NBER WORKING PAPER SERIES

BUSINESS APPLICATIONS AS A LEADING ECONOMIC INDICATOR?

Jose Asturias
Emin Dinlersoz
John C. Haltiwanger
Rebecca J. Hutchinson
Alyson Plumb

Working Paper 33224
http://www.nber.org/papers/w33224

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2024

Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau or the National Bureau of Economic Research. All results in this paper use publicly available data from Census Bureau websites. The Census Bureau has reviewed the public domain products for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release (Approval IDs: CBDRB-FY24-0148, CBDRB-FY24-0150). John Haltiwanger was also a Schedule A employee of the Bureau of the Census at the time of the writing of this paper. We thank John Abowd, Lucia Foster, Ron Jarmin, Shawn Klimek, Nick Orsini, Amy Newman Smith, Stephanie Studds, Kirk White, and Erin Wrona for useful comments and suggestions.

Business Applications as a Leading Economic Indicator?
Jose Asturias, Emin Dinlersoz, John C. Haltiwanger, Rebecca J. Hutchinson, and Alyson Plumb
NBER Working Paper No. 33224
December 2024
JEL No. C82, E01, L26

ABSTRACT

How are applications to start new businesses related to aggregate economic activity? This paper explores the properties of three monthly business application series from the U.S. Census Bureau's Business Formation Statistics as economic indicators: all business applications, business applications that are relatively likely to turn into new employer businesses ("likely employers"), and the residual series -- business applications that have a relatively low rate of becoming employers ("likely non-employers"). Growth in applications for likely employers significantly leads total nonfarm employment growth and has a strong positive correlation with it. Furthermore, growth in applications for likely employers leads growth in most of the monthly Principal Federal Economic Indicators (PFEIs). Motivated by our findings, we estimate a dynamic factor model (DFM) to forecast nonfarm employment growth over a 12-month period using the PFEIs and the likely employers series. The latter improves the model's forecast, especially in the years following the turning points of the Great Recession and the COVID-19 pandemic. Overall, applications for likely employers are a strong leading indicator of monthly PFEIs and aggregate economic activity, whereas applications for likely non-employers provide early information about changes in increasingly prevalent self-employment activity in the U.S. economy.

Jose Asturias
U.S. Census Bureau
4600 Silver Hill Rd
Suitland, MD 20746
jose.asturias@census.gov

Emin Dinlersoz
U.S. Census Bureau
4600 Silver Hill Rd
Suitland, MD 20746
emin.m.dinlersoz@census.gov

John C. Haltiwanger
Department of Economics
University of Maryland
College Park, MD 20742
and NBER
haltiwan@econ.umd.edu

Rebecca J. Hutchinson
U.S. Census Bureau
4600 Silver Hill Rd
Suitland, MD 20746
rebecca.j.hutchinson@census.gov

Alyson Plumb
U.S. Census Bureau
4600 Silver Hill Rd
Suitland, MD 20746
alyson.r.plumb@census.gov

1. Introduction

Employer business startups (new firms with paid employees) have a critical role in job creation and productivity growth in the U.S. economy.[1] Most startups either fail or remain small, but some grow into large and productive firms transforming industries and the economy in the process. The economic impact of startups is not limited to the jobs created by them; they also include many positive externalities, pecuniary and non-pecuniary, such as innovation spillovers as well as new demand for suppliers and construction.

Startup activity and the performance of young firms are particularly sensitive to business cycles and economic conditions.[2] For example, the pace of business startups during the Great Recession declined substantially. However, this decline was only visible in hindsight as a result of a lack of high-frequency, timely, and up-to-date information on business formation in the economy based on data available at the time.

The availability of timely data on business formation can enhance our understanding of changes in entrepreneurial activity when aggregate economic conditions worsen or improve. Moreover, high-frequency movements in new business formations themselves have the potential to foretell changing economic conditions, as entrepreneurs may react to early signs of such changes and reassess their business plans. Recent research also indicates that early characteristics and initial conditions of businesses tend to be influential in post-entry business growth and dynamics.[3] Measuring the volume and nature of early-stage business activity is therefore critical in assessing the contribution of likely new businesses to the U.S. economy.

Until recently, comprehensive, timely, and high-frequency data on new business initiations has not been available. The Census Bureau's Business Formation Statistics (BFS), initially released in 2018 as a quarterly data product, fills this gap by providing monthly data on new business applications, and actual and projected employer business formations originating from these applications.[4] The monthly BFS is typically released within two weeks of the end of the reference month.

The BFS has been particularly useful in real-time tracking of new business formation activity during the COVID-19 recession and its recovery, and at sectoral and state-level detail. Following a sharp drop in business applications during the early phases of the pandemic, a substantial surge took place that resulted in an all-time high in business applications for the period starting in 2004. The timely BFS data has been widely followed during the pandemic, as policymakers, government agencies, and the business community monitored the progression of business applications to assess the impact of the pandemic and its aftermath on entrepreneurship. The

[1] See Haltiwanger, Jarmin, and Miranda (2013).
[2] See Davis and Haltiwanger (2024), Dinlersoz et al. (2021), and Buffington et al. (2021).
[3] See, for example, Bayard et al. (2018), Brown et al. (2017), Guzman and Stern (2020), and Sterk, Sedláček, and Pugsley (2021).
[4] See Bayard et al. (2018) on the construction of the BFS series.

strong surge in new business applications has persisted through October 2024.[5] Evidence has shown that this surge in applications has led to a rise in new employer businesses. Moreover, spatial variation in business applications exhibits patterns consistent with changes in work and lifestyle in the pandemic and its aftermath.[6]

Despite the emerging research that shows a connection between business applications and aggregate and local economic activity, no systematic work has been done to establish the properties of the BFS series as potentially leading indicators of aggregate economic activity. In particular, an open question is whether the BFS series contain early information on changing aggregate economic conditions and how strong this information is compared to that provided by various existing Principal Federal Economic Indicators (PFEIs). The objective of this paper is to examine the behavior of monthly business application series from the BFS in relation to other monthly PFEIs and assess their ability to capture the month-to-month growth in these PFEIs.

While the BFS series consists of four business application and eight business formation series, the analysis in this paper focuses on a parsimonious subset of the application series. The two main application series studied are Business Applications (BA) and High-Propensity Business Applications (HBA). BA is the broadest set of business applications, including applications that may result in either new employer or non-employer businesses. Data provided in the business applications offer information on the likely outcome of the application. Using this data, HBA is defined as the set of applications with characteristics that make them more likely to transition into an employer business. These characteristics include information that the new business plans to hire workers, will be incorporated, or are in industries where new businesses are more likely to be employers. As shown in Bayard et al. (2018), Haltiwanger (2021), and Decker and Haltiwanger (2024), HBA tracks actual employer startups over the next four to eight quarters very closely.

The analysis also separately considers the difference between BA and HBA, or Non-High-Propensity Business Applications (NHBA), which is referred to here as the set of likely non-employer applications. Haltiwanger (2021) and Decker and Haltiwanger (2024) show that NHBA closely tracks fluctuations in non-employers at an annual frequency.[7] The increasing

[5] See https://www.census.gov/econ/bfs/index.html for the latest BFS statistics. We use data through December 2023 in our statistical analysis in this paper.

[6] Decker and Haltiwanger (2024) shows that the surge in applications has contributed substantially to an increase in new employer businesses and job creation in the post-pandemic economy. They also show the spatial pattern is related to changes in work and lifestyle. Dinlersoz et al. (2021) highlight the difference in the dynamics of business applications and formations during the Great Recession vs. the pandemic contraction. Dinlersoz et al. (2023) provide further analysis of the sources of spatial variation in applications and startup application using the BFS micro data.

[7] Research in progress by Dinlersoz, Kroff, and Luque (2024) focuses on estimating the rate of non-employer business formation from business applications from the microdata. This work will help refine our understanding of the relationship between business applications and new non-employers, much like the Bayard et al. (2018) research provides guidance about the relationship between business applications and employer startups.

prevalence of self-employment and gig jobs in the U.S. economy makes this series critical for understanding the trends in self-employment.

As will be further discussed in Section 2, these three series contain the core of the information about business applications. We do not use the actual business formation series for the current analysis since these series are not timely. Similarly, we do not use the projected business formation series, which, while insightful, are largely derived from the application series (specifically HBA) that are the focus of the current paper. Given our interest in understanding the timing of fluctuations in business applications with other timely high-frequency key economic indicators, we focus on the core application series that provide guidance about the forward-looking plans of potential future new businesses.

The analysis first examines, using cross-correlations and correlograms, pairwise correlations between the growth in monthly application series and the growth in various monthly PFEIs for several lags and leads. We focus on the relationship between business applications with nonfarm employment as the latter is a critical coincident indicator of economic activity. The correlograms indicate that growth in BA, HBA, and NHBA all lead growth in nonfarm employment, but HBA has the highest correlation and the largest lead. Furthermore, HBA strongly leads most of the monthly PFEIs considered. In terms of correlation strength with nonfarm employment, HBA is around the middle of the PFEIs studied, but BA and NHBA rank at the bottom. HBA's lead of nonfarm employment is quite large (11 months), and it surpasses all PFEIs on this front except for New Single-Family Homes Sold. BA and NHBA, on the other hand, are more mildly leading (5 months and 1 month, respectively). Overall, HBA has the highest correlation, in absolute terms, with nonfarm employment among all other PFEIs. The sectoral components of HBA also lead the sector-specific PFEIs. For instance, manufacturing HBA leads the manufacturing-related PFEIs. Similarly, retail HBA leads the retail-related PFEIs.

To provide perspective, we contrast the relationship between HBA and other PFEIs with Advance Monthly Sales Retail and Food Services (or retail sales for short). Retail sales is a strong leading indicator for most other PFEIs but with a shorter lead time compared to HBA. For example, retail sales lead nonfarm employment by 4 months, while the lead for HBA is 11 months. On the other hand, retail sales have much stronger correlations with other PFEIs than HBA.

Motivated by these findings, we estimate a dynamic factor model (DFM) to forecast nonfarm employment growth over a 12-month period using the indicators that we identified as leading nonfarm employment. A DFM allows researchers to incorporate information from many indicators into a forecast model. We estimate a baseline model using all the leading indicators. Next, we reestimate a similar model, except that we drop one leading indicator from the baseline model. This exercise allows us to study how that indicator affects the model's forecast and its mean squared forecast error (MSFE). We repeat this exercise for each leading indicator.

We find that removing HBA from the baseline model increases its MSFE by 8 percent. This result, however, masks substantial heterogeneity in terms of how HBA can significantly affect the model's forecasts. For example, during the years 2013 to 2015, the baseline model predicted substantially lower nonfarm employment growth rates than the model without HBA. The pessimistic predictions from the model with HBA were more consistent with the actual trajectory of growth rates of nonfarm employment. In fact, during this period, the removal of HBA from the baseline model results in an increase of 69 percent in the model's MSFE.

Overall, the analysis suggests that HBA is a novel leading indicator for many PFEIs that receive much attention in economic measurement. The fact that HBA has a substantial lead over many prominent PFEIs examined here suggests it provides a very early signal of changing economic conditions. While BA and NHBA are relatively less informative indicators of overall economic activity, they are of particular interest as they contain unique information on an increasingly important component of economic activity, not as well captured by the existing PFEIs: self-employment. Work is in progress to establish further properties of the NHBA series.[8]

The idea of using information on new business formation as an economic indicator is not entirely new. An earlier monthly series on new business activity was released by the Bureau of Economic Analysis (BEA) starting in 1947 until the mid-1970s. This series was based on the number of new business incorporations by Dun and Bradstreet and other confidential information reported to the BEA. However, the coverage was limited to various states and new corporate entities, and also included expansions of existing corporations.[9] This series was designated as a leading economic indicator by the National Bureau of Economic Research (NBER). Compared to this earlier series, the BFS series has a number of advantages. It is based on federal administrative data, it covers all legal forms of organizations, not just corporations, it includes all states and the District of Columbia, and it allows for sectoral analysis.[10]

Despite the extended presence of many monthly PFEIs, little systematic work has been done to assess their properties and behavior, especially in recent decades. We contribute to the analysis of these indicators by examining how they relate to nonfarm employment and business applications in the last two decades and through two major recessions. We also examine the relative contribution of each leading indicator to the forecasting of nonfarm employment within a DFM that incorporates these indicators and business applications.

The rest of the paper proceeds as follows. Section 2 describes the BFS series and discusses the reasoning behind why we consider specific business application series for analysis. Section 3

[8] Dinlersoz, Kroff, and Luque (2024) aim to provide evidence on the properties of NHBA series and non-employer business formations from the applications included in NHBA.
[9] Specifically, the statistics included new businesses that incorporated, existing businesses changed from the noncorporate form to the corporate form, existing corporations given authority to operate in another state, and existing corporations transferred to a new state. See Business Conditions Digest, November-June (1968-1969).
[10] While not included in the U.S. totals, the BFS statistics now include separate tabulations for Puerto Rico as an experimental product.

provides a background on PFEIs. Section 4 establishes some key properties of the business application series with respect to the existing PFEIs based on correlation analysis. Section 5 presents the results of the DFM analysis. Section 6 concludes.

2. Business Formation Statistics

This section gives some background on the BFS and describes its constituent series. It then motivates our focus on three specific series from the BFS, which we study in more detail in relation to the PFEIs.

2.1. The origins and definitions of BFS series

The Business Formation Statistics (BFS) is a Census Bureau data product that offers timely, high-frequency, and comprehensive information on new business applications, as well as actual and projected employer business startups in the United States. The BFS started as a research project at the Center for Economic Studies in collaboration with economists at the Federal Reserve Board, Federal Reserve Bank of Atlanta, and the universities of Maryland and Notre Dame. Research behind the construction of the BFS data series is documented in detail in Bayard et al. (2018).

The BFS relies on applications for Employer Identification Numbers (EINs) made to the Internal Revenue Service (IRS) using IRS Form SS-4 and delivered on a weekly basis to the Census Bureau.[11] EIN applications data are matched with the Census Bureau's Longitudinal Business Database (LBD), which provides information on the incidence and timing of new employer business formation. This match identifies whether an EIN application becomes an employer business following application and, if so, the year and quarter of the first payroll on record in the LBD. The combined data are used to construct nationwide, industry, and state level time series for business applications and business formations originating from these applications.[12]

The BFS includes four business application series and eight business formation series. The core series, Business Applications (BA), contains all EIN applications excluding those made without a business intent, such as trusts, estates, public entities, and others related to personal finance and household tasks. Among BA, there is a high degree of heterogeneity in terms of the likelihood of becoming an employer business in the future. Certain broad application characteristics revealed in IRS Form SS-4 during the EIN application process are correlated with employer business formation. Using these characteristics, applications are further classified into a subset of BA that represents "likely employers" – labeled as High-Propensity Business Applications (HBA). HBA is the set of applications that satisfy one or more of the following criteria: (i) from corporations, (ii) indicate plans to hire an employee, (iii) indicate a first wages paid/planned date, and (iv) in

---

[11] Detailed contents of IRS Form SS-4 are available at https://www.irs.gov/pub/irs-pdf/fss4.pdf.

[12] The BFS also releases a business application series annually at the county level.

certain industries: accommodation and food services or portions of construction, manufacturing, retail, professional, scientific and technical services, educational services, and health care.[13] Overall, HBA have about a 27% likelihood of becoming employer businesses within eight quarters after application, as opposed to the non-HBA (NHBA), which become employer businesses with a probability of only 3.8%.[14]

Two subsets of HBA exhibit particularly high rates of employer business formation. One is the set of applications that indicate a first wages paid/planned date, which have nearly a 40% likelihood of transitioning into an employer business within eight quarters of application. The other is the applications from corporate entities, which turn into employer businesses at a rate of approximately 30%. These two subsets are tracked over time as separate series in the BFS in addition to HBA: Business Applications with Planned Wages (WBA), and Business Applications from Corporations (CBA). Altogether, HBA, WBA, and CBA are informative about "likely employer" business formations, whereas the NHBA series captures "likely non-employer" business activity.[15]

In addition to the business application series described above, the BFS contains eight business formation series. Business Formations within 4 Quarters (BF4Q) provides the count of actual employer businesses that originate from business applications within four quarters of the time of application. Because the LBD used in identifying business formations has a typical lag of two years, actual formations for recent applications are not observed in a timely fashion. To fill this gap, the BFS uses an econometric model to estimate business formations from a given set of applications and provides projections of business formations for recent periods based on this model.[16] Projected Business Formations within 4 Quarters (PBF4Q) is the resulting series.

Combining BF4Q and PBF4Q, an up-to-date series that includes both actual formations and projections for recent periods is also provided: Spliced Business Formations within 4 Quarters (SBF4Q). In addition, the BFS provides information on the duration between business applications and formation. The duration series is called Average Duration (in Quarters) from Business Application to Formation within 4 Quarters (DUR4Q). The BFS also contains another set of business formation series, which are defined analogously to the four-quarter window series but use an eight-quarter window to measure business formations from applications (these series are formally labeled in the BFS as BF8Q, PBF8Q, SBF8Q, and DUR8Q, respectively). The BFS

---

[13] The HBA definition was updated in November 2021. The previous definition is available at https://www.census.gov/econ/bfs/technicaldocumentation/methodology.html.
[14] See Bayard et al. (2018) for the rates of employer business formation associated with different application characteristics.
[15] Appropriate caution is needed in interpreting the term "likely." Less than half of the HBA applications yield an employer startup, but as discussed below, fluctuations in HBA track fluctuations in employer startups very closely. For likely non-employers, we know less about the time-series patterns of new non-employers, but the analysis below shows that when projected non-employers are determined using a constant exit rate from Davis et al. (2009) and NHBA as a proxy for entrants, actual non-employers can be tracked closely.
[16] See Bayard et al. (2018) for the details of the econometric model. In a nutshell, a linear probability model is used to relate the indicator of business formation from an application to application characteristics.

application and formation series together offer a rich picture of early-stage business formation activity in the U.S. economy.

The BFS was originally released with quarterly data in 2018. The release of BFS continued at a quarterly frequency between 2018 and 2020, with the earliest data pertaining to 2004q3. The quarterly BFS aggregated applications to a quarterly frequency, even though applications are transmitted to the Census Bureau from the IRS at a weekly frequency. Up until 2020, the BFS time series included data on only one recession: the Great Recession that officially started in 2007q4 and ended in 2009q2. As economic conditions rapidly deteriorated with the onset of the pandemic in the first quarter of 2020, the demand by policymakers for more timely data sources for measuring economic activity surged. This demand presented a unique opportunity for BFS to monitor business formation activity at a higher frequency. As a result, the BFS initiated a monthly release starting in January 2021 with the complete set of business application and formation series. The monthly BFS replaced the quarterly BFS and includes additional critical components including all of the series at the two-digit NAICS level. The monthly frequency is crucial for the purpose of comparing the BFS series with other economic indicators, as most of the PFEIs are monthly. A few others are quarterly, one is semi-annual, and only one is weekly, as discussed in Section 3.

2.2. The focus on BA and HBA as economic indicators

While all 12 series in BFS provide critical information on different aspects of business applications and formations, a parsimonious subset consisting of the BA and HBA series, and the series for their difference, NHBA, can serve as indicators of the different aspects of aggregate economic activity.[17] BA, HBA, and NHBA transparently capture key dimensions of early-stage business formation for the purposes of PFEI analysis.

First, for the task of tracking overall business initiation activity in the economy, BA is the most comprehensive series, as it contains information on both "likely employer" and "likely non-employer" applications. This broad coverage makes it appealing as a potential indicator of both future new job creation and new non-employer business activity, especially as the gig economy continues to grow and the new employer business formation rate continues to decline.

[17] Note that NHBA is not a series separately released as part of the monthly BFS, but is rather an implicit series derived from BA and HBA. As such, for the purposes of this study, the NHBA series is not separately seasonally adjusted (direct adjustment) but is obtained as the difference between the seasonally adjusted BA and HBA series (indirect adjustment). However, NHBA obtained this way may have residual seasonality. An alternative is to seasonally adjust separately the two components of BA: HBA and NHBA. In the future, if a decision is made to release NHBA as a stand-alone series in BFS, this alternative may be pursued. However, the literature offers no consensus as to whether direct or indirect adjustment is superior. See, for example, Hood and Findley (2003) and Evans (2009). Future work may study the differences between the two approaches in the context of NHBA.

Second, HBA is particularly useful in understanding the behavior of applications that can turn into job creators in the near future with a relatively high likelihood. While the two subsets of HBA (WBA and CBA) have a higher likelihood of transitioning into employer status than HBA, they also leave out a large set of applications that still have a considerable probability of turning into job creators. HBA make up nearly 50% of applications, whereas WBA and CBA each constitute about 24% of the applications with a large overlap between the two.[18] Thus, HBA is a more comprehensive measure of potential new employer businesses that can form in the future. Furthermore, HBA tends to follow the trends in actual and predicted business formation series (SBF4Q, SBF8Q) relatively closely. Figure 1 presents the series BA, HBA, WBA, CBA, SBF4Q, and SBF8Q. The correlation between HBA and SBF8Q is 0.93. The high correlation between HBA and SBF8Q reflects the close correspondence between HBA and employer startups. Since projected startups are based on the relationship between application characteristics and actual startups, it is instructive to focus on the correlation between HBA and SBF8Q through 2019 so that only actual employer startups are used. The correlation between HBA and actual startups for the 2004 through 2019 period is 0.92.[19]

The positive correlation between BA and SBF8Q in Figure 1 (0.59) largely reflects the surge in BA and SBF8Q during the pandemic and its aftermath. The correlation between BA and actual startups from 2004 through 2019 is close to zero. This reflects the pattern of BA including non-employer business applications; during the period of the Great Recession and its aftermath, non-employer applications were robust, whereas actual employer startups were on the decline.

Third, the difference between BA and HBA, NHBA, is relevant in terms of tracking non-employer business activity. Drawing from Haltiwanger (2021) and Decker and Haltiwanger (2024), Figure 2 shows a close correspondence between the actual non-employer series and that projected from a simple transition equation using NHBA as a proxy for non-employer entrants.[20]

The rise in likely non-employer business applications can be a particularly important indicator of worsening economic conditions and recessions, during which some displaced workers are pushed into self-employment. In fact, the surge in NHBA during the COVID-19 recession may be partly due to this dynamic.[21] In addition, a rising trend in NHBA is consistent with the change in employment patterns, consistent with the increasing prevalence of gig economy jobs. Note that NHBA cannot account for all new non-employer business activity, as most of such activity can be conducted using a Social Security number as a tax ID rather than an EIN. Even then, non-

[18] See Bayard et al. (2018).
[19] For this paper, the last year of the LBD we use is 2021, implying that startups eight quarters ahead are available through 2019.
[20] The transition equation is: $NES_t = (1 - ExitRate_t) * NES_{t-1} + NHBA_t$. The exit rate is estimated from Davis et al. (2009). See Haltiwanger (2021) for more details. Appropriate caution is needed in using this measurement approach since the Census Bureau's Nonemployment Statistics series (NES) include sole proprietor non-employers not captured in NHBA. The index number approach in Figure 2 alleviates implied level differences, but the lack of information on the entry of sole proprietor non-employers still suggests caution.
[21] See Dinlersoz et al. (2021).

employers with an EIN tend to have a larger scale (in terms of revenue) compared to those without an EIN, so NHBA could be informative about likely non-employer businesses that tend to perform better than the average non-employer.[22]

Finally, application series have critical advantages over business formation series as timely indicators. The actual formation series BF4Q and BF8Q have a lag of two years, so they are not timely. While the projected formation series PBF4Q and PBF8Q fill the gap in the actual formation series, they are model-based estimates. The model uses various application characteristics to estimate the likelihood of employer business formation from an application. The most critical of the characteristics used in the model are the ones that underlie the definition of HBA. As such, while the model uses additional information that is not contained in the definition of HBA, most of the key information is already captured by HBA. This is evident in Figure 1, where HBA tracks the trends in SBF4Q and SBF8Q relatively closely.[23] The remaining two series, DUR4Q and DUR8Q, are informative about the delay in employer business formation, which has been trending up over the BFS sample period. Changes in the delay in hiring can be tied to the underlying economic conditions; for instance, increasing uncertainty. However, just like the actual business formation series they are based on, the duration series have a lag of two years, and they also do not provide information on the volume of business formation activity.

Figure 3 contains the three BFS series that are the focus of the rest of the analysis. The HBA and NHBA series display very different patterns. The former declines substantially as the Great Recession hits and stays relatively flat until the COVID-19 recession. In contrast, after a slight decline, NHBA grows persistently between the Great Recession and the COVID-19 recession before dropping sharply and subsequently reaching all-time highs. As a result, the gap between the NHBA and HBA series opens up over time. The time-path of the BA series, on the other hand, is influenced mainly by the trends in the NHBA series. As noted above, HBA tracks actual employer startups closely, while NHBA exhibits a negative correlation with actual employer startups (−0.58).

The rest of the analysis focuses on the properties of the monthly BA, HBA, and NHBA series as economic indicators. An attractive feature of these series is their economic relevance, high frequency, timeliness, not being subject to significant revisions, and availability of a long time series.[24] The analysis compares these three series with existing monthly PFEIs and seeks to understand whether any of these series lead, coincide with, or lag the PFEIs.

---

[22] See Davis et al. (2009).
[23] In addition, the model parameters are estimated using actual formations from at least two years prior to the current year.
[24] These criteria are stated in the 2019 guidelines on producing leading, composite, and sentiment indicators by the United Nations Economic Commission for Europe (UNECE). See United Nations (2017).

3. Principal Federal Economic Indicators (PFEIs)

Policy directives issued by the Office of Management and Budget (OMB) provide guidelines for the production and release of a Principal Federal Economic Indicator (PFEI). As defined by Statistical Policy 3, PFEIs are those "statistical series that are widely watched and heavily relied upon by government and the private sector as indicators of the current condition and direction of the economy."[25] Statistical Policy 3 also controls the release of and methodological updates to these PFEIs. The OMB has granted PFEI status to 36 statistical data products that are published by eight different federal statistical agencies, including the Census Bureau, the Bureau of Labor Statistics, the Federal Reserve, and the Bureau of Economic Analysis. Table 1 provides a summary of all existing PFEIs. The Census Bureau currently publishes 13 PFEIs—more than any other agency. The last new PFEI to be designated by the OMB was the Energy Information Administration's Weekly Natural Gas Storage Report in 2007.

As Table 1 shows, PFEIs are published at a variety of frequencies: weekly, monthly, quarterly, and semi-annually. For reasons stated earlier, for our analysis we focus on the monthly PFEIs and a further subset to those PFEIs with characteristics that would be particularly useful in the analysis of potential BFS indicator properties. This subset of indicators used in the analysis are denoted with an asterisk (*) in Table 1 and represent measurements that capture economic activity in personal and business consumption, construction, manufacturing, and employment. The data series from each of the PFEIs used in our analysis are summarized in Table 2.

Some of the monthly PFEIs in Table 1 are not considered in the analysis for a variety of reasons. We focus on monthly indicators that track nonfarm private sector business activity where new business applications are potentially more relevant. In general, the PFEIs that focus on international trade, agriculture, prices, and natural resources have less relevance BFS series. International trade is dominated by large, established firms, and only a tiny fraction of new employer businesses are engaged in international trade.[26] Agriculture is a relatively small part of the economy, and business applications in the agriculture sector are a very small fraction of total applications.[27] The fluctuations in natural gas storage are likely less related to new firm formation and more to the inventories held by existing natural gas producers. The natural gas market reacts to changes in inventory levels, which inform trading decisions that move natural gas prices. Price indices (Consumer Price Index and Producer Price Index) are driven by many of the same economic conditions affecting overall business activity but have complicated dynamics given the adjustment process of prices. Such adjustment dynamics are interesting in their own

---

[25] U.S. Office of Management and Budget (1985). Statistical Policy Directive No. 3: Compilation, release, and evaluation of principal federal statistical indicators. 50 Federal Register 38932 (September 25, 1985). Available: Federal Register Notice (archives.gov)

[26] See, for example, Bernard, Jensen, and Schott (2009) and Bernard, et al. (2018) for evidence that U.S. traders tend to be very large and trade volume is highly concentrated among the top traders.

[27] See sector-level business applications data: (https://www.census.gov/econ/bfs/csv/bfs_monthly.csv).

right but are not the focus of this study. There may be, for example, a connection between firm entry and markups, but price indices do not provide direct indicators of markups.[28]

There is relatively little formal documentation of the processes or research behind the determination of whether a statistical data product is or is not PFEI. Furthermore, there is limited research on understanding the nature of PFEIs in terms of whether they lead, lag, or coincide with economic activity. Our analysis below documents some properties of the monthly PFEIs alongside the properties of the BFS series we consider. However, this analysis is not meant to be a comprehensive look at unique properties of each PFEI, which is an endeavor beyond the scope of this article.

In addition to the PFEIs, many other economic indicators and indices are derived from the PFEIs or their components, as well as from additional data such as stock prices and the interest rate spread. These are regularly released by private entities, non-profit institutions, and other government agencies. For instance, the Federal Reserve banks and the Conference Board maintain a set of economic indicators/indices that are released on a regular basis. It is worth noting that several of the monthly PFEIs we analyze are included in the Conference Board's leading, coincident, and lagging indicators.[29] The Conference Board's leading indicators include subcategories of manufacturing new orders and building permits for new private residences, and coincident indicators include nonfarm payroll employment, industrial production and manufacturing, and trade sales.

4. Properties of business application series in relation to PFEIs

To explore the relationship between business applications and the PFEIs, we focus initially on the sample period from July 2004 (the start of the monthly BFS) through December 2019, which includes the Great Recession and the subsequent recovery. We then analyze the period July 2004 to December 2023, which includes the COVID-19 pandemic and the subsequent recovery. As is already evident from Figures 1 and 3, the dramatic fluctuations in business applications during the COVID-19 pandemic are unprecedented in magnitude and also quite distinct from the patterns in the Great Recession, so it is useful to separately evaluate these periods.

4.1. Using cross-correlations to study properties of business application series

In this section, we aim to understand the relationship between the BFS series and the selected monthly PFEIs. In particular, we explore to what extent the BFS series correlate with other PFEIs and whether they lead, coincide with, or lag the PFEIs in monthly growth.

---

[28] For recent work on the connection of markups and entry in a macroeconomic context, see, for example, Cavallari (2013) and Lewis and Stevens (2015).

[29] The Conference Board U.S. Business Cycle Indicators, "The Conference Board Leading Economic Index (LEI) for the United States and Related Composite Economic Indexes for February 2021," released March 18, 2021, https://www.conference-board.org/pdf_free/press/US%20LEI%20-%20Tech%20Notes%20MARCH%202021.pdf.

We start with a concrete example of the relationship between HBA and nonfarm employment. We ask the following questions:

1. Does HBA lead nonfarm employment? We say that HBA leads nonfarm employment if movements in its growth rate typically precede those of nonfarm employment. If movements in these two series tend to take place at the same time, then we say that they are coincident. If HBA moves after nonfarm employment does, then it lags nonfarm employment.
2. What are the direction and magnitude of the correlations between the growth rate of HBA and growth in nonfarm employment?

We calculate the cross-correlations between the monthly year-over-year growth rates of these two series. This cross-correlation analysis provides a simple way of addressing the questions posed above. We first define the cross-correlation function between variables x and y, which in our case are monthly year-over-year growth rates of the economic series of interest, as

$$\rho_{xy}(k) = corr(x_t, y_{t+k}), , \tag{1}$$

where k is an indicator of the lag time of y and $corr$ is a function that calculates the correlation between two variables. Thus, we calculate the correlation between the contemporaneous growth rate of nonfarm employment (x above) and the growth rate of HBA (y above) for lags and leads that range from -12 to 12 (k above).

The results of this exercise using monthly data up to December 2019 can be seen in the cross-correlogram in Figure 4a. The x-axis of the graph indicates the value of k used to calculate the correlation. Note that for $k < 0$, indicated by the values to the left of zero on the x-axis, the analysis uses the growth rates of y from previous periods. If the correlations are highest in this region, then y is a leading indicator. On the other hand, when $k > 0$ (the values to the right of zero), the analysis uses the growth rates from future periods of y. If the correlations are highest in this region, then it is a lagging indicator. For $k = 0$, the analysis uses contemporaneous values of y. Thus, if the correlation is highest at $k = 0$, then it is a coincident indicator. Backus, Routledge, and Zin (2010) and Stock and Watson (1998) conduct the same type of cross-correlation analysis to determine the properties of macroeconomic series.

The figure shows that the growth rates of the two series are positively correlated for all values of k. Furthermore, the highest correlation is for $k = -11$. Thus, HBA leads nonfarm employment growth by 11 months, and the correlation at that lead is 0.64. The interpretation is that the HBA growth rate from 11 months prior is the best predictor of the current growth rate in nonfarm employment. Thus, we conclude that growth in HBA is a useful indicator in predicting future changes in the growth in nonfarm employment.

For comparison, Figure 4b repeats the analysis with retail sales instead of HBA. Consumption generally reflects consumers' expectations about the employment outlook and real wage growth.

Thus, retail sales are a key PFEI, since many standard macroeconomic models would imply an early consumption response to a changing economic outlook.[30] Retail sales indeed lead nonfarm employment by four months, and the correlation at that lead is 0.85. While HBA has a lower correlation with nonfarm employment compared to retail sales, it has a substantially larger lead.

4.2. Results of cross-correlation analysis

In this section, we apply the cross-correlation analysis described above to analyze the relationship between nonfarm employment and the BFS series along with other PFEIs using monthly data up to December 2019. For all variables, we focus on the monthly year-over-year growth rates.

Table 3 shows a summary of the results of the cross-correlation analysis for the three BFS series (HBA, BA, and NHBA) and all of the PFEIs considered.[31] The table indicates whether the series is leading or lagging, the magnitude of the correlation at the lead/lag with the highest (absolute) correlation, and the number of periods that it leads/lags by (timing). The table is sorted by the absolute value of the correlation from the highest to the lowest correlation.

Table 3 has a few key findings. First, almost all the PFEIs lead nonfarm employment. Furthermore, the PFEIs that lag nonfarm employment do so by only 1 or 2 months. This result suggests that the existing PFEIs we consider are mainly economic indicators that are leading or coincident. Second, HBA leads nonfarm employment by 11 months, which is one of the highest lead times in the table (exceeded only by the lead time of New Single-Family Homes Sold). The correlation between the growth in HBA and the growth in nonfarm employment is 0.64, which is in the middle range of the PFEIs considered. Third, the growth rates of BA and NHBA are positively correlated with that of nonfarm employment, although these correlations are among the weakest of the series analyzed. In terms of timing, BA is mildly leading by 5 months and NHBA leads by only 1 month.

Table 3 shows that retail sales have the highest correlation with nonfarm employment among all the PFEIs we consider. One important difference between HBA and retail sales is that the latter tends to lead nonfarm employment by less (4 months compared to 11 months for HBA).

Table 4 reports the results of a similar exercise in which we calculate the correlation of the growth rate of the other PFEIs and the leading/lagging growth rates of HBA. This table gives us information about whether HBA leads or lags the other PFEIs. First, we find that HBA leads

[30] See, for example, Breeden (2012).

[31] Table 3 does not include total private sector employees and the unemployment rate because of the high correlations with nonfarm employment and because these two series are either coincident or lead nonfarm employment by only a few months. Total private sector employment growth rates are coincident with nonfarm employment growth rates (column 2) and these growth rates have a correlation of 0.996 (column 3). The unemployment rate leads nonfarm employment by only 2 months (column 2) and these growth rates have a correlation of −0.967 (column 3).

almost all the other PFEIs. This result is consistent with Table 3 in which HBA leads nonfarm employment by 11 months, substantially more than almost all of the other series.

To put the results from Table 4 into perspective, in Table 5 we repeat the same exercise, except that we use the leading/lagging growth rates of retail sales along with the growth rates of the other PFEIs. Comparing the results of Table 4 and 5, we find that while the retail sales series tends to have a higher correlation than HBA, it tends to lead by less, implying that this series is more coincident with economic activity than HBA.

4.3. Cross-correlation analysis using retail and manufacturing HBA

One strength of the BFS series is that they are also available by sector. It is natural to wonder about the relationship between the sectoral measures of HBA (i.e., manufacturing HBA and retail HBA) and the PFEIs that are associated with these sectors. Appendix Table A3 reports the results of the same analysis as in Table 4, except that we use the manufacturing HBA instead of the overall HBA. Furthermore, we focus the analysis on the PFEIs that are most closely related to manufacturing (i.e., new orders, shipments of manufacturing, and sales in manufacturing and trade). We find that manufacturing HBA leads these three series by one to two months. For comparison, we include the results from Table 4, in which we use the overall HBA. The results show that the correlations tend to be similar. The main difference is that the overall HBA leads the manufacturing PFEIs by significantly more than the manufacturing HBA.

Appendix Table A4 reports the results of an exercise similar to the one in Table 4, except that we use the retail HBA and focus on retail sales (we include the results of the overall HBA in the table as a reference). In this case, retail HBA leads retail sales by more than the overall HBA. Furthermore, the retail HBA also has a slightly higher correlation.

4.4. Including the 2020–2023 time period

The period from 2020 to 2023 exhibits extraordinary fluctuations not just in HBA but in all of the economic indicators we have been considering. To help illustrate this, Figure 5 depicts the growth rates of HBA, retail sales, and nonfarm employment through time for the period 2004–2019, and Figure 6 includes the 2020–2023 period. It is evident that the volatility is extraordinary in the latter period, especially in the March–June 2020 period. Table 6 shows the standard deviation and autocorrelation of the growth rates for HBA, nonfarm employment, and retail sales for all years and for the 2004–2019 period. The 2020–2023 period exhibits substantially higher standard deviations, lower autocorrelations for retail sales and nonfarm employment, and higher autocorrelations for HBA.

We now repeat our cross-correlation analysis with the 2020–2023 time period included. Given the very large and unexpected movements in economic indicators during the spring and early summer of 2020 evident in Figure 6, we exclude the growth rates for March through June 2020 in the main text (but show the patterns including all months in the appendix). In addition, because we are using year-over-year growth rates, we also exclude the growth rates for March through June 2021.

Table 7 reports the results analogous to Table 3. We find that the cross-correlations tend to be weaker across the PFEIs. For example, the cross-correlation of retail sales with nonfarm employment declines from 0.85 to 0.64, and the cross-correlation for HBA declines from 0.64 to 0.44. The lead times remain reasonably stable at 4 months and 12 months for retail sales and HBA, respectively. We note that if we include all months including March through June 2020 the cross-correlations have similar magnitudes, but the lead times (see appendix Table A1) are reduced to 1 month and 9 months for retail sales and HBA, respectively. The decline in lead times reflects the extraordinary short-term collapse in economic activity over this period. It is also interesting to note that the cross-correlation with BA and NHBA becomes negative with the inclusion of 2020–2023 (with or without the March–June 2020 months).

Table 8 reports the results analogous to Table 4 where we extend the analysis to include 2020–2023 (again excluding the March to June 2020/2021 periods, with results using all months in Appendix Table A2). We find that HBA still strongly leads most of the other PFEIs, except that the strength of the correlation declines.

To sum up, the 2020–2023 period exhibits extraordinary fluctuations in economic activity especially in the March through June 2020 period. We find that if we conduct our cross-correlation analysis, findings are largely robust to the inclusion of this period as long as we exclude the March through June 2020 period. Inclusion of the latter with its enormous volatility tends to drive the PFEIs to being coincident given the enormous short-term decline in economic activity during this period. Given these challenges, the following dynamic factor model analysis shows results for 2004–2019 and all years through 2023 but excluding the March through June 2020 period (and lagged growth rates linked to those periods).

5. Dynamic Forecasting Model

Section 4 used bivariate cross-correlations to identify leading indicators of nonfarm employment growth, as summarized in Tables 3 and 7. In this section, we introduce a forecasting model and study the model's forecasts of nonfarm employment growth when we jointly consider multiple leading economic indicators. Furthermore, we analyze how these forecasts change if we exclude a particular economic indicator, such as HBA, from the analysis. This exercise will allow us to study the contribution of a particular leading indicator to forecasts of nonfarm employment growth. For example, in the case of HBA, it would allow us to study how its inclusion in a forecasting model would affect predictions of nonfarm employment growth relative to the existing set of PFEIs that we classify as leading indicators. Among the BFS series explored earlier, we focus on HBA in this section because it has the highest correlations as a leading indicator of nonfarm employment relative to BA or NHBA, as documented in Tables 3 and 7.

5.1. Forecasting model

For our forecasting model, we use the methodology developed by Stock and Watson (2002) based on dynamic factor models (DFMs). The methodology supposes that many macroeconomic series have comovements that are driven by a few latent factors. These latent factors can first be estimated using principal component analysis (PCA), and then forecasts can be created using these estimated factors. Among the several advantages of this methodology are that it incorporates information from many series in a parsimonious set of latent factors, and it performs well relative to other methodologies. For example, Stock and Watson (2002) include as many as 215 series in a DFM, and yet it remains computationally tractable. In contrast, vector autoregressive (VAR) models are limited in the number of series that can be incorporated into the model given the number of parameters that must be estimated.

We use the following multistep-ahead version of the DFM implemented by Stock and Watson (2002):

$$y_{t+h}^{h} = \alpha_h + \beta_h(L)F_t + \gamma_h(L)y_t + \varepsilon_{t+h}^{h}, \tag{4}$$

where y_{t+h}^{h} is the variable to be forecast at time t using the h-step-ahead method, α_h is a constant, $\beta_h(L)$ and $\gamma_h(L)$ are lag polynomials in nonnegative powers of L, F_t are the factors, and ε_{t+h}^{h} is an error term.

We follow a two-step strategy to estimate our model. First, we estimate the factors, F_t, using principal components. Given these estimated factors, $\hat{F}_t$, we estimate α_h, $\beta_h(L)$, and $\gamma_h(L)$ by regressing y_{t+h}^{h} on $\hat{F}_t$, y_t, and their corresponding lags.

5.2. Principal component analysis (PCA)

We apply PCA to the leading indicators described in Table 3, in which we conduct the cross-correlation analysis on the period before the pandemic. That table contains 12 leading indicators that are PFEIs and HBA, which yields a total of 13 indicators that we will focus on. The indicators that we use are listed in the first column of Appendix Table A5.

We apply PCA after standardizing the 13 indicators, a common practice in the literature. To better understand the results from this PCA analysis, we report the results using the period up to 2019 (i.e., the last growth rate analyzed was from December 2018 to December 2019). For comparison, we also report the results based on the entire period including the pandemic. Figure 7 shows the scree plots of the PCA, which are related to the fraction of variation explained by a given principal component. Panel (a) shows the scree plot when we use the entire span, and panel (b) shows the scree plot when we use the growth rates up to and including 2019. First, we see that both scree plots are very similar. Second, in both cases, there is a sharp drop-off after the third principal component in the fraction of variation explained by the principal components.

Thus, given the sharp drop-off in the fraction of variation explained by the fourth and subsequent principal components, we use the first three principal components for our analysis.

Appendix Table A5 shows the weights assigned by the PCA to the 13 indicators. The columns labeled "All Data" indicate that we used all the year-over-year growth rates in the analysis. The columns labeled "Data through 2019" indicate that we used growth rates up to December 2019. Note that the weights in a particular column in the table do not sum to one; rather, the sum of the squared weights does. Regardless of the span used, the weights are similar, which indicates that weights tend to be stable even when we include growth rates from the pandemic.

Interestingly, the highest weight for HBA is in the third principal component. For example, when we use all of the data, the HBA weight is 0.79, and the indicator with the next highest weight in this third principal component has a weight of 0.37. At the same time, HBA has low weights relative to many of the other indicators in the first and second principal components. These results suggest that the information contained in HBA is being reflected primarily in the third principal component.

5.3. Forecast model for nonfarm employment

We focus on the multistep-ahead forecast of nonfarm employment. We define

$$y_{t+h}^{h} = \ln\left(\frac{NFE_{t+h}}{NFE_t}\right), \tag{5}$$

where NFE_t is nonfarm employment (in levels) at t. Similarly, we define

$$y_t = \ln\left(\frac{NFE_t}{NFE_{t-1}}\right). \tag{6}$$

Given the estimates of $\hat{F}_T$, $\hat{\alpha}_h$, $\hat{\beta}_h(L)$, and $\hat{\gamma}_h(L)$, let $\hat{y}_{T+h|T}^{h}$ be the forecast for $T+h$ when we only use data through period T:

$$\hat{y}_{T+h|T}^{h} = \hat{\alpha}_h + \sum_{j=1}^{m} \hat{\beta}_{hj}' \hat{F}_{T-j+1} + \sum_{j=1}^{p} \hat{\gamma}_{hj} y_{T-j+1}, \tag{7}$$

where $\hat{F}_{T-j+1}$ is a vector of k factors that were estimated using PCA, and m and p are the number of lags used for the factors and y_t, respectively. In our analysis, we use $m = 1$ and $p = 1$, as we use year-over-year growth rates, which already includes information about the growth rates over the previous 12 months.

We focus on the model's ability to forecast nonfarm employment growth over a 12-month horizon (i.e., $h = 12$). We compute the model's estimated MSFE using a pseudo-out-of-sample forecasting methodology. In our case, we estimate the initial model using all the information available in July 2008. Thus, our first out-of-sample forecast uses all available information in July 2008 and forecasts the 12-month ahead nonfarm employment growth (i.e. the year-over-year growth between July 2008 and July 2009). Note that nonfarm employment growth exhibits a turning point in the middle of 2009, as shown in Figure 8. Thus, this initial model allows us to evaluate our forecast models during the turning point of the Great Recession. These results should be taken cautiously, however, as the initial model is estimated using only 25 observations.[32]

Next, we reestimate the model using all available information as of August 2008 and then use this reestimated model to predict the 12-month ahead nonfarm employment growth (i.e. the year-over-year growth between August 2008 and August 2009). In this process, we recompute the weights using PCA and the parameters to generate a new forecast. We repeat this process as for all subsequent months in the estimation sample. We use the out-of-sample predictions to calculate a model's estimated MSFE.

We estimate a model that uses all 13 leading indicators in the analysis, which we refer to as the baseline model. To better understand each indicator's contribution relative to all the other indicators, we redo the exercise, except that we exclude one indicator at a time. We then compare the MSFE of the baseline model and the model in which we exclude a given indicator. Suppose that the dropping of a particular indicator leads to a large increase in the MSFE. We then conclude that this indicator provides additional information that is useful in forecasting nonfarm employment growth, relative to the other indicators considered in the analysis.

5.4. Results

Table 9 shows the results of our analysis. The first column contains the indicator that was removed from the baseline model. The second column reports the ratio of the MSFE when a particular indicator has been removed and the baseline model when we consider forecasts of the 12-month nonfarm employment predictions up to December 2019. A number above 1 indicates that removing a particular indicator increases the model's MSFE. We find a range of 0.77 to 1.23. Removing HBA from the baseline model leads to an increase of 9 percent in the MSFE over this period, a number that is only exceeded by New Single-Family Homes Sold.

The third column uses the whole sample for the analysis, excluding the pandemic period. As before, for this time period we excluded growth rates during March–June 2020 in our analysis, given the large swings in the indicators during this period. We also excluded March–June 2021 as we are using year-over-year growth rates, and these equally exhibit large swings. We find that the results are generally similar to those in column 2.

To further study these results, Figure 8 plots the predicted nonfarm employment growth rates from the baseline model and the actual growth rates for the out-of-sample period. We also plot the predicted nonfarm employment growth rates in the model that excludes HBA to understand how the exclusion of HBA affects the baseline model. We see that both models predict similar

[32] Note that in the first observation of the initial model, we use July 2005 year-over-year growth rate to forecast July 2006 year-over-year growth rate. In the last observation of the initial model, we use July 2007 year-over-year growth rate to forecast July 2008 year-over-year growth rate.

turning points for the Great Recession. Furthermore, the most striking difference between the two sets of predicted nonfarm employment growth rates occurs for the period between 2012 and 2015. In this period, the model without HBA predicts substantially higher nonfarm employment growth rates than the baseline. This finding is consistent with the slow recovery of HBA after the Great Recession, as seen in Figure 3.

In Appendix Table A6, we report the MSFE of the baseline model and the model without HBA during different subperiods until 12/2019. We find that the most significant differences between the baseline and model without HBA occur during two periods. First, during 1/2013–12/2015 (fourth column), removing HBA from the baseline model increases the MSFE by 69 percent. This finding is consistent with the evolution of nonfarm employment and the predictions from these two models after the Great Recession, as shown in Figure 8. Second, during 1/2016–12/2019 (fifth column), removing HBA from the baseline model leads to a decline of 34 percent in the MSFE. We see in Figure 8 that the predictions from the two models are similar and tend to converge to actual nonfarm employment growth rates. This latter period is a stable period for nonfarm employment growth, as is evident in the much lower baseline MSFE. While HBA does not exhibit a systematic trend over this period either, it is more volatile than other series. We interpret these results as suggesting appropriate caution in using a very short stable period to draw inferences. Put differently, these subperiod exercises capture variation within these subperiods without considering the variation between subperiods. During a stable short period, there is not much systematic variation to forecast.

We next turn to the period beginning in 2020, which coincides with the beginning of the COVID-19 pandemic. Figure 9 plots the out-of-sample predictions of the baseline model and the model without HBA, along with the actual 12-month growth rates of nonfarm employment. As mentioned before, we exclude the growth rates between March and June of 2020 and 2021 from our analysis. The lack of growth rates in March–June of 2021 implies that we cannot calculate forecasted growth rates for March–June of 2022. For example, if we are missing the 12-month nonfarm employment growth rate in March 2021, we are missing a necessary input to forecast growth over the subsequent 12 months.

Figure 9 shows that large decreases in the 12-month nonfarm employment growth rates took place during the July 2020–February 2021 period. We also see that both models predicted positive growth rates, which is consistent with the patterns of the economy before the pandemic. Second, for July 2021–February 2022, the nonfarm employment growth rates were approximately 5 percent, which are large relative to historical growth rates (see, for example, a comparison with Figure 8). The baseline model tends to predict higher growth rates relative to the model without HBA, which is consistent with the large increases in HBA after the initial sharp decline in 2020, as seen in Figure 1. Second, the predictions from the baseline model tend to be more volatile than the model without HBA. Third, after July 2022, both models tend to predict slower growth rates relative to the realized outcomes. We do not see large differences between the levels of predictions of the two models. The baseline model, however, tends to be more volatile than the model without HBA.

Appendix Table A7 reports the MSFE of the models for various subperiods. Interestingly, for the 7/2021–2/2022 period, the MSFE of the model without HBA is 37 percent higher than that of the baseline model. This finding is consistent with the predicted growth rates of nonfarm employment from the baseline model being closer to actual growth rates over this period, as seen in Figure 9. In addition, the model without HBA has an MSFE that is 8 percent higher than the baseline model over the period 7/2022–12/2023.

6. Conclusion

Business applications from the BFS provide timely and high-frequency information on early-stage entrepreneurial activity -- a precursor of both employer and non-employer business formations in the future. This paper has examined the properties of three business application series from the BFS in relation to existing monthly PFEIs. The analysis indicates that applications for likely employers (HBA) are particularly useful as early indicators of aggregate economic activity. The HBA series leads most PFEIs and appears to be an especially strong early indicator of total nonfarm employment and has a high leading positive correlation with it in terms of year-over-year monthly growth rates. Compared with the 16 monthly PFEIs analyzed, HBA ranks second in terms of the size of its lead for growth in nonfarm employment and at that lead has a high correlation with nonfarm employment growth. In retrospect, the HBA series would have been a highly useful leading indicator during the Great Recession, if the series were available and its properties were known at that time. Figure 3 indicates that the HBA series starts to decline substantially in the months leading up to and during the early months of the Great Recession, signaling a strong worsening of economic conditions ahead.

The business application series as a whole (BA) is also a leading indicator of total nonfarm employment, but its correlation with nonfarm employment is weaker and its lead is smaller, potentially owing to the fact that it contains a large volume of applications that are not likely to become employer businesses. These applications do not generate employment for others and have little to contribute to the statistics on business activity as captured by PFEIs.

The set of likely non-employer applications (NHBA) is of distinct interest, as it contains information on non-employer business formation and self-employment trends. This series could be especially useful for tracking the rise of self-employment and gig jobs in the U.S. economy (see Abraham et al. (2021) for challenges in measuring the latter).[33]

We use a dynamic factor model to provide more systematic insights about the value added of HBA as an economic indicator. We find that there is a significant increase in the mean squared forecast error when excluding HBA from a model that seeks to forecast nonfarm employment

[33] Work is in progress to understand the nature and timing of non-employer business formation from business applications, in comparison to employer business formation (Dinlersoz, Kroff, and Luque (2024)).

from the existing set of leading PFEIs. The reduction in forecasting error variance is especially large in economic recoveries.

Finally, further work is needed to understand the mechanisms underlying the dynamic empirical relationships documented here. Specifically, does the strong leading nature of HBA resulting from its ability to closely track future employer startups have important implications for economic activity? Alternatively, is this pattern driven by the fact that nascent entrepreneurs are inherently forward looking so that changes in their behavior are a good indicator for future activity? These mechanisms as well as others may be at work. Even without having sorted out the driving forces, the analysis presented here suggests HBA is a useful novel leading indicator of economic activity.

References

Abraham, Katharine G., John C. Haltiwanger, Kristin Sandusky, and James R. Spletzer. 2021. "Measuring the Gig Economy Current Knowledge and Open Issues" in NBER/CRIW Conference Volume on *Measuring and Accounting for Innovation in the 21st Century*.

Backus, David K., Bryan R. Routledge, and Stanley E. Zin. 2010. "The Cyclical Component of US Asset Returns," NYU Working Paper.

Bayard, Kimberly, Emin Dinlersoz, Timothy Dunne, John Haltiwanger, Javier Miranda, and John Stevens. 2018. "Early Stage Business Formation: An Analysis of Applications for Employer Identification Numbers," NBER Working Paper, 24364.

Bernard, Andrew B., Jensen, J. Bradford, and Peter K. Schott. 2009. "Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods", in *Producer Dynamics: New Evidence from Micro Data,* University of Chicago Press, pp. 513-552.

Bernard, Andrew B., Jensen, J. Bradford, Redding Stephen J., and Peter K. Schott. 2018. "Global Firms," *Journal of Economic Literature*, 56(2): 565-619.

Breeden, Douglas T. 2012. "Consumption as a Leading Indicator", Working Paper, Fuqua School of Business, Duke University.

Brown, David, Earle, John, Kim, Meejung, and Lee, Kyungmin. 2017. "High-Growth Entrepreneurship", CES Working Paper, 17-53.

Buffington, Catherine, Daniel Chapman, Emin Dinlersoz, Lucia Foster, and John Haltiwanger. 2021. "High Frequency Business Dynamics in the United States During the COVID-19 Pandemic" Center for Economic Studies Working Paper, 21-06.

Cavallari, Lilia. 2013. "A note on firm entry, markups and the business cycle", *Economic Modelling*, 35: 528-535.

Davis, Steven, John Haltiwanger, C.J. Krizan, Ron Jarmin, Javier Miranda, Al Nucci, and Kristin Sandusky. 2009. "Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Non-Employer Businesses," in *Producer Dynamics: New Evidence from Micro Data*, (Dunne, Jensen and Roberts, eds.) NBER/University of Chicago Press, 329-366.

Davis, Steven and John Haltiwanger. 2024. "Dynamism Diminished: The Role of Housing Markets and Financial Conditions", *American Economic Journal: Macroeconomics*, April, 29-61.

Decker, Ryan and John Haltiwanger. 2024. "Surging Business Formation in the Pandemic: Causes and Consequences?" *Brookings Papers on Economic Activity*, 3-24.

Dinlersoz, Emin, Timothy Dunne, John Haltiwanger, and Veronika Penciakova. 2021. "Business Formation: A Tale of Two Recessions," *AEA Papers and Proceedings*, 111: Pages 253–257.

Dinlersoz, Emin, Timothy Dunne, John Haltiwanger, and Veronika Penciakova. 2023. "The Local Origins of Business Formation," Center for Economic Studies Working Paper, 23-34.

Dinlersoz, Emin, Can Dogan, and Nikolas Zolas. 2024. "Starting Up AI," Working Paper CES-WP-24-09, Center for Economic Studies, US Census Bureau.

Dinlersoz, Emin, Zachary Kroff, and Adela Luque. 2024. "The Modes of Entry: Employer versus Non-employer Business Formation," Work in progress, Center for Economic Studies, U.S. Census Bureau.

Evans, Thomas. 2009. "Direct vs. Indirect Seasonal Adjustment for CPS National Labor Force Series," Bureau of Labor Statistics, Office of Survey Methods and Research, Statistical Survey Paper.

Guzman, Jorge, and Scott Stern. 2020. "The State of American Entrepreneurship? New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988-2014," *American Economic Journal: Economic Policy*, 12(4): 212-243.

Haltiwanger, John. 2021. "Entrepreneurship During the COVID-19 Pandemic: Evidence from the Business Formation Statistics," Paper for presentation at the NBER Entrepreneurship, Innovation and Economic Policy Conference, April.

Haltiwanger, John, Ron Jarmin and Javier Miranda. 2013. "Who Creates Jobs? Small versus Large versus Young", *Review of Economics and Statistics*, 95(2): 347-361.

Hood, Catherine, and David Findley. 2003. "Comparing Direct and Indirect Seasonal Adjustments of Aggregate Series", in *Seasonal Adjustment*, eds. Michele Manna and Romana Peronaci, pp. 9-21.

Lewis, Vivien, and Arnaud Stevens. 2015. "Entry and markup dynamics in an estimated business cycle model", *European Economic Review*, 74:14-35.

Sterk, Vincent, Petr Sedláček, and Benjamin Pugsley. 2021. "The Nature of Firm Growth." *American Economic Review*, 111 (2): 547-79.

Stock, James H., and Mark W. Watson. 1998. "Business Cycle Fluctuations in U.S. Macroeconomic Time Series", *NBER Working Paper Series Working Paper 6528.*

Stock, James H., and Mark W. Watson. 2002. "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business & Economic Statistics* 20 (2): 147–162.

United Nations, Eurostat. 2017. "Handbook on Cyclical Composite Indicators for Business Cycle Analysis," available at https://ec.europa.eu/eurostat/documents/3859598/8232150/KS-GQ-17-003-EN-N.pdf/868e9a5b-9bae-450b-a2d9-c668524c738d?t=1506070558000.

United States. Bureau of the Census. (1968-1990). Business conditions digest: BCD. Washington, D.C.: U.S. Dept. of Commerce, Bureau of the Census.

Table 1. Principal Federal Economic Indicators (PFEIs) by frequency and federal agency

Indicator Frequency	Federal Agency	PFEI Data Product
Weekly	Energy Information Administration	• Weekly Natural Gas Storage Report
Monthly	Bureau of the Census	• Value of Construction Put in Place* • New Residential Construction* • New Residential Sales* • Monthly Wholesale Trade* • Advance Monthly Sales for Retail and Food Services* • U.S. International Trade in Goods and Services • Manufacturing and Trade: Inventories and Sales* • Manufacturers' Shipments, Inventories, and Orders* • Advance Report on Durable Goods - Manufacturers' Shipments, Inventories, and Orders*
	Bureau of Labor Statistics	• The Employment Situation* • Producer Price Indexes • Consumer Price Index • Real Earnings* • Employment Cost Index • U.S. Import and Export Price Indexes
	National Agriculture Statistics Service	• Agricultural Prices • Crop Production • Cattle on Feed
	Bureau of Economic Analysis	• Personal Income and Outlays • U.S. International Trade in Goods and Services
	Federal Reserve Board	• Industrial Production and Capacity Utilization* • Consumer Credit
	World Agricultural Outlook Board	• World Agricultural Supply and Demand Estimates
	Foreign Agricultural Service	• World Agricultural Production
Quarterly	Bureau of the Census	• Quarterly Financial Report - Manufacturing, Mining, Wholesale Trade, and Selected Service Industries • Quarterly Financial Report - Retail Trade • Housing Vacancies • Quarterly Services
	Bureau of Labor Statistics	• Productivity and Costs
	National Agriculture Statistics Service	• Grain Stocks • Hogs and Pigs
	Bureau of Economic Analysis	• Corporate Profits • U.S. International Transactions • Gross Domestic Product
Semiannual	National Agriculture Statistics Service	• Plantings

Notes: (*) indicates series analyzed in relation to the BFS series.

Table 2. PFEI data series used in the analysis

PFEI	Series used in the analysis
Advance Monthly Sales for Retail and Food Services	U.S. Total Retail Trade and Food Services Monthly Sales (Millions of Dollars)
Construction Spending	U.S. Total Annual Rate for Total Construction (Millions of Dollars)
Manufacturers' Shipments, Inventories, and Orders	U.S. Total Manufacturing Value of Shipments (Millions of Dollars)
	U.S. Total Manufacturing Value of New Orders (Millions of Dollars)
Manufacturing and Trade Inventories and Sales	U.S. Total Business Monthly Sales (Millions of Dollars)
	U.S. Total Business Monthly Inventories (Millions of Dollars)
Monthly Wholesale Trade: Sales and Inventories	U.S. Total Merchant Wholesalers (excluding Manufacturers' Sales Branches and Offices) Monthly Sales (Millions of Dollars)
	U.S. Total Merchant Wholesalers (excluding Manufacturers' Sales Branches and Offices) Monthly Inventories (Millions of Dollars)
New Home Sales	U.S. Annual Rate for New Single-Family House Sold (Thousands of Units)
	U.S. New Single-Family Houses for Sales (Thousands of Units)
New Residential Construction	U.S. Annual Rate for Housing Units Authorized in Permit-Issuing Places (Thousands of Units)
	U.S. Annual Rate for Housing Units Completed (Thousands of Units)
	U.S. Annual Rate for Housing Units Started (Thousands of Units)
Advance Monthly Manufacturers' Shipments, Inventories and Orders	U.S. Total of New Orders for Durable Goods (Millions of Dollars)
The Employment Situation	U.S. Total Nonfarm Employees
Real Earnings	Real average hourly earnings of production and nonsupervisory employees
Industrial Production and Capacity Utilization	Industrial Production Index

Notes: All series are seasonally adjusted.

Table 3. Cross-correlations: PFEI growth rate and growth rate in nonfarm employment (all months up to December 2019)

Series name	Timing	Correlation	# of Periods
Advance Monthly Sales Retail and Food Services	Leading	0.845*** (0.041)	−4
Construction Spending	Lagging	0.844*** (0.041)	1
New Single-Family Homes for Sale	Lagging	0.798*** (0.046)	1
Manufacturing and Trade Inventories	Lagging	0.795*** (0.046)	1
New Residential Construction Permits	Leading	0.784*** (0.046)	−8
Industrial Production	Leading	0.782*** (0.048)	−4
New Residential Construction Units Started	Leading	0.774*** (0.048)	−7
New Single-Family Homes Sold	Leading	0.737*** (0.050)	−12
Wholesale Inventories	Lagging	0.701*** (0.054)	2
New Residential Construction Units Completed	Leading	0.680*** (0.056)	−6
Durable Goods New Orders	Leading	0.677*** (0.056)	−4
HBA	Leading	0.643*** (0.057)	−11
Manufacturing and Trade Sales	Leading	0.640*** (0.059)	−3
Manufacturing Shipments	Leading	0.603*** (0.061)	−3
Manufacturing New Orders	Leading	0.591*** (0.061)	−3
Wholesale Sales	Leading	0.551*** (0.064)	−3
Real Hourly Earnings Production and Nonsupervisory Employees	Leading	−0.463*** (0.068)	−2
BA	Leading	0.446*** (0.068)	−5
NHBA	Leading	0.232** (0.074)	−1

Notes: Column 1 reports the series. Column 2 reports whether the series in column 1 is leading, coincident, or lagging. Column 3 reports the largest correlation in terms of magnitude for $k = -12, \ldots, 12$ using the cross-correlation function described in Section 4.1, and column 4 indicates the value of k. * p<0.05, ** p<0.01. *** p<0.001. Standard errors in parentheses (calculated using the formula for the standard error of Pearson's correlation coefficient). Table is sorted by the absolute value of the correlation.

Table 4. Cross-correlations: HBA growth rate and growth rate of other PFEIs (all months up to December 2019)

Series name	Timing	Correlation	# of Periods
Total Nonfarm Employees	Leading	0.643***	−11
		(0.057)	
Total Private Sector Employees	Leading	0.637***	−10
		(0.059)	
Unemployment Rate	Leading	−0.597***	−7
		(0.061)	
Advance Monthly Sales Retail and Food Services	Leading	0.552***	−6
		(0.064)	
Manufacturing and Trade Inventories	Leading	0.551***	−12
		(0.064)	
New Single-Family Homes for Sale	Leading	0.535***	−8
		(0.064)	
Industrial Production	Leading	0.507***	−8
		(0.066)	
New Single-Family Homes Sold	Lagging	0.487***	5
		(0.067)	
Construction Spending	Leading	0.479***	−12
		(0.067)	
New Residential Construction Units Started	Coincident	0.467***	0
		(0.067)	
Wholesale Inventories	Leading	0.463***	−12
		(0.068)	
New Residential Construction Units Completed	Leading	0.459***	−1
		(0.068)	
New Residential Construction Permits	Lagging	0.459***	3
		(0.068)	
Durable Goods New Orders	Leading	0.439***	−5
		(0.068)	
Manufacturing and Trade Sales	Leading	0.422***	−10
		(0.068)	
Manufacturing Shipments	Leading	0.408***	−10
		(0.07)	
Manufacturing New Orders	Leading	0.392***	−5
		(0.07)	
Wholesale Sales	Leading	0.376***	−9
		(0.071)	
Real Hourly Earnings Production and Nonsupervisory Employees	Leading	−0.270***	−10
		(0.072)	

Notes: See notes to Table 3. Column 1 reports the series. Column 2 reports whether HBA is leading, coincident with, or lagging the series in column 1. Column 3 reports the largest correlation in terms of magnitude for $k = -12, \dots, 12$ using the cross-correlation function, and column 4 indicates the value of k. * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses (calculated using the formula for the standard error of Pearson's correlation coefficient). Table is sorted by the absolute value of the correlation.

Table 5. Cross-correlations: Retail sales growth rate and growth rate of other PFEIs (all months up to December 2019)

Series name	Timing	Correlation	# of Periods
Manufacturing and Trade Inventories	Leading	0.908***	−5
		(0.032)	
Manufacturing and Trade Sales	Coincident	0.898***	0
		(0.034)	
Industrial Production	Leading	0.897***	−1
		(0.034)	
Manufacturing Shipments	Leading	0.869***	−1
		(0.037)	
Wholesale Inventories	Leading	0.864***	−6
		(0.037)	
Total Private Sector Employees	Leading	0.859***	−4
		(0.039)	
Unemployment Rate	Leading	−0.855***	−3
		(0.039)	
Manufacturing New Orders	Leading	0.851***	−1
		(0.039)	
Durable Goods New Orders	Leading	0.847***	−1
		(0.041)	
Total Nonfarm Employees	Leading	0.845***	−4
		(0.041)	
Wholesale Sales	Leading	0.844***	−1
		(0.041)	
Real Hourly Earnings Production and Nonsupervisory Employees	Leading	−0.729***	−1
		(0.052)	
New Residential Construction Units Started	Coincident	0.688***	0
		(0.055)	
New Residential Construction Permits	Lagging	0.682***	1
		(0.056)	
Construction Spending	Leading	0.577***	−6
		(0.061)	
New Single-Family Homes Sold	Lagging	0.544***	6
		(0.064)	
New Single-Family Homes for Sale	Leading	0.499***	−5
		(0.066)	
New Residential Construction Units Completed	Leading	0.446***	−2
		(0.068)	

Notes: Column 1 reports the series. Column 2 reports whether retail sales is leading, coincident with, or lagging the series in column 1. Column 3 reports the largest correlation in terms of magnitude for $k = -12, \dots, 12$ using the cross-correlation function, and column 4 indicates the value of k. * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses (calculated using the formula for the standard error of Pearson's correlation coefficient). Table is sorted by the absolute value of the correlation.

Table 6. Standard deviation and autocorrelation of growth rates for HBA, nonfarm employment, and retail sales

	HBA		Nonfarm employment		Retail and food services sales	
Years	**Standard deviation**	**Auto correlation**	**Standard deviation**	**Auto correlation**	**Standard deviation**	**Auto correlation**
All	12.50	0.61	2.94	0.87	5.77	0.85
2004–2019	8.26	0.40	1.72	0.99	4.04	0.94
All (excluding March–June 2020/21)	10.34	0.50	2.37	0.93	4.37	0.93

Table 7. Cross-correlations: PFEI growth rate and growth rate in nonfarm employment (all months excluding March–June 2020/2021)

Series name	Timing	Correlation	# of Periods
Manufacturing and Trade Inventories	Coincident	0.774***	0
		(0.044)	
New Single-Family Homes for Sale	Coincident	0.698***	0
		(0.048)	
Wholesale Inventories	Coincident	0.687***	0
		(0.050)	
Industrial Production	Coincident	0.665***	0
		(0.050)	
Advance Monthly Sales Retail and Food Services	Leading	0.639***	−4
		(0.052)	
Real Hourly Earnings Production and Nonsupervisory Employees	Coincident	−0.617***	0
		(0.054)	
Construction Spending	Coincident	0.608***	0
		(0.055)	
Manufacturing Shipments	Leading	0.569***	−2
		(0.056)	
Manufacturing and Trade Sales	Leading	0.560***	−3
		(0.057)	
Manufacturing New Orders	Leading	0.526***	−3
		(0.057)	
Wholesale Sales	Leading	0.514***	−2
		(0.059)	
Durable Goods New Orders	Leading	0.514***	−4
		(0.059)	
New Residential Construction Permits	Leading	0.485***	−8
		(0.059)	
New Residential Construction Units Completed	Leading	0.453***	−6
		(0.061)	
New Residential Construction Units Started	Leading	0.446***	−4
		(0.061)	
HBA	Leading	0.436***	−12
		(0.061)	
NHBA	Coincident	−0.407***	0
		(0.063)	
New Single-Family Homes Sold	Leading	0.358***	−12
		(0.064)	
BA	Coincident	−0.339***	0
		(0.064)	

Notes: See notes to Table 3. Column 1 reports the series. Column 2 reports whether the series in column 1 is leading, coincident, or lagging. Column 3 reports the largest correlation in terms of magnitude for $k = -12, \ldots, 12$ using the cross-correlation function, and column 4 indicates the value of k. * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses (calculated using the formula for the standard error of Pearson's correlation coefficient). Table is sorted by the absolute value of the correlation.

Table 8. Cross-correlations of HBA growth and growth of other PFEIs (all months excluding March–June 2020/2021)

Series name	Timing	Correlation	# of Periods
Unemployment Rate	Leading	−0.505***	−12
		(0.057)	
New Single-Family Homes Sold	Coincident	0.502***	0
		(0.057)	
Advance Monthly Sales Retail and Food Services	Leading	0.500***	−12
		(0.057)	
New Single-Family Homes for Sale	Leading	0.478***	−12
		(0.057)	
Total Private Sector Employees	Leading	0.451***	−12
		(0.059)	
Total Nonfarm Employees	Leading	0.436***	−12
		(0.059)	
Manufacturing and Trade Sales	Leading	0.433***	−12
		(0.059)	
Wholesale Sales	Leading	0.421***	−12
		(0.059)	
Manufacturing and Trade Inventories	Leading	0.417***	−12
		(0.059)	
Construction Spending	Leading	0.412***	−12
		(0.061)	
Manufacturing Shipments	Leading	0.398***	−12
		(0.061)	
New Residential Construction Units Completed	Lagging	0.383***	8
		(0.061)	
Manufacturing New Orders	Leading	0.370***	−12
		(0.061)	
Industrial Production	Leading	0.370***	−10
		(0.061)	
Wholesale Inventories	Leading	0.368***	−12
		(0.061)	
New Residential Construction Permits	Leading	0.365***	−4
		(0.061)	
New Residential Construction Units Started	Coincident	0.357***	0
		(0.061)	
Durable Goods New Orders	Leading	0.351***	−11
		(0.061)	
Real Hourly Earnings Production and Nonsupervisory Employees	Coincident	0.225***	0
		(0.064)	

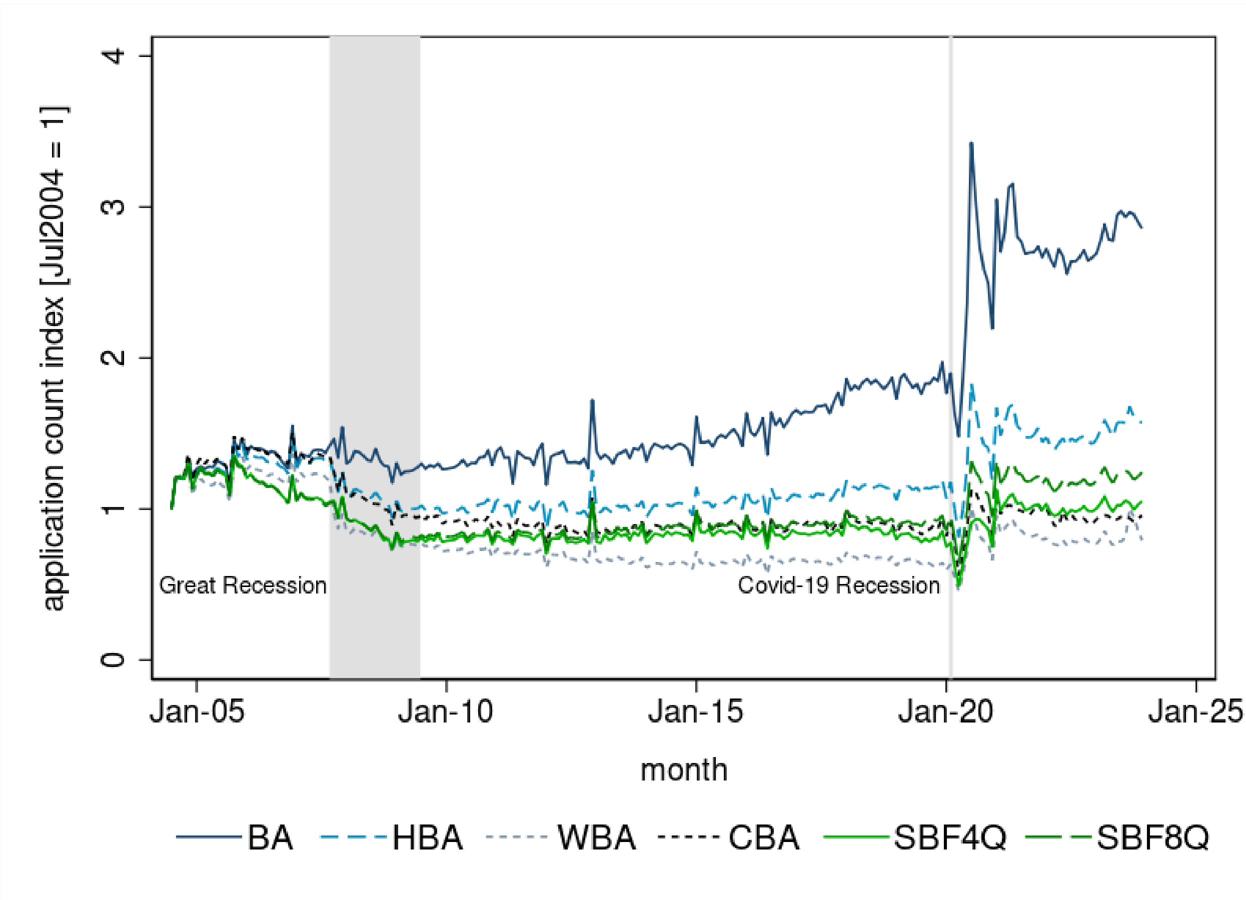
Notes: See notes to Table 4. Column 1 reports the series. Column 2 reports whether HBA is leading, coincident with, or lagging the series in column 1. Column 3 reports the largest correlation in terms of magnitude for $k = -12, \ldots, 12$ using the cross-correlation function, and column 4 indicates the value of k. * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses (calculated using the formula for the standard error of Pearson's correlation coefficient). Table is sorted by the absolute value of the correlation.

Table 9. MSFE of models relative to baseline using PCA (12-month forecast)

Series removed	July 2009 – December 2019	July 2009 – December 2023 (excluding March–June 2020/2021/2022)
Advance Monthly Sales Retail and Food Services	0.96	0.98
Industrial Production	0.82	0.88
New Residential Construction Permits	1.08	1.06
New Residential Construction Units Started	1.00	1.00
New Single-Family Homes Sold	1.23	1.15
Durable Goods New Orders	1.05	1.03
New Residential Construction Units Completed	0.77	0.84
Manufacturing and Trade Sales	1.00	1.00
Real Hourly Earnings Production and Nonsupervisory Employees	1.08	1.06
Manufacturing Shipments	0.97	0.98
Manufacturing New Orders	1.06	1.04
Wholesale Sales	1.09	1.06
HBA	1.09	1.08
MSFE baseline model	13.71	15.10
Variance dependent variable	3.32	5.48

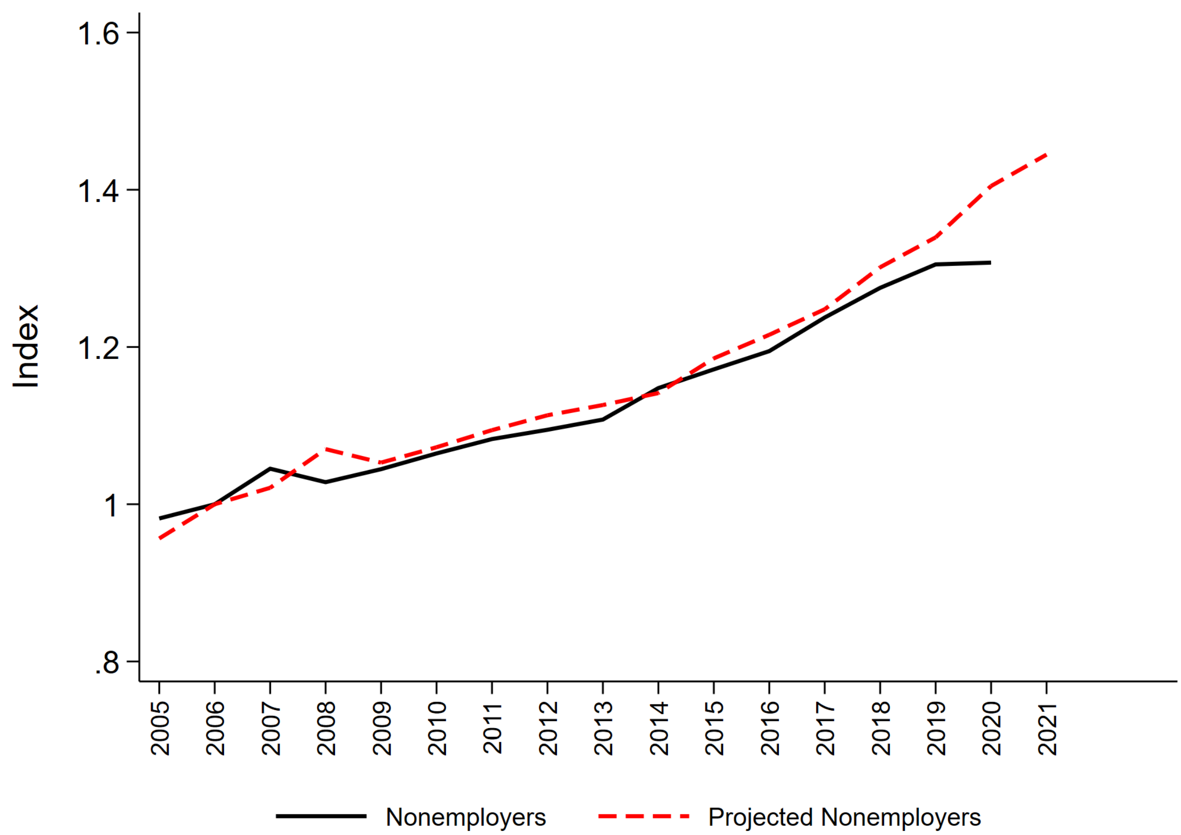
Notes: This table reports the ratio of the MSFE of the model without the indicator listed in column 1 and the MSFE of the baseline model. Note that the baseline model includes all 13 leading indicators. A ratio greater than 1 indicates a worsening of the MSFE when the indicator in column 1 is removed from the baseline model. Column 2 reports the results when we only consider predicted growth rates up to December 2019 (i.e., the last predicted growth rate from the model was the 12-month nonfarm employment growth rate from December 2018 to December 2019). Column 3 reports the statistics from a similar analysis, except that the analysis includes predicted growth rates up to June 2023. The analysis in column 3 also excludes predicted growth rates for March–June of 2020, 2021, and 2022.

Figure 1. Monthly business application and formation series from the BFS



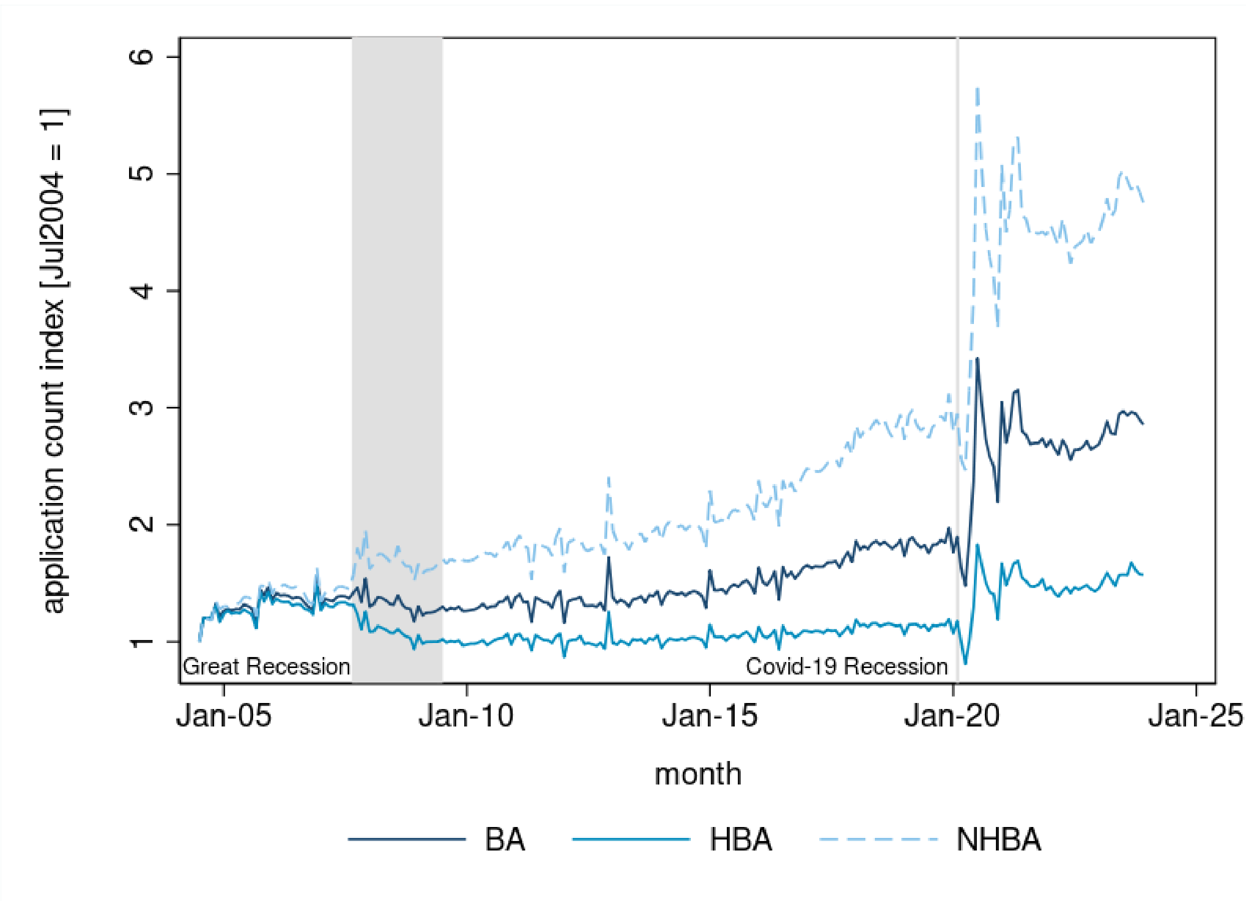
Notes: Selected BFS series. All series seasonally adjusted.

Figure 2. NES series and projected NES using NHBA



Notes: BFS and NES statistics from Census Bureau. Projected non-employers using NHBA uses exit rates from non-employers from Davis et al. (2009) and entry rates based on NHBA. See Haltiwanger (2021) and Decker and Haltiwanger (2024) for further discussion.

Figure 3. Monthly BFS business application series used in the analysis



Notes: Selected BFS series. All series seasonally adjusted.

Figure 4. Cross-correlograms
(all months up to December 2019)

(a) Cross-correlogram of nonfarm employment and HBA | **(b) Cross-correlogram of nonfarm employment and retail sales**

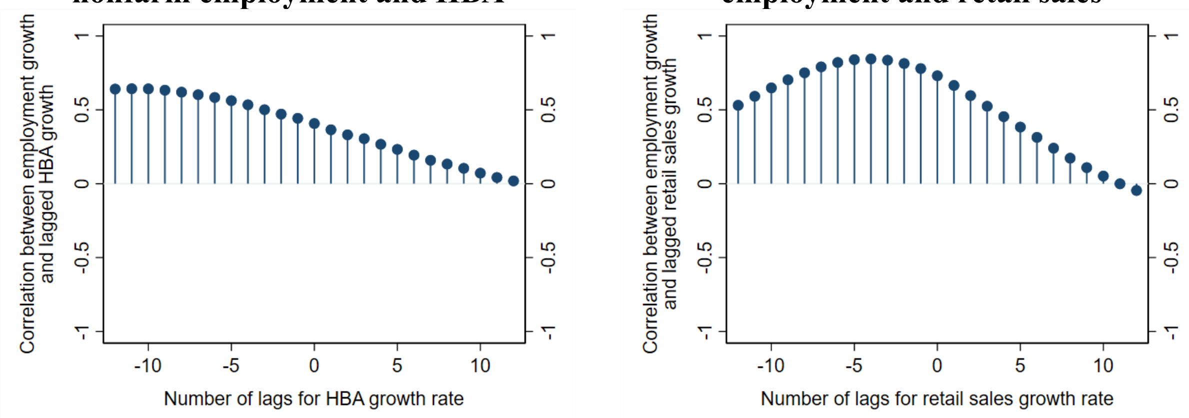


Figure 5. Growth rate of HBA, retail sales, and nonfarm employment (all months up to December 2019)

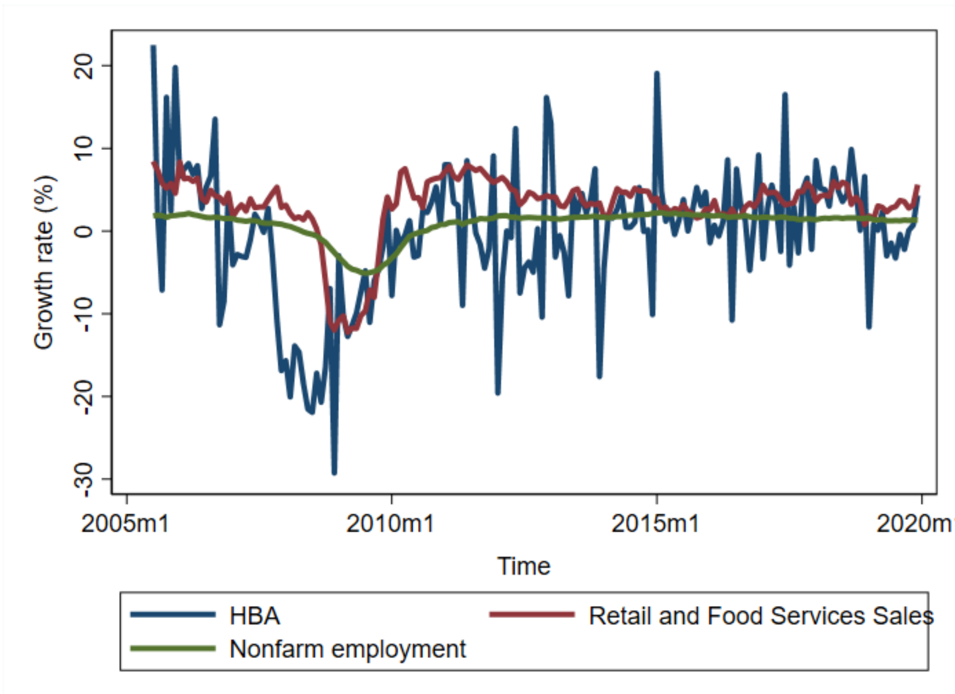


Figure 6. Growth rate of HBA, retail sales, and nonfarm employment (all years)

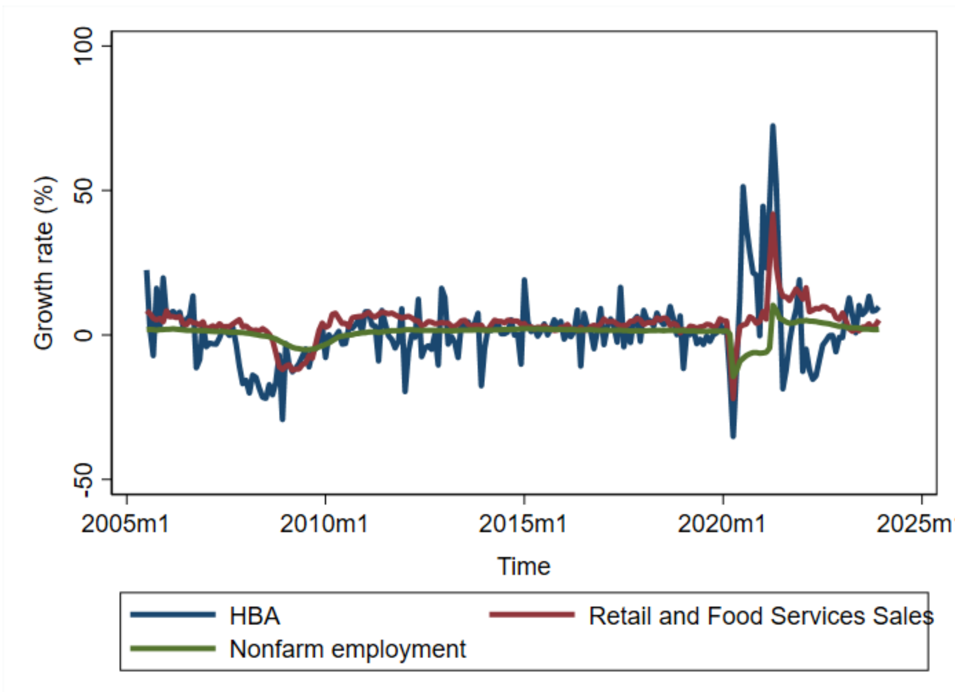
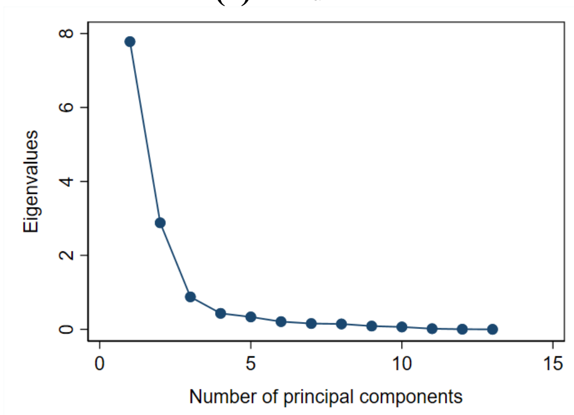


Figure 7. Scree plot

(a) All data



(b) Data through 2019

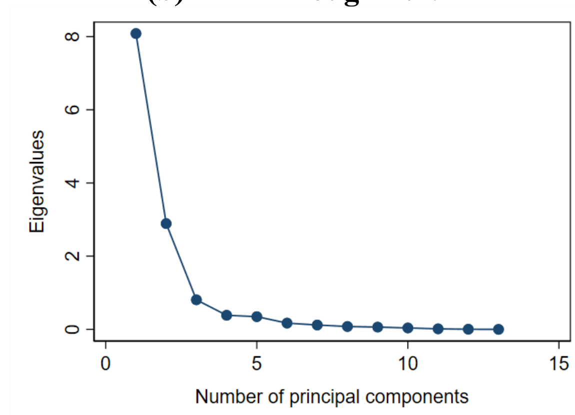


Figure 8. Predicted and actual 12-month changes in nonfarm employment (up to December 2019)

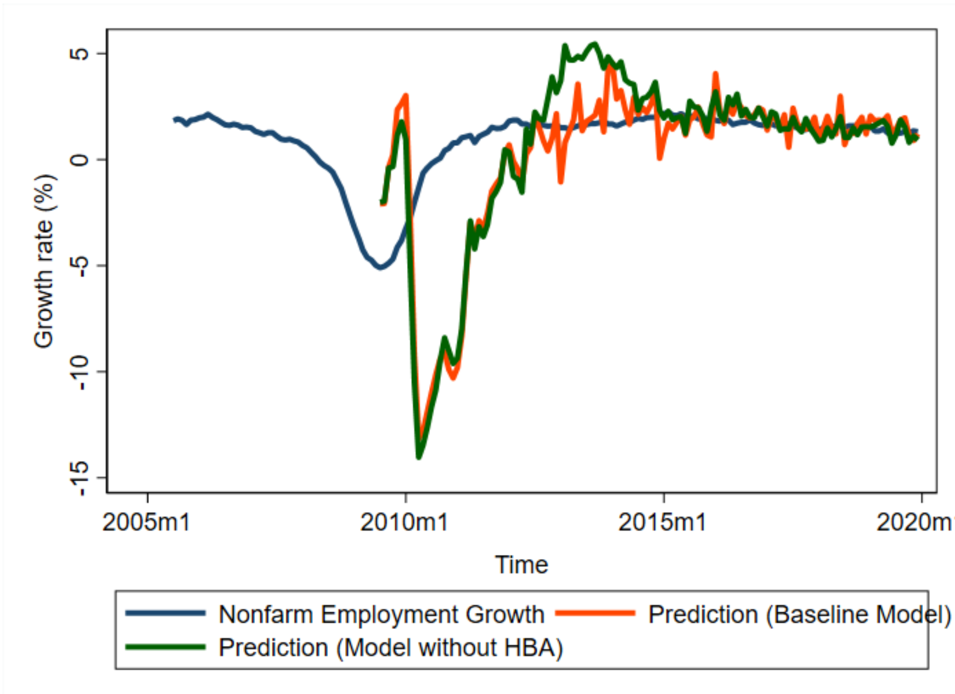
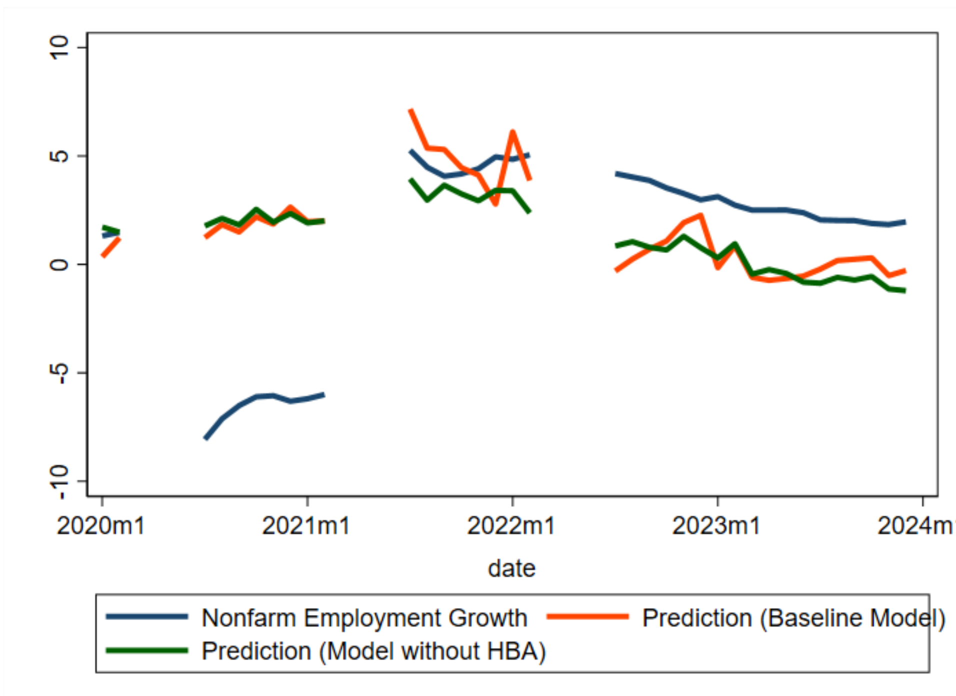


Figure 9. Predicted and actual YOY changes in nonfarm employment, excluding March–June 2020/2021 (2020 and after)



Appendix: Additional Tables

Table A1. Cross-correlations: PFEI growth rate and growth rate in nonfarm employment (all months)

Series name	Timing	Correlation	# of Periods
Industrial Production	Coincident	0.747*** (0.045)	0
Manufacturing and Trade Inventories	Lagging	0.713*** (0.046)	1
Real Hourly Earnings Production and Nonsupervisory Employees	Coincident	−0.712*** (0.046)	0
Advance Monthly Sales Retail and Food Services	Leading	0.679*** (0.048)	−1
Manufacturing Shipments	Leading	0.643*** (0.052)	−1
Wholesale Inventories	Lagging	0.642*** (0.052)	1
Manufacturing and Trade Sales	Leading	0.632*** (0.052)	−1
New Single-Family Homes for Sale	Lagging	0.603*** (0.054)	2
Wholesale Sales	Leading	0.602*** (0.054)	−1
Manufacturing New Orders	Leading	0.586*** (0.055)	−1
NHBA	Lagging	−0.561*** (0.056)	3
Durable Goods New Orders	Leading	0.548*** (0.056)	−1
BA	Lagging	−0.526*** (0.057)	3
Construction Spending	Lagging	0.483*** (0.059)	1
HBA	Leading	0.474*** (0.059)	−9
New Residential Construction Permits	Leading	0.399*** (0.061)	−10
New Residential Construction Units Started	Leading	0.387*** (0.061)	−9
New Residential Construction Units Completed	Leading	0.370*** (0.063)	−2
New Single-Family Homes Sold	Leading	0.309*** (0.064)	−11

Notes: See notes to Table 7. This table reports the same type of information, but statistics are computed using all months.

Table A2. Cross-correlations of HBA growth and growth of other PFEIs (all months)

Series name	Timing	Correlation	# of Periods
Advance Monthly Sales Retail and Food Services	Leading	0.583***	−9
		(0.055)	
Manufacturing and Trade Inventories	Leading	0.507***	−12
		(0.057)	
New Single-Family Homes for Sale	Leading	0.498***	−12
		(0.057)	
Manufacturing and Trade Sales	Leading	0.492***	−9
		(0.059)	
Unemployment Rate	Leading	−0.491***	−9
		(0.059)	
Total Private Sector Employees	Leading	0.485***	−9
		(0.059)	
New Single-Family Homes Sold	Lagging	0.485***	1
		(0.059)	
Wholesale Sales	Leading	0.475***	−9
		(0.059)	
Total Nonfarm Employees	Leading	0.474***	−9
		(0.059)	
Wholesale Inventories	Leading	0.466***	−12
		(0.059)	
Manufacturing Shipments	Leading	0.438***	−10
		(0.061)	
Industrial Production	Leading	0.435***	−9
		(0.061)	
Manufacturing New Orders	Leading	0.435***	−9
		(0.061)	
New Residential Construction Units Started	Coincident	0.422***	0
		(0.061)	
Durable Goods New Orders	Leading	0.413***	−9
		(0.061)	
Construction Spending	Leading	0.409***	−12
		(0.061)	
New Residential Construction Permits	Coincident	0.409***	0
		(0.061)	
Real Hourly Earnings Production and Nonsupervisory Employees	Lagging	0.366***	3
		(0.063)	
New Residential Construction Units Completed	Lagging	0.357***	1
		(0.063)	

Notes: See notes to Table 8. This table reports the same type of information, but statistics are computed using all months.

Table A3. Cross-correlations: Manufacturing-HBA growth rate and growth rate of PFEIs related to manufacturing (all months up to December 2019)

Series name	HBA Series	Timing	Correlation	# of Periods
Manufacturing New Orders	Manufacturing	Leading	0.427*** (0.069)	−2
Manufacturing Shipments	Manufacturing	Leading	0.412*** (0.069)	−2
Manufacturing and Trade Sales	Manufacturing	Leading	0.414*** (0.069)	−2
Manufacturing and Trade Sales	All Sectors	Leading	0.422*** (0.068)	−10
Manufacturing Shipments	All Sectors	Leading	0.408*** (0.07)	−10
Manufacturing New Orders	All Sectors	Leading	0.392*** (0.07)	−5

Notes: Column 1 reports the series. Column 2 reports whether manufacturing-HBA is leading, coincident with, or lagging the series in column 1. Column 3 reports the largest correlation in terms of magnitude for $k = -12, \ldots, 12$ using the cross-correlation function, and column 4 indicates the value of k. * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses (calculated using the formula for the standard error of Pearson's correlation coefficient). Table is sorted by the absolute value of the correlation within HBA manufacturing and HBA all sectors.

Table A4. Cross-correlations: Retail-HBA growth rate and growth rate of PFEIs related to retail (all months up to December 2019)

Series name	HBA Series	Timing	Correlation	# of Periods
Advance Monthly Sales Retail and Food Services	Retail	Leading	0.581*** (0.062)	−6
Advance Monthly Sales Retail and Food Services	All Sectors	Leading	0.552*** (0.064)	−6

Notes: Column 1 reports the series name. Column 2 reports whether retail-HBA is leading, coincident with, or lagging the series in column 1. Column 3 reports the largest correlation in terms of magnitude for $k = -12, \ldots, 12$ using the cross-correlation function, and column 4 indicates the value of k. * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses (calculated using the formula for the standard error of Pearson's correlation coefficient).

Table A5. Weights of the first three principal components

	All Data			Data through 2019		
Series	**PC 1**	**PC 2**	**PC 3**	**PC 1**	**PC 2**	**PC 3**
Advance Monthly Sales Retail and Food Services	0.33	0.03	0.23	0.34	0.03	0.05
Industrial Production	0.32	−0.08	−0.07	0.32	−0.06	0.16
New Residential Construction Permits	0.23	0.39	−0.20	0.24	0.38	−0.18
New Residential Construction Units Started	0.24	0.38	−0.20	0.25	0.37	−0.15
New Single-Family Homes Sold	0.09	0.53	−0.03	0.14	0.50	−0.15
Durable Goods New Orders	0.33	−0.04	0.11	0.32	−0.10	0.14
New Residential Construction Units Completed	0.14	0.41	−0.28	0.15	0.44	−0.13
Manufacturing and Trade Sales	0.34	−0.15	0.06	0.33	−0.19	−0.01
Real Hourly Earnings Production and Nonsupervisory Employees	−0.27	0.19	0.37	−0.28	0.14	0.43
Manufacturing Shipments	0.33	−0.20	−0.00	0.32	−0.22	0.01
Manufacturing New Orders	0.34	0.16	0.05	0.32	−0.20	0.05
Wholesale Sales	0.33	−0.20	0.01	0.32	−0.24	−0.05
HBA	0.15	0.28	0.79	0.15	0.26	0.82

Notes: This table reports the PCA weights for the first three principal components for 13 leading indicators. The columns labeled “All Data” indicate that we used all the year-over-year growth rates in the analysis. The columns labeled “Data through 2019” indicate that we used growth rates up to December 2019 (i.e., the last growth rate was the logged difference between nonfarm employment levels in December 2019 and December 2018). Note that the sum of each column’s squared weights equals one.

Table A6. MSFE by subperiod (up to December 2019)

	7/2009–12/2010	1/2011–12/2012	1/2013–12/2015	1/2016–12/2019
MSFE Baseline	71.48	15.74	1.07	0.38
MSFE no HBA	72.41	16.35	1.80	0.25
MSFE Ratio (no HBA/baseline)	1.01	1.04	1.69	0.66

Notes: This table reports the ratio of the MSFE of the model without the indicator listed in column 1 and the MSFE of the baseline model. Note that the baseline model includes all 13 leading indicators. A ratio greater than 1 indicates a worsening of the MSFE when the indicator in column 1 is removed from the baseline model. Columns 2–5 report the results when we consider the predicted growth rates indicated by each column.

Table A7. MSFE by subperiod (pandemic to present)

	7/2020–2/2021	7/2021–2/2022	7/2022–12/2023
MSFE Baseline	71.6	1.73	7.25
MSFE no HBA	74.4	2.36	7.80
MSFE Ratio (no HBA/baseline)	1.04	1.37	1.08

Notes: This table reports the ratio of the MSFE of the model without the indicator listed in column 1 and the MSFE of the baseline model. Note that the baseline model includes all 13 leading indicators. A ratio greater than 1 indicates a worsening of the MSFE when the indicator in column 1 is removed from the baseline model. Columns 2–4 report the results when we consider the predicted growth rates indicated by each column.