Real-Time Economics: A New Platform to Track the Impacts of COVID-19 on People, Businesses, and Communities Using Private Sector Data*

Raj Chetty, John N. Friedman, Nathaniel Hendren, Michael Stepner,
and the Opportunity Insights Team†

May 7, 2020

Abstract

We build a new, publicly available economic tracker that measures economic activity at a high-frequency, granular level. Using anonymized data from several large businesses – credit card processors, payroll firms, job posting aggregators, and financial services firms – we construct statistics on consumer spending, employment rates, incomes, business revenues, job postings, and other key indicators. We report these statistics in real time using an automated pipeline that ingests data from these businesses and reports statistics publicly on the data visualization platform, typically with a lag of three days or less after the relevant transactions occur. We present fine disaggregations of the data, reporting each statistic by county and by industry and, where feasible, by initial (pre-crisis) income level and business size. We illustrate how the tracker can be used by measuring the economic impacts of the COVID-19 crisis on people, businesses, and communities and estimate the causal effects of recent local policy decisions. Going forward, we hope this tracker will serve as a public good that facilitates more precise targeting of policies and rapid diagnosis of the root causes of economic crises.

*We thank Gabriel Chodorow-Reich, John Grigsby, Erik Hurst, Lawrence Katz, James Stock, and Ludwig Straub for helpful comments. We also thank the corporate partners who provided the underlying data used in the Economic Tracker, who as of this version include: Affinity Solutions, Burning Glass, Earnin, Facteus, Homebase, Second Measure, Womply, and Zearn. We are very grateful to Ryan Rippel of the Gates Foundation for his support in launching this project and to Gregory Bruich for early conversations that helped spark this work. The work was funded by the Chan-Zuckerberg Initiative, Bill & Melinda Gates Foundation, Overdeck Family Foundation, and Andrew and Melora Balson. The project was approved under Harvard University IRB 20-0586.

†The Opportunity Insights Economic Tracker Team consists of Matthew Bell, Gregory Bruich, Tina Chelidze, Lucas Chu, Westley Cineus, Sebi Devlin-Foltz, Michael Droste, Shannon Felton Spence, Dhruv Gaur, Federico Gonzalez, Rayshauna Gray, Abby Hiller, Matthew Jacob, Tyler Jacobson, Margaret Kallus, Laura Kincaide, Cailtin Kupsc, Sarah LaBauve, Maddie Marino, Kai Matheson, Kate Musen, Danny Onorato, Sarah Oppenheimer, Trina Ott, Lynn Overmann, Max Pienkny, Jeremiah Prince, Daniel Reuter, Peter Ruhm, Emanuel Schertz, Kamelia Stavreva, James Stratton, Elizabeth Thach, Nicolaj Thor, Amanda Wahlers, Kristen Watkins, Alanna Williams, David Williams, Chase Williamson, Shady Yassin, and Ruby Zhang.

I Introduction

Economic policy decisions have traditionally been made based on employment and business activity data collected from national surveys of households and businesses. Although such statistics have great value in understanding the economy, they have two limitations that have become apparent in the face of the fast-moving COVID-19 pandemic. First, such data are often available only with a significant time lag. For instance, the Employment Situation Summary (i.e., jobs report) released by the Bureau of Labor Statistics on May 8 presents information on employment rates as of the week ending April 12; the next update will not come for another month. Second, due to limitations in sample sizes, such statistics typically cannot be used to assess granular variation across geographies or subgroups; most statistics are typically reported only at the state level and breakdowns by demographic subgroups or sectors are often unavailable.

In this paper, we address these challenges by building a new, freely accessible platform that tracks economic activity in real time at a granular level. Using anonymized and aggregated data from several large businesses – credit card processors, payroll firms, job posting aggregators, and financial services firms – we construct statistics on consumer spending, employment rates, incomes, business revenues, job postings, and other key indicators. We report these statistics in "real time" using an automated pipeline that ingests data from these businesses and reports statistics publicly on the data visualization platform, typically with a lag of three days or less after the relevant transactions occur. We present fine disaggregations of the data, reporting each statistic by county and by industry and, where feasible, by initial (pre-crisis) income level and business size.

Many firms already analyze their own data internally to inform their business decisions and some firms have begun sharing aggregated data with policymakers and researchers during the current crisis. Our contribution is to (1) collect these disparate data sources into a single, publicly accessible platform that eliminates the need to write contracts with specific companies to access relevant data; (2) systematize these data sources by documenting the samples they cover and adjusting for selection biases, seasonal fluctuations, and other statistical issues; and (3) provide the combined series in an interactive data visualization tool that facilitates comparisons across outcomes, areas, and subgroups. The key technical problem that our platform solves is that it allows companies to share data without disclosing non-public information about their business or clients by combining data from multiple sources into a single series. Furthermore, we use standard techniques from the privacy literature – such as masking small cells, outliers, and reporting changes relative to a base period (typically early January 2020) rather than levels – to protect privacy. In short, the platform serves as a coordination device for the use of private sector data to inform public policy, one that we hope will expand over time and be a useful resource in this crisis and beyond.

We illustrate the value of the tracker by analyzing the impacts of recent policy decisions in the COVID-19 crisis, focusing in particular on state shutdowns and re-openings. Perhaps surprisingly, we find these policies have little or no impact on economic activity. The decline in economic activity – consumer spending, the number of small business open, and employment – occurred in most cases *before* states "shut down," consistent with other recent work examining data on hours of work and movement patterns (Bartik et al. 2020, Villas-Boas et al. 2020). Moreover, we show that recent policies ending these shut-downs in certain states such as Georgia and South Carolina have not been associated with significant increases in economic activity. These findings suggest that the primary barrier to economic activity is the threat of COVID-19 itself as opposed to legislated economic shutdowns. This simple analysis illustrates the utility of the tracker: this finding would not be evident in traditional government survey data for several months, but is easily observed in private sector data a few days after policy changes are made.

Although the data compiled here contain considerable information, we caution that they are not explicitly designed to provide information that is representative of individuals or businesses in the U.S. Unlike official government statistics, which are typically based on sampling frames that are designed to provide representative information, the private sector data sources we aggregate reflect the activity of the clients of the set of firms who provide data. There is no guarantee that the statistics from such data sources reflect total economic activity perfectly. To mitigate such biases, we sought data from sources that have large samples (e.g., at least one million individuals) and span well-defined, broad sectors of the economy (e.g., small businesses, hourly workers, etc.). We then characterize the samples that each series represents and assess the extent to which they match historical government statistics, to the extent feasible by the privacy requirements set forth by companies and restrictions on the disclosure of material nonpublic information. Going forward, it will be valuable to compare our indices to public statistics to understand their biases and construct aggregates that are more representative of overall economic activity – underscoring the value of having both traditional government statistics as benchmarks and private sector data as leading indicators. Despite these limitations, we believe the raw statistics published here contain useful information because the scale of the shocks induced by the COVID-19 crisis is very large relative to plausible biases that may be induced by non-representative sampling, as shown e.g. by Cajner et al. (2020) in the context of payroll data.

Our work builds on and contributes to a rapidly evolving literature on the economic impacts of COVID-19 as well as a long literature in macroeconomics on the measurement of economic activity at business cycle frequencies. Several recent papers have used private sector data analogous to what we assemble here to analyze labor market trends (e.g., Bartik et al. 2020, Kurmann, Lalé, and Ta 2020, Kahn, Lange, and Wiczer 2020), spending patterns (e.g., Baker et al. 2020, Chen, Qian, and Wen 2020), business revenues (e.g., Alexander and Karger 2020), and social distancing (e.g., Allcott et al. 2020, Chiou and Tucker 2020, Goldfarb and Tucker 2020, Mongey, Pilossoph, and Weinberg 2020). These papers have identified a number of important patterns that we observe in our data as well, such as larger reductions in income and employment for lower-income workers and concentrated impacts in certain industries such as food and accommodation (e.g., Cajner et al. 2020). Each of these papers typically analyzes one or two of these data sources, obtained through a data use agreement with the relevant firm. We combine many of these datasets into a unified, freely accessible platform that is automatically updated to pull the most recent data from companies' internal databases. This approach eliminates the need to obtain specific permissions to use data from each company, thereby democraticizing data access and providing a public good that we hope will support the work of researchers, policy makers, and the general public.[1]

Going forward, we envision two roles for such a platform to support macroeconomic policy. First, the data can permit precise targeting of policies to subgroups and areas that are most affected by a crisis by directly revealing which groups have been impacted most. Second, the data can be used to learn rapidly from heterogeneity across areas, as different places are often hit by differential shocks and pursue different local policy responses. This approach can permit rapid diagnosis of the root factors underlying an economic crisis, potentially facilitating more effective macroeconomic policy responses.

The paper is organized as follows. The next section describes our data and the methods we use to construct the indices. In Section 3, we illustrate how the tracker can be used by presenting some simple illustrative event studies of key outcomes around legislated shutdowns and re-openings at the state level. Section 4 concludes by discussing policy implications and potential future applications

1. Additionally, building on work by Gupta et al. (2020), we build a systematic, quantitative list of key policy changes made in the COVID-19 crisis at the federal, state, and local levels that can be used by researchers and policymakers to identify causal effects of policy changes and uncover mechanisms underlying economic outcomes.

of the tracker. Technical details on data, methods, and supplementary analyses are available in an online appendix. This paper will be updated to include further analysis as we obtain and analyze more data, both from the data providers described below and other companies whose data are currently being incorporated into the platform.

II Data and Methods

We use anonymized data from several private companies to construct indices of spending, employment, and other metrics. In this section, we describe how we construct each series. To facilitate comparisons between series, we adopt the following set of principles when constructing each series (wherever feasible given data availability constraints).

First, the central challenge in using private sector data to measure economic activity is that they capture information exclusively about the customers each company serves, and thus are not necessarily representative of the full population. Instead of attempting to adjust for this non-representative sampling, we characterize the portion of the economy that each series captures by comparing the characteristics of each sample we use to national benchmarks.[2]

Second, we clean each series to remove artifacts that arise from changes in the data providers' coverage or systems. For instance, firms' clients often change discretely, sometimes leading to discontinuous jumps in series, particularly in small cells. We systematically search for large jumps in series (e.g., >80%), seek to understand their root causes, and address such discontinuities by imposing continuity as described below.

Third, many series exhibit substantial periodic fluctuations across days. We address such fluctuations through aggregation, e.g. reporting 7-day moving averages to smooth daily fluctuations. Certain series – most notably consumer spending and business revenue – exhibit strong weekly fluctuations that are autocorrelated across years (e.g., a surge in spending around the holiday season). We de-seasonalize such series by normalizing each week's value in 2020 relative to corresponding values for the same week in 2019 in our baseline analysis, but also report raw values for 2020 for researchers who prefer to make alternative seasonal adjustments.

Fourth, to protect confidentiality of business market shares, we do not report levels of the series.

2. An alternative approach is to reweight samples based on observable characteristics – e.g., industry – to match national benchmarks. We do not pursue such an approach here because the samples we work with track relevant national benchmarks – at least for the scale of shocks induced by the COVID crisis – without such reweighting. However, the disaggregated data we report by industry and county can be easily reweighted as desired in future applications.

Instead, we report indexed values that show percentage changes relative to mean values in January 2020.[3] We also suppress small cells and exclude outliers to protect the privacy of individuals and businesses, with thresholds that vary across datasets as described below.

Finally, we seek to release data series at the highest possible frequency. To limit revisions, we permit a sufficient lag to adjust for reporting delays (typically one week). We disaggregate each series by two-digit NAICS industry code; by county, metro area, and state; and by income quartile where feasible.[4]

We now describe each of the series in turn, discussing the raw data sources, construction of key variables, and cross-sectional comparisons to publicly available benchmarks.[5] All of the data series described below can be freely downloaded from the Economic Tracker website: www.tracktherecovery.org.

II.A Consumer Spending: Affinity Solutions

We measure consumer spending using aggregated and anonymized consumer purchase data collected by Affinity Solutions Inc, a company that aggregates consumer credit and debit card spending information to support a variety of financial service products.

We obtain raw data from Affinity Solutions at the county-by-ZIP code income quartile-by-industry-by-day level starting from January 1, 2019. Industries are defined by grouping together similar merchant category codes. ZIP code income quartiles are constructed at the national level using Census data on population and median household income by ZIP. Cells with fewer than five unique card transactions are masked.

The raw data include several discontinuous breaks caused by entry or exit of credit card providers from the sample. We identify these breaks using data on the total number of active cards in the cell. We then estimate the discontinuous level shift in spending resulting from the break (using a standard regression discontinuity estimator). At the state level (including Washington, DC), we adjust the series within each cell by adding the RD estimate back to the raw data to obtain a smooth series. At the county-level, there is too much noise to implement a reliable correction, so we exclude counties that exhibit such breaks from the sample. After cleaning the raw data in

3. We always norm after summing to a given cell (e.g. geographic unit, industry, etc.) rather than at the firm or individual level. This dollar-weighted approach overweights bigger firms and higher-income individuals, but leads to smoother series and is arguably more relevant for certain macroeconomic policy questions (e.g., changes in aggregate spending).

4. We construct metro area values for large metro areas using a county to metro area crosswalk described in the Appendix.

5. We benchmark trends in each series over time to publicly-available data in the context of our analysis in the next section.

this manner, we construct daily values of the consumer spending series using a seven-day moving average of the current day and previous six days of spending. We then seasonally adjust the series by dividing each calendar date's 2020 value by its corresponding value from 2019.[6] Finally, we index the seasonally-adjusted series relative to pre-COVID-19 spending by dividing each day's value by the mean of the seasonally-adjusted seven-day moving average from January 8-28.

Comparison to QSS and MRTS. Total debit and credit card spending in the U.S. was $7.08 trillion in 2018 (Board of Governors of the Federal Reserve System 2019), approximately 50% of total personal consumption expenditures recorded in national accounts. Affinity Solutions captures nearly 10% of debit and credit card spending in the U.S. To assess which categories of spending are covered by the Affinity data, Appendix Figure 1 compares the spending distributions across sectors to spending captured in the nationally representative Quarterly Services Survey (QSS) and Monthly Retail Trade Survey (MRTS). Affinity has broad coverage across industries. However, as expected, it over-represents categories where credit and debit cards are used for purchases. In particular, accommodation and food services and clothing are a greater share of the Affinity data than financial services and motor vehicles. We therefore view Affinity as providing statistics that are representative of total card spending (but not total consumer spending). We assess whether Affinity captures changes in card spending around the crisis in Section 3.1 below.

II.B Small Business Revenue: Womply

We obtain data on small business transactions and revenues from Womply, a company that aggregates data from several credit card processors to provide analytical insights to small businesses and other clients. In contrast to the Affinity series on consumer spending, which is a cardholder-based panel covering total spending, Womply is a firm-based panel covering total revenues of small businesses. The key distinction is that location in Womply refers to the location where the business transaction occurred as opposed to the location where the cardholder lives.

We obtain raw data on small business transactions and revenues from Womply at the ZIP-industry-day level starting from January 1, 2019.[7] Small businesses are defined as businesses with annual revenue below Small Business Administration thresholds. To reduce the influence of outliers, firms outside twice the interquartile range of firm annual revenue within this sample are excluded and the sample is further limited to firms with 30 or more transactions in a quarter and more than

6. We divide the daily value for February 29, 2020 by the average value between the February 28, 2019 and March 1, 2019.

7. We crosswalk Womply's transaction categories to two-digit NAICS codes using an internally generated Womply category-NAICS crosswalk, and then aggregate to NAICS supersectors.

one transaction in 2 out of the 3 months.

We aggregate these raw data to form two publicly available series at the county by industry level: one measuring total small business revenue and another measuring the number of small businesses open. We measure small business revenue as the sum of all credits (generally purchases) minus debits (generally returns). We define small businesses as being open if they have a transaction in the last three days. We exclude counties with a total average revenue of less than $250,000 during the pre-COVID-19 period (January 4-31).

For each series, we construct daily values in exactly the same way that we constructed the consumer spending series. We first take a seven-day moving average, then seasonally adjust by dividing each calendar date's 2020 value by its corresponding value from 2019. Finally, we index relative to pre-COVID-19 by dividing the series by its average value over January 4-31.

Comparison to QSS and MRTS. Appendix Figure 1 shows the distribution of revenues observed in Womply across industries in comparison to national benchmarks. Womply revenues are again broadly distributed across sectors, particularly those where card use is common. A larger share of the Womply revenue data come from industries that have a larger share of small businesses, such as food services, professional services, and other services, as one would expect given that the Womply data only cover small businesses.

II.C Employment and Earnings: Earnin and Homebase

We use two data sources to obtain information on employment and earnings for low-income workers: Earnin and Homebase.

Earnin is a financial management application that provides its members with access to their income as they earn it. Workers sign up for Earnin individually using a cell phone app, which tracks their hours using GPS location information and records payroll information from bank accounts. Many lower-income workers across a wide spectrum of firms – ranging from the largest firms and government employers in the U.S. to small businesses – use Earnin; we discuss the characteristics of these workers further below. We obtain raw data from Earnin at the worker-day level with information on home ZIP, workplace ZIP, industry and firm size decile from January 2020 to present.[8] We use these data to measure hours worked, total payroll, and hourly wage rates for low-income employees. We assign workers to locations using their workplace ZIP codes. We suppress estimates for ZIP codes with fewer than 50 worker-days observed in Earnin over the period January

8. We map each firm to a NAICS code using firm names and a custom-built crosswalk constructed by Digital Divide Data. We obtain data on firm sizes from Reference USA.

4-31.

Homebase provides scheduling tools for small businesses (on average, 8.4 employees) such as restaurants (64% of employees for whom sectoral data are available) and retail stores (15% of employees for whom sectoral data are available). Unlike Earnin, Homebase provides a complete roster of workers at a given firm, but only covers workers at small businesses. We obtain de-identified individual-level data on hours and total pay for employees at firms that contract with Homebase at the establishment-worker-day level, starting on January 1, 2018. We restrict this sample to non-salaried employees. We then form each aggregate series at the county and industry level, assigning location based on the ZIP code of establishment. To protect confidentiality, we suppress estimates for cells with fewer than 10 Homebase clients in January 2020.

In both datasets, we measure hours worked as a seven-day moving average of total hours worked, expressed as a percentage change relative to hours worked between January 4-31 and total employment as a seven-day moving average of total number of active employees, expressed as a percentage change relative to January. In the Homebase data, we measure hourly wage rates using the change in the first reported hourly wage rate in the current week and the average reported wage between January 4-31, 2020, divided by that average. Finally, we measure the total earnings of workers using a seven-day moving average of earnings divided by the average daily total earnings of those workers between January 4-31. In the Earnin data, where we do not observe individual identifiers, we measure wages as the seven-day moving average of daily mean wages, expressed as a percentage change from daily mean wages between January 4-31. In addition to hours worked, we also observe receipt of paychecks in the Earnin data. We calculate total daily worker earnings by distributing each worker's earnings at the end of their pay period over each day in their pay period. We then measure the change in worker earnings as the seven-day moving average of total worker earnings, expressed as a percentage change relative to January 4-31.

Comparisons to OES and QCEW. Appendix Figure 2 compares the industry composition of the Earnin and Homebase samples to nationally representative statistics from the Quarterly Census of Employment and Wages (QCEW). The Earnin sample is fairly representative of the broader industry mix in the U.S., although high-skilled sectors (such as professional services) are under-represented. Homebase has a much larger share of workers in food services, even relative to small establishments (those with fewer than 50 employees) in the QCEW, as expected given its client base.

Overall, annualizing January earnings would imply median earnings of roughly $23K per year

($11-12 per hour). In Appendix Table 1, we compare the median wage rates of workers in Earnin and Homebase to nationally representative statistics from the BLS's Occupational Employment Statistics. Workers enrolled in Earnin have median wages that are at roughly the 10th percentile of the wage distribution within each NAICS code. The one exception is the food and drink industry, where the median wages are close to the population median wages in that industry (reflecting that most workers in food services earn relatively low wages). Homebase exhibits a similar pattern, with lower wage rates compared to industry averages, except in sectors that have low wages, such as food services and retail.

We conclude based on these comparisons that Earnin and Homebase provide statistics that may be representative of low-wage (bottom-quintile) workers. Earnin provides data covering such workers in all industries, whereas Homebase is best interpreted as a series that reflects workers in the restaurant and retail sector.

II.D Job Postings: Burning Glass

We obtain data on job postings from 2007 to present from Burning Glass Technologies. Burning Glass aggregates nearly all jobs posted online from approximately 40,000 online job boards in the United States. Burning Glass then removes duplicate postings across sites and assigns attributes including geographic locations, required job qualifications, and industry.

We obtain raw data on job postings at the industry-week-job qualification-county level from Burning Glass. Industry is defined using select NAICS supersectors, aggregated from 2-digit NAICS classification codes assigned by a Burning Glass algorithm. Job qualifications are defined using ONET Job Zones. These job zones are mutually exclusive categories that classify jobs into five groups: needing little or no preparation, some preparation, medium preparation, considerable preparation, or extensive preparation. We also obtain analogous data broken by educational requirements (e.g., high school degree, college, etc.).

Comparison to JOLTS. Burning Glass data have been used extensively in prior research in economics; for instance, see Hershbein and Kahn (2018) and Deming and Kahn (2018). Carnevale, Jayasundera, and Repnikov (2014) compare the Burning Glass data to government statistics on job openings and characterize the sample in detail. In Appendix Figure 3, we compare the distribution of industries in the Burning Glass data to nationally representative statistics from the Bureau of Labor Statistics' Job Openings and Labor Market Turnover Survey (JOLTS) in January 2020. In general, Burning Glass is well aligned across industries with JOLTS, with the one exception that it under-covers government jobs. We therefore view Burning Glass as a sample representative of private sector jobs in the U.S.

II.E Education: Zearn

Zearn is an education nonprofit that partners with schools to provide a math program, typically used in classrooms, that combines in-person instruction with digital lessons. Many schools continued to use Zearn as part of their math curriculum after COVID-19 induced schools to shift to remote learning.

We obtain data on the number of students using Zearn Math and the number of lessons they completed at the school-grade-week level. The data we obtain are masked such that any county with fewer than two districts, fewer than three schools, or fewer than 50 students on average using Zearn Math during the pre-period is excluded. We fill in these masked county statistics with the commuting zone mean whenever possible. We winsorize values reflecting an increase of greater than 300% at the school level. We exclude schools who did not use Zearn Math for at least one week from January 6 to February 7 and schools that never have more than five students using Zearn Math during our analysis period. To reduce the effects of school breaks, we replace the value of any week for a given school that reflects a 50% decrease (increase) greater than the week before or after it with the mean value for the three relevant weeks.

We measure online math participation as the number of students using Zearn Math in a given week. We measure student progress in math using the number of lessons completed by students each week. We aggregate to the county, state, and national level, in each case weighting by the average number of students using the platform at each school during the base period of January 6-February 7, and we normalize relative to this base period to construct the indices we report.

Comparison to American Community Survey. In Appendix Table 2, we assess the representativeness of the Zearn data by comparing the demographic characteristics of the schools for which we Zearn data (based on the ZIP codes in which they are located) to the demographic characteristics of K-12 students in the U.S. as a whole. In general, the distribution of income, education, and race and ethnicity of the schools in the Zearn sample is similar to that in the U.S. as a whole suggesting that Zearn likely provides a fairly representative picture of online learning for public school students in the U.S.

II.F Public Data Sources: UI Records, COVID-19 Incidence, and Google Mobility Reports

Unemployment Benefit Claims. We collect county-level data by week on unemployment insurance claims starting in January 2020 from state government agencies since no weekly, county-level national data exist. Location is defined as the county where the filer resides. We use the initial claims reported by states, which sometimes vary in their exact definitions (e.g., including or excluding certain federal programs). In some cases, states only publish monthly data. For these cases, we impute the weekly values from the monthly values using the distribution of the weekly state claims data from the Department of Labor (described below). We construct an unemployment claims rate by dividing the total number of claims filed by the 2019 Bureau of Labor Statistics labor force estimates. Note that county-level data are available for 22 states, including the District of Columbia.

We also report weekly unemployment insurance claims at the state level from the Office of Unemployment Insurance at the Department of Labor. Here, location is defined as the state liable for the benefits payment, regardless of the filer's residence. We report both new unemployment claims and total employment claims. Total claims are the count of new claims plus the count of people receiving unemployment insurance benefits in the same period of eligibility as when they last received the benefits.

COVID-19 Data. We report the number of new COVID-19 cases and deaths each day using publicly available data from the New York Times available at the county, state and national level.[9] We also report daily state-level data on the number of tests performed per day per 100,000 people from the COVID Tracking Project.[10] For each measure - cases, deaths, and tests – we report two daily series per 100,000 people: a seven-day moving average of new daily totals and a cumulative total through the given date.

Google Mobility Reports. We use data from Google's COVID-19 Community Mobility Reports to construct measures of daily time spent at parks, retail and recreation, grocery, transit locations, and workplaces.[11] We report these values as changes relative to the median value for the corresponding day of the week during the five-week period from January 3rd - February 6, 2020. Details on place

9. See the New York Times data description for a complete discussion of methodology and definitions. Because the New York Times groups all New York City counties as one entity, we instead use case and death data from New York City Department of Health data for counties in New York City.

10. We use the Census Bureau's 2019 population estimates to define population when normalizing by 100,000 people. We suppress data where new counts are negative due to adjustments in official statistics.

11. Google Mobility trends may not precisely reflect time spent at locations, but rather "show how visits and length of stay at different places change compared to a baseline." We call this "time spent at a location" for brevity.

types and additional information about data collection is available from Google. We use these raw series to form a measure of time spent outside home as follows. We first use the American Time Use survey to measure the mean time spent inside home (excluding time asleep) and outside home in January 2018 for each day of the week. We then multiply time spent inside home in January with Google's percent change in time spent at residential locations to get an estimate of time spent inside the home for each date. The remainder of waking hours in the day provides an estimate for time spent outside the home, which we report as changes relative to the mean values for the corresponding day of the week in January 2018.

III Illustrative Application: Government Shut-Downs

We illustrate the value of the tracker by analyzing the impacts of state shutdowns and re-openings during the COVID-19 crisis using event studies.

Figure 1 plots several outcomes around the day on which a state-level shutdown was implemented. Figure 1A plots the average series for states that shut down "early," defined as those that issued a stay-at-home order and non-essential business closure in the week of March 19-26. Figure 1B plots the average series across states that issued stay at home orders after this period. Consistent with previous analyses (Bartik et al. 2020, Villas-Boas et al. 2020), hours of work begins to fall prior to the formal date of the state-level shut-down. We also see similar declines in other series such as consumer spending, small business spending, and time spent at work, all well before the shutdowns. Broadly, these patterns suggests that the decline in economic activity was not driven directly by the formal shut-downs themselves, but rather a general response to the onset of the national COVID-19 epidemic.

Several states began efforts to "re-open" their economies by ending shut-down orders in late April and early May. Figure 2 evaluates the early impacts of these policy changes by plotting the same set of outcomes shown in Figure 1 for four states that have implemented re-opening policies: Georgia, Oklahoma, South Carolina, and Alaska.

Figure 2A shows that we do not see an increase in economic activity in Alaska after it lifted its stay-at-home order on April 24. Consumer spending, employment and hours at small businesses, the number of small businesses that are open, and time spent at work all remain relatively similar to their levels prior to April 24. We find similar patterns in other states as well, as illustrated in Panels B-D. This suggests the primary factor limiting economic activity are choices being made by individuals and businesses in response to the threat of COVID-19 itself, as opposed to government policies that impose restrictions on economic activity.

The simple analysis in Figures 1 and 2 illustrates the utility of the tracker. These findings would not be evident in traditional government survey data for several months, but are easily observed in private sector data a few days after policy changes are made.

IV Conclusion

Data held by private companies provide an unprecedented capacity to measure economic activity at a granular level very rapidly. These data have become increasingly integral for corporations in improving business decisions. In this paper, we have constructed a freely available platform that harnesses the same data with the aim of supporting policymakers, non-profits, and the public seeking to make better decisions.

In these uncertain and unprecedented times, we hope this real-time economic tracker provides valuable information for understanding the state of the economy and facilitating the national recovery. We look forward to expanding upon and improving this tracker as additional companies contribute data and reporting analyses that emerge from these data. More broadly, we hope the approach proposed here will serve as a template to permit real-time responses to changes in economic conditions going forward.

References

Alexander, Diane, and Ezra Karger. 2020. "Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior." *Federal Reserve Bank of Chicago Working Paper No. 2020-12* (April).

Allcott, Hunt, Levi Boxell, Jacob C Conway, Matthew Gentzkow, Michael Thaler, and David Y Yang. 2020. "Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic." *NBER Working Paper No. 26946* (April).

Baker, Scott R, R. A Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis. 2020. "Income, Liquidity, and the Consumption Response to the 2020 Economic Stimulus Payments." *NBER Working Paper No. 27097* (May).

Bartik, Alexander W., Marianne Bertrand, Feng Li, Jesse Rothstein, and Matt Unrath. 2020. "Labor market impacts of COVID-19 on hourly workers in small- and medium-sized businesses: Four facts from Homebase data" (April).

Board of Governors of the Federal Reserve System. 2019. *The 2019 Federal Reserve Payments Study.*

Cajner, Tomaz, Leland D. Crane, Ryan A. Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz, and Ahu Yildirmaz. 2020. "The U.S. Labor Market during the Beginning of the Pandemic Recession." *Working Paper* (May).

Carnevale, Anthony P, Tamara Jayasundera, and Dmitri Repnikov. 2014. "Understanding online job ads data." *Georgetown University, Center on Education and the Workforce, Technical Report (April).*

Chen, Haiqiang, Wenlan Qian, and Qiang Wen. 2020. "The Impact of the COVID-19 Pandemic on Consumption: Learning from High Frequency Transaction Data." *Working Paper* (April).

Chiou, Lesley, and Catherine Tucker. 2020. "Social Distancing, Internet Access and Inequality." *NBER Working Paper No. 26982* (April).

Deming, David, and Lisa B. Kahn. 2018. "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals." *Journal of Labor Economics* 36 (S1): S337–S369.

Goldfarb, Avi, and Catherine Tucker. 2020. "Which Retail Outlets Generate the Most Physical Interactions?" *NBER Working Paper No. 27042* (April).

Gupta, Sumedha, Thuy D Nguyen, Felipe Lozano Rojas, Shyam Raman, Byungkyu Lee, Ana Bento, Kosali I Simon, and Coady Wing. 2020. "Tracking Public and Private Responses to the COVID-19 Epidemic: Evidence from State and Local Government Actions." *NBER Working Paper No. 27027* (April).

Hershbein, Brad, and Lisa B. Kahn. 2018. "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings." *American Economic Review* 108, no. 7 (July): 1737–72.

Kahn, Lisa B, Fabian Lange, and David G Wiczer. 2020. "Labor Demand in the Time of COVID-19: Evidence from Vacancy Postings and UI Claims." *NBER Working Paper No. 27061* (April).

Kurmann, André, Etienne Lalé, and Lien Ta. 2020. "The Impact of COVID-19 on U.S. Employment and Hours: Real-Time Estimates with Homebase Data" (May).

Mongey, Simon, Laura Pilossoph, and Alex Weinberg. 2020. "Which Workers Bear the Burden of Social Distancing Policies?" *NBER Working Paper No. 27085* (May).

Villas-Boas, Sofia B, James Sears, Miguel Villas-Boas, and Vasco Villas-Boas. 2020. "Are We #StayingHome to Flatten the Curve?" *UC Berkeley: Department of Agricultural and Resource Economics CUDARE Working Papers* (April).

Supplementary Appendix

In this appendix, we describe additional details about information reported in the tracker: key dates in the COVID-19 crisis and geographic definitions.

Key Dates for COVID-19 Crisis. The Economic Tracker also includes information about key dates relevant for understanding the impacts of the COVID-19 crisis. At the national level, we highlight three key dates:

- First U.S. COVID-19 Case: 1/20/2020
- National Emergency Declared: 3/13/2020
- CARES Act Signed in to Law: 3/27/2020

At the state level we highlight dates when:

- Schools closed statewide: Sourced from COVID-19 Impact: School Status Updates by MCH Strategic Data, available here. Compiled from public federal, state and local school information and media updates.
- Nonessential businesses closed: Sourced from the Institute for Health Metrics and Evaluation state-level data (available here), who define a non-essential business closure order as: "Only locally defined 'essential services' are in operation. Typically, this results in closure of public spaces such as stadiums, cinemas, shopping malls, museums, and playgrounds. It also includes restrictions on bars and restaurants (they may provide take-away and delivery services only), closure of general retail stores, and services (like nail salons, hair salons, and barber shops) where appropriate social distancing measures are not practical. There is an enforceable consequence for non-compliance such as fines or prosecution."
- Stay-at-home order goes into effect: Sourced from the New York Times stay at home order data, available here.
- Stay-at-home order ends: Sourced from the New York Times reopening data, available here. Defined as the date at which the state government lifted or eased the executive action telling residents to stay home.
- Partial business reopening: Sourced from the New York Times reopening data, available here. Defined as the date at which the state government allowed the first set of businesses to reopen.

These dates are updated as of 5/4/2020.

Geographic Definitions. For many of the series we convert from counties to metros and zip codes to counties. Unless mentioned as otherwise the crosswalks are as follows:

ZIP Codes to County We use the HUD-USPS ZIP Code Crosswalk Files to convert from zip code to county. When a zip code corresponds to multiple counties, we assign the entity to the county with the highest business ratio, as defined by HUD-USPS ZIP Crosswalk.

County to Metro Areas We generate metro values for a selection of large cities using a custom metro-county crosswalk, available in Appendix Table 3. We assigned metros to counties and ensured that a significant portion of the county population was in the metro of interest. Some large metros share a county, in this case the smaller metro was subsumed into the larger metro.

Appendix Table 1
Hourly Wage Rates By Industry

		2019 BLS Wages			Median in Private Datasets	
		10th Percentile (Pre Tax)	25th Percentile (Pre Tax)	Median (Pre Tax)	Earnin (Post Tax)	Homebase (Pre Tax)
NAICS Code	NAICS Description	(1)	(2)	(3)	(4)	(5)
22	Utilities	18.56	26.82	38.06	15.00	
55	Management of Companies and Enterprises	16.09	22.42	34.74	12.34	
54	Professional, Scientific, and Technical Services	14.85	21.62	34.00	12.63	13.00
51	Information	12.90	19.56	32.13	12.49	
52	Finance and Insurance	14.25	18.40	27.42	12.77	
21	Mining, Quarrying, and Oil and Gas Extraction	15.36	19.11	25.82	15.69	
61	Educational Services	11.54	16.18	24.47	13.25	11.50
23	Construction	13.78	17.51	23.92	13.94	
42	Wholesale Trade	12.30	15.73	22.05	11.79	
48-49	Transportation and Warehousing	12.07	15.49	20.89	13.20	15.00
31-33	Manufacturing	12.36	15.35	20.77	12.66	
53	Real Estate and Rental and Leasing	11.31	14.14	19.31	12.64	
62	Health Care and Social Assistance	11.18	13.59	19.27	11.68	14.00
81	Other Services (except Public Administration)	9.73	12.02	16.57	10.97	14.00
56	Administrative Support	10.33	12.26	15.71	11.82	
71	Arts, Entertainment, and Recreation	9.21	11.17	14.09	10.38	12.00
11	Agriculture, Forestry, Fishing and Hunting	11.28	11.89	13.38	11.56	
44-45	Retail Trade	9.49	11.18	13.36	9.76	12.00
72	Accommodation and Food Services	8.68	9.61	11.81	9.26	11.00

Notes: This table reports wages at various percentiles for two-digit NAICS sectors. 2019 BLS Wages (1-3) come from the May 2019 Occupational Employment Statistics and are inflated to 2020 dollars using the Consumer Price Index. Columns (4) and (5) report median wages in two private employment datasets, Earnin and Homebase. In Earnin and Homebase, the median wage is the 50th percentile of hourly wages for workers of the given industry during the pre-COVID period (January 8th - March 10th). In Earnin (4), wages are calculated by dividing the payment deposited in the individual's bank account by hours worked and are thus post-tax. Homebase wages are pre-tax. Industries missing from the Homebase data are left blank.

Appendix Table 2
Demographic Characteristics of Zearn Users

	Zearn Users (1)	US Population (2)
Panel A: Income		
ZIP Median Household Income		
25th Percentile	43,766	45,655
Median	54,516	57,869
75th Percentile	70,198	77,014
Number of ZIP codes	5,148	33,253
Number of People	803,794	322,586,624

	Zearn Users	US K-12 Students
Panel B: School Demographics		
Share of Black Students		
25th Percentile	1.4%	1.5%
Median	5.6%	5.8%
75th Percentile	21.3%	19.1%
Share of Hispanic Students		
25th Percentile	4.3%	5.6%
Median	10.9%	15.0%
75th Percentile	35.7%	40.6%
Share of Students Receiving FRPL		
25th Percentile	33.8%	28.2%
Median	55.5%	50.1%
75th Percentile	78.5%	74.8%
Number of Schools	8,801	88,459
Number of Students	767,310	49,038,524

Notes: This table reports demographic characteristics for US schools. Household income percentiles are calculated using the 2017 median household income in each school's ZIP code. The share of students who are Black, Hispanic, or receive Free or Reduced Price Lunch (FRPL) in a given school are calculated using school demographic data from the Common Core data set from MDR Education, a private education data firm. Percentile distributions for each demographic variable are calculated separately and weighted by the number of students in each school. Column (1) reports school characteristics for students using Zearn, while Column (2) reports income data for the entire US population and shares of students who are Black, Hispanic, or receive FRPL for all US elementary school students.

Appendix Table 3
City to County Crosswalk

City Name	State Name	County	County Fips Code
Los Angeles	California	Los Angeles	6037
New York City	New York	Richmond	36085
New York City	New York	Kings	36047
New York City	New York	Queens	36081
New York City	New York	New York	36061
New York City	New York	Bronx	36005
Chicago	Illinois	Cook	17031
Houston	Texas	Harris	48201
Phoenix	Arizona	Maricopa	4013
San Diego	California	San Diego	6073
Dallas	Texas	Dallas	48113
Las Vegas	Nevada	Clark	32003
Seattle	Washington	King	53033
Fort Worth	Texas	Tarrant	48439
San Antonio	Texas	Bexar	48029
San Jose	California	Santa Clara	6085
Detroit	Michigan	Wayne	26163
Philadelphia	Pennsylvania	Philadelphia	42101
Columbus	Ohio	Franklin	39049
Austin	Texas	Travis	48453
Charlotte	North Carolina	Mecklenburg	37119
Indianapolis	Indiana	Marion	18097
Jacksonville	Florida	Duval	12031
Memphis	Tennessee	Shelby	47157
San Francisco	California	San Francisco	6075
El Paso	Texas	El Paso	48141
Baltimore	Maryland	Baltimore	24005
Portland	Oregon	Multnomah	41051
Boston	Massachusetts	Suffolk	25025
Oklahoma City	Oklahoma	Oklahoma	40109
Louisville	Kentucky	Jefferson	21111
Denver	Colorado	Denver	8031
Washington	District of Columbia	District Of Columbia	11001
Nashville	Tennessee	Davidson	47037
Milwaukee	Wisconsin	Milwaukee	55079
Albuquerque	New Mexico	Bernalillo	35001
Tucson	Arizona	Pima	4019
Fresno	California	Fresno	6019
Sacramento	California	Sacramento	6067
Atlanta	Georgia	Fulton	13121
Kansas City	Missouri	Jackson	29095
Miami	Florida	Dade	12086
Raleigh	North Carolina	Wake	37183
Omaha	Nebraska	Douglas	31055
Oakland	California	Alameda	6001
Minneapolis	Minnesota	Hennepin	27053
Tampa	Florida	Hillsborough	12057
New Orleans	Louisiana	Orleans	22071
Wichita	Kansas	Sedgwick	20173
Cleveland	Ohio	Cuyahoga	39035
Bakersfield	California	Kern	6029
Honolulu	Hawaii	Honolulu	15003
Boise	Idaho	Ada	16001
Salt Lake City	Utah	Salt Lake	49035
Virginia Beach	Virginia	Virginia Beach City	51810
Colorado Springs	Colorado	El Paso	8041
Tulsa	Oklahoma	Tulsa	40143

Notes: This table shows our metro area (city) to county crosswalk. We assigned metros to counties and ensured that a significant portion of the county population was in the metro of interest. Some large metros share a county, in this case the smaller metro was subsumed into the larger metro.

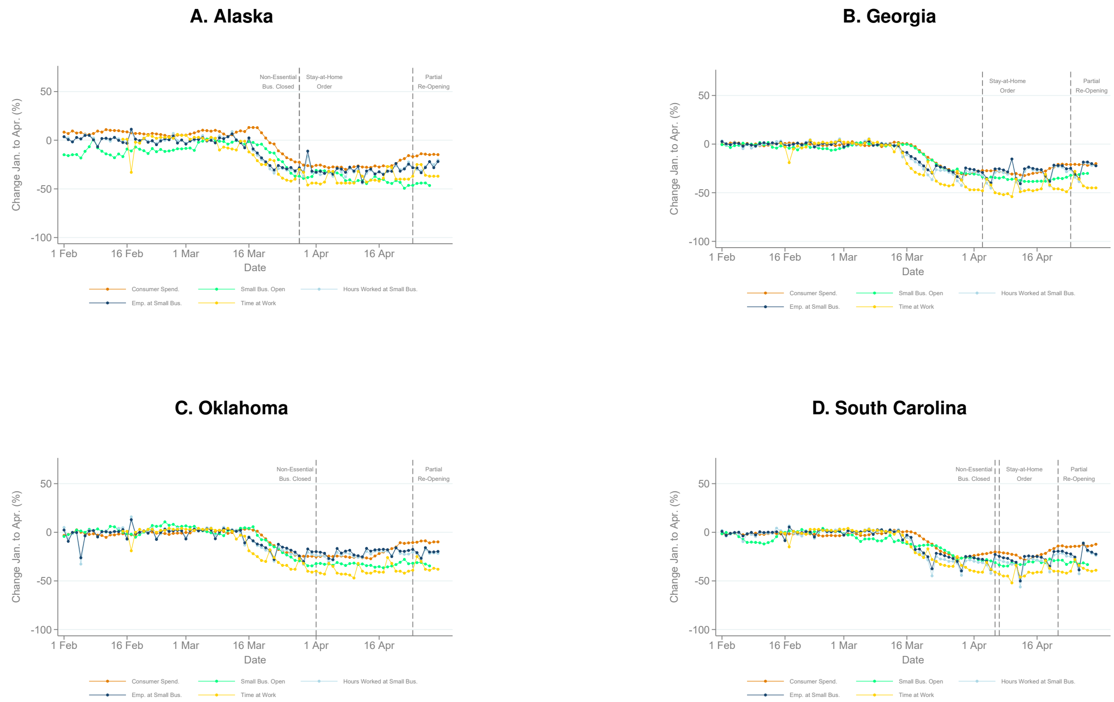
FIGURE 1: Change in Consumer Spending, Small Business and Hours Worked Measures Around Stay-At-Home Order

A. States Issuing Stay-at-Home Order and Business Closure Order on Same Day in Week 19-26 March

B. States Issuing Stay-at-Home Order or Business Closure Order After 26 March

Notes: This figure studies the effect of Stay-at-Home and Business Closure orders on Consumer Spending, Small Bussiness Open, Hours Worked at Small Businesses, Employment at Small Businesses and Time at Work relative to January 2020. Panel A plots the average series for states that issued the Stay-at-Home orders in the week between March 19th and 26th, while Panel B does the same for states that issued the order after March 26th. The points are residualised by day of week fixed effects. Data sources: Affinity Solutions, Google Mobility, HomeBase, Womply

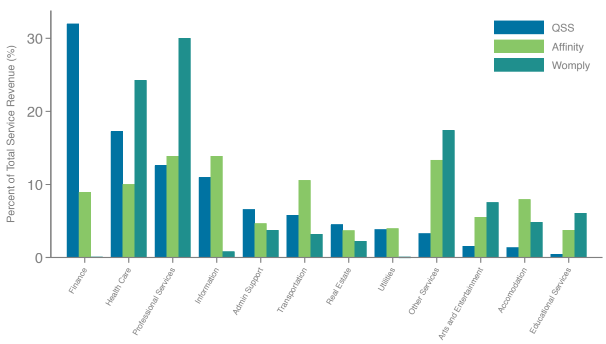
FIGURE 2: Change in Consumer Spending, Small Business and Hours Worked Measures Around Partial Re-Opening



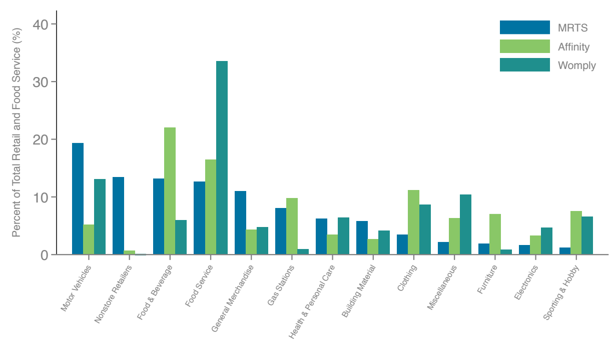
Notes: This figure studies the effect of partial re-openings of the economy in Consumer Spending, Small Bussiness Open, Hours Worked at Small Businesses, Employment at Small Businesses and Time at Work relative to January 2020. Panel A plots the series for Alaska, Panel B, C and D do the same for Georgia, Oklahoma and South Carolina respectively. See notes to Figure 1 for details. Data sources: Affinity Solutions, Google Mobility, HomeBase, Womply

APPENDIX FIGURE 1: Industry Shares of Consumer Spending and Business Revenues Across Datasets

A. Compared to QSS

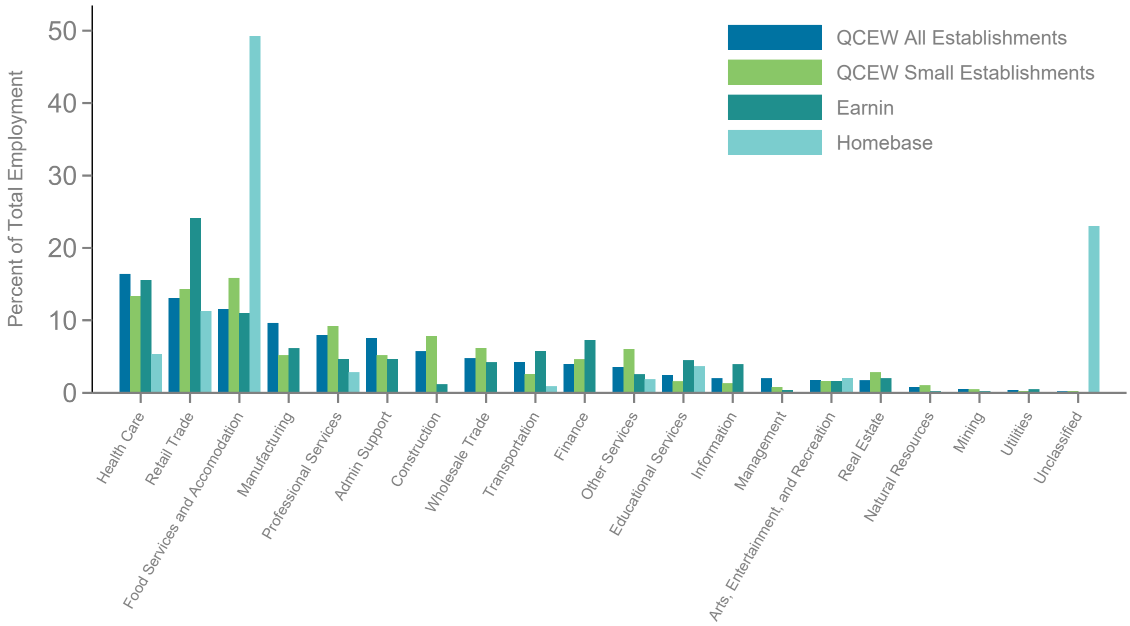


B. Compared to MRTS



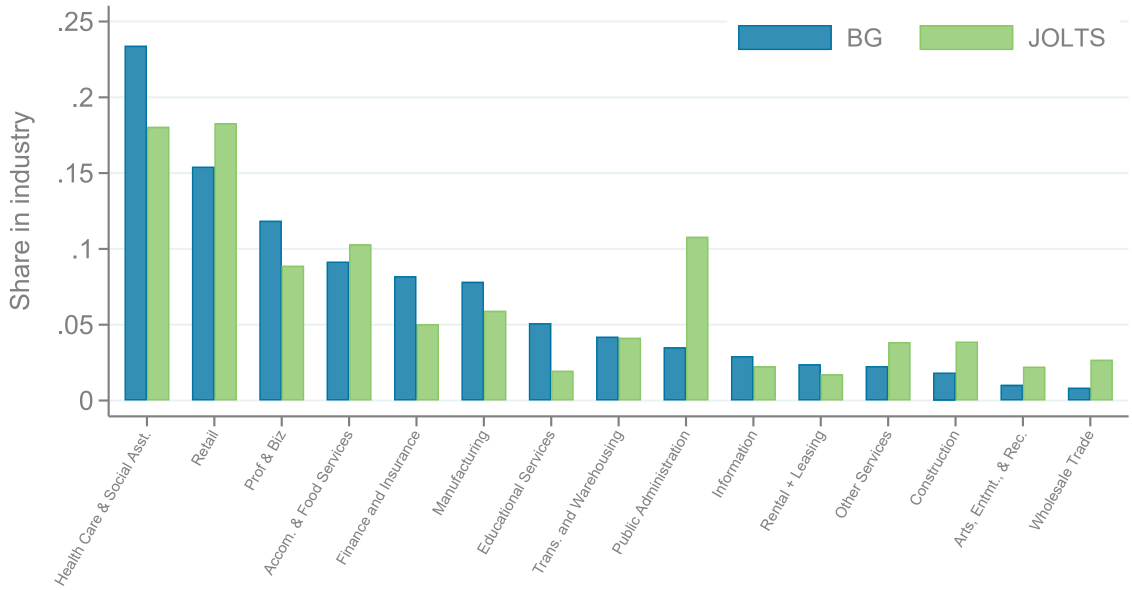
Notes: Panel A shows the NAICS two-digit industry mix for two private business credit card transaction datasets compared with the Quarterly Services Survey (QSS), a survey dataset providing timely estimates of revenue and expenses for selected service industries. Subsetting to the industries in the QSS, each bar represents the share of revenue in the specified sector during Q1 2020. We construct spending and revenue shares for the private datasets, Affinity and Womply, by aggregating firm revenue (from card transactions) in January through March of 2020. Panel B shows the NAICS three-digit industry mix for the same two private datasets compared with the Monthly Retail Trade Survey (MRTS), another survey dataset which provides current estimates of sales at retail and food services stores across the United States. Subsetting to the industries in the MRTS, each bar represents the share of revenue in the specified sector during January 2020. We construct revenue shares for the private datasets, Affinity and Womply, by aggregating firm revenue (from card transactions) in January 2020. Data sources: Affinity Solutions, Womply

APPENDIX FIGURE 2: Industry Shares of Employment Across Datasets



Notes: This figure shows the NAICS two-digit industry mix for two private employment-based datasets compared with the Quarterly Census of Employment and Wages (QCEW), an administrative dataset covering the near-universe of firms in the United States. Each bar represents the share of employees in the given dataset who work in the specified sector. We construct data for all establishments and small establishments using employment data from the Q1 2019 QCEW. Small establishments are defined as having fewer than 50 employees. We construct employment shares for the private datasets, Earnin and Homebase, using January 2020 employment. We define employment in Earnin as the total number of worker-days in the month. We define employment in Homebase as the number of unique individuals working a positive number of hours in the month. Data sources: Earnin, HomeBase

APPENDIX FIGURE 3: Industry Shares of Job Postings in Burning Glass and Job Openings in JOLTS



Notes: This Figure displays the NAICS two-digit industry mix of job postings in Burning Glass and job openings in JOLTS, the Job Openings and Labor Turnover Survey data provided by the U.S. Bureau of Labor Statistics, in January 2020. Data source: Burning Glass