weekly. “PC Wt” denotes first principal component weights after they have been standardized to sum to 100. “Monthly Corr” denotes correlation of the transformed monthly series with the labor (PC) after the latter has been approximately converted to monthly with the same aggregation pattern as the paired series. “Weekly Corr” denotes correlation of weekly interpolated (when applicable) and possibly backcasted values of series with the weekly activity PC.

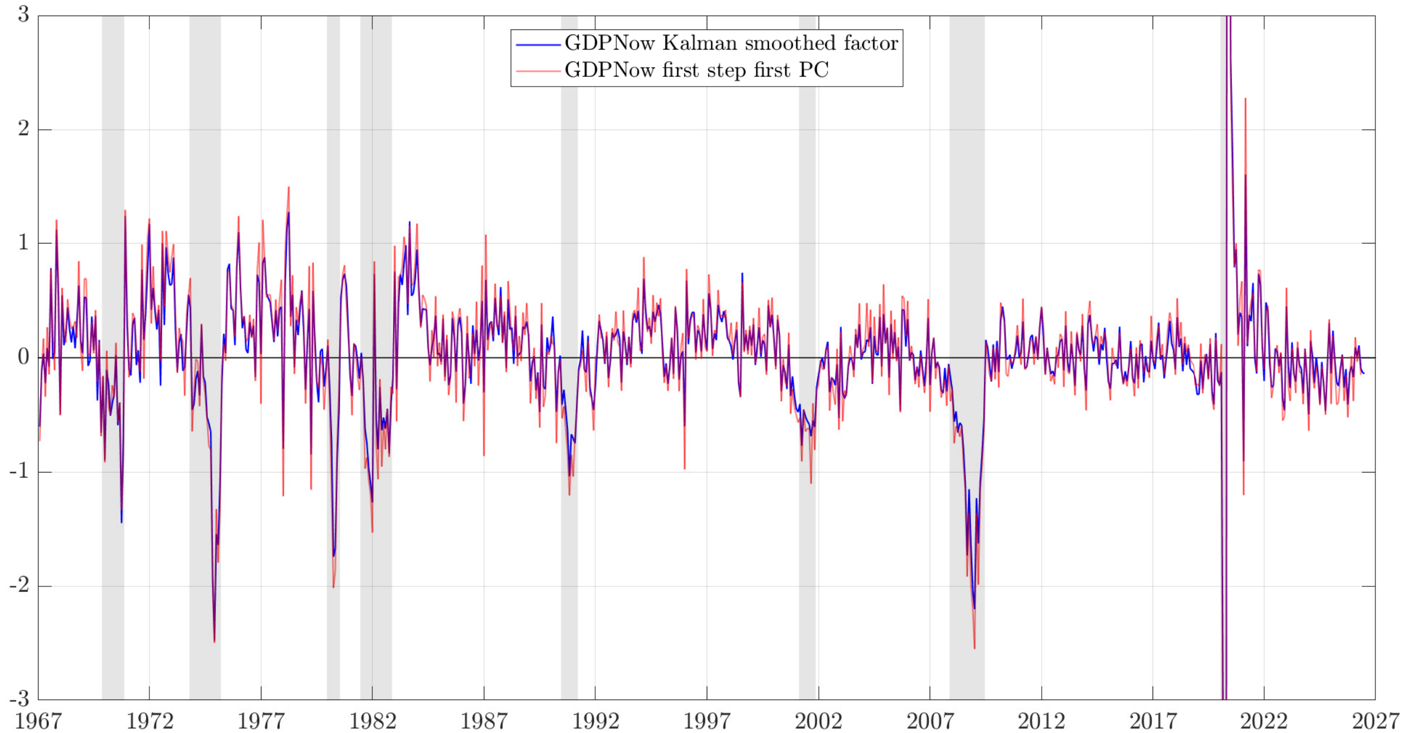

FIGURE 1. The blue line is the smoothed factor used by the standard GDPNow model on August 7, 2026 used by the model on that date. Forecast values beyond 2026m7 have been removed. The red line is the August 7, 2026 EM algorithm estimate of first principal component that is used estimate the parameters of the Kalman filter/smoother in GDPNow's factor model. This series ends in 2026m5. Both series are standardized to have mean 0, standard deviation 1.

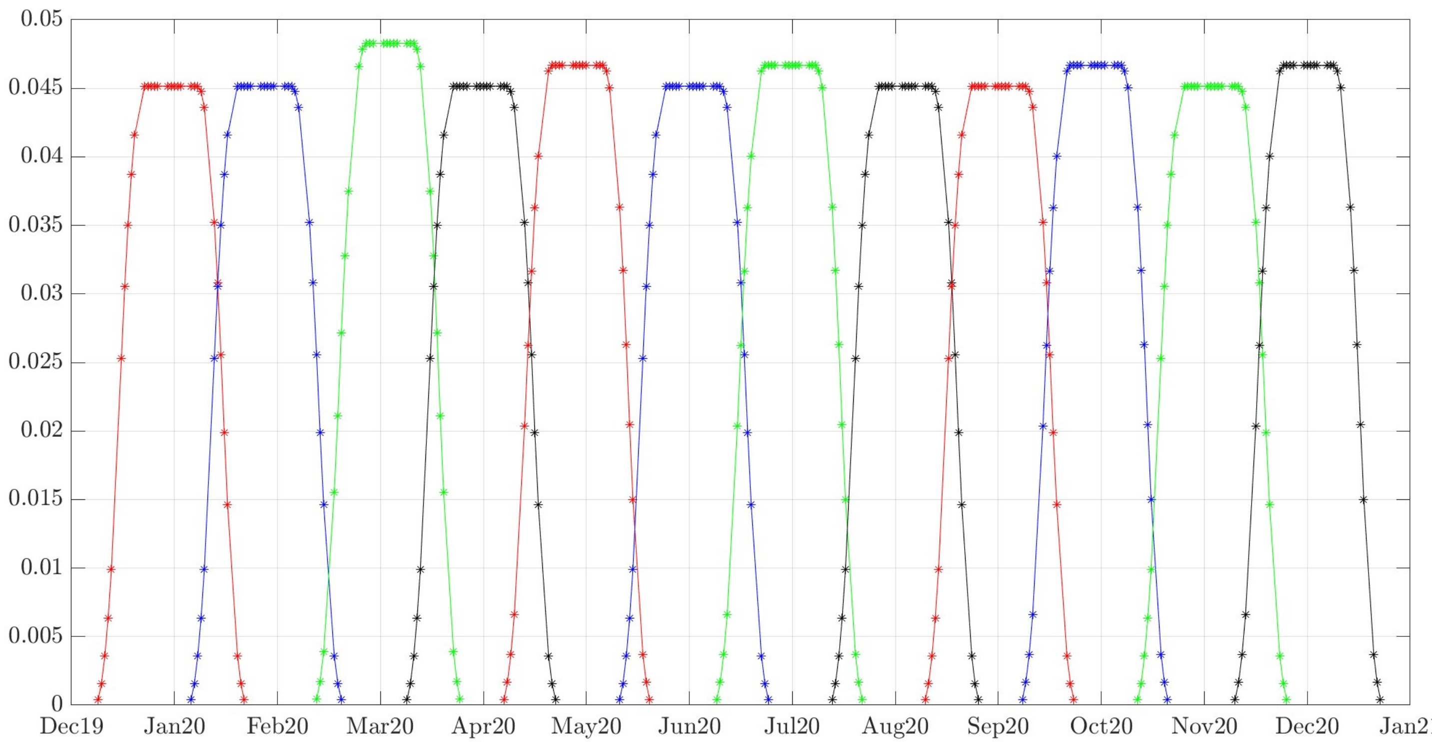

FIGURE 2. Dallas Fed's Texas Manufacturing Outlook Survey weights for each business day for each of January 2020 through December 2020. Each distinctly colored hump shaped curve corresponds to the weights for a particular survey month. The x-axis labels correspond to the first calendar of each month day and the y-axis label denotes the weight of a particular business day for a particular survey month. For each curve, the values of the hatch-marked weights sum to 1.0.

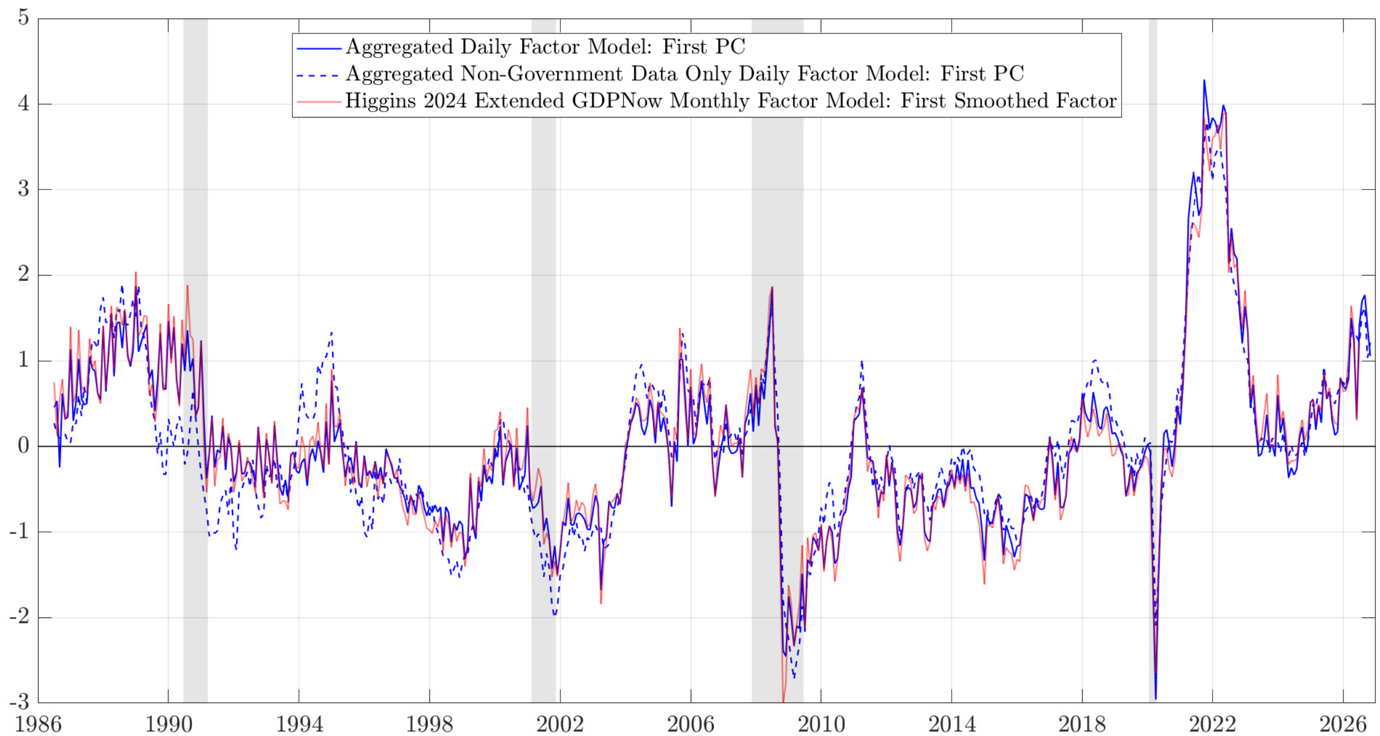

Figure 3. The red solid line is the first smoothed factor estimated using a monthly dynamic factor very similar to the one in Higgins [2024] that ends in 2026m7. The blue solid line is its first principal component analog for the daily PC-based factor model developed here that has been aggregated to be consistent with inflation rates for monthly averages. The blue dashed line is the first PC estimated using the daily and non-government price series only. Both blue lines are extended to end in 2026m11. All series are standardized to have mean 0, standard deviation 1 with data collected on August 7, 2026.

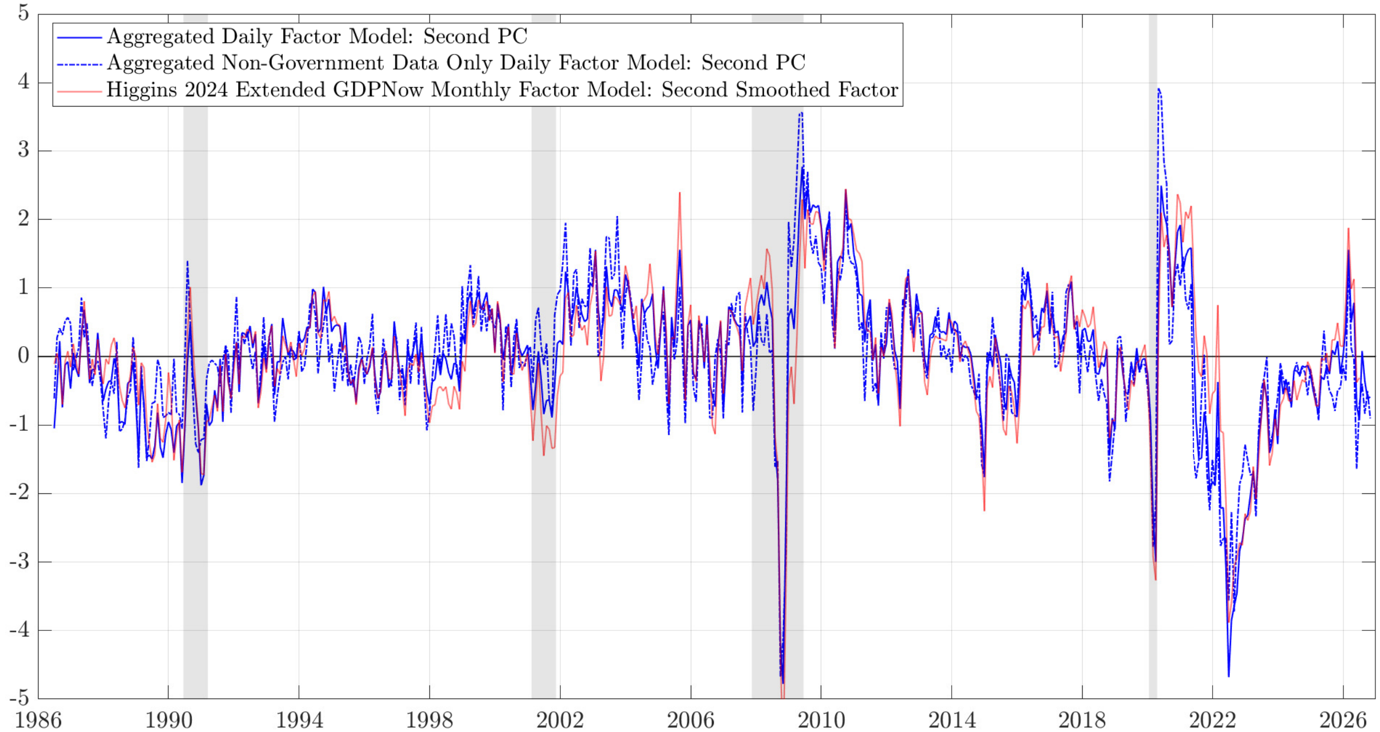

Figure 4. The red solid line is the second smoothed factor estimated using a monthly dynamic factor very similar to the one in Higgins [2024] that ends in 2026m7. The blue solid line is its second principal component analog for the daily PC-based factor model developed here that has been aggregated to be consistent with inflation rates for monthly averages. The blue dashed line is the second PC estimated using the daily and non-government price series only. Both blue lines are extended to end in 2026m11. All series are standardized to have mean 0, standard deviation 1 with data collected on August 7, 2026.

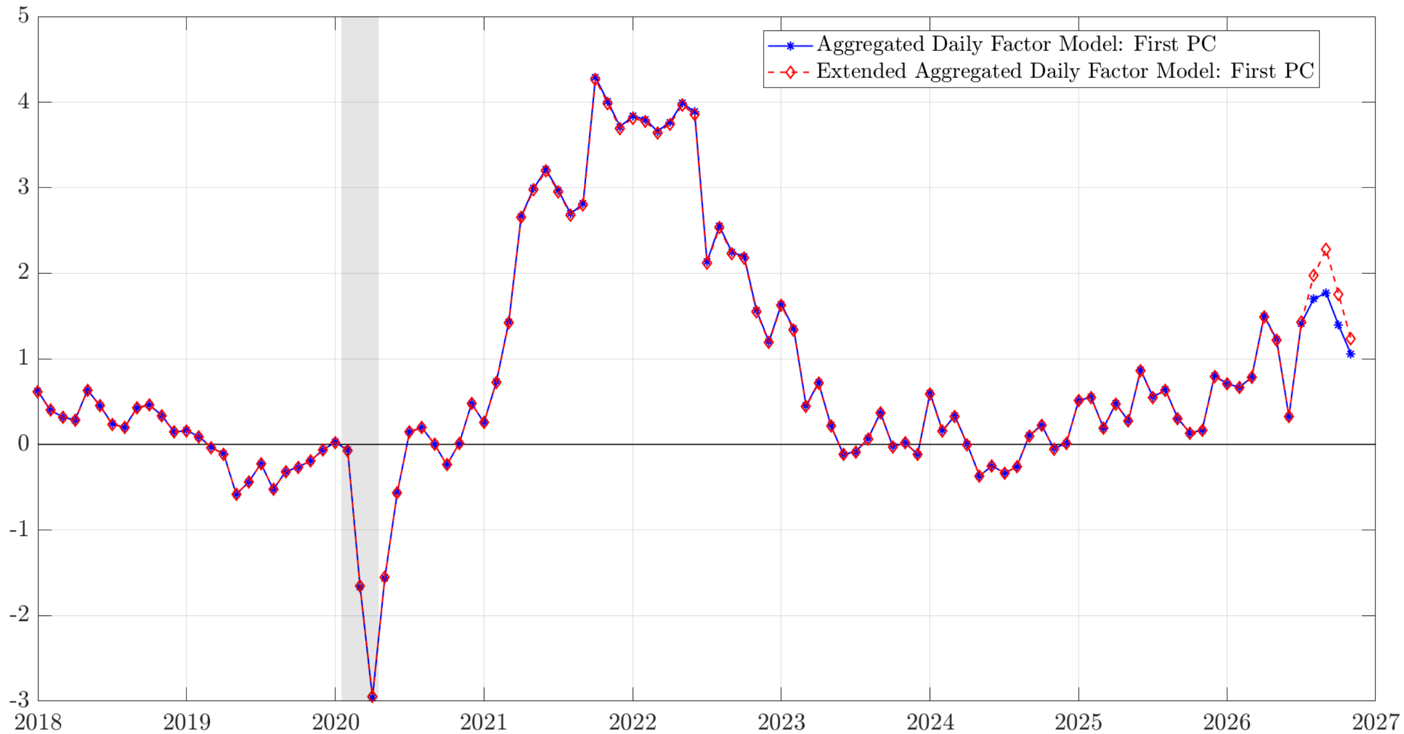

FIGURE 5. The blue solid line is the first principal component in the daily-monthly mixed frequency based price factor model developed here that has been aggregated to be consistent with inflation rates for monthly averages. It ends with data collected on August 7, 2026. The red dashed line is a counterfactual extension of that series that imposes shocks that puts daily inflation for the 8 daily commodity prices included in the model 7 standard deviations above their mean on August 10, 2026 and sets them to 0 for the remaining 4 days of that week. The two series plotted above are standardized to have mean 0, standard deviation 1 over the 1986m7-2026m11 (extended) sample.

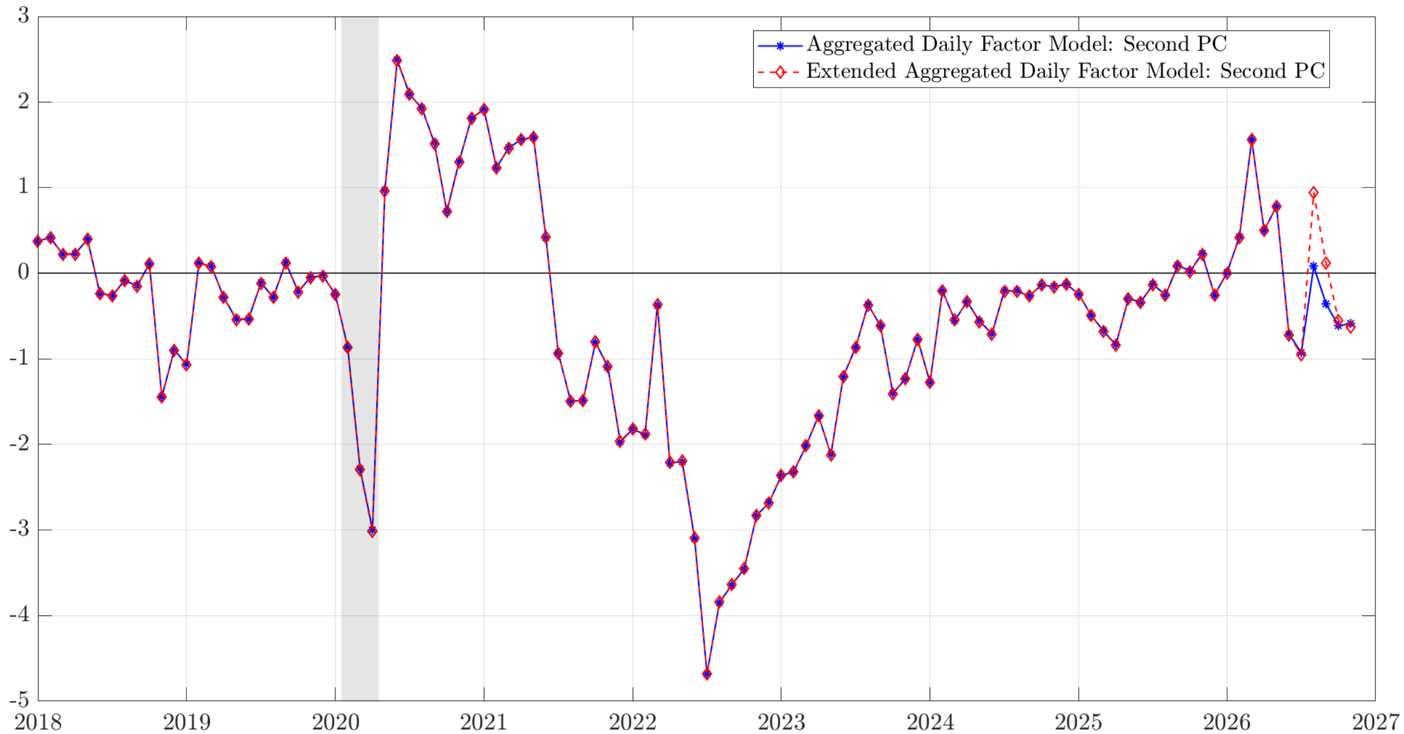

Figure 6. The blue solid line is the second principal component in the daily-monthly mixed frequency based price factor model developed here that has been aggregated to be consistent with inflation rates for monthly averages. It ends with data collected on August 7, 2026. The red dashed line is a counterfactual extension of that series that imposes shocks that puts daily inflation for the 8 daily commodity prices included in the model 7 standard deviations above their mean on August 10, 2026 and sets them to 0 for the remaining 4 days of that week. The two series plotted above are standardized to have mean 0, standard deviation 1 over the 1986m7-2026m11 (extended) sample.

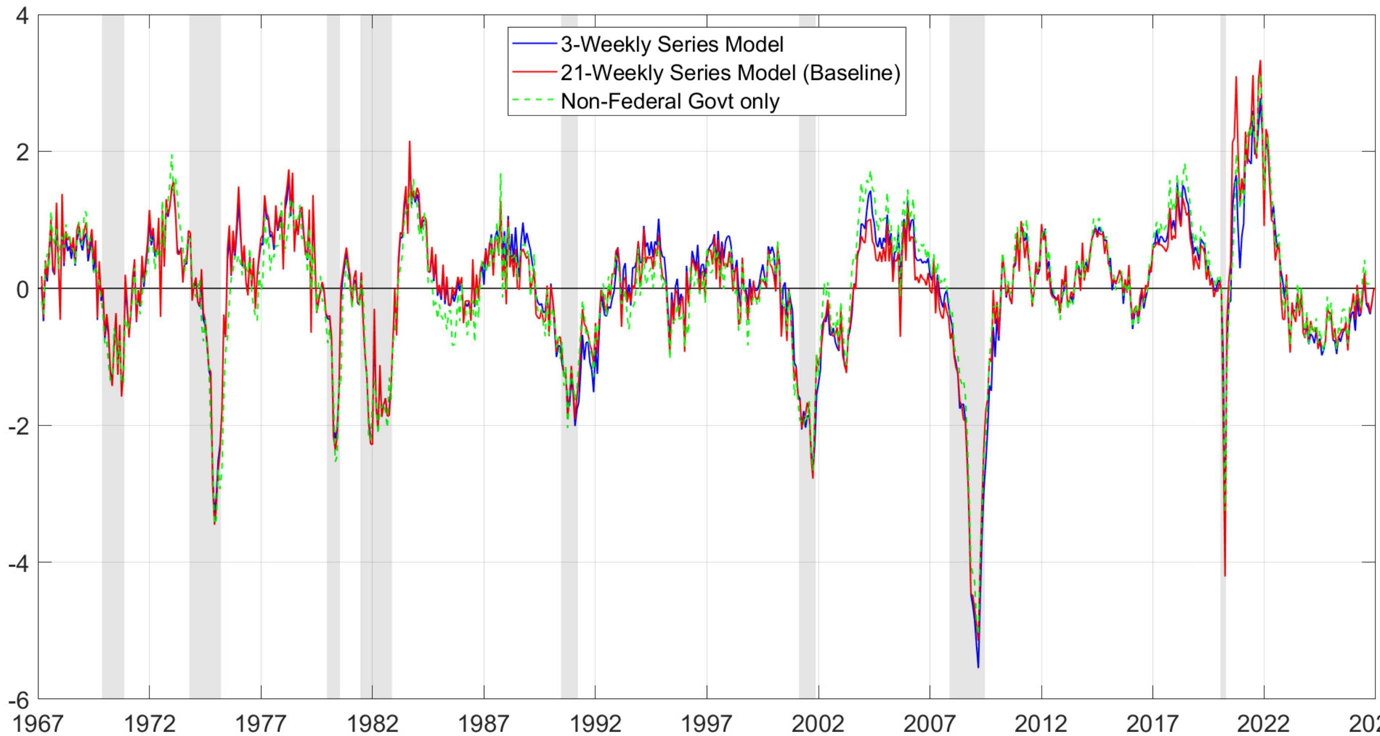

FIGURE 7. The blue solid line is the factor/PC from the monthly plus 3-weekly variable labor market variables factor model while the red line factor/PC also incorporates weekly initial and continuing claims for each of the 9 Census Divisions. The green dashed line is from a factor/PC model using only those variables – excluding unemployment insurance claims since those are largely derived from state labor agency data – that are not sourced from the federal government. Rather than 18 regional series on initial and continuing claims, this model only has both claims series for each of the 4 Census Regions. All series are converted from weekly to monthly-average aggregation as in equation 4 and are standardized to have mean zero and standard deviation 1 over the 1967m3 to 2026m10 forecast padded sample period and are estimated with data through the morning of August 11, 2026.

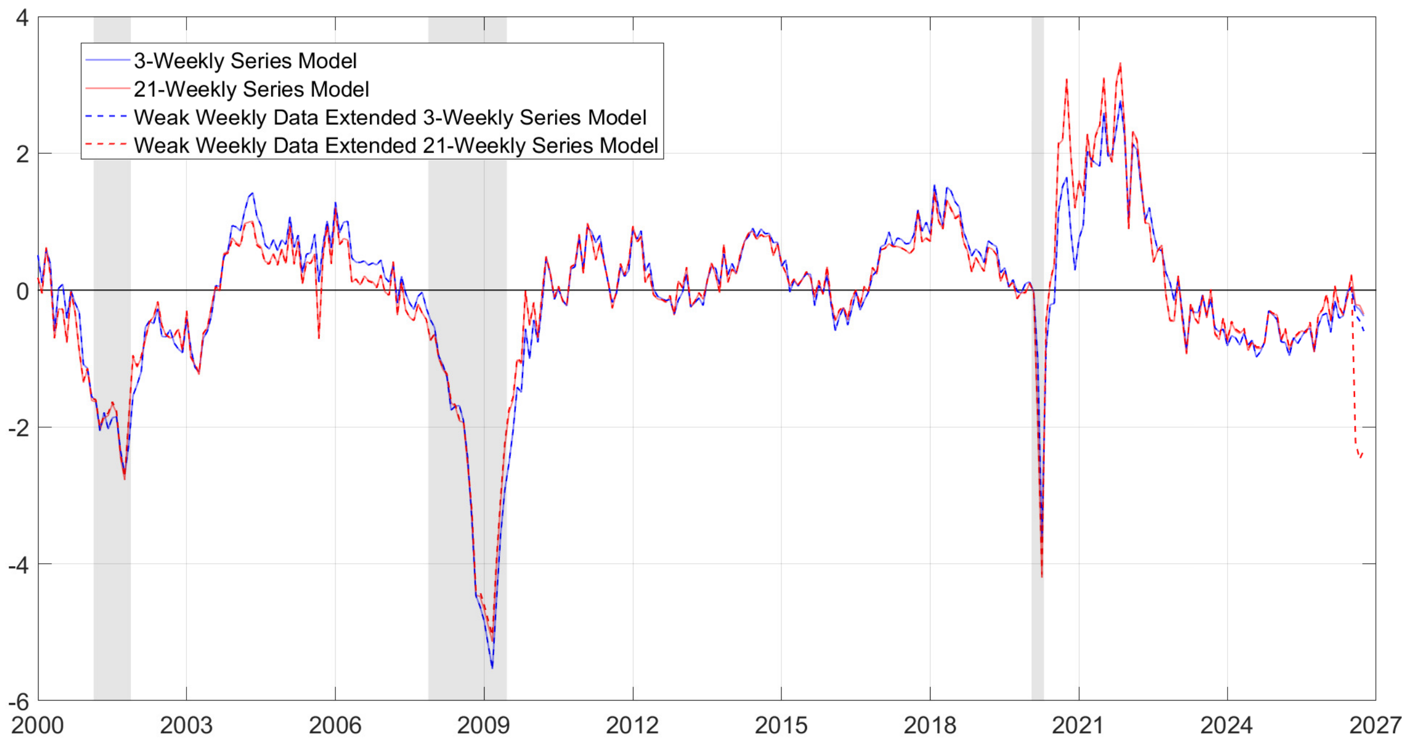

FIGURE 8. The blue solid line is the factor/PC from the monthly plus 3-weekly variable labor market variables factor model while the red solid line factor/PC also incorporates weekly initial and continuing claims for each of the 9 Census Divisions. The dashed lines are estimating by extending the weekly series through November so that – as described in the text – initial claims ratchet up from about 200k per week at the beginning of August to 400k per week at the end of August and remain there for the rest of the month. All series are converted from weekly to monthly-average aggregation as in equation 4 and are standadized to have mean zero and standard deviation 1 over the 1967m3 to 2026m10 forecast padded sample.

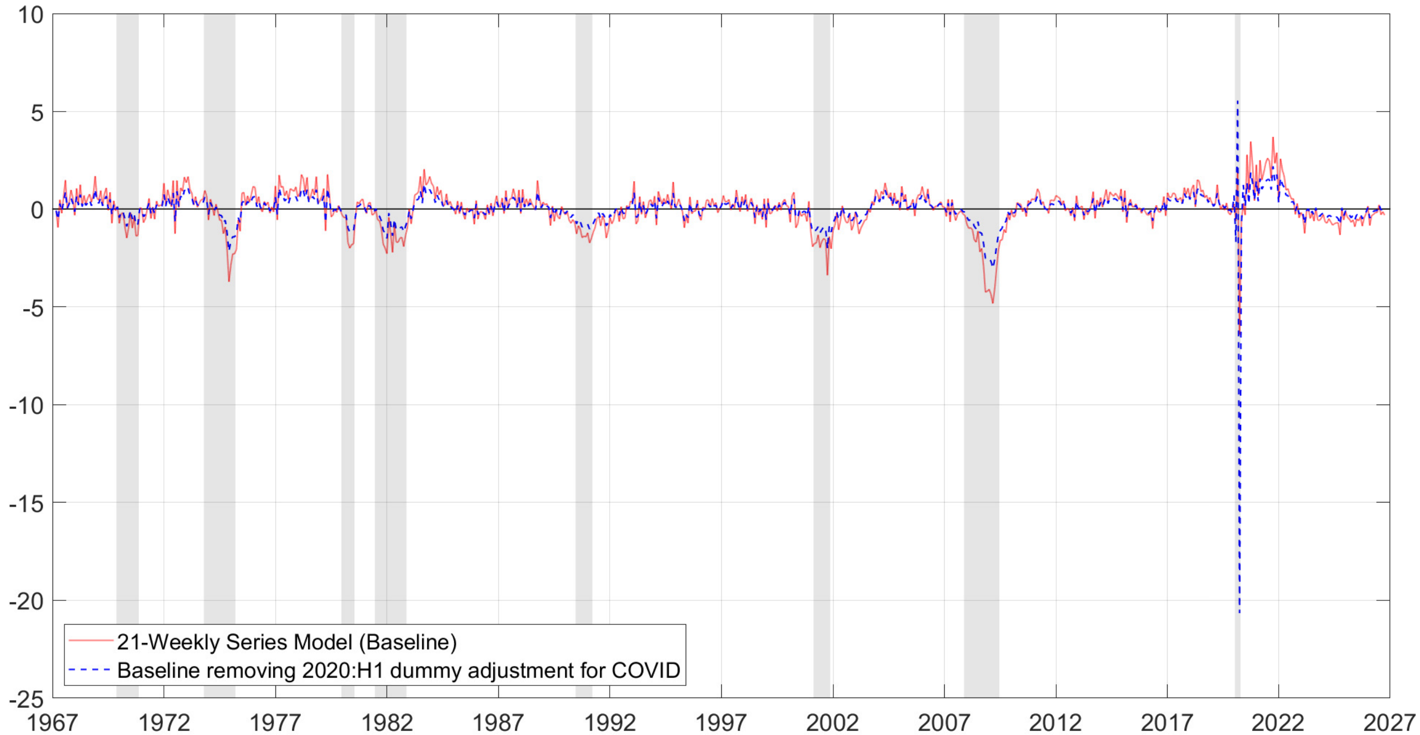

FIGURE 9. The red solid line is the factor/PC from the benchmark standard monthly plus 21-weekly variable labor market variables factor model. As described in the text, the blue dashed line maintains the PC factor coefficients from the benchmark model but removes the dummy variable interacted scaling factors for 2020:H1 used to standardize the series. Both series are converted from weekly to monthly-average aggregation as in equation 4, standadized to have mean zero and standard deviation 1 over the sample 1967m3 to 2026m10 forecast padded sample period, and are estimated with data through the morning of August 11, 2026.

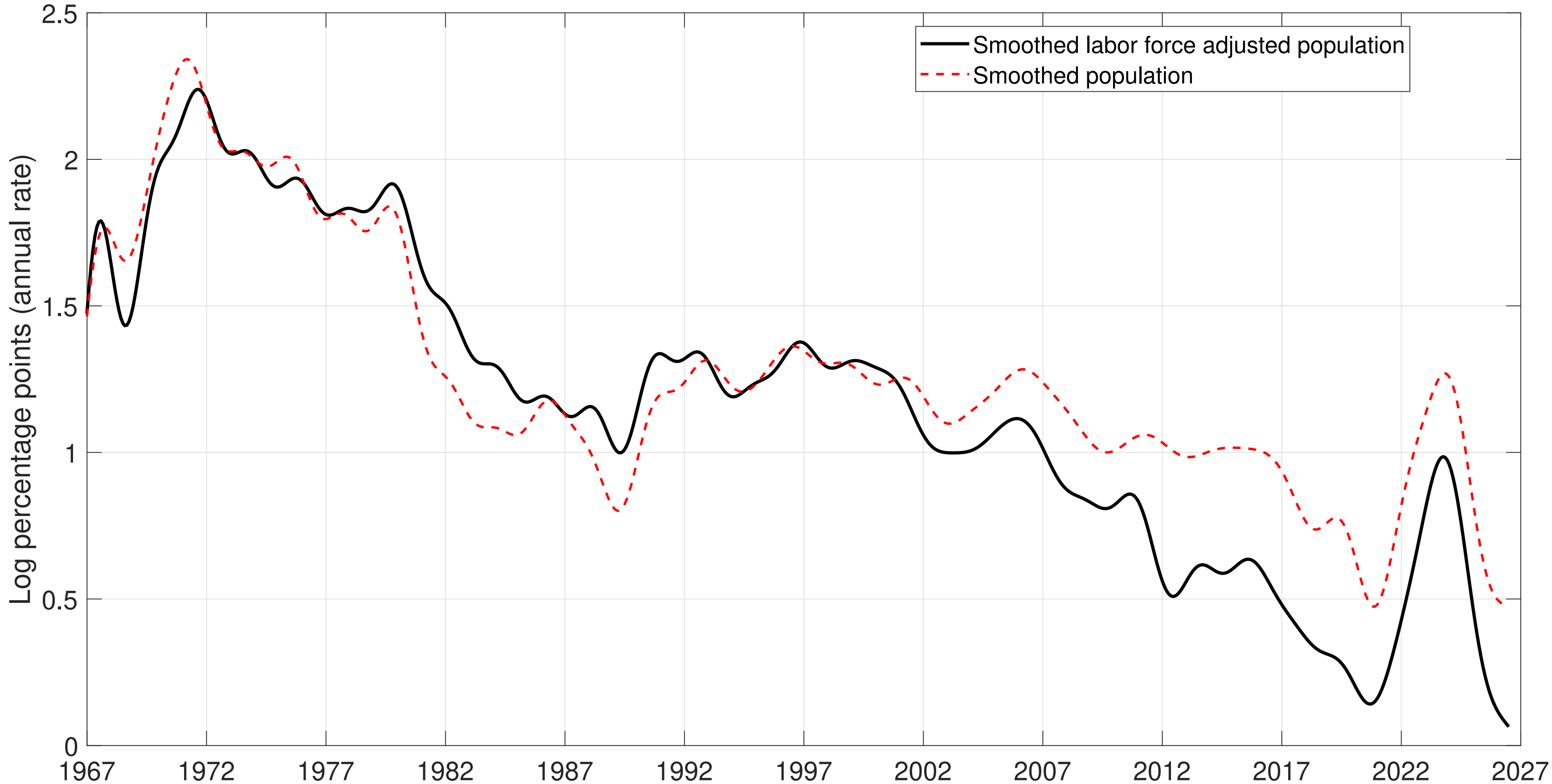

FIGURE 10. The red dashed line is a smoothed estimate of population growth derived from Coglianese et al. [2025] as described in the text. The black line further adjusts by weighting population growth for 16 age-sex cells for the age 16+ population using labor force shares. It is used to scale the person/aggregate hours series used to estimate the thick solid blue line in Figure 11.

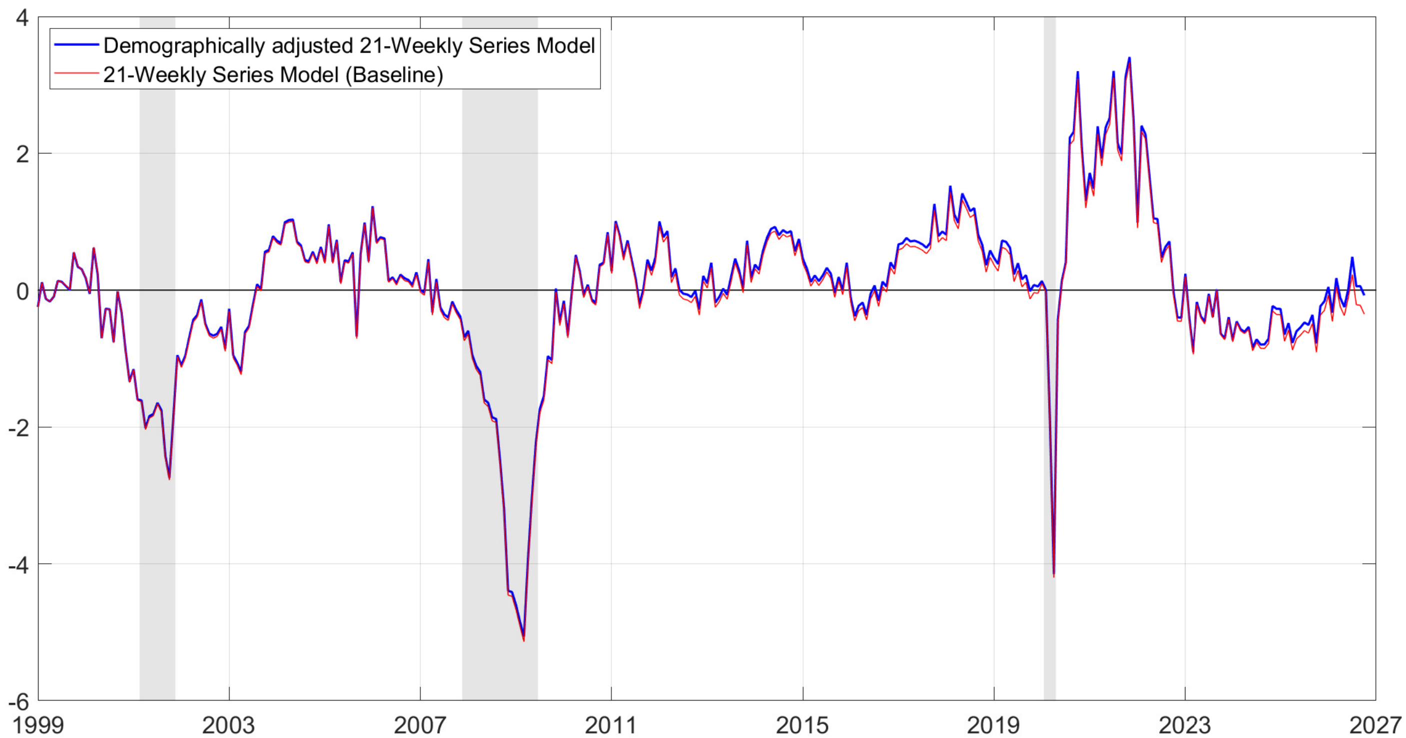

Figure 11. The red solid line is the factor/PC from the benchmark standard monthly plus 21-weekly variable labor market variables factor model. The thicker blue solid estimates that same collection of labor market variables after scaling a subset of the person/hours count data series with the measure of smoothed labor force attachment adjusted population growth shown as the black line in Figure 10. Both series are converted from weekly to monthly-average aggregation as in equation 4 and are standadized to have mean zero and standard deviation 1 over the sample 1967m3 to 2026m10 forecast padded sample period and are estimated with data through the morning of August 11, 2026.

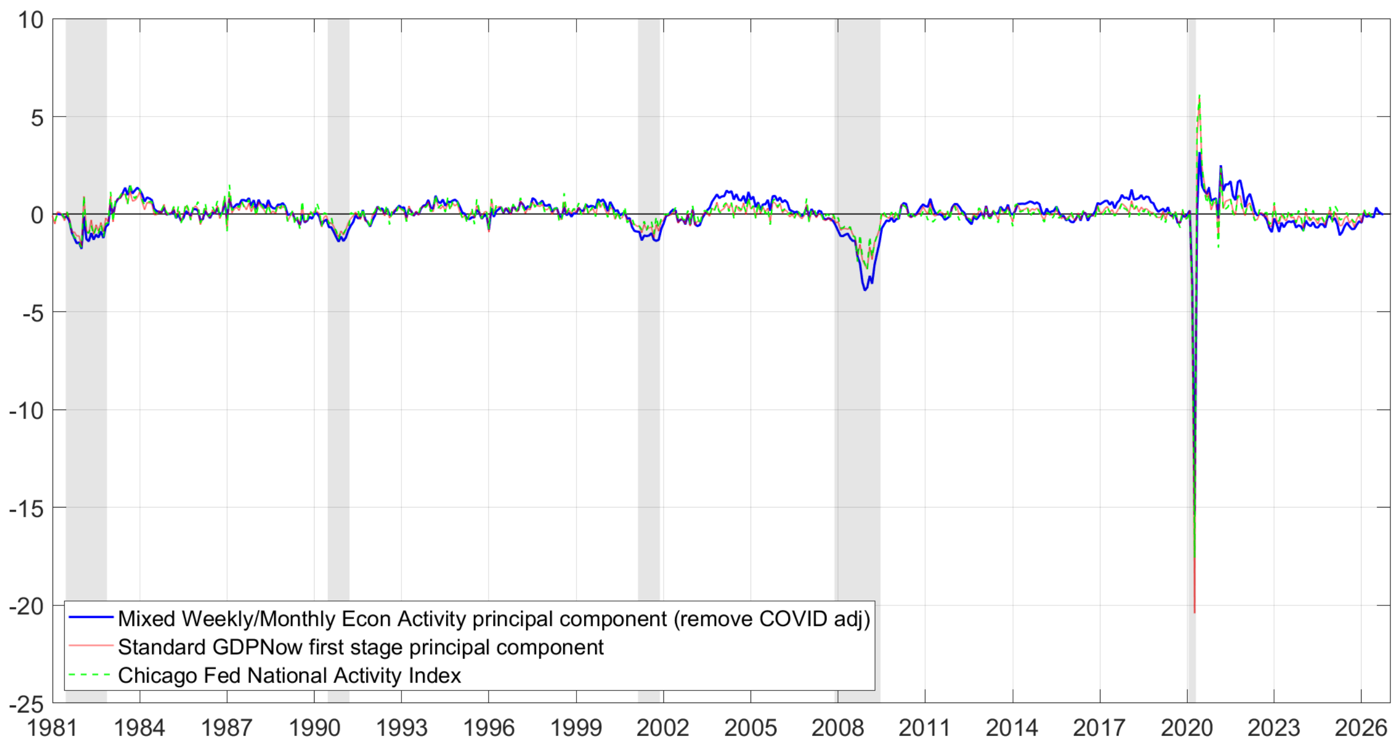

FIGURE 12. The thick solid blue line is the aggregated weekly to monthly principal component like indicator estimated by using the PC coefficient loadings from the COVID 2020:H1 dummy adjusted model and applying them to the data standardized without the COVID dummy adjustment. The thin red line is the first stage PC estimated on August 7, 2026 in the standard GDPNow factor model. The dashed green line is the Chicago Fed National Activity. All three series have been standardized to have mean 0, standard deviation 1 over the sample from 1981m2 through their last observation in 2026.

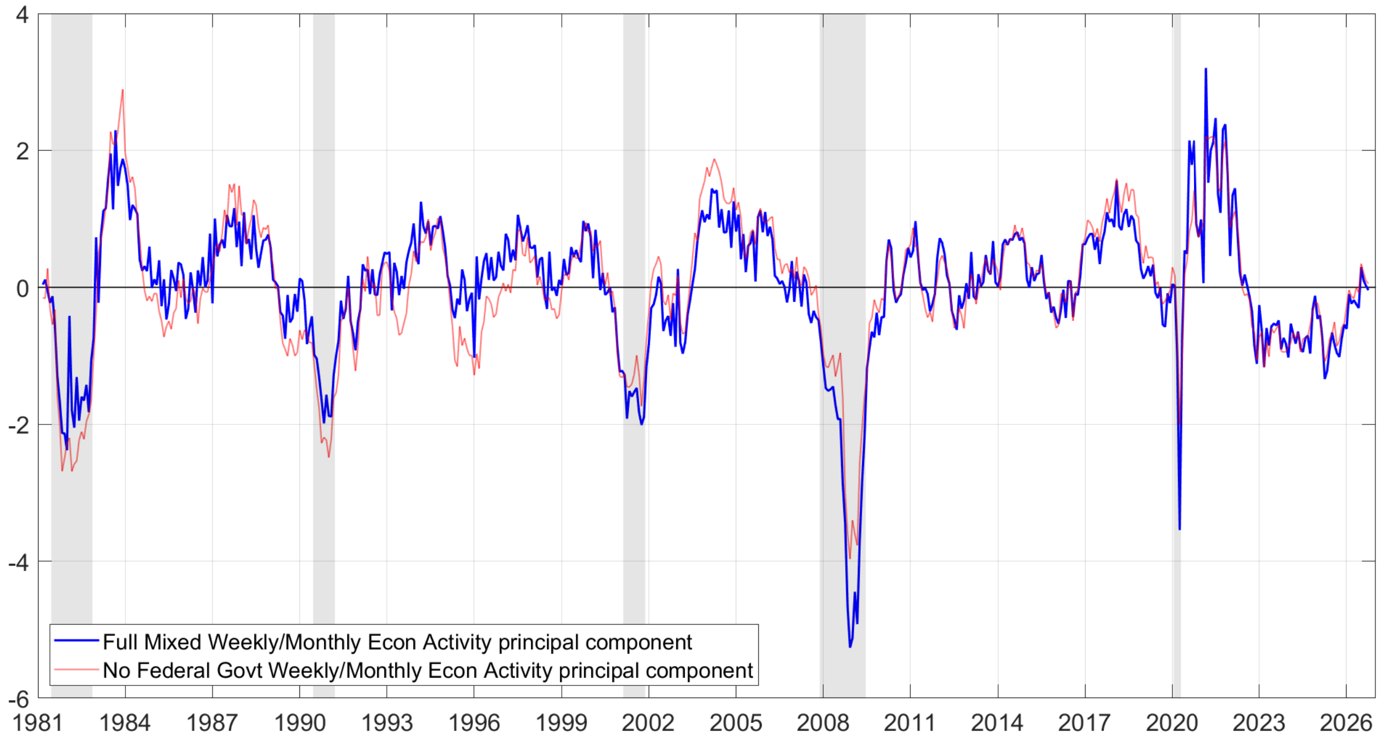

FIGURE 13. The thick solid blue line is the (standard) aggregated weekly to monthly principal component like indicator estimated by using dummy variables in 2020:H1 both to standardize the data and to estimate factor dynamics during and around that period. The thinner solid red line is an alternative measure that removes all but 47 of the 152 monthly indicators that do not come from the federal government. In both models, the 9 weekly indicators are retained including initial and continuing claims because they can largely be reproduced by data from state agencies nearly all of which continued reporting in the 2025 government shutdown.

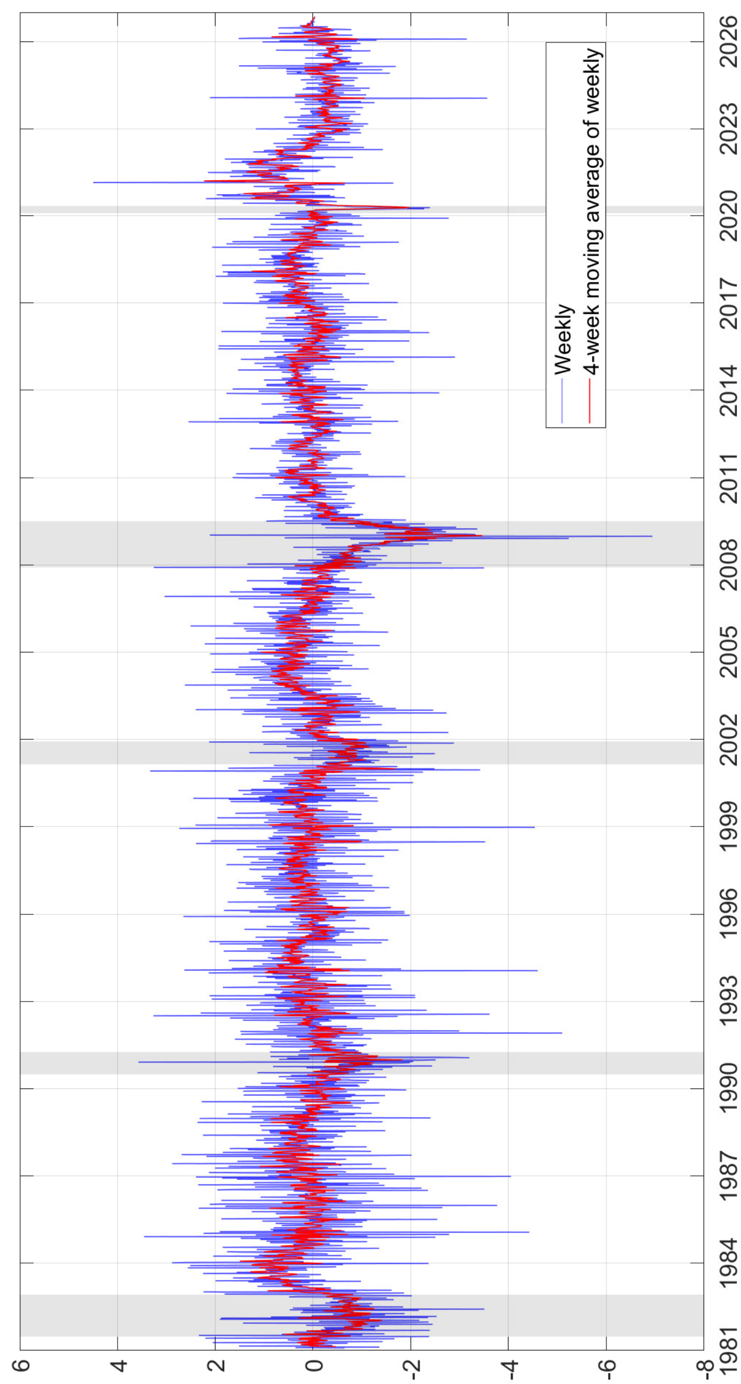

FIGURE 14. The solid semi-transparent blue line is the standard full activity model mixed monthly and weekly principal component at its highest (weekly) frequency that has been standardized to have mean 0, standard deviation 1. The red line on top of it is the (un-standardized) 4-week trailing average of that series.