Assessing Labor Market Conditions Using High-Frequency Data

Maximiliano Dvorkin and Maggie Isaacson

When the COVID-19 pandemic struck in March 2020, the U.S. economy experienced a sharp, unexpected recession with large employment losses. The information on employment available from traditional data sources arrives with a lag and does not promptly reflect sudden changes in labor market conditions. In this article, we discuss how new high-frequency data from Homebase and Ultimate Kronos Group can offer critical information on the state of labor markets in real time. Using these datasets, we construct coincident employment indices to assess employment at a high frequency. Employment during the pandemic reacted to changes in the number of infections and the restrictions imposed by government officials (see, e.g., the discussion in Dvorkin and Bharadwaj, 2020). Our latest data suggest that employment has recently increased and will continue to increase as the pandemic wanes. (JEL J2, E27)

Federal Reserve Bank of St. Louis *Review*, Fourth Quarter 2021, *103*(4), pp. 461-76.
https://doi.org/10.20955/r.103.461-76

INTRODUCTION

The coronavirus pandemic severely affected economic activity in the United States and the rest of the world. In the United States, gross domestic product contracted by almost 33 percent in annual terms in the second quarter of 2020, creating the deepest and sharpest recession since World War II. Economic activity has substantially recovered since, but this plunge in economic activity translated into unprecedented job losses, and the level of U.S. employment in April 2020 showed 25 million fewer people employed relative to January 2020. Since April 2020, employment has recovered at a fast pace, but employment levels are still below those seen at the beginning of 2020.

The study of U.S. labor market conditions typically relies on two monthly surveys. The first is the Current Population Survey (CPS), which surveys around 60,000 U.S. households and is the main source of official unemployment measures.[1] The second is the Current Employment Statistics survey (CES), which surveys around 145,000 U.S. businesses and government agencies and is the main source of changes in payroll, employment, and wages.[2]

Maximiliano Dvorkin is a senior economist and Maggie Isaacson is a research associate at the Federal Reserve Bank of St. Louis.

In normal times, labor market conditions change slowly, and monthly information from the CPS and CES on employment and unemployment is sufficient to gauge the state of labor markets and predict their evolution. However, the rapidly changing conditions during the pandemic called for higher-frequency data to evaluate the state of labor markets in real time. In a recent working paper, Chetty et al. (2020) developed a granular real-time publicly available dataset to address this challenge. We also use high-frequency employment data from two companies with ample coverage of U.S. labor markets, Homebase[3] and Ultimate Kronos Group[4] (UKG). Kurmann, Lalé, and Ta (2020) use the Homebase data to demonstrate how business closures by firm size impacted employment throughout the pandemic. In this article, we use the Homebase and UKG data to create and track useful measures of aggregate and regional employment conditions.

HIGH-FREQUENCY DATA

Homebase is a private company delivering payroll, scheduling, and timesheet tools to businesses. They offer both a free version of their software with basic capabilities and a fee-based version with premium services. Homebase currently provides services to over 100,000 businesses, per their website, across many industries. However, the bulk of Homebase customers are centered in leisure and hospitality, especially in the food and drink industry. After the pandemic started in March 2020, Homebase began releasing daily information on key labor market variables for the companies using their products. These updates took two forms: (i) a summary dataset detailing how much national- and state-level employment, hours, and businesses changed relative to January 2020 and (ii) firm- and worker-level microdata.

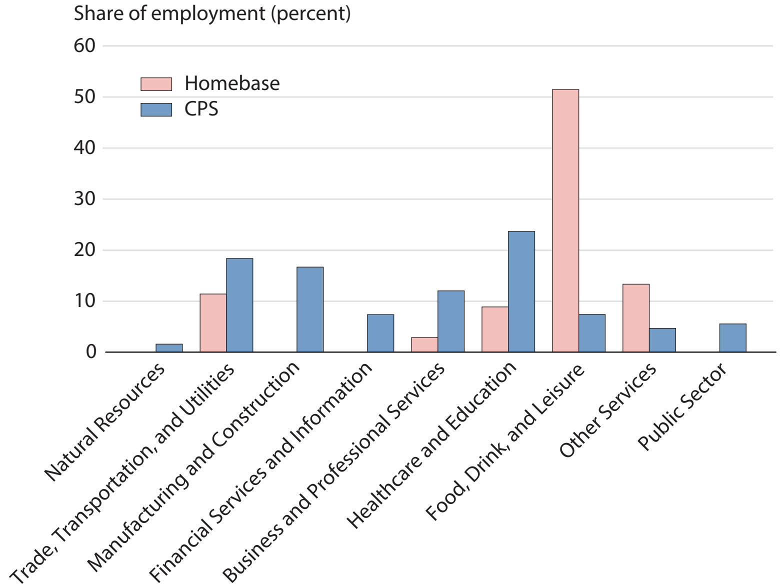
The summary datasets, published daily, use a cohort of roughly 60,000 firms that were active on the Homebase platform in January 2020 to create benchmark numbers of employees, the hours they log, and active businesses in the data. For example, on July 12, 2020, across the entire United States, the number of employees working was down 26.3 percent relative to January, the number of hours worked was down 27.0 percent, and the number of active businesses was down 22.0 percent. These numbers are available also by state and industry.

The microdata contain similar information on hours, employees, and businesses, but this information is not limited to the cohort of firms active in January 2020—it is available for all firms in the sample. For 2020, there are about 100,000 unique companies represented in the data, with an average of 12 employees and about 4,800 hours logged per company between January and December.

Figure 1 compares the distribution of employment across industries for the Homebase data with that of the CPS data. Homebase uses its own industry classification system, which we translate into nine broad sectors according to the North American Industry Classification System (NAICS) to be able to compare the Homebase data with the CPS data.[5] Some major industries, such as manufacturing and construction, are not present in the Homebase data at all, while professional employment is severely underrepresented. The figure shows that the Homebase data are heavily weighted toward establishments that serve food and drink and provide recreational activities. Employees in the food, drink, and leisure industry make up over half of the entire dataset.

Figure 1

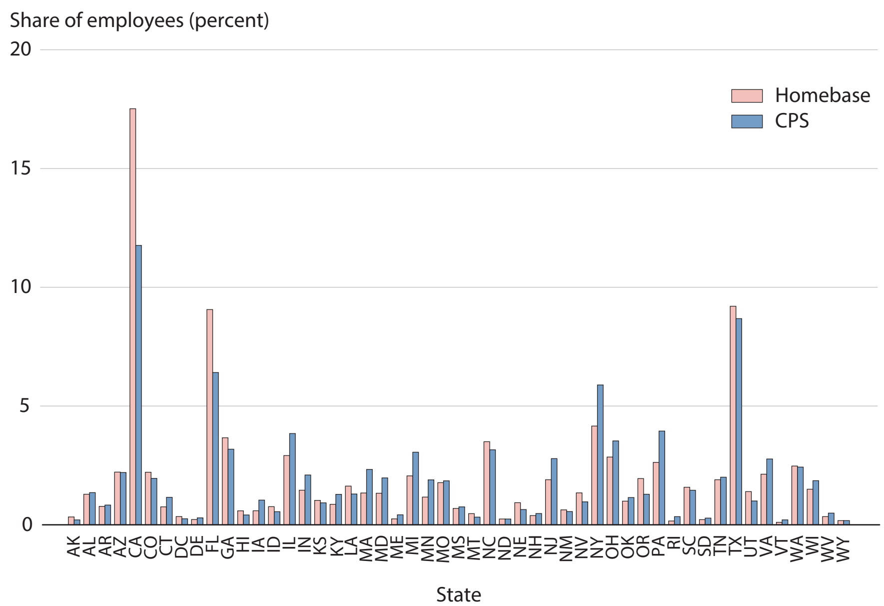
Distributions of Employment in the Homebase and the CPS Data, by Industry



SOURCE: Homebase, CPS, and authors' calculations.

Figure 2

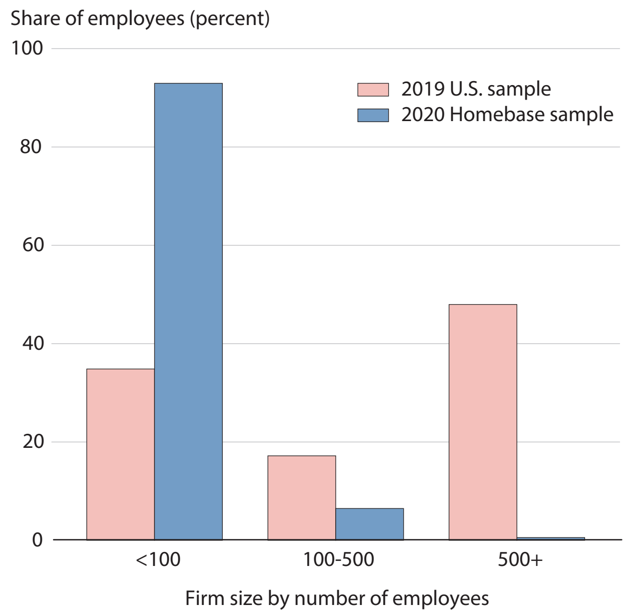
Distributions of Employment in the Homebase and the CPS Data, by State



SOURCE: Homebase, CPS, and authors' calculations.

Figure 3

Distributions of Employment in the Homebase Data and the U.S. Economy, by Firm Size



SOURCE: Homebase, BLS Business Employment Dynamics (https://www.bls.gov/bdm/bdmfirmsize.htm), and authors' calculations.

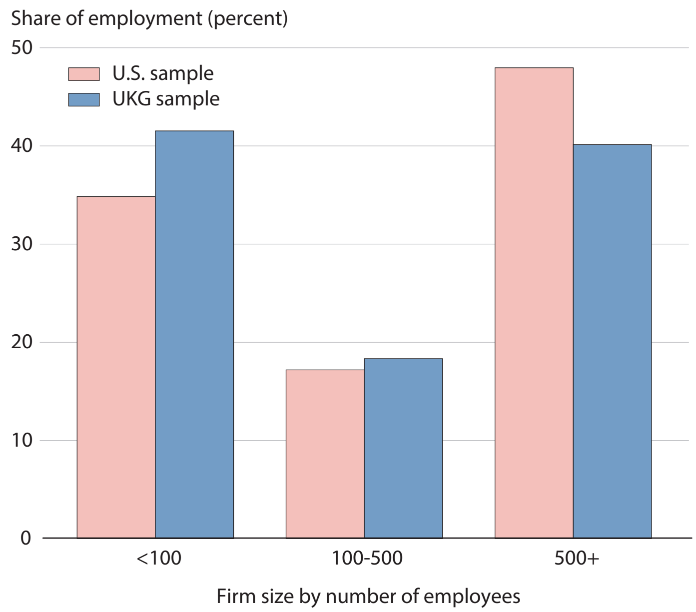
Figure 2 compares the distribution of employment in the Homebase data and in the CPS data across U.S. states. We can see that states such as California, Florida, and Texas are overrepresented in the Homebase data compared with the CPS data, and states such as New York and Illinois are underrepresented. However, all states and Washington, D.C., are present in the Homebase data. The overly large number of firms in the food, drink, and leisure industry might overrepresent certain areas of the country. Other key industries, such as manufacturing, are not present in the Homebase data at all, which would underrepresent areas of the country in which those industries are prevalent.

Many Homebase clients are small firms, which is reflected in the Homebase sample. Almost 90 percent of the businesses in the Homebase data have fewer than 100 employees. Large firms, those with more than 500 employees, represent less than 1 percent of the firms in the Homebase sample. Figure 3 compares the shares of employment by firm size for the Homebase data with those for the U.S. economy in the first quarter of 2020.[6] As the figure shows, most of the employment in the Homebase data corresponds to firms with fewer than 100 employees, while for the U.S. economy this share is close to 37 percent. This is a caveat of the data that we need to consider in our analysis.

The second dataset of interest comes from UKG, a company that also provides payroll and time management tools to businesses and their employees. UKG provides services to over half of the Fortune 1000 firms and has data on millions of employees. The UKG sample, used for

Figure 4

Distributions of Employment in the UKG Dataset and the U.S. Economy, by Firm Size



SOURCE: UKG, BLS Business Employment Dynamics (https://www.bls.gov/bdm/bdmfirmsize.htm), and authors' calculations.

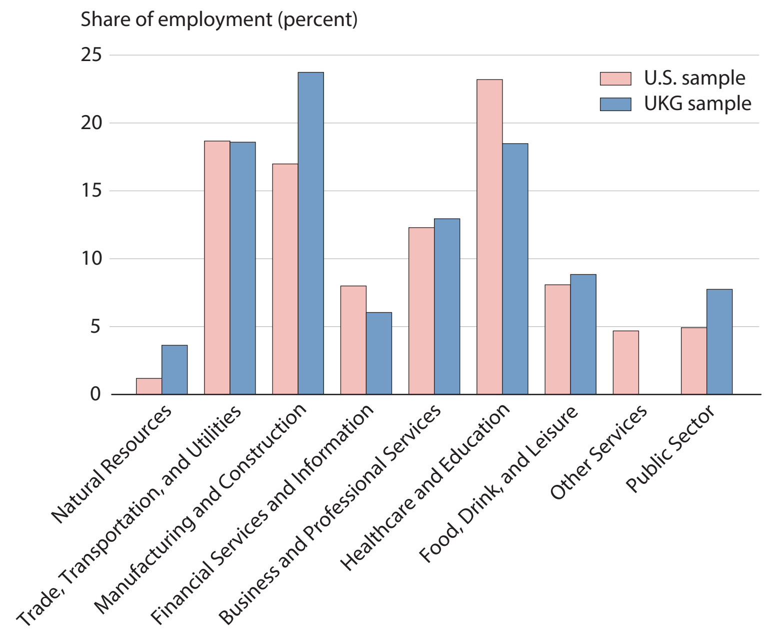
January 2020 to March 2021, contains weekly information on total employee "clock punches" at the national, state, and firm levels. It contains information on over 30,000 businesses across all 50 states.[7] The punch data record each time an employee clocks in or clocks out. For most companies, one employee working one shift would record one punch in and one punch out. However, some require their employees to clock out and back in again for breaks. For those companies, a single employee shift would generate four punches. While the punch data cannot perfectly capture the exact number of employees working in a week, the changes in the number of punches are indicative of the changes in employment in a state or industry.

Figure 4 compares the distribution of employment by firm size for the UKG sample with that for the U.S. economy. As the figure shows, the UKG data include firms of all sizes and the distribution of employment is similar to that for the U.S. economy, with only a slightly larger share of employment in smaller firms and slightly smaller share in very large firms in the UKG sample.

UKG classifies industries into six major sectors: healthcare; manufacturing; retail, hospitality, and food service; services and distribution; public sector and non-profit work; and unclassified. They also categorize industries by NAICS, which in Figure 5 we group into nine broad sectors and compare with the CPS data. Unlike the Homebase dataset, most of the major industries in the UKG dataset have positive employment, including natural resources, financial occupations, and construction, and the distribution of employment is similar to that for the U.S. economy.[8]

Figure 5

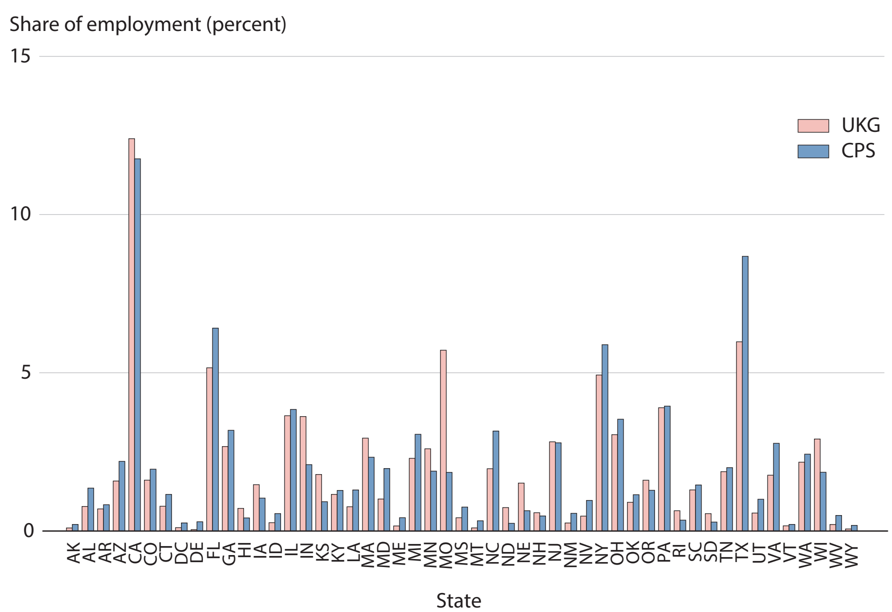
Distributions of Employment in the UKG Dataset and the U.S. Economy, by Industry



SOURCE: UKG, CPS, and authors' calculations.

Figure 6

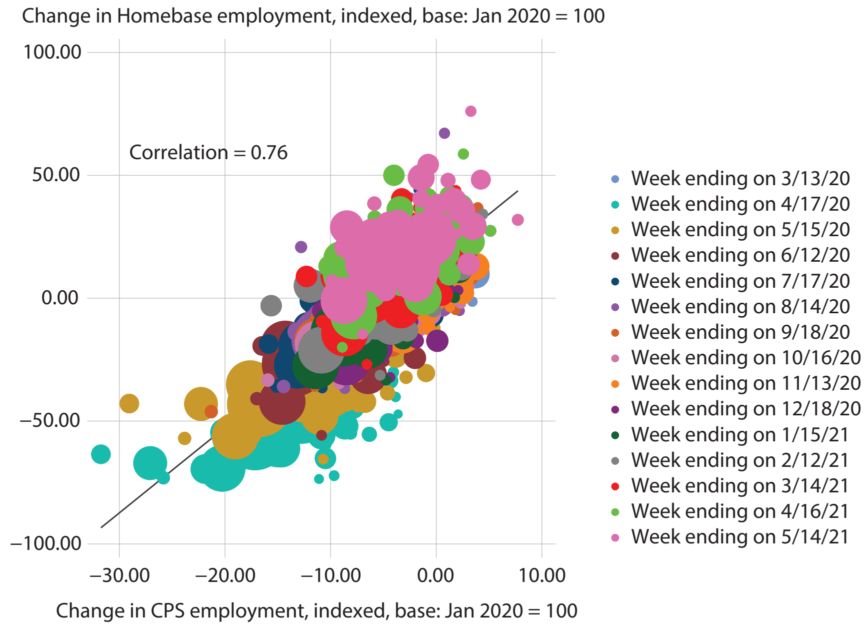
Distributions of Employment in the UKG Dataset and the U.S. Economy, by State



SOURCE: UKG, CPS, and authors' calculations.

Figure 7

Employment Changes in the Homebase Dataset vs. in the CPS Data, by Month and State



SOURCE: Homebase, CPS, and authors' calculations.

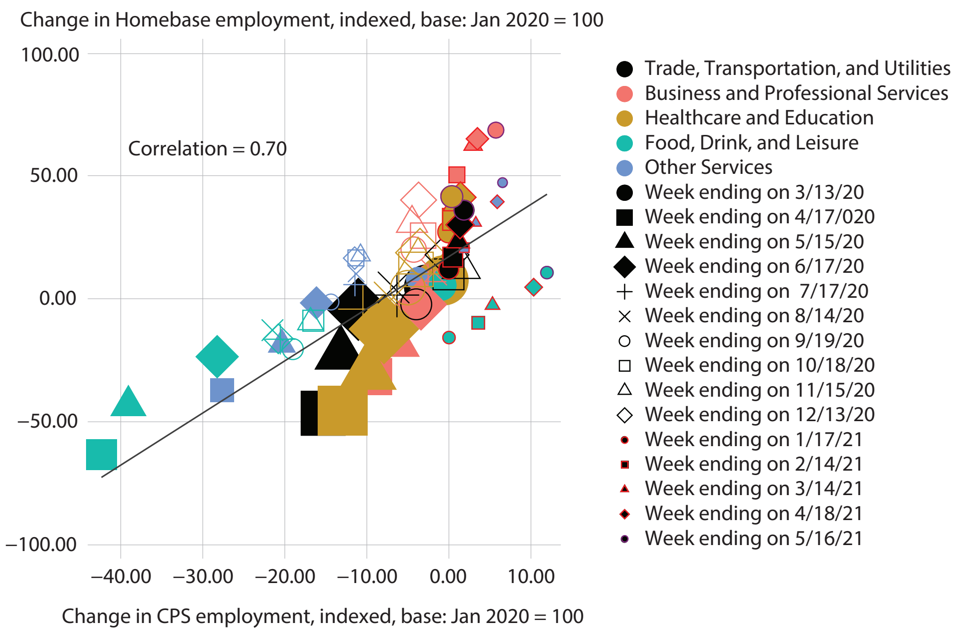
Last, Figure 6 compares the distribution of employment by state for the UKG sample with that for the U.S. economy. The UKG sample has all states represented in the data, and the distribution of employment by state is similar to that for the U.S. economy. However, some individual states, notably Florida, Texas, and Missouri, are over or underrepresented by a few percentage points.

EVOLUTION SINCE MARCH 2020

We focus on employment dynamics in the two datasets since the beginning of the pandemic in March 2020 and compare them with the dynamics in the U.S. economy as measured by the CPS. In order to compare similar time periods, we use data from Homebase and UKG for the week in which the CPS survey was conducted.[9] The scatterplot in Figure 7 shows employment changes relative to January 2020 in the Homebase dataset versus in the CPS data, by month and state. Each circle represents the change in employment in a given month for a U.S. state, the size of each circle corresponds to the state's share of employment in January 2020, and the color of the circle represents the given month. The April 2020 data (teal circles) are among the lowest data on the chart, indicating that the level of employment in April 2020 was much lower relative to January 2020 than in other months. Employment in May 2020 and June 2020 was also low relative to January 2020, but as employment recovered, the values become less negative on average. The points since June 2020 have stayed relatively close together.

Figure 8

Employment Changes in the Homebase Dataset vs. in the CPS Data, by Month and Industry



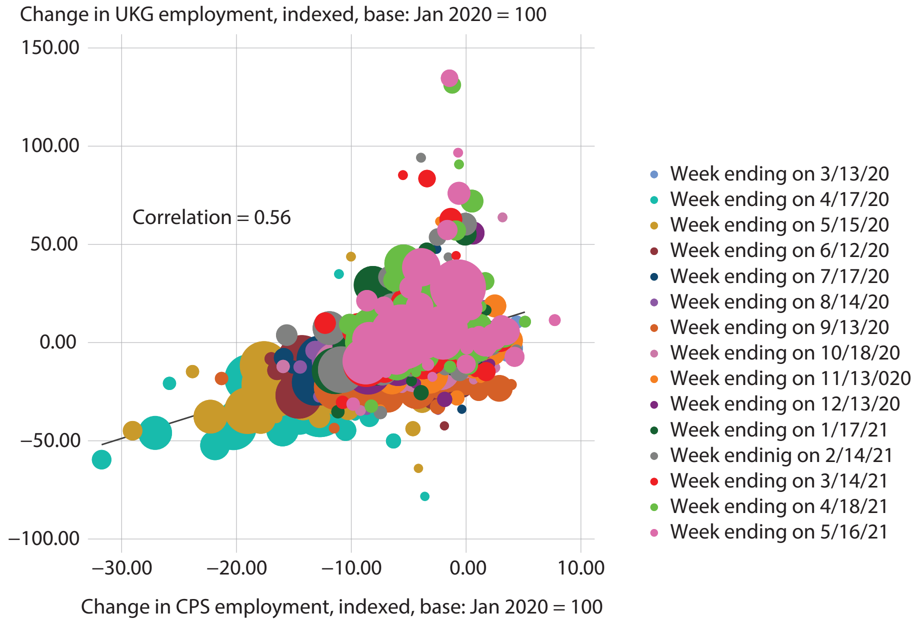
SOURCE: Homebase, CPS, and authors' calculations.

In a similar way, the scatterplot in Figure 8 shows employment changes relative to January 2020 in the Homebase dataset versus in the CPS data, by month and industry. Each marker color represents one of the NAICS sectors, the size of the marker corresponds to the size of the industry in January 2020, and the shape of the marker indicates the month. April 2020 and May 2020 employment across almost all industries dropped sharply relative to January 2020, just like in the state data. Employment in the food, drink, and leisure industry dropped the most out of all the industries in the data, as the restrictions imposed due to the pandemic affected these industries (which include restaurants, hotels, and so on) more than others.

The UKG punch data can be used to map national, state, and industry employment across UKG customers starting in January 2020. In this case, the UKG data track the overall number of punches across a week instead of the number of active employees each day. Figure 9 compares the indexed values of UKG employment to those of CPS employment, by state and by month, using the same markings as in Figure 7. The UKG dataset is highly correlated with the CPS index, and while the correlation is lower than that for the Homebase data with the CPS index, both datasets display similar patterns. The levels of employment in April 2020 and May 2020 compared with those in January 2020 are consistently lower than those for other months. Some small states in terms of employment in the UKG data display large departures from the average changes in the data, which suggests that for these cases the employment changes in the UKG data do not align well with the changes in the CPS data. This could be

Figure 9

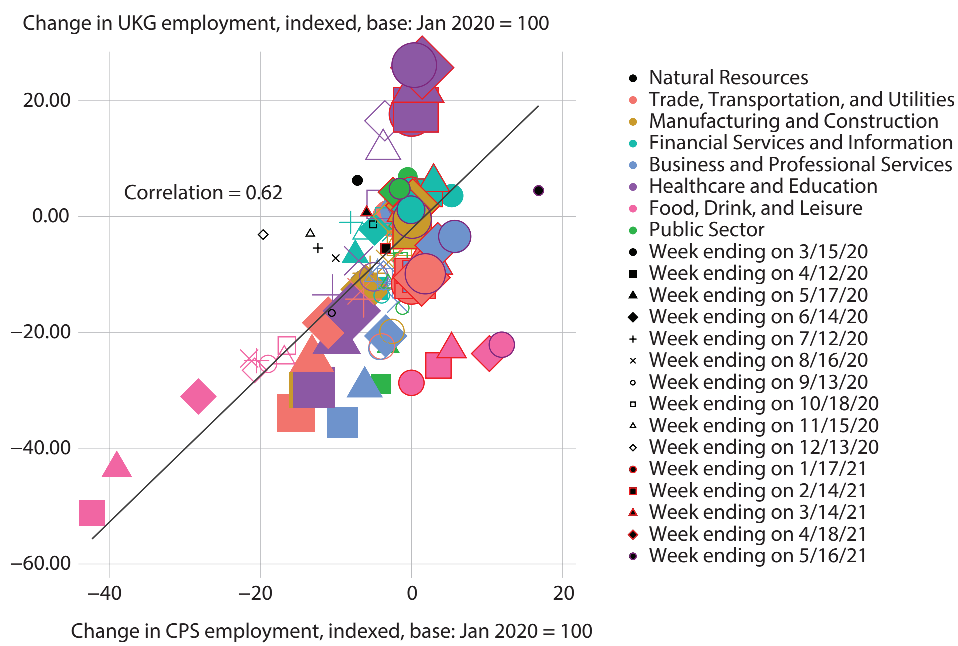
Employment Changes in the UKG Dataset vs. in the CPS Data, by Month and State



SOURCE: UKG, CPS, and authors' calculations.

Figure 10

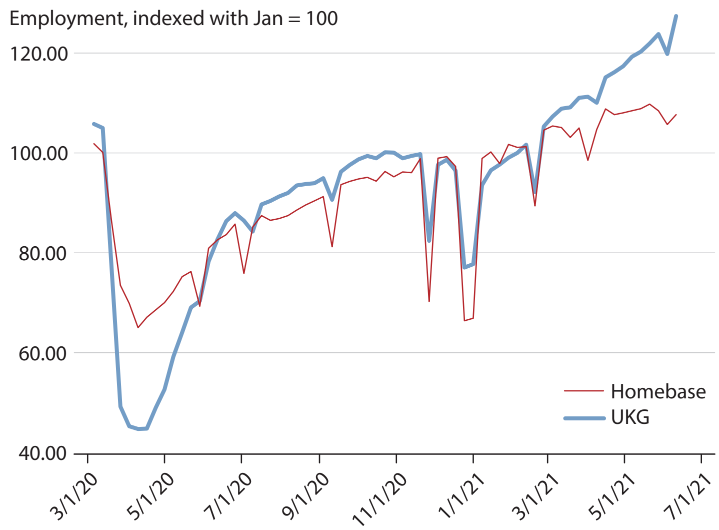
Employment Changes in the UKG Dataset vs. in the CPS Data, by Month and Industry



SOURCE: UKG, CPS, and authors' calculations.

Figure 11

UKG and Homebase Aggregate CEI for the U.S. Economy



SOURCE: Homebase, UKG, CPS, and authors' calculations.

the result of a small sample (a small number of firms and employees in these states in the UKG dataset) leading to an outsized impact by a small number of firms.

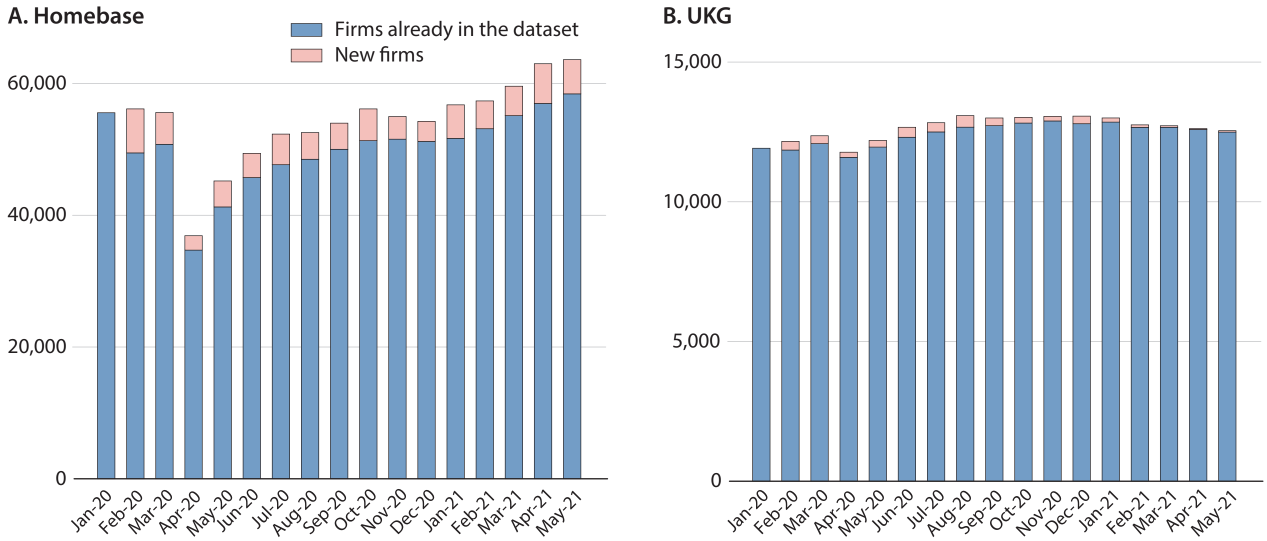
Figure 10 compares employment changes in the UKG dataset and the CPS for the major industries, using the same markings as in Figure 8. The UKG data have a larger set of industries than the Homebase data, and employment in the UKG data is not heavily concentrated in a small set of sectors. As the figure shows, for both datasets, the food, drink, and leisure industry displays the sharpest decrease in employment early in the pandemic.

An advantage of the UKG and Homebase datasets is the high frequency of their data, which is available almost in real time. Figure 11 shows the weekly evolution of aggregate employment in both the UKG and Homebase datasets. They follow similar patterns across the given period, although the magnitudes of the changes vary, with the Homebase dataset having a much larger contraction in employment in April 2020. This finding reflects differences in the compositions of the sectors and firms by size in these datasets. As mentioned, the Homebase dataset has a large number of firms in the food, drink, and leisure industry. In the early weeks of the pandemic, restaurants and bars quickly shut down, leaving many employees with no work. Once the economy began to slowly open again in May and June, many restaurants reopened and were able to bring their employees back to work or hire new staff. Industries in the UKG data are less concentrated in terms of employment, thus the evolution of employment in that dataset does not present such a large initial drop.

In addition, both series present sudden temporary drops in employment, which are due to holidays, such as the Fourth of July, Thanksgiving, and Christmas, when a large number of employees are not at work and thus not clocking in and out.

Figure 12

Number of Old and New Firms in the Homebase Dataset and the UKG Dataset



SOURCE: UKG, Homebase, and authors' calculations.

COINCIDENT EMPLOYMENT INDICES

The Homebase and UKG employment datasets are strongly correlated with the CPS employment data, suggesting that we can use these datasets to predict national employment at a higher frequency and in real time. In this section we discuss how we do this.

As we showed previously, there are potential caveats with the use of these datasets since the composition of the sample of firms in terms of geography, industry, and firm size may differ in some dimensions from that of the U.S. economy. There is one additional problem we must tackle. Every period, new firms are added to the dataset. These new firms are new customers of Homebase and UKG that start using their products, but they are not necessarily newly created firms. While we would like to capture employment growth due to new firm creation in our indicators, using these new firms in the datasets would be inappropriate. At the same time, completely excluding these new firms would bias downward our estimates since the sample would have business closures and firm shutdowns but not any new businesses.

The number of firms in the datasets dropped precipitously in April 2020, as seen in Figure 12. The blue portions of the bars represent the firms that were in the dataset for at least one month prior, while the pink portions of the bar represent any firms entering the dataset for the first time. If we were to use a fixed cohort of firms in the construction of our indicators, our measures of employment would be affected by having a reduction in the number of firms but no firms entering to make up for the loss. Unless the remaining firms in the dataset significantly increased their employment, it would take a long time for any measure of employment to fully return to the levels of January 2020, if it returned fully at all. All of the firms in pink

Table 1

Regression Output from Regressing CPS Index on Homebase Chain Index and UKG Chain Index

Variables	(1) Homebase index	(2) UKG index
Change in Homebase index	0.183*** (0.006)	
Change in UKG index		0.166*** (0.011)
Observations	765	765
R^2	0.636	0.316

NOTE: Each regression also includes a constant term. Standard errors are in parentheses. ***$p < 0.01$.

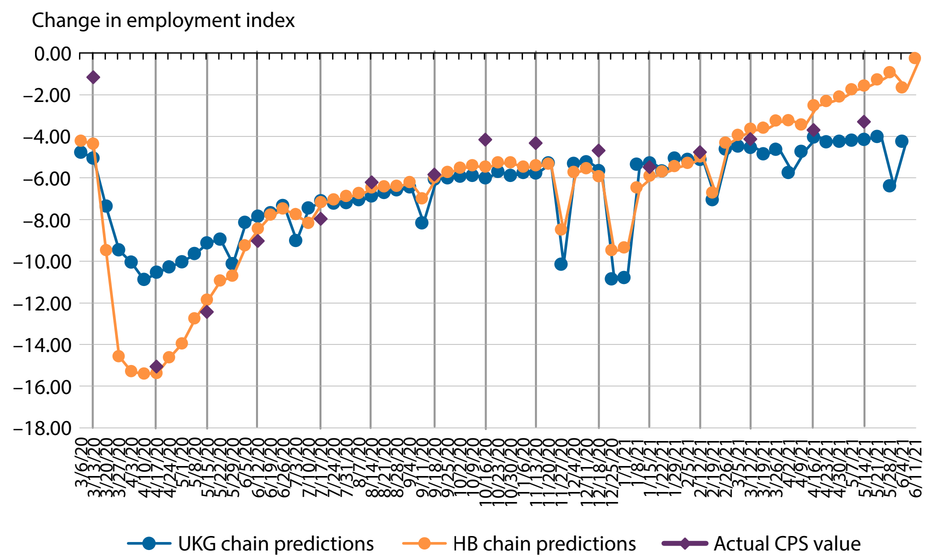
enter the dataset throughout the months we analyze, and any increase in employment that they provide goes uncounted if a fixed set of firms is used. Instead, our approach incorporates information from the new firms entering the dataset, but with a lag. In this way, we allow for the cohort of firms to vary over time, and compute changes in employment from each cohort by period. In other words, we construct a chain index for each of the Homebase and UKG datasets.

We construct a chain index for each state as well as the overall United States on a weekly basis. The Homebase data is first collapsed to weekly frequency and measured from Friday to Friday, while the UKG data is already reported at a weekly level from Sunday to Sunday. The first value in the index in January 2020 is normalized to 100, while the values of the index for the weeks in February use the employment growth between January and February for all the firms present in the dataset in January. Similarly, the index for March is computed using employment growth for the firms present in the dataset in February and so on throughout the year. The base value of employment for each month is the median value of weekly employment in each state in the given month. The raw value for the number of employees in a week is divided by the base value drawn from the month before and multiplied by 100. This creates the relative links for each week. The relative links are then multiplied by the previous month's index and divided by 100 to create the chain index values.

Then, using our chain indices for UKG and Homebase at the state level, we regress a state-level index of employment from the CPS for each month on the state values of each of our indices separately, where the regression is weighted by the level of employment in the CPS in January 2020. In this regression, we use the weekly data from UKG and Homebase that correspond to the survey week of the CPS. The coefficients from the regression, found in Table 1, are used to predict values of employment for the overall United States for each chain index. The R-squared of the regressions are 0.636 for Homebase and 0.316 for UKG, so our indices have substantial forecasting ability. The coefficient of the UKG index is slightly smaller than that of the Homebase index, which implies that our forecasts of the CPS values tend to react more to changes in the Homebase data when compared with similar changes in the UKG data.

Figure 13

CPS Employment and CEI-Predicted Employment



NOTE: Vertical gray lines indicate the selected CPS survey weeks. HB, Homebase.

SOURCE: CPS, UKG, Homebase, and authors' calculations.

Using these regression coefficients, we project the employment indices from the Homebase and UKG data onto CPS values and create a real-time predictor of CPS employment. Figure 13 shows the values of these two predictors over time and compares them to the actual CPS values for employment as deviations from January 2020. We label our predictors the coincident employment indices (CEI).

As the figure shows, both indices track well the evolution of the CPS. This is not surprising since the regression is meant to deliver that. However, over time, the regression coefficients tend to change only slightly; thus, it is the underlying data from Homebase and UKG that are responsible for the good in-sample fit.

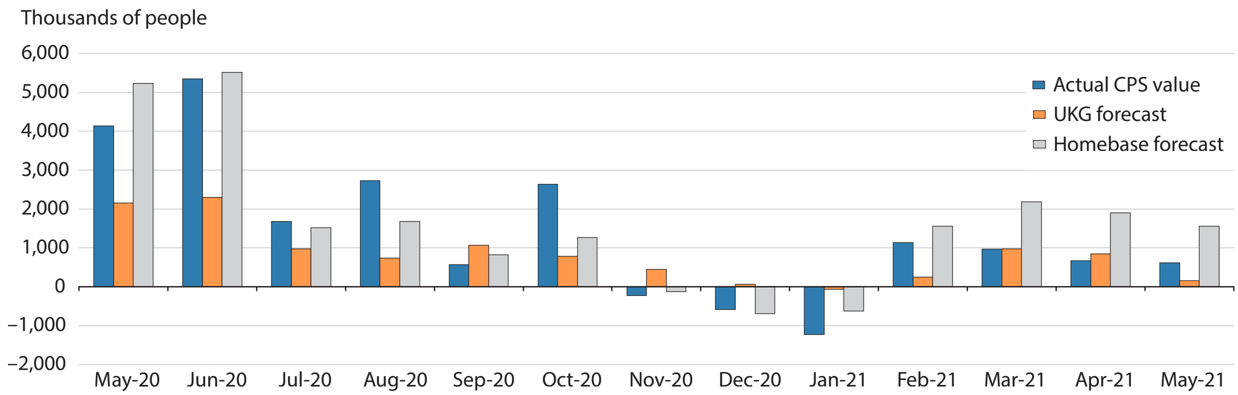
CEI FORECASTING PERFORMANCE

We now ask whether our CEI are good at predicting labor market conditions and compare them with another indicator. We focus on the ADP Research Institute (ADP; 2021) employment forecasts from May 2020 onward (published in their "ADP Employment Reports"). Panel A of Figure 14 compares the Homebase and UKG CEI forecasted changes to the actual changes in employment as measured by the CPS (not seasonally adjusted). To generate the values for the CEI forecasted changes in employment, we computed smoothed versions of each index by taking the average between the value of the index a week before the survey and a week after the survey. The reason for this is that some holidays tend to lower the number of clock punches in the Homebase and UKG data, as some employees take some time off. Whether

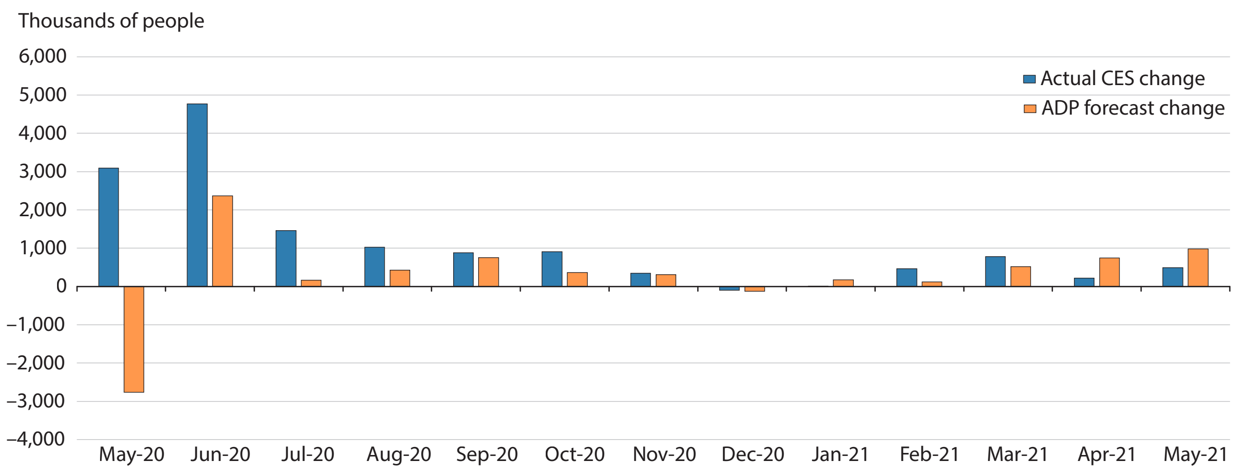
Figure 14

Forecasting Performance During the Pandemic

A. CEI forecasts vs. actual CPS employment (not seasonally adjusted)



B. ADP forecast vs. actual CES private employment (seasonally adjusted)



SOURCE: CPS, UKG, Homebase, CES, ADP, and authors' calculations.

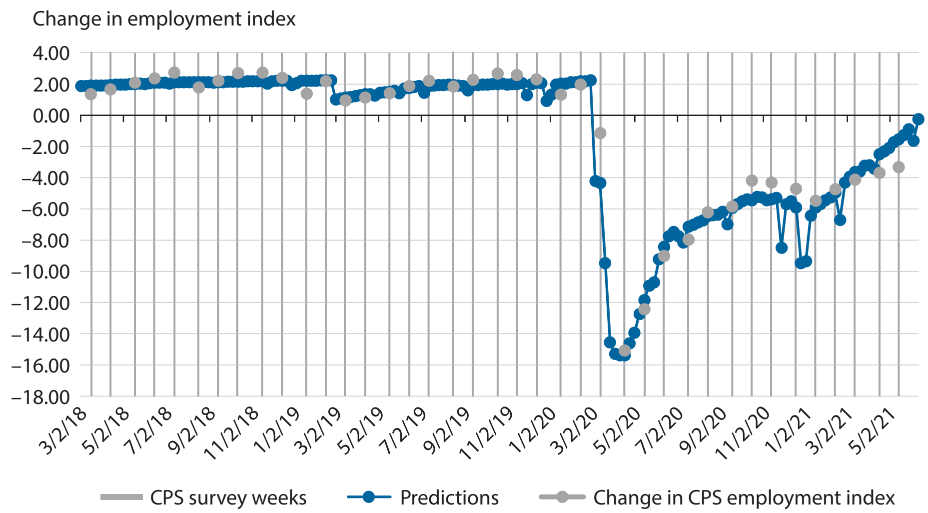
we use the smoothed version in our forecast depends on when holidays fall. For most of the data in Figure 14, we used the smoothed version of the index and then multiplied the index value by the value for employment that month.

ADP forecasts U.S. nonfarm private sector employment as measured by the CES (seasonally adjusted). Thus, Panel B of Figure 14 compares ADP forecasts to the CES employment numbers as published in the first release.[10] The ADP value is the value for the current month minus the value for the previous month (reported in the same report) to generate the predicted change.

While all three indices displayed above are not perfect predictors, the ADP forecast and the CEI values all do a good job predicting CPS employment levels. Toward the beginning of the pandemic, the Homebase and UKG indices do a better job than the ADP forecast.

Figure 15

Spliced Homebase Index Employment Predictions and CPS Employment



SOURCE: Homebase and authors' calculations.

PRE-RECESSION INDICES

Although the different high-frequency indices generate accurate predictions for employment in the current recession, that does not necessarily imply that the same methodology will work for non-pandemic and non-recession times. Using Homebase microdata from 2018 onward, we generate chain index predictions starting in March 2018. Indexing all time periods to January 2018 produces a relatively large constant and a relatively small coefficient on index changes. Thus, instead of indexing everything on January 2018, we separate the data into time periods from March to March of each available year, resulting in three separate indices for three blocks of time. The data from March 2018 to March 2019 are indexed to January 2018, the data from March 2019 to March 2020 to January 2019, and the data from March 2020 to March 2021 to January 2020. Then we replicate the chain index process from above in each time period. Splicing the separate indices together, we generate Figure 15, which compares the separated indices to the CPS unemployment number in each month.

The predictions track the CPS unemployment values extremely well for both pre-recession and recession employment values. The high-frequency nature of the data even tracks the initial employment drop, which is difficult to track with less-frequent datasets.

CONCLUSION

In the second quarter of 2020, the U.S. economy was facing one of the most severe and sudden recessions, with unprecedented employment loss. Yet, the information available up to the early days of May about employment and unemployment only reflected labor market conditions in mid-March before much of the restrictions and lockdowns were set.

Policymakers need accurate and timely data to adopt optimal policy actions. Traditional data sources for labor market data, such as the CPS and CES, provide valuable information but face important lags. In this article, we discuss how new high-frequency data can offer critical information on the state of labor markets in real time. Using data from Homebase and UKG, we construct coincident employment indices to assess labor market conditions at a high frequency.

To a large extent, the number of COVID-19 cases has been steadily falling since January 2021, and state and local governments have lifted most restrictions on daily activities. The recent evolution of our indices suggest that employment has increased and most likely will continue to do so as the pandemic wanes. ■

NOTES

1 The CPS is co-sponsored by the Census Bureau and the Bureau of Labor Statistics (BLS).

2 The CES is sponsored by the BLS.

3 The Homebase data are available at https://joinhomebase.com/.

4 The Ultimate Kronos Group data are available at https://www.ukg.com/.

5 Eleven percent of the firms in the Homebase data do not have a classification and are not included in the distribution of firms by industry.

6 The distribution of employment by firm size for the U.S. economy comes from BLS data and the authors' calculations.

7 In addition, the UKG sample contains information on payrolls, which provides a count on the number of employees receiving paychecks in a given week. The payroll data record the number of paychecks issued, which, for most companies, occurs every two weeks. The payroll data keep weekly records; as such, many companies have payroll levels that oscillate between zero employees recorded and the actual number of paychecks issued.

8 The distribution of employment excludes the 16.7 percent of employees who work for a firm without any industry classification. Of the 29,642 firms in the data, 11,415 firms, or 38.5 percent, do not have an industry classification.

9 The BLS conducts the CPS every month on the week containing the 12th day of the month. We use the same week in our comparison using Homebase and UKG data.

10 The CES employment numbers are subject to revisions. Here we use the first release, which is the first (unrevised) number the BLS publishes, typically on the first Friday of each month.

REFERENCES

ADP Research Institute. "ADP Employment Reports." April 2021; https://adpemploymentreport.com.

Chetty, R.; Friedman, J.N.; Hendren, N.; Stepner, M. and the Opportunity Insights Team. "The Economic Impacts Of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." NBER Working Paper 27431, National Bureau of Economic Research, 2020.

Dvorkin, M. and Bharadwaj, A. "The Recent COVID-19 Spike and the U.S. Employment Slowdown." Federal Reserve Bank of St. Louis *On the Economy Blog*, August 4, 2020; https://www.stlouisfed.org/on-the-economy/2020/august/recent-covid19-us-employment-slowdown.

Kurmann, A.; Lalé, E. and Ta, L. "The Impact of COVID-19 on Small Business Employment and Hours: Real-Time Estimates with Homebase Data." ESG UQÀM, Département des sciences économiques, École des sciences de la gestion, Université du Québec à Montréal, 2020.