I. Summary[1]

The ZIP file, located at `https://reopeningdata.github.io/`, contains the replication materials for "Reopening the Economy: What Are the Risks, and What Have States Done?" by Enghin Atalay, Shigeru Fujita, Sreyas Mahadevan, Ryan Michaels, and Tal Roded. Along with the aforementioned list of co-authors, Ryan Kobler produced the revisions to the initial version of the reopening dataset.

The three main do files that need to be run, in order, are `create_essential.do`, `create_dataset.do`, and `reopening_figures.do`. The first two of the three files constructs the main datasets (`reopening_dataset.dta` and `reopening_dataset_combined.dta`). The final of the three do files produces the figures and tables that are included in the note and accompanying website.

II. On Measuring States' and Counties' Initial Designations of Essential Industries

Between the sixth and seventh (final) versions of our dataset, we have substantially modified the way in which we measure restrictions at the beginning of the pandemic. See Section V.F.1 for a discussion of our previous methodology to measure closures at the initial stages of the pandemic period. These changes not only allow us to more accurately capture states' and counties' restrictions on economic activity during the initial stages of the pandemic, but also allow us to measure the evolution of these restrictions over March and April of 2020. Previously, we had only recorded closures as of mid-April 2020 in one variable.

In the initial stages of the pandemic, CISA (the Cybersecurity and Infrastructure Security Agency) released three versions of their essential industries guidance (version 1.0 on March 19, 2020, version 2.0 on March 28, 2020, and 3.0 on April 17, 2020), that some states and local governments followed in the initial stages of the pandemic. Previous versions of our essential designations were based off the third version of CISA. However, for the updated version, we give the variation in between the three CISA versions. So, across all locality guidance we include the version of CISA as specified in the order, or if not specified, the active CISA version at the time of the order's release.

We also bring in initial gatherings restrictions for 38 states. Such gatherings restrictions tended to precede a main stay-at-home order. Incorporating these changes exposes a gradual increase in restrictions in March, and a decrease as guidance tended to expand to more industries in April and May.

Additional orders

The original set of local orders include Anchorage, Alaska, Austin, Texas, Bexar County, Texas, Broward County, Florida, Clay County, Missouri, Dallas County, Texas, Harris County, Texas, Hillsborough County, Florida, Jackson County, Missouri, Miami-Dade County, Florida, Orange County, Florida, Salt Lake City, Utah, St. Louis City, Missouri, St. Louis County, Missouri,

[1] Research results and conclusions expressed throughout this project are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve System, or the Federal Reserve Board of Governors.

Tarrant County, Texas. We've retained the aforementioned orders and added counties in California (Alameda, Contra Costa, Los Angeles, Sacramento, San Francisco, Santa Clara, and Ventura) between March 17, 2020 to May 4, 2020. For Florida, we include Duval, Palm Beach, and Pinellas counties; and for Kansas--Douglas, Johnson, Sedgewick, Wyandotte counties. For Missouri we add Greene and Jackson counties; for Texas, we add Denton and Hidalgo counties. Each county released a textual description of essential industries that are then matched to the prior keyword matching algorithm. We've also added 18 state-level orders of text-based essential industries.

In the previous version of our dataset, we separated states into those that released essential industry guidance and those that provided a nonessential list. However, several states indeed released both. Often, the inessential lists were released in early to mid-March prior to a stay-at-home or essential order. In this dataset update, we capture this additional level of detail for all states and Washington D.C., with exception of South Dakota, Hawaii, Idaho, and Pennsylvania.

Inclusion of multiple versions of CISA

For states that took guidance from CISA alone or referred to CISA in addition to a textual description of industries, we use the CISA order active at the time or specified in the order, updating the vector of essential industries across several states. For a complete list of the executive orders and their corresponding CISA versions (if applicable), see order_outline.xlsx.

Gatherings restrictions

We code gatherings restrictions – size thresholds that certain localities imposed – as generic or specific. A list is generic if an order prohibited gatherings above a certain size but does not explicitly mention certain industries. Gatherings are then mapped by size to three different sets of generic industries: above 250 people (NAICS 512131, 56192, 7111, 7112, 7113, 71329), above 50 people (NAICS 512131, 56192, 7111, 7112, 7113, 71211, 71212, 71213, 7131, 71329, 71392, 71394, 71395, 71399, 71329), or above 10 people (all 50+ people NAICS industries and 51912). If specific industries are mentioned, we include them instead of the generic list.

County-industry level employment weights

For aggregation in previous versions of the data, we used national-level employment by industry from the March 2019 QCEW to weight the shares into coarser industry classifications. We include additional detail in this update with QCEW county-industry-level weights, imputing where employment is nondisclosed. Industries are aggregated when determining the essential status of wholesale industries (described as a weighted average of manufacturing industries in the following section). They are also provided should the user want to aggregate to an industry and locality level coarser than the full detail.

Manufacturing and wholesale

For manufacturing, we use the 2012 input-output tables from BEA, which show domestic requirements for inputs to a destination industry, and 1997 Capital Flow table, that shows which industries purchase what types of equipment, to construct essential shares for all industries not mentioned explicitly but are necessary to other essential industries. We calculate the share of

essential employment in the origin sector by summing up over essential industry destinations times the share of origin employment in the destination sector—the essential input-output share. For each commodity, we also calculate the share of essential capital sales then translate back into units of employment and map them to industries--the essential capital share. The sum of the essential input-output and capital shares is then used as the essential share.

For wholesale industries, we assign essential status based upon an average of related manufacturing industries weighted by industry-county employment is nondisclosed.

III. On Measuring Reopening

We follow a three-step procedure to construct our dataset of reopening paths by industry and county. The first step involves hand collecting information from states' reopening orders, storing this information in excel files. The second step, executed in STATA, involves compiling this information into a structured dataset; in this dataset, observations are NAICS industry-county pairs and variables indicate whether the county-industry pair should be marked as "open" or "closed" on a given week (for each week between May 1 and the most recently available period.) We describe the first steps in the remaining portions of Section III. In a third step, we apply a sequence of do-files that update industries' open or closed status beginning in the fall of 2020. These .do files compile changes discussed in Section V.

III.A Collecting Information from States' Reopening Orders

Our data collection efforts center on reading through states' executive orders and local news stories.

There are two groups of states in this step. First, certain states have clearly defined "phases" and within-state regions, whereby the movement into successive phases varies across regions. For instance, Pennsylvania's reopening plan involves a "red" (initial) phase, a "yellow" phase, and a (terminal) "green" phase. The state's reopening process largely involves each of the 67 counties passing through the three phases. Other states (e.g., New York) have grouped counties into regions that move through the reopening process in tandem.[2] For seven states – Illinois, Michigan, New York, Oregon, Pennsylvania, Virginia, and Washington – we have compiled two worksheets: the movement of counties into different phases, and the list of industries that can be open in each county.

For the second set of states, we instead directly list the NAICS industries that are re-opened on each date. Even for this second group of states, there are certain instances of within-state across-county variation in when industries reopen. For counties that are exceptions to the rest of their respective states, we list these exceptions in our excel file as well. The file that collects information on reopening for these states is `reopening_other.xlsx`. Within this file, there are four worksheets. The first worksheet includes the main information that will be read in by STATA. The second and third worksheets list the counties (and associated FIPS county codes) and the titles of industries (and associated NAICS industry codes). The final worksheet provides a partial list of the sources for this worksheet.

[2] Within New York, for example, eight counties – Albany, Columbia, Green, Renssalaer, Saratoga, Schenectady, Warren, and Washington – comprise the "Capital District" region. These eight counties moved into phase 1 of New York's plan on May 20, phase 2 on June 3, phase 3 on June 17, and phase 4 on July 1.

Regarding the mapping between text and NAICS industries, there are three exceptional cases to keep in mind. First, certain states issued blanket reopening orders as part of their reopening. For instance, on June 1 Oklahoma entered "phase 3" of its reopening plan, whereby all industries are allowed to operate (albeit with certain restrictions). In cases like these, we assign all industries to be open. Second, certain states have indicated that "office-based businesses" (or similar phrases) are allowed to reopen. (An example of this would be Rhode Island, in phase 1 of its plan.) In these instances, we take NAICS sectors beginning with "5" (with some exceptions) in addition to NAICS sectors 8132-8139.[3] Third, in their reopening orders, many states mention reopening manufacturing without discussing wholesale and related distributing activities, while other states mention reopening these two sectors in tandem. In our view, it is unlikely that wholesale is still a restricted activity (especially given that manufacturing and retail have been allowed to open in every state.) So, for each state that (i) lists manufacturing as a reopening and (ii) has not, to the current date, explicitly mentioned wholesale as a sector that has re-opened, we designate the re-opening date for wholesale sector (plus distribution support activities and warehousing and storage, NAICS codes 488 and 493 respectively) to be that of the manufacturing sector's reopening date.[4]

III.B Storing Information in a STATA database

The code to produce a STATA database of states' reopening paths is given in steps 2 to 6 of `create_dataset.do`.

- Step 2 reshapes the excel files corresponding to reopening in Illinois, Michigan, New York, Oregon, Pennsylvania, Virginia, Washington so that it has the same format as all other states.
- Step 3 takes the reshaped files from step 2 and the excel file from other states and stores local macro variables with information on opening dates by counties and industries.
- Step 4 uses the local macro variables from step 3 to construct a database with county-industry pairs as the observation and reopening status (on a weekly basis, from early May to the current period) as different variables. This step also constructs a variable identifying when the industry was closed in its state.
- Step 5 modifies the dataset to account for recent (re-)closures (e.g., Texas' late June closure of bars). The output of this step is `reopening_dataset.dta`
- Step 6 collapses up to a coarser industry definition, the industry definition that allows us to merge to datasets like O*NET and the CPS. The output of this step is `reopening_dataset_combined.dta`

Step 4 takes several hours to run. However, to circumvent running the time-consuming portion of the do file, one may download `reopening_dataset.dta` (from `https://reopeningdata.github.io/reopening_dataset.zip`) and then set the local macro variable skip_step_4_5 equal to 1.

---

[3] The exceptions, within the NAICS 5 sector, are those industries that have been explicitly listed in states' reopening orders. These include 51211 (Motion Picture and Video Production), 51213 (Motion Picture and Video Exhibition), 51912 (Libraries and Archives), 54192 (Photography Services), 54194 (Veterinary Services), and 5617 (Services to Buildings and Dwellings).

[4] Of course, this change only applies to states that had at least partially classified wholesale and relate distribution activities as nonessential in their initial closing orders.

Finally, recently (beginning in the last week of June) certain states have re-imposed restrictions on economic activity. Most of these new closures are targeted to individual industries: bars, movie theaters. However, in other instances (primarily California) restrictions cover a wide swath of industries. For instance, on July 13 California Governor Gavin Newsom ordered nonessential offices in certain "watch list" counties to close. The word "nonessential" refers back to the state's initial designation of essential and nonessential industries. In states' re-closure orders we assume that no industry that was initially classified as essential will be forced to close.

IV. Figures and Tables

We produce the figures and tables in our note and on `https://reopeningdata.github.io/figures.html` using the `reopening_figures.do` do file. In addition to the data on reopening, whose construction we describe above, this do file employs data from the American Time Use Survey (to compute the share of workers who report the ability to work from home in each industry), the County Business Patterns (to compute employment in each industry and county), the Current Population Survey (to compute the age of workers in each industry), O*NET (to compute contact intensity in each occupation), the National Employment Matrix (to compute the number of workers in each industry), and the Quarterly Census of Employment and Wages (to compute the number of workers in each industry). The input files are all collected in in the replication ZIP file.

V. Changes Between Versions

V.A Changes between the first and second versions of our dataset and code

There are four main differences between the first (July 8) and second (July 21) version of our dataset and code. The first three are fixes to errors in our coding of states' reopening orders. The final is an update to include the most recent data.

First, on May 13, 2020 a Wisconsin State Supreme Court decision effectively ended its statewide restrictions on economic activity. In response, individual counties enacted replacement restrictions. We have coded up the restrictions for Dane County, Milwaukee County, and Sauk County. Unfortunately, the code in our July 8 version did not capture the statewide change (Wisconsin minus the three aforementioned counties).

Second, in the first version of our code and database, we had initially assumed that Louisiana had fully reopened by May 15. In reality, the following industries should have marked closed until the week of May 29: Beauty Salons; Nail Salons; and Other Personal Care Services. In addition, the following industries are still closed: Child Daycare Services; Promoters of Performing Arts, Sports; and Amusement, Gambling and Recreation. We have updated our code and dataset to reflect this reality.

Third, within create_dataset.do, there are the following two lines of code:

```
sort statefips countyfips naics month day
by statefips countyfips naics: keep if _n==1
```

The idea, here, is to make sure that each NAICS industry only appears once for a given state-county (and that we take the first mention of the NAICS industry as the date of reopening). Within the July 8 version of our do file, the following line of code was missing.

```
sort statefips countyfips month day naics
```

To see why this line of code is necessary, consider a hypothetical state-county which opens restaurants on May 1 and bars on May 15. In the excel spreadsheet, we would have NAICS=7225 open on May 1 and NAICS=722 on May 15. Without the final extra line of code described above, the macro variable corresponding to NAICs=722 is read *before* the macro variable for 7225. This is wrong. What will happen is that all of 722 (including 7225) will be reported as closed on May 1 and May 8. When adding the extra line of sorting, 7225 is correctly read as open on May 1 and May 8. The error that we have identified leads us to classify the pace of reopening as slower than it actually was. This error was pertinent for slightly less than 0.1 percent of the state-county-industry triples in our dataset.

Finally, we have added two new variables, indicating whether industries were open as of July 17 or as of July 24. Within New York, New York City moved into the final phase of the state's reopening plan. In the other direction, California has increased the scope of its re-closures. Pennsylvania and New York have closed bars statewide. Nevada and Colorado had restrictions on bars that began before July 10, and that we have now included in our dataset.

V.B. Changes between the second and third versions of our dataset and code

There are three main corrections and two speed enhancements between the second (July 21) and third (August 5) versions of the code. Together, the corrections resulted in changes to 0.34% of observations, with respect to the dates at which industries were reopened.

V.B.1 Corrections:

1. Reformatting month and day variables as numeric within Step 3 of `create_dataset.do`.

After updating the sorting procedure in the second version with

```
sort statefips countyfips naics month day
```

The following line of code was missing immediately preceding the sort:

```
destring month day, replace force
```

Without this line of code, month and day would string variables, meaning that, for example, May 22 as erroneously sorted before May 9.

2. Add an if-condition (month==.) to reopening date replacements.

Reopening dates are associated with either one, two, three, four, five, or six digit NAICS industry codes. We assume that coarser codes map onto several finer codes, if no finer reopening date exists. So, the reopening date associated with 448, clothing and accessory stores in Worcester County, Maryland should become the reopening date for finer NAICS 448320 (leather and luggage stores), 448150 (tie shops), 448210 (shoe stores), since they are within the category of clothing and accessory stores and were not mentioned explicitly within Worcester County's reopening order.

However, in Worcester County industry code 448320 was coded as reopening on June 19, 2020. This date originates from the reopening date for the 2-digit NAICS industry 44, or retail overall. But, in this county, there exists a finer 3-digit NAICS 448 for clothes and accessories with a

reopening date of May 15, 2020, a month before all of retail. The sorting in section V.A. orders these two NAICS codes properly as follows:

NAICS MONTH DAY

1. 448 5 15

2. 44 6 19

But the line

```
replace open_may_21 = (date< if substr(naics,1,`qq')=="`naics_`x'"" & fipst==`state_`x'' & `county_`x''==fipsc
```

first sets open_may_21 to 1, from row 1 in the table above. Then row 2 flips open_may_21 to 0 in the subsequent iteration due to the later reopening date.

To resolve this issue, the phrase in bold is added to the end of the line of code

```
replace open_may_21 = (date< if substr(naics,1,`qq')=="`naics_`x"" & fipst==`state_`x" & `county_`x"==fipsc & open_may_21==.
```

which in this example allows information to be written from the first row, but skips the second row of the above table. Thus, the reopening date for all of retail will only be written to 6-digit industries starting with 44 that have not yet been filled by finer NAICS codes. In general, this addition prevents overwriting reopening dates from finer NAICS codes with those of coarser codes.

3. Merge reopening dates to (state, county, six-digit NAICS code) triples, in place of NAICS codes of various digits. At the beginning of the step 4_5 loop within `create_dataset.do`, we replace

```
use naics_essential_list, replace
```

with

```
use naics_list, replace
```

The dataset records 1,057 6-digit NAICS codes, resulting in 3,371,830 county-state-NAICS triples when crossed with all 3,190 U.S. counties. Previously, naics_essential_list comprised 604 NAICS codes of various digits that aligned with each industry mentioned in counties' initial essentiality lists, resulting in 1.9M observations when crossed with all U.S. counties. However, there were some cases in which reopening orders were more finely coded, i.e. required more digits, than their initial essentiality designations.

For example, Winchester County, VA initially closed the performing arts sector with and without facilities (NAICS 7113). The then county reopened performing arts without facilities (71132) on June 5, 2020, and the remainder of performing arts (711) on July 1, 2020. Therefore, the line of code:

```
substr("`naics_`x'"",1, qq2)==naics | substr(naics,1,`qq')=="`naics_`x'""
```

was needed to match longer reopening NAICS codes to their shorter counterparts when not present in the 604 industries of naics_essential_list. In this case, this technique matches the reopening date from performing arts without facilities to all performing arts, which is incorrect.

To resolve this issue, we expand to the 6-digit naics_list and remove the first half of the or-statement in the previous line of code, replacing it with:

```
substr(naics,1,`qq')=="`naics_`x'"
```

V.B.2 Speed enhancements:

1. Partition the local macros into 2 lists.

In the most recent version of the code, within the end of Step 3 of `create_dataset.do`, we partition the original local macro lists based on whether the variable `countyfips' is missing. countyfips determines whether a certain industry's reopening date is recorded at the state or the county-state level. Because observations in our reopening_dataset are ultimately county-state-NAICS triples, this partitioning is necessary to define whether reopening dates are county-specific and should map to one observation, or should be written to several observations, counties without county-specific orders within a state.

a. After partitioning, we define two sets of local macro variables: `county_1', `county_2', … and `county2_1', `county2_2',… for county-state and state level reopening orders respectively.

b. Rather than iterating over the all pooled county and state-level orders twice, partitioning the orders beforehand results substantially fewer iterations.

2. Generate open_may_1, open_may_7, … , open_aug_7 variables outside of the macro loop.

V.B.3 Updates to reopening orders:

Note, these updates are also listed at `https://reopeningdata.github.io/updates`

- We have added variables indicating whether industries were open as of July 31 or as of August 7.
- We have added movie theaters (open at 50% capacity) and gyms to Washington's Phase III essential industry list
- Kentucky closed bars in the last week of July. Santa Cruz County and San Mateo County in California were added to the state's watch list in that week.
- South Carolina has movie theaters, concert halls, and sport arenas to open, as of August 3.
- South Carolina had allowed retail to re-open in late April, something that we had previously missed.
- Previously, we had categorized Allegheny County, Pennsylvania as having closed restaurants and bars from the weeks of July 3 to July 17. In fact, outdoor service was allowed beginning July 8.

V.C. Changes between the third and fourth versions of our dataset and code

There are four main corrections between the third (August 5) and fourth (October 5) versions of the code.

- We modified our classification of Kansas' initial closure orders. Initially, Kansas was among the list of states for which we applied an R program to vectorize the text in the executive orders, and then mapped certain words/phrases to NAICS sectors. On further

consideration, since the text within the Kansas order was so sparse, we have recognized that the automated text search did not work as well for Kansas as for other states. Instead, we have hand classified Kansas' initial order to categorize which industries were initially closed.

- New Hampshire fully reopened on June 29. In the previous version of our dataset, we had missed this.
- In previous versions of our code, we had classified lodging firms (NAICS 721) as being closed in New Mexico. While lodging was not listed as essential within New Mexico's list closure orders, there is circumstantial evidence (here: www.currentargus.com/story/news/local/2020/03/26/coronavirus-spread-challenges-new-mexico-tourism-industry/5077501002/) that lodging was allowed to be open, at least partially, throughout the spring of 2020.
- Between August 5 and October 2, states had made a number of changes to their orders. We summarize the post-August 5 changes here: reopeningdata.github.io/updates.html .

V.D. Changes between the fourth and fifth versions of our dataset and code

There are three main corrections between the fourth (October 5, 2020) and fifth (January 6, 2021) versions of the code.

- An error in previous versions of our code had led us to erroneously miss changes in certain industries open closed status. The error is as follows. In previous version of our `create_dataset.do` file the following line

  ```
  collapse (mean) open_* initial, by(fipstate fipscty naics_qcew)
  ```

  was placed directly after we had created our indicator variables of state-county designations by week, and directly before the lines of code that modify these designations for the latter half of 2020.
  The problem is that the collapse statement leads industries to be coarser than the industries that are mentioned in modifying lines of code. For instance, naics_qcew has Other Amusement and Recreational Activities (NAICS 7139) listed as a single industry, yet the modifying lines of code refer to NAICS 71394 (gyms) and NAICS 71395 (bowling alleys).
  In our code, now, this collapse statement has now been moved down, directly below the lines of code that refer to the `sept25_updates.do` and `jan1_updates.do` files.
- In previous versions of our dataset, with a few exceptions, we had hard-coded public schools (NAICS 6111) as essential in the early stages of the pandemic (in the spring of 2020). Now, drawing on data collected by Christopher Adolph, Kenya Amano, Bree Bang-Jensen, Nancy Fullman, and John Wilkerson (see `https://github.com/COVID19StatePolicy/SocialDistancing`), we have updated our measures of K-12 schools open/closed status. In future versions of our dataset, we plan on incorporating information on restaurants and bars from the Adolph et al. (2020) database.

- Between October 2, 2020 and January 1, 2020 states had made a number of changes to their orders. We summarize the post-October 5 changes here: `reopeningdata.github.io/updates.html` .

V.E. Changes between the fifth and sixth versions of our dataset and code

There are six main corrections between the fifth (January 6, 2021) and sixth (May 13, 2021) versions of the code.

- We have completely recoded the way in which restaurants (NAICS 7225) are coded. We begin with data collected by Christopher Adolph, Kenya Amano, Bree Bang-Jensen, Nancy Fullman, and John Wilkerson (see `https://github.com/COVID19StatePolicy/SocialDistancing`). For observations with StatePolicy="RestaurantRestrict", we record the indoor capacity, outdoor capacity, and distancing requirements. These capacity levels are stored in a separate dataset: restaurant_panel_weekly.csv. Now, for our main dataset, record restaurants as closed for business if the maximum outdoor capacity and maximum indoor capacity are both equal to 0.
- In previous versions of our code, we had coded – within do files – changes to California's counties' restrictions. In this version, we download counties' restrictions from `https://data.chhs.ca.gov/dataset/covid-19-blueprint-for-a-safer-economy` week by week, store these restrictions in a csv file, then write a separate do file (`ca_updates.do`).
- We added several industries to phase 3 and phase 4 of New York's summer 2020 reopening plan. These are industries which we now believe we had missed previously.
- In October 2020, several counties within Washington state moved from phase 1 to phase 2: Benton, Chelan, Douglas, Franklin, and Yakima. We had missed these amendments in previous version of our dataset.
- For the fall 2020 to the present, Colorado had closed (on a county-by-county basis) only bars and a subset of entertainment and personal services industries. In previous version of our code, we had listed additional industries as closed during this period. Now, we mark all of these industries as open beginning when Colorado moved to its "dial" system in September 2020.
- We added data (extra_reopenings.csv) with additional reopenings that were not initially recorded, most often affecting Civic and Social Organizations (NAICS ID 8134).

V.F. Changes between the sixth and seventh (final) versions of our dataset and code

V.F.1 Treatment of closures in the initial weeks of the pandemic (up to April 2020)

Between the sixth and seventh (final) versions of our dataset, we have substantially modified the way in which we measure restrictions at the beginning of the pandemic.

The appendix of "Reopening the Economy: What Are the Risks, and What Have States Done?" provides some background on our old procedure to construct an industry-by-county dataset of closures (as of April 17, 2020). Here, we review this material, referring specifically to the replication files (corresponding to the older versions of our dataset).

We followed two different procedures, one for a set of states (and counties) with relatively concise closing orders, and a second for states with lengthier orders. For the states with more concise orders, we hand coded the NAICS industries that we identify as closed or open. Our hand-coded industries are collected in `essential_industries_initial.xlsx`. For states with lengthier orders, we developed a list of keywords (collected in `finalkey.csv`) associated with each NAICS industry. Then, we searched for these keywords in the text of states' and counties' closing orders. The output of this searching procedure is collected in `initial_other_states_based_on_cisa3.csv`. The code (and raw text from states' orders) to produce `itial_other_states_based_on_cisa3.csv` is given in the folder Create_Keyword_Based_Designations. Using `essential_industries_initial.xlsx` and `initial_other_states_based_on_cisa3.csv` as inputs, the code to produce a STATA database of states' initial closures is given in step 1 of `create_dataset.do`.

V.F.2 Treatment of closures during the later stages (May 2020 and beyond) of the pandemic

As we have discussed, above, we have modified the way in which we measure closures in the initial stages of the pandemic. Compared to previous version of our dataset, these changes lead us to measure some industries as closed when we had previously described them as open (and some industries as open when we had previously measured them as closed.) For a subset of the industries which we had previously classified as open and now classify as closed in the initial stages of the pandemic, running the rest of our old code would lead us to classify them as closed even in the later stages of the pandemic. But we know that – as of the end of April 2021 – very few if any industries were actually closed.

What is likely happening is that (i) states' and counties' initial designations listed industries as being closed in March and April (ii) the same states and counties never explicitly mention reopening of those industries. For these industries, we have recorded them as having reopened when during the initial stages of the jurisdiction's reopening plans. To give two examples: certain transportation and warehousing industries were closed in New York state's initial stay-at-home orders, and were never explicitly mentioned in any of the 4 phases in their spring/summer 2020 reopening plan. Based on our own judgement, we have placed this sector within phase 1 of New York's reopening plan. A second example, is the wholesale sector, which initially closed and never mentioned in Michigan's reopening plan's. Again, based on our judgement, we have placed this sector as reopening in phase 3 – the same phase as the manufacturing sector – of the state's reopening plan. While imperfect, our ex post adjustment of these industries' statuses accurately depicts the initial and final closure statuses of these industries.