Opportunity Insights Economic Tracker Data Revisions

last updated on 2023-09-01

1 Overview

This document provides a description of major revisions to the data posted by the Opportunity Insights Economic Tracker. The document is organized sequentially by series in the tracker, among series that have had substantive data revisions since June 30, 2021 due to changes in data processing or data sources over time.

This document is updated regularly and the following information is subject to change.

For further information or if you have any questions please feel free to reach out to info@opportunityinsights.org and someone on our team will be in touch.

2 Data Series

2.1 Consumer Spending

Revisions on March 15th 2022

The consumer spending data was revised with three changes:

1. We have improved our modeling of spending levels over time to better control for growth (or shrinkage) in payment providers' customer bases. Previously we controlled for the sharp discontinuities generated by entries and exits of payment providers, but did not control for steady growth or shrinkage in individual payment providers' customer bases. We have detected an increase in the customer base of existing payment providers during the year 2021, which was biasing our estimates of spending upward. The new processing, described in detail below, causes a downward revision in our estimates of the change in consumer spending—primarily in the second half of 2021.

   Our objective is to simultaneously (1) isolate and remove variation in spending driven by the number of debit/credit cards in our sample changing due to growth or shrinkage in a payment provider's customer base, and (2) incorporate variation in spending driven by the number of debit/credit cards in our sample changing due to changes in the utilization of cards, which can reflect underlying economic conditions. To account for both, we first estimate a two-state switching model of the relationship between Personal Consumption Expenditure as measured by the U.S. Bureau of Economic Analysis and the level of credit and debit spending and number of cards in usage we received from Affinity Solutions. Our estimates imply a state change between February 2020 and August 2020: the number of cards in use is predictive of Personal Consumption Expenditures during this period, but statistically insignificant outside this period. This is consistent with a pattern where

people exhibited extensive margin changes in card spending during the initial waves of COVID-19, but outside of that period changes in the number of cards being used reflect changes in the number of people in-sample as opposed to changes in the propensity to consume of a given set of people.

We therefore construct an estimate of spending that reflects total spending during February to August 2020 and spending per card outside this period. For each location (l), industry (i) and time (t) we compute:

$$\text{adjusted_spending}_{l,i,t} = \begin{cases} \frac{spending_{l,i,t}}{cards_{l,\cdot,t}} \times \overline{cards_{l,\cdot,t' \in \text{Jan 2020}}} & \text{if } t \leq \text{January 2020} \\ \frac{spending_{l,i,t}}{cards_{l,\cdot,t}} \times \overline{cards_{l,\cdot,t' \in \text{Jan 2020}}} \times \frac{\overline{cards_{l,\cdot,t' \geq t | t' \in \text{Feb 2020}}}}{\overline{cards_{l,\cdot,t' \in \text{Feb 2020}}}} & \text{if } t \in \text{February 2020} \\ \frac{spending_{l,i,t}}{cards_{l,\cdot,t}} \times \overline{cards_{l,\cdot,t' \in \text{Jan 2020}}} \times \frac{cards_{l,\cdot,t}}{\overline{cards_{l,\cdot,t' \in \text{Feb 2020}}}} & \text{if } t \in [\text{March 2020, July 2020}] \\ \frac{spending_{l,i,t}}{cards_{l,\cdot,t}} \times \overline{cards_{l,\cdot,t' \in \text{Jan 2020}}} \times \frac{\overline{cards_{l,\cdot,t' \leq t | t' \in \text{Aug 2020}}}}{\overline{cards_{l,\cdot,t' \in \text{Feb 2020}}}} & \text{if } t \in \text{August 2020} \\ \frac{spending_{l,i,t}}{cards_{l,\cdot,t}} \times \overline{cards_{l,\cdot,t' \in \text{Jan 2020}}} \times \frac{\overline{cards_{l,t' \in \text{Aug 2020}}}}{\overline{cards_{l,\cdot,t' \in \text{Feb 2020}}}} & \text{if } t \geq \text{September 2020} \end{cases}$$

2. We have improved the internal mapping between Merchant Category Codes (MCCs) to corresponding industry sectors to be more accurate. This causes a noticeable revision in the Transportation series.

3. We have improved the amount of state-level data available by changing the aggregation of masked cells. Previously data in masked county-level cells were included in the national-level data but excluded from the state-level data. Now data in masked cells is included in the state-level data, which increases the set of state-level series available.

Revisions on June 15 2022

We have revised the method implemented on March 15, 2022 to control for steady growth or shrinkage in individual payment providers' customer bases.

- Our seasonally adjusted estimates of spending for 2020 and beyond are obtained by dividing by 2019 estimates of spending. The method implemented on March 15, 2022 measures spending using spending per card in all periods except February to August, 2020: during that period, we use total spending for the reasons discussed in the section above. This generated an upward bias in the seasonally adjusted data for summer 2020, since the 2019 reference period reflected only spending per card while the February to August 2020 period incorporated extensive margin increases in the number of cards in the sample.

- We have therefore constructed an estimate of spending that reflects total spending during February to August, 2019 and spending per card outside this period in 2019, mimicking the transition between total spending and spending per card previously implemented for 2020. To create seasonally adjusted estimates of spending in 2020, we use this newly constructed analogous data for 2019. To create seasonally adjusted estimates for 2021 and later, we use spending per card in 2019.

- This adjustment reduced the estimated spending level in summer 2020. The new estimates align better with public benchmarks (NIPA, MARTS).

Revisions on August 12th 2022

As of June 5, 2022 we only receive consumer spending data at a weekly level of aggregation. We will continue to publish weekly data on consumer spending, but we will not publish daily variation past June 5, 2022.

Revisions on September 9th 2022

From August 12 to September 9, 2022, an error in the processing code resulted in small but significant biases due to incorrect chain-weighting around the sharp discontinuities generated by entries and exits of payment providers. For more information about the procedure for correcting those discontinuities, see the "Revisions on March 15th 2022" above.

During this period, the number of cards was correctly chain-weighted at break dates, but the spending per card was not being chain-weighted. This has been corrected in the latest data release so that both are chain weighted at break dates. Before and after this revision, the trends are similar. However, there are noticeable level changes in a handful of states: FL, ID, KY, ME, MO, OK and SD.

Revisions on December 2nd 2022

We corrected a bug in the implementation of our chain-weighting methodology. This methodology adjusts counties with sharp discontinuities generated by entries and exits of payment providers, as explained in Appendix B.2 of the companion paper:

> we correct our estimates as follows. We first compute the state-level week-to-week percent change in spending excluding all counties with a structural break (using the national series for DC and states for which all counties have a structural break). If we identify a structural break in week t, we impute spending levels in weeks $t-1$, t, and $t+1$, as we cannot ascertain the precise date when the structural break occurred (e.g., it may have occurred on the 2nd day of week $t-1$ or the 6th day of week t). When there is a change in coverage we adjust the series to be in line with the lower level of coverage.

From March 15 to December 2, 2022, the series was always chain-weighted forwards, so that it was always in line with the level of coverage prior to the discontinuity — rather than the lower level of coverage. This has been corrected so that the series is now chain-weighted forwards or backwards to be in line with the lower level of coverage. This has resulted in a noticeable revision to a few counties, which now have lower levels after a discontinuous break. At the state level, this adjustment has resulted in only minor revisions.

Revisions on June 14th 2023

We added four supplemental files:

- *Affinity Income Shares - National - 2019.csv*: share of total spending in January 2019 by income quartile.
- *Affinity Income Shares - National - 2020.csv*: share of total spending in January 2020 by income quartile.

- *Affinity Industry Composition - National - 2020.csv*: share of total spending in January 2020, and the share of the decline in total spending during the first wave of the COVID-19 pandemic, by industry.
- *Affinity Daily Total Spending - National - Daily.csv*: daily total spending indexed to January 2019 by income quartile, without smoothing using a 7-day moving average.

For more details, please see the data documentation.

Revisions on July 21st 2023

We refined the date handling in our holiday adjustment methodology at the weekly level and corrected a processing error at the daily level. As described in Appendix B.2 of the companion paper, we intend to seasonally adjust the data by calculating, for each week and day, the year-on-year change relative to the 2019 value. To account for holidays, the year-on-year change for dates with holidays after 2019 is calculated relative to the day or week of the same holiday in 2019. In the latest release we have made two changes. First, we have updated how we define and align weeks across years during our adjustments of the weekly level data, and as a result, there are small revisions around some holidays. Second, due to an error in our processing scripts our holiday adjustments were being incorrectly applied to the daily level data. We have now fixed this error and the adjustments are being applied appropriately resulting in smoothing of the series around holidays.

Revisions on August 15th 2023

As of August 15, 2023 we receive a larger sample of consumer spending data from 2022 to present, resulting in small revisions to the series.

2.2 Small Business Revenue & Small Businesses Open

Revisions on March 4th 2022

We now derive our published datasets using a county-level panel of small businesses. In each calendar year, we follow the sample of businesses operating during the first week of the year (i.e. we start following a new panel each calendar year). No new businesses enter our panel during the calendar year. Businesses may exit because they stop operating or because the underlying payment processors ceased providing data.

We detect cases where a payment processor disappears by detecting sharp drops in businesses operating, at the national and the state level. We then adjust the series to remove these drops at the State x Industry level, using the following assumptions:

- In the initial Covid period (March to July 2020) we assume momentum, and impute the value of merchants/sales for the week to continue the rate of change we observe in the 4 weeks prior.
- In the rest of the series (before March 2020 and after July 2020) we assume the series stays constant for the adjusted week.

After performing these adjustments, we aggregate up to the State level, the National x Industry level, and the National level.

Since the no-entry panel does not cover the ZIP code level, we impute the ZIP code level data to reconcile it with the rest of the geographic levels. We perform an additive adjustment on the ZIP level series so that the weighted sum of the ZIP series aligns with the county level same store

series. In doing this, we get the levels from the County no-entry panel, and the within-county variation from the ZIP level data.

Revisions on May 17th 2022

We are now releasing county-level and city-level data derived from the same panel of small businesses described in the March 4, 2022 data revision.

Revisions on June 26th 2023

We added a supplemental file:

- *Womply - ZCTA - 2020.csv*: Small business revenue levels in April and July 2020 at the ZIP-code level.

For more details, please see the data documentation.

2.3 Job Postings

Revisions on December 15th 2021

At the request of Lightcast (formerly known as Burning Glass Technologies) we've suppressed the historical data we publish for the following Job Postings series:

- The state aggregate series, state by education level series, and county aggregate series are now only published for the present and previous 12 months
- The state by supersector series, county by education level series, and county by supersector series are only published for the present and previous 6 months

Revisions on August 30th 2022

Lightcast detected and corrected for a data anomaly related to a major national job board. This correction retroactively revises the data from May 2022 onward: a downward revision in May, an upward revision in June, and a downward revision in July. Lightcast issued the following statement outlining the data anomaly observed and corrective actions taken:

> As you may know, Lightcast has over a billion job postings, career profiles, and other data points. As the labor market shifts through changes in the economy, our team actively reviews and manages our data set, working daily to refine our taxonomy and review data sources.
>
> That's why we set the global standard on commitment to data excellence - ensuring our data is consistent with other sources and benchmarked to the BLS Job Openings and Labor Turnover Survey (JOLTS) report.
>
> One example of why this is important is a recent anomaly that our data team monitored in job posting data. The team found an uptick in posting numbers for May from one of the major national job boards. This trend then dropped in June and increased again in July.
>
> The Lightcast team determined that this anomaly was the result of a change in job posting methodology at this particular source, and to compensate, the Lightcast data team changed how that data is reflected in our overall job posting numbers. In this case, users will see adjustments in the reported numbers and trendlines for May, June, and July, and in limited cases, may want to re-run analyses for these months.

Revisions on December 6th 2022

We have changed the file names of the data files corresponding to our job postings series: they are now prefixed `Job Postings - *` instead of `Burning Glass - *`. The data provider, the data received, and the data processing steps we perform remain otherwise unchanged.

We have renamed the files in light of the recent name change by our data provider, which now operates as Lightcast and was formerly known as Burning Glass Technologies.

Revisions on June 26th 2023

We added a supplemental file:

- *Job Postings Industry Shares - National - 2020.csv*: share of job postings by industry in January 2020.

For more details, please see the data documentation.

Revisions on September 1st 2023

We have updated the Lightcast datastream and API used to pull the job postings data in response to the deprecation of the existing datastream and its corresponding API we have used up to this point. As part of the change to the new datastream, the existing scraped raw job postings data has undergone revamped and improved classification and coding by Lightcast resulting in both some small retroactive changes in the aggregate counts of postings and noticeable retroactive changes in the counts of postings by super sector and required education level as posts in the raw data were reclassified. As such, small retroactive changes in the total postings series have been made as well as moderate retroactive changes have been made to the industry and education level subgroup series.

2.4 Employment

Revisions on June 30th 2021

The Employment data was revised on June 30, 2021 due to three independent changes in methodology.

- **Revisions to address end-of-year "churn" in Paychex client base:** over time, some firms enter and exit Paychex's client base. This is especially concentrated at the end of each calendar year, where there is significant churn as firms renew their payroll processing contracts. This creates two sources of error in the tracker series. First, due to this seasonal pattern, the raw Paychex data displays a downwards trend in employment at the end of each calendar year as some clients leave Paychex, followed by an upward trend in employment at the very beginning of each calendar year as new clients join. Second, we take steps to avoid firm entry and exit that are more responsive to firm entry and exit at finer levels of geography, so that we can minimize the number of discontinuous changes in employment. The firm entry and exit occurring due to end-of-year "churn" in Paychex's client base resulted in a discrepancy between the national employment series and the corresponding series at the state level.

  To avoid these sources of error, we have changed the way in which we process employment data around the end of each calendar year. We now adjust for the end-of-year pattern in the Paychex data using data from the end of 2019. For each date between December 10, 2020 and January 10, 2021, using Paychex data on employment at the national level, we

compute the change in employment relative to December 10, 2020 at the two-digit NAICS code x income quartile level. We also compute the change in employment between the corresponding day in the previous year and December 10, 2019. We divide the change in employment relative to December 10, 2020 by the corresponding change in employment the previous year relative to December 10, 2019. At the national level, as of June 2021, this change results in an upwards revision of the employment series by between 1 and 3 percentage points from December 2020 through May 2021. We then apply the same adjustment to each two-digit NAICS code x income quartile cell at the state, county and city levels. For this reason, the state, county and city-level series have also been adjusted upwards by between 1 and 3 percentage points from December 2020 through May 2021.

- **Revisions arising from changes to adjustment for firm entry/exit in Paychex data**: over time, the Paychex sample changes as clients begin to use or stop using Paychex's payroll processing services. We previously adjusted for firm entry and exit separately at the national and state levels; for details on this adjustment, see Appendix D of Chetty, Friedman, Hendren and Stepner (November 2020). This introduced the possibility of discrepancies between the national-level and the (employment-weighted) mean of the state-level employment series. Empirically, these discrepancies were small throughout most of 2020, but began to grow in December 2020 due to increased churn in Paychex's client base at the end of 2020, as described above. Since the firm entry/exit adjustment was applied retrospectively, this led to discrepancies throughout the series.

  We have changed our approach to adjusting for firm entry/exit to avoid these discrepancies. In each county x industry (two-digit NAICS code) x firm size x income quartile cell, we now compute the change in employment relative to January 4-31 2020, and the change in employment relative to July 1-31 2020. For county x industry x firm size x income quartile cells with over 50 employees at any point between January 2020 and the end of the series, we reduce the weight we place on the series if we observe changes in employment that indicate firm entry or exit. In particular, we compute the weight on the cell for county c, industry i, firm size s, and income quartile q as:

$$\begin{aligned}\text{Weight}_{c,i,s,q} = \max\Big\{ & 1 - \mathbf{1}\{\text{Min Normed July}_{c,i,s,q} \leq 50\} \times (50 - \text{Min Normed July}_{c,i,s,q}) \times 0.02 \\ & - \mathbf{1}\{\text{Max Normed January}_{c,i,s,q} \geq 150\} \times (\text{Max Normed January}_{c,i,s,q} - 150) \times 0.02, \\ & 0\Big\}\end{aligned}$$

  where Min Normed July$_{c,\ i,\ s,\ q}$ is the smallest value of indexed employment we observe at each date relative to its mean level over the period July 1 to 31, 2020, and Max Normed January$_{c,\ i,\ s,\ q}$ is the largest value of indexed employment we observe at each date relative to its mean level over the period January 4 to 31, 2020.

  That is, we reduce the weight we place on the cell by two percentage points for each percentage point of growth we observe above 150 percentage points relative to January 2020. We then further reduce the weight we place on each cell by two percentage points of its January 2020 level for each percentage point of decline we observe below 50 percentage points relative to July 2020.

  At the national level, this change revises the aggregate employment series between -2 (in October 2020) and +2 (in January 2021) percentage points, as of June 2021. This change also substantially revises the state-level employment series. The mean revision (in absolute value) for aggregate employment at the state level is 5.6 percentage points as of June 2021.

This revision is largest in March 2021, when the mean revision (in absolute value) is 9.1 percentage points.

- **Revisions to address changes in minimum wages**: we use hourly wage thresholds when constructing employment by income quartile in the Paychex data. The threshold for the bottom quartile of employment is $13; that is, workers who earn below $13 are assigned to the bottom quartile, whereas workers earning above (or exactly) $13 are allocated to other wage quartiles. On January 1, 2021, minimum wage changes came into force in CA, MA, AZ and NY, which caused the minimum wage for some workers to move from below $13 to above $13. This resulted in a decline in employment for workers earning below $13, and a corresponding increase in employment for workers earning above $13, as firms increased workers' wages in response to the minimum wage change. This was reflected in a decrease in low-income employment and an increase in middle-income employment in the tracker data for states in which the minimum wage had increased above $13.

  Though the tracker data for these states accurately represented trends in employment among workers earning below $13, the movement of workers around the minimum wage threshold created difficulties when comparing trends in low-wage employment across states. We have taken four steps to address this problem. First, we have created additional series for below-median-income and above-median-income employment, which can be downloaded from this repository. As the national median wage threshold in the Paychex data is $18.18, the below-median-income series is not affected by the shifting of workers induced by the minimum wage change. Second, we have added annotations to the tracker data indicating minimum wage changes in the relevant states. Third, we have suppressed series cut by income quartile in these states after December 1, 2020, to avoid displaying series that are substantially affected by minimum wage changes. Finally, the impacts of the minimum wage on low-income employment in these states also affected trends at the national level: as workers' wages increased above the $13 threshold, national employment fell. When computing the national-level trend in employment, we now exclude trends in CA, MA, AZ and NY between December 10, 2020 and February 10, 2021. We continue to use trends in these states when computing national-level employment from February 11, 2021 forward.

After making these changes, the (population-weighted) RMSE (root mean square error) of the state-level employment series relative to the CPS is 4.55 percentage points as of April 2021, after removing public sector and furloughed workers and expressing employment in seasonally-unadjusted terms relative to January 2020. Though we will continue to assess our series relative to the CPS, users should note that some amount of noise remains in both our state-level series estimates and in the CPS estimates. Particularly in instances where these two measures of employment differ, users may consider both the Opportunity Insights series and the CPS as helpful inputs in identifying local patterns.

Revisions on May 21st 2022

The Employment data was revised on May 21, 2022 due to four independent changes in methodology.

- **Revisions to combination process:** We are currently constructing the combined Employment series using Paychex and Intuit data.
- **Revisions to income quartile definitions:** The original income quartile thresholds were set using annualized wages of up to $13/hr (Q1), $18/hr (Q2), and $29/hr (Q3). These thresholds were derived from the wage distribution of employees in Paychex's client base in 2019. However, wage growth and inflation trends over time will mechanically move workers across these fixed thresholds into higher income quartiles. For example, if a Q1

worker's wages were to grow such that they were now over the Q2 threshold, then the series would mechanically show an decrease in Q1 employment and an increase in Q2 employment, despite aggregate employment remaining the same. Eventually, with fixed thresholds, every worker would be placed in the Q4 category.

To address this issue, we have implemented moving thresholds that are based on the federal poverty line. The income categories are now defined using multiples of the poverty threshold, normally a yearly measure that we extend monthly using the CPI. The upper thresholds for each income quartile are now 1x the poverty line (Q1), 1.5x the poverty line (Q2), and 2.5x the poverty line (Q3). A secondary issue that arises from using moving thresholds, however, are discontinuities in the series when a threshold crosses a whole number wages where there is bunching in the wage distribution, e.g. $15/hr. When the Q1 threshold crosses $15, for example, many workers who were previously a part of Q2 will now be defined as Q1, causing a discontinuity in the series. To address this issue, we "spread" workers out from the whole number wages as if we added a random number between -0.5 and 0.5 to their wages, creating a uniform distribution between (whole number wage - 0.5, whole number wage + 0.5).

Implementing these moving thresholds results in upwards revisions in employment estimates for quartiles that experienced wage growth, such as Q1. These revisions are particularly notable in the second half of 2021.

- **Revisions to address changes in the minimum wage:** We have updated our previous methodology released on June 30, 2021. Instead of supressing states in which the minimum wage is increased past one of our wage thresholds and removing these states from the national trends, we replace the county, industry, firm size, and quartile cells in the affected states with the national trend in the same industry, firm size, and quartile. This does not change the national series, but allows us to present an estimate for employment in states where the minimum wage has increased.
- **Revisions arising from changes to adjustment for firm entry/exit in Paychex data:** We have updated our previous methodology released on June 30, 2021. Previously we applied weights solely to cells with over 50 employees; we now apply weights to all cells. However, we apply a different weighting scheme to cells below 50 employees to account for small firm births.
  - For cells with over 50 employees:
    - We reduce the weight by two percentage points for each percentage point of decline we observe below 50 percentage points relative to July 2020.
    - We reduce the weight by two percentage points for each percentage point of growth we observe above 150 percentage points relative to January 2020.
  - For cells with 50 employees or less:
    - We reduce the weight by two percentage points for each percentage point of decline we observe below 50 percentage points relative to July 2020
    - We reduce the weight by 0.1 percentage points for each percentage point of growth we observe above 4000 percentage points relative to January 2020.

We adopt this alternative weighting scheme for cells with 50 or fewer employees to balance the fact that small firm births account for a sizable portion of the economic recovery in the second half of 2020 against the fact that growth of small cells can also be spuriously driven by entry into our sample. This new weightng scheme leads to better estimates of employment relative to public benchmarks (CES, QCEW, CPS, etc).

Revisions on October 17th 2022

The Employment data was revised in several ways.

- **Revisions to income quartile definitions**: The methodology behind the moving income thresholds, implemented on May 21, 2022, has been slightly updated. Previously, the income thresholds were defined using the yearly poverty guidelines and extended monthly using the CPI. In periods with high inflation, this methodology resulted in a discontuinity in the threshold before and after the release of the annual poverty guidelines (historically in January). This has been updated so that when the annual poverty guidelines are released, the previous year's monthly thresholds are linearly revised such that the CPI adjusted poverty threshold is equal to the official poverty guideline.

  Revising these moving thresholds results in slight differences in employment estimates for each quartile across the entire sample, since revising the thresholds redistributes some workers between quartiles. For some years, the thresholds have been revised slightly downward, while for others they have increased. These changes are particularly notable from the start of 2021 to present, when the revisions have led to lower thresholds.

- **Revisions to the methodology that adjusts for firm entry/exit in Paychex data**: We have updated our previous methodology released on May 21, 2022. We have revised the weighting scheme for cells over 50 employees, in order to better align with public benchmarks:
  - For cells with over 50 employees:
    - We reduce the weight by two percentage points for each percentage point of decline we observe below 50 percentage points relative to July 2020.
    - *We reduce the weight by 0.5 percentage points for each percentage point of growth we observe above 600 percentage points relative to January 2020.*
  - For cells with 50 employees or less:
    - We reduce the weight by two percentage points for each percentage point of decline we observe below 50 percentage points relative to July 2020
    - We reduce the weight by 0.1 percentage points for each percentage point of growth we observe above 4000 percentage points relative to January 2020.

- **Revisions to the adjustment for minimum wage changes**: We revised how we adjust the series in states that increased their minimum wage at the end of 2020 (CA, NY, MA). We begin by calculating the percentage change in the number of employees in each state x industry (2-digit NAICS) cell from December 4 2020 onwards for the first two wage quartiles combined. From December 4, 2020 onwards, for the first two wage quartiles in minimum wage change states, we impute the trend in each county x industry x wage quartile cell using the below-median income employment trend in their own state x industry cell. Since employment in the first wage quartile recovered less than employment in the second wage quartile in 2021, an imputation that aggregates the trends in the first and second wage quartiles tends to overstate the recovery in the first quartile and understate the recovery in the second quartile. As such, we further rescale the state x industry below-median income trend separately for each wage quartile, using the coefficient from the regression of that quartile's employment change on the below-median income employment change in non-minimum wage-change states without a constant between December 4, 2020 and December 3, 2021. We then aggregate the adjusted employment counts to the relevant geographies (e.g. state; national) before calculating the change in employment since January 2020.

Revisions on April 14th 2023

The Employment data was revised on April 14, 2023 to correct a data processing error.

From December 16, 2022 to April 14, 2023, there was a bug in the code adjusting for discontinuities in the Paychex data when quartile thresholds cross integer wages. For more details on this adjustment see Online Appendix E.2 of Chetty, Friedman, Hendren and Stepner (April 2023). During this period, the adjustment for the mass of employees at wage $X was incorrectly being made using the mass of employees at the lower wage $X-1. These discontinuities were therefore incompletely smoothed.

After correcting the code, the employment trends are smoother around these discontinuities. The overall trends in employment changes over time are unchanged.

Revisions on June 26th 2023

We added a supplemental file:

- *Earnin - ZCTA - 2020.csv*: Employment levels in April and July 2020 at the ZIP-code level.

For more details, please see the data documentation.

2.5 Unemployment Claims

Revisions on July 29th 2021

The unemployment data was revised on July 29, 2021 to correct a data processing error. Previously we assigned continued PEUC and PUA claims to the end of week date indicated by the Report Date rather than the Reflect Date in the Department of Labor data. Effectively, this meant continued PEUC and PUA claims were offset by one week into the future in our data. We've corrected this, and claims now align to the appropriate week.

Revisions on July 20th 2023

The unemployment data was revised on July 20, 2023 to change the frequency of the county level Iowa initial claims data from the weekly frequency to the monthly frequency in light of changes in the underlying data's publication at the Iowa Workforce Development - Labor Market Information Division.

2.6 Online Math Participation and Student Progress in Math

Revisions on December 14th 2021

As schools have attritioned from the sample we've revised our methodology to impute missing values for an otherwise active school. Previously school retention in the sample was very high and we imputed across the sample observations missing in a given week for a given school as zeros as the likelihood the omission of data was a sign of no usage but continuity in the sample was quite high. Overtime, and particularly during the 2021 fall semester, a subset of schools naturally attritioned from the sample of schools using Zearn prior to the onset of the COVID-19 pandemic and we paused our imputation procedure for the newest sample. As such we've now revised our methodology to impute missing values within semester based on semester activity. For this subset of schools that are missing observations for a given week within a semester and are active Zearn users at any point within a semester, we impute these values as zeros within

the given semester. As such we avoid imputation of schools observations where a school has dropped from the sample in a given semester and identify weeks missing for schools that remain within the sample for a given semester and impute them as zeros appropriately. This change has resulted in a small downward revision to the estimates post summer 2020.

In addition, previously in the case when correcting for transitory anomalous spikes, week to week changes in the underlying levels of greater than 50 percentage points that immediately revert by an equally sized change in the subsequent week, for a given week for a given school we have changed how both the procedure overall operates and how one sided spikes bounded by a missing value are handled. Overall previously the procedure for spike correction detected and applied the correction sequentially for positive and the negative value spikes. This has been revised to apply these changes simultaneously to a series. Given the size of anomalous change needed for a spike to be identified this has a small impact on the overall series. In the case where a one sided spike was bounded by a missing observation we previously in the 2021 fall semester imputed the transitory anomalous spike as a missing observations as well. As we now impute within semester active schools' missing observations as zeros, we now no longer impute these one sided spikes as missing and instead impute them with the 3 week moving average identically to all other transitory spike corrections in the series. This applies to a small number of cases within certain states and has minimal impact.

2.7 COVID-19 Infections and Vaccinations

Revisions on March 17th 2021

Previously we pulled reported cases and deaths from the New York Times' COVID-19 database and reported tests from the COVID Tracking Project. At the conclusion of the COVID Tracking Project's efforts in order to collect testing data we instead began pulling reported cases, deaths, and tests at the county level from the Centers For Disease Control and Prevention's COVID Data Tracker and aggregated to other geographies.

Revisions on August 4th 2021

Previously we pulled reported cases, deaths, and tests at the county level from the Centers For Disease Control and Prevention's COVID Data Tracker and aggregated to other geographies. On July 17th the Centers For Disease Control and Prevention began suppressing reported cases and deaths making aggregations across counties no longer feasible and we began to instead pull reported cases and deaths from the New York Times' COVID-19 database, state level reported tests from the Johns Hopkins Coronavirus Resource Center's U.S. testing database, and county level reported tests from The Centers for Disease Control and Prevention.

Revisions on April 11th 2023

On March 23, 2023 the New York Times discontinued its COVID tracking efforts and published the last updates to its database tracking COVID cases and deaths in light of increased tracking and reporting by government agencies like The Centers for Disease Control and Prevention. For COVID-19 cases and deaths data published after March 23, 2023, we now make use of weekly data published by The Centers for Disease Control and Prevention. In order to integrate the new weekly sums published by The Centers for Disease Control and Prevention with the previous daily values published by New York Times we scale all cases and deaths values to a 7 day rolling sum rather than a 7 day rolling average as previously reported.

Revisions on June 23rd 2023

1. On May 11, 2023, the Centers for Disease Control and Prevention (CDC) ceased to update their COVID-19 vaccination tracking data series and existing COVID-19 cases and deaths tracking data series. On June 1, 2023, the CDC published the final update to their weekly data series tracking COVID-19 cases and deaths.

   The Economic Tracker data on weekly COVID-19 cases, deaths and vaccinations up to May 10, 2023 reflects these final updates from the CDC, which will not receive further updates.

   We continue to pull and collate the state-level death counts from the data available on the updated COVID Data Tracker and state level counts of hospitalizations available from the Department of Health and Human Services.

2. We no longer compute a rolling 7 day sum of the *cumulative* COVID-19 cases and deaths. Starting on April 11, 2023 we switched from a 7-day rolling average to a 7-day rolling sum in order to ensure the units remain consistent across historical daily and weekly data. But the 7-day rolling sum was erroneously applied to *cumulative* COVID-19 cases and deaths, which multiplies the cumulative counts by 7. We have corrected the error and restored a 7-day rolling average for *cumulative* cases and deaths.

2.8 Time Outside Home

Revisions on October 20th 2022

On October 15, 2022, Google discontinued its updates to the COVID-19 Community Mobility Reports, posting the last revision on October 17, 2022. The Economic Tracker data on Time Outside Home reflects this final update from Google, and will not receive further updates.