A New Claims-Based Unemployment Dataset: Application to Postwar Recoveries Across U.S. States*

Andrew Fieldhouse† Sean Howard‡

Christoffer Koch§ David Munro¶

September 29, 2022

Abstract

Using newly digitized unemployment insurance claims data we construct a historical monthly unemployment series for U.S. states going back to January 1947. The constructed series are highly correlated with the Bureau of Labor Statics' state-level unemployment data, which are only available from January 1976 onwards, and capture consistent patterns in the business cycle. We use our claims-based unemployment series to examine the evolving pace of post-war unemployment recoveries at the state level. We find that faster recoveries are associated with greater heterogeneity in the recovery rate of unemployment and slower recoveries tend to be more uniformly paced across states. In addition, we find that the pace of unemployment recoveries is strongly correlated with states' manufacturing share of output.

Keywords: State-Level Unemployment Rates, Unemployment Insurance, Economic Recoveries, Regional Business Cycles.

JEL Codes: C82, E24, E32, J64, J65, R11

---

*We thank seminar and conference participants at Middlebury College, the University of Alberta, the Liberal Arts Macroeconomics Labor group, the Midwest Macroeconomics Conference, and the Liberal Arts Macroeconomics Conference for constructive feedback and comments. We are especially grateful to Erica Cooper of the Department of Labor for assistance in providing digital scans of historical reports. We also thank Jeremy Piger for providing data. David Munro gratefully acknowledges funding received from MiddData during the completion of this project. The views expressed in this paper are the views of the authors only and do not necessarily reflect the views of the Federal Reserve Bank of Dallas, the Federal Reserve System, or the International Monetary Fund, its Executive Board, or its Management.

†Department of Finance, Mays Business School, Texas A&M University, *afieldhouse@mays.tamu.edu*.

‡Wood Mackenzie, *sphoward@middlebury.edu*.

§International Monetary Fund, *ckoch2@imf.org*.

¶Department of Economics, Middlebury College, *dmunro@middlebury.edu* and GLO Fellow.

1 Introduction

Macroeconomists are increasingly leveraging panel datasets and regional heterogeneity to identify economic relationships.[1] There is also an increasing awareness that the unemployment rate is one of the best indicators of economic activity, particularly for business cycle analysis (Romer and Romer, 2019).[2] Unfortunately, official U.S. unemployment rate data at the state level only begins in 1976, which greatly hampers historical state-level analyses. For instance, a rich literature on state-level labor market recoveries, regional business cycles, and state coincident economic indexes has largely been limited to starting around 1978 because of the unavailability of state unemployment data for earlier years (e.g., Blanchard and Katz (1992); Crone and Clayton-Matthews (2005); Owyang et al. (2005); Brown (2017)).

In this paper we present a newly constructed monthly unemployment dataset for U.S. states beginning in January 1947. Our novel unemployment series is constructed from a large dataset of newly digitized monthly unemployment insurance claims from a series of historical reports previously published by the Department of Labor (DOL) and Social Security Administration (SSA). Together with available monthly data on total nonfarm payroll employment we compute an alternative claims-based unemployment rate and show that it is highly correlated with official state-level unemployment rates for 1976–2022, capturing consistent patterns in both national and state business cycles.

We use our claims-based unemployment series to examine the evolving pace and heterogeneity of labor market recoveries across postwar U.S. recessions from 1948 through 2022.[3]

[1]For instance, Nakamura and Steinsson (2014) and Chodorow-Reich (2019) exploit regional heterogeneity to identify cross-sectional fiscal multipliers and Hazell et al. (2022) exploit regional heterogeneity to study the slope of the Phillips curve. See Glandon et al. (2022) for a broader overview of the recent shift in empirical macro toward panel data and micro data.

[2]This is particularly true for state-level business cycle analysis. Blanchard and Katz (1992) find that, following an adverse employment shock, the level of state employment is permanently altered, largely driven by labor mobility across states and out migration, rendering mean-reverting unemployment rates better suited than employment for studying state-level economic recoveries.

[3]More accurately, we analyze trough-to-peak recovery rates following the 1948–49 through 2007-09 recessions and recovery rates to date following the pandemic recession because of limited data availability.

Our analysis of unemployment recoveries closely follows the recent work of Hall and Kudlyak (2022) and Dupraz et al. (2019), but does so at the state level. Hall and Kudlyak (2022) document that recoveries in the U.S. unemployment rate have been quite stable since the early 1960s, but the pace of recovery has decelerated markedly since the recoveries from the earlier post-war recessions of 1948-49, 1953-54, and 1957-58.[4] We corroborate this stylized fact at the state level, and find that faster recoveries are associated with greater heterogeneity in the pace of recoveries across states: Relatively faster national recoveries reflect very dissimilar experiences across states, whereas states tend to experience more similar paces of recovery in slower national recoveries. This is particularly true for the faster labor market recoveries in the 1940s and 1950s, but the correlation remains significant when excluding these earlier postwar recoveries. In addition, we find that the faster recoveries at the state level tend to be associated with states with larger manufacturing sectors. The strength of this association has declined over time, but remains positive and significant in the most recent recoveries during the 1990s, 2000s, and 2010s, before the pandemic.

2 Dataset Construction

In this section we first overview the digitization process for historical state-level unemployment insurance claims, including how outliers and possibly spurious data entries were handled. We discuss the construction of a novel claims-based unemployment series from this newly digitized data. And we conclude by discussing the relationship between our claims-based unemployment series with official measures of the unemployment rate and the insured unemployment rate.

[4]They find that, on average, the unemployment rate falls by 0.1 log points—or one tenth of the starting unemployment rate—each year after recovery begins, until this stable rate of recovery is upended by the next recession or crisis. They interpret their results as casting doubt on the existence of a "constant natural rate of unemployment around which unemployment oscillates" (Hall and Kudlyak, 2022).

2.1 Digitizing Historical Unemployment Claims

Monthly state-level unemployment insurance claims are available in digital form dating back to January 1971 from the Department of Labor's website. Using scanned versions of printed reports previously published by the DOL and SSA, we backdate the publicly data available by digitizing monthly data on Initial Claims (IC) and Continued Claims (CC) back to December 1946 for all 50 states and the District of Columbia.[5] December 1946 is chosen as the start of sample so that a three-month centered moving average of claims is available back to January 1947 (discussed below). The historical claims data originate from one of a series of periodical reports: *Employment Security Activities*, *The Labor Market and Employment Security*, *Unemployment Insurance Statistics*, and the *Unemployment Insurance Review*. We were able to access most of these primary sources via HathiTrust or Google Books, and supplemented missing publications with Interlibrary Loan requests or scans from the Department of Labor's internal library. The image quality of the scans we were able to locate was generally quite good. When merited and feasible, we retrieved duplicate copies of scanned reports to resolve uncertainties relating to poorer scan quality. We also used data in the reports on national aggregates as a crosscheck with the sum of state claims and, in cases where image quality presented legibility issues, data on the percentage change from the prior (or subsequent) month to guide digitization; see Appendix A.1 for details.

To construct a full time series backdated to December 1946, newly digitized claims data were merged with DOL's publicly available state-level IC and CC data for regular state programs only, to be consistent with the historical claims data (see Appendix A.1 for details). We digitized IC data for up through December 1972 and CC data through December 1977 to ensure and validate a smooth merge between historical claims data and DOL's claims data; the two datasets align seamlessly for IC throughout 1971–72 and align for CC somewhere

[5]Federal unemployment insurance programs and the Bureau of Employment Security were transferred from the SSA–at the time part of the Federal Security Agency–to DOL in the late 1940s, and publication of the *Economic Security Activities* report flipped over from SSA to DOL in October 1949.

between 1971–1977, depending on the state (see Appendix A.1). In total, just over 36,000 monthly observations were digitized.

After merging the two claims series, we seasonally adjusted the full IC and CC series for December 1946 through June 2022 using the Census Win-X-13 seasonal adjustment program; see Appendix A.1 for details. We also used Win-X-13 to run a series of outlier tests, which identified roughly 200 potential outliers from approximately 91,000 observations (newly digitized and existing data combined). These outliers were roughly evenly distributed between our newly digitized historical claims data and the existing DOL data. We manually checked each potential outlier to assess whether it represented a legitimate change in claims (e.g., a surge in Louisiana following Hurricane Katrina) or a "fat thumb" data coding issue (e.g., an implausible IC spike in Missouri in June 1974 that exceeded the state's population). We used several verification processes in checking and cleaning outliers; see Appendix A.1.

2.2 Claims-Based State Unemployment Rates

Using these unemployment claims data, we construct monthly claims-based unemployment rates for all 50 states and the District of Columbia. Our claims-based unemployment rate is conceptually similar to the official unemployment rate estimated by the Bureau of Labor Statistics (BLS)—the ratio of unemployed workers to the labor force—but uses initial and continued unemployment insurance claims as a measure of unemployed workers and uses employed workers plus unemployed workers receiving these benefits as a measure of the labor force. We use nonfarm payroll employment from the Current Employment Statistics (CES) as our measure of state-level employment, a choice simply motivated by data availability; to our knowledge it is the only monthly state-level employment series going back to 1947.[6] As with the claims data, we seasonally adjust nonfarm payroll employment for each state using

[6] A few states do not have nonfarm employment data going all the way back to January 1947: Data for MN begins in January 1950, data for MI begins in January 1956, data for HI begins in January 1958, and data for AK begins in January 1960. Our claims-based unemployment rate is constrained to these later start dates for these four states.

the Census Win-X-13 seasonal adjustment program.[7]

Our claims-based unemployment rate for state i in month t is computed as

$$UR_{i,t}^{Claims} = \frac{IC_{i,t} + CC_{i,t}}{NP_{i,t} + IC_{i,t} + CC_{i,t}} \tag{1}$$

from December 1946 onwards, where $NP_{i,t}$ denotes nonfarm payroll employment. The digitized monthly IC and CC data reflect all claims filed with the state unemployment office in that month, so an individual can show up as an initial claimant as well as a continued claimant and/or show up repeatedly as a continued claimant in a given month. To avoid such double-counting of individuals, we convert monthly claims to average weekly claims for our measure of $IC_{i,t}$ and $CC_{i,t}$, an approach conceptually similar to the reference week used by the BLS in sampling labor force activity.[8]

We construct a similar claims-based unemployment rate for the United States, aggregating the seasonally adjusted weekly average of claims and nonfarm payroll employment data for all 50 states and the District of Columbia,

$$UR_{US,t}^{Claims} = \frac{IC_{US,t} + CC_{US,t}}{NP_{US,t} + IC_{US,t} + CC_{US,t}} \tag{2}$$

where $IC_{US,t} = \sum_{i=1}^{51} IC_{i,t}$ and $CC_{US,t}$ is defined analogously. The claims data, particularly IC, can be rather noisy, so we smooth the data using a three-month centered moving average of each claims-based unemployment rate series in all figures and results below.

Figure 1 plots our claims-based unemployment rate (red) and the BLS's official unemployment rate (blue) for Ohio, meant as an illustrative and representative large state.[9] The two unemployment series identify nearly identical features of Ohio's business cycle when both

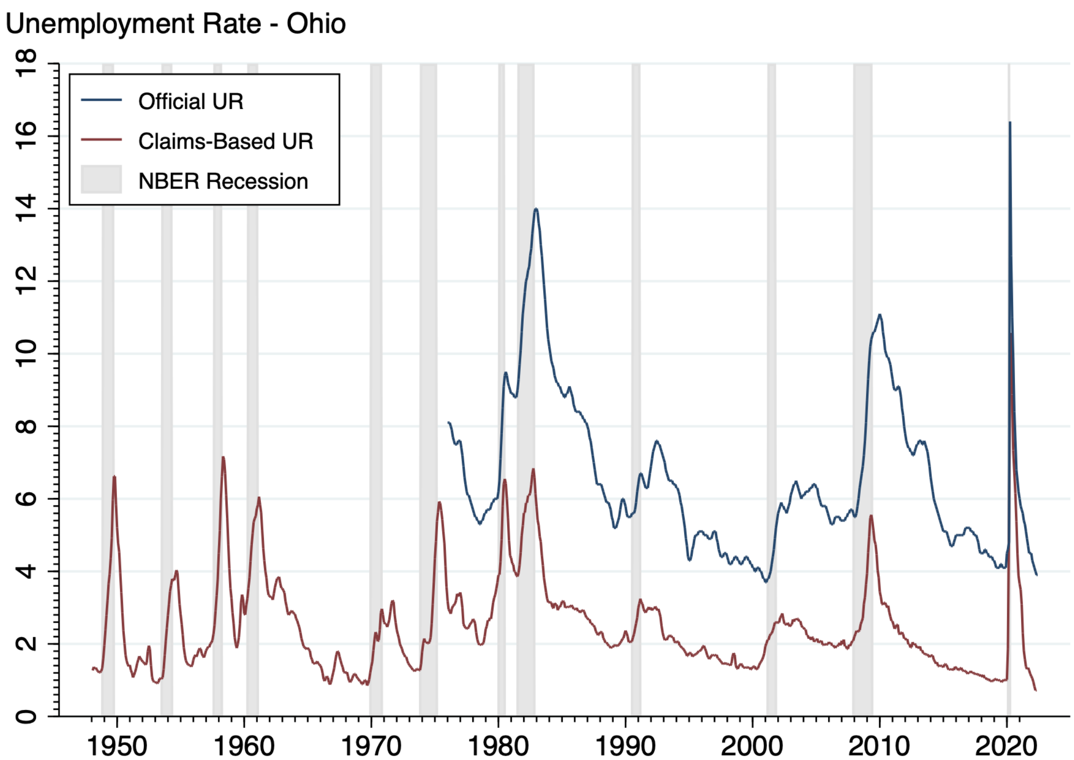
[7] Seasonally adjusted state-level nonfarm payroll employment data is only available from BLS for January 1990 onwards.

[8] In keeping with the DOL data for average weekly insured unemployment in a given month, monthly data are weighted by the split number of five-day workweeks in the month. We calculate the weights as the sum or workdays in each given month, ignoring holidays, divided by five days for the workweek.

[9] Appendix B.2 plots our claims-based unemployment rates for all 50 states along with gray recession bars derived from those unemployment rates.

are available over January 1976–May 2022; the correlation between the two series is 0.81. Figure 1 also underscores the practical benefit of our claims-based unemployment rates as a measure of labor market slack: Relative to the official BLS series, our claims-based series offer nearly three additional decades of monthly data at the state level, a period spanning the first six post-war national recessions identified by the National Bureau of Economic Research (NBER).

Figure 1: Comparison of the Official and Claims-based Unemployment Rates for Ohio



Notes: The claims-based unemployment rate is smoothed with a three-month centered moving average. Sample: January 1948–May 2022.

The difference in levels between the two unemployment series is to be expected. Our claims-based measure should be strictly lower than the official unemployment rate because of the narrower pool of workers eligible for benefits and because of the exhaustion of benefits for workers unemployed beyond the maximum duration for regular state UI programs, discussed below. The absolute difference in levels, however, is not a concern when using claims-based unemployment rates for business cycle analysis or informational content about labor market slack, as the two series are highly correlated.

2.3 Comparison with Official Unemployment Measures

It is important to emphasize at the outset that, while conceptually similar, our claims-based unemployment rates measure labor market slack differently than the official unemployment rate. It is also the case that there is no single objective measure of unemployment. Even the official unemployment rate must take stand on job search activity requirements for individuals to be counted as unemployed, which can be an important source of bias when measuring slack in the labor market from the headline unemployment rate, e.g., the unemployment rate being pushed down by discouraged workers dropping out of the labor force during the recovery from the Great Recession. Our claims-based unemployment rates are also conceptually similar to the Insured Unemployment Rate (IUR) produced by the DOL's Employment and Training Administration (ETA), which could be characterized as an "official" claims-based unemployment rate. Indeed our claims-based unemployment rates lie conceptually between the headline unemployment rate and the IUR. All three series provide economically significant and highly correlated information about labor market slack at the state or national level. The key advantage of our claims-based unemployment rate series is that they can be consistently calculated at the state level going as far back as 1943, whereas, at the state level, official unemployment rates are only available starting in January 1976 and IUR data are only publicly available starting in January 1986.[10]

To be counted as unemployed by the BLS, a worker must not have been employed but have been available for work during the surveyed reference week, and must have either actively searched for work in the four weeks ending in the surveyed reference week or been expecting to be recalled to work following a temporary layoff. The official unemployment rate

[10]Monthly state-level IC and CC data are available going to back to February 1943 in the *Economic Security Activities* reports. At the national level, the official IUR is publicly available starting in 1971.

is then calculated as the ratio of unemployed workers to the labor force, defined as the sum of employed and unemployed workers. Cognizant that there is no single objective measure of unemployment, the BLS calculates five alternative measures of labor underutilization in addition to the official, or "headline," unemployment rate ("U-3"), with the iteratively broader U-4, U-5, and U-6 measures loosening active search and work requirements to include discouraged workers, marginally attached, and those work part-time for economic reasons.[11] All six measures are calculated solely from data collected in the Current Population Survey (CPS), a monthly survey of roughly 60,000 households inquiring about their employment status over the seven-day reference week.

Our claims-based definition of unemployment is instead restricted to those individuals actively claiming regular state UI benefits as reported weekly by state unemployment offices to ETA or its preceding federal agencies. There are both conceptual drawbacks and advantages to our approach relative to BLS's household survey methodology. Using actual claims as a proxy for unemployment leaves no margin for bias from respondents misunderstanding definitions or misreporting their circumstances, or from time-varying non-response rates to surveys, a growing concern with the CPS of late (Bernhardt et al., 2021). On the other hand, state unemployment offices could misunderstand ETA's data definitions or incorrectly transcribe data. Another key difference arises from benefit duration limits: Unemployed workers who exhaust regular state benefits will drop from our measure of unemployment, whereas they would continue to be counted as unemployed by BLS provided they meet active search requirements. Official unemployment rates instead see workers drop from their headline measure if they do not report having searched for work in the previous four weeks. State unemployment programs also typically exclude certain workers from benefit eligibility, e.g., agricultural workers, domestic workers, and independent contractors, all of whom could be counted as unemployed if surveyed by the CPS. Regular state UI programs also exclude

[11] The U-1 and U-2 measures are more narrow than U-3, looking at long-term unemployment and job losers plus temporary workers, respectively.

federal workers and veterans who would instead qualify for the federally funded Unemployment Compensation for Federal Employees (UCFE) and Unemployment Compensation for Ex-service members (UCX) programs; the much larger pool of state and local government employees is, however, reflected in regular state UI claims data.

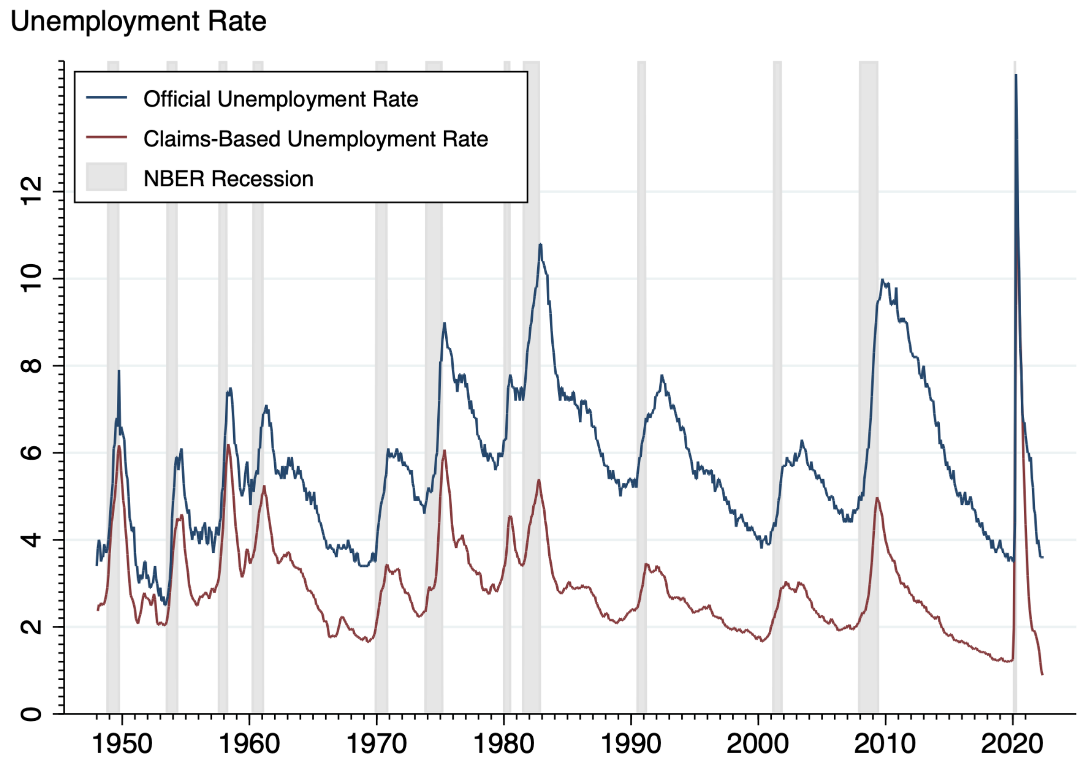
While the official unemployment rate and claims-based unemployment rate measure somewhat different things, it is important to emphasize that they are highly correlated measures of labor market slack, and identify similar inflection points in the business cycle (discussed below). Indeed, official measures of state-level unemployment rates are a statistical construct which are, in part, derived from unemployment claims data.[12] Figure 2 plots our U.S. claims-based unemployment rate (red) and the official U.S. unemployment rate (blue) over January 1948–May 2022, underscoring that the two series are highly correlated and identify broadly consistent features of the business cycle for the nation as a whole, not just a representative state. The expected level difference between the two series is less pronounced in the 1940s and 1950s because of the larger share of workers employed in agriculture; our use of nonfarm payroll employment downwardly biases the size of the labor force relative to that used for the official unemployment rate, pushing the claims-based unemployment rate higher when more workers are employed in agriculture.[13] But the secular trend of a declining share of workers employed in agriculture has a minimal effect on business cycle inflection points, as evidenced by Figure 2, and this time-varying trend would be absorbed by most detrending exercises.

While the official unemployment rate and claims-based unemployment rate capture similar features of the business cycle, they may be telling different stories about the relative degree of labor market slack—particularly how much slack persists today. Figure 2 shows the official unemployment rate nearing but not quite reaching pre-pandemic levels during

[12]The BLS's Local Area Unemployment Statistics (LAUS) program uses data from the CES, CPS, and state UI programs to estimate state unemployment rates; see BLS's LAUS program webpage for more details.

[13]The ratio of total farm employment (from the Historical Statistics of the United States, K-179) to nonfarm payroll employment has steadily fallen from 23.6% in 1947 to 13.0% in 1960 and 6.4% in 1970.

Figure 2: Comparison of the Official and Claims-based Unemployment Rates for the U.S.



Notes: The claims-based unemployment rate is smoothed with a three-month centered moving average. Sample: January 1948–May 2022.

the recovery to date from the pandemic recession, whereas the claims-based unemployment rate has fallen well below it's pre-pandemic levels, hitting a record low in May 2022.

In terms of creating a consistent measure across time, one possible concern about our claims-based unemployment rate is that maximum duration of benefits have, to some degree, changed over time. To gauge this potential concern, we examine how the maximum duration of regular state benefits have evolved over time using the State Unemployment Insurance Laws dataset compiled by Massenkoff (2021) for 1970–2018, which we extend back to 1947 from DOL reports.[14] Appendix Figure A.1 shows the mean of the maximum duration of benefits for all 50 states, plotted with one standard deviation bands. The average maximum

[14]The Massenkoff (2021) dataset is digitized from the *Significant Provisions of State Unemployment Insurance Laws* reports issued by the DOL ETA every January and July since 1965. We extend this dataset back to 1947 using earlier annual issues of the DOL reports, which are available online from DOL.

duration of benefits begins at approximately 22 weeks in 1950, and rises to approximately 26 weeks by 1960. From 1960 to 2011 it remains quite stable around 26 weeks and declines slightly to 25 weeks when a handful states began to reduce benefits during the recovery from the Great Recession. Overall, the mean maximum duration of benefits is quite stable. The share of unemployed workers who are out of work for 27 weeks or longer, slightly exceeding these mean maximum benefit durations, is also generally quite low; see Appendix Figure A.2.

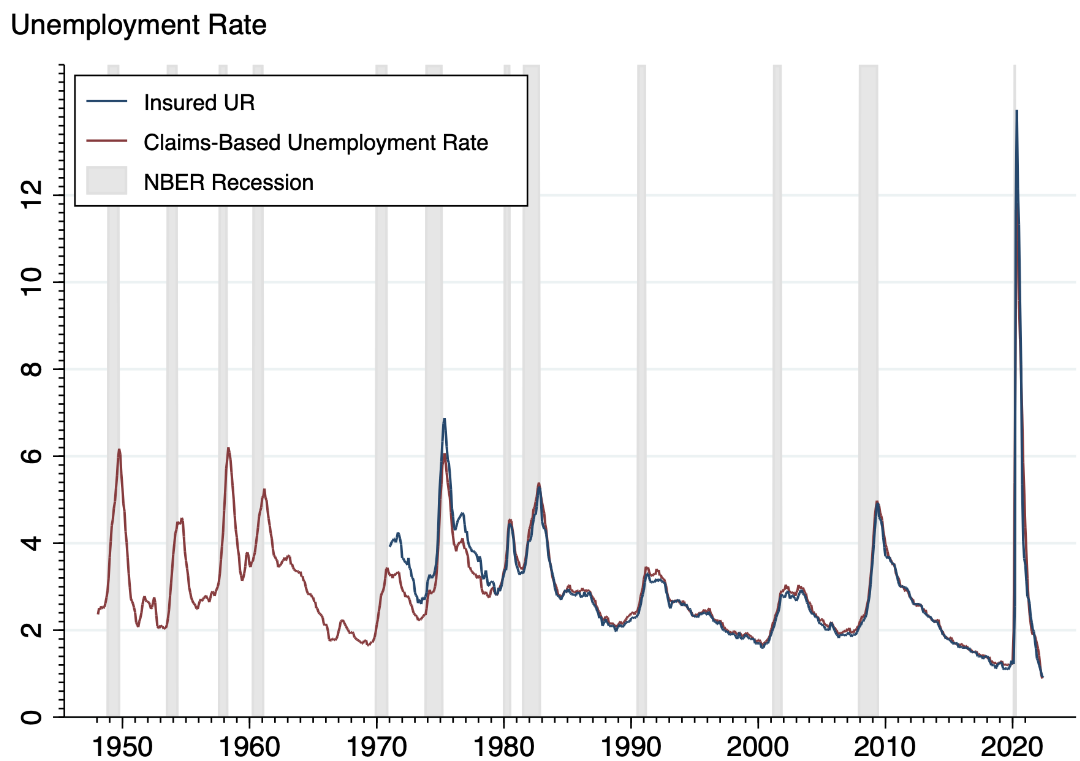
Given the relative stability of the maximum duration of benefits for regular state programs combined with the typically small share of long-term unemployed workers, legislative changes to maximum duration of regular programs should have a limited influence over time variation in total claims.[15] Relatedly, our choice to calculate claims-based unemployment rates from regular state programs is intended in part to avoid the influence of policy variation in duration stemming from standing or ad hoc temporary benefit extensions (e.g., Extended Benefits, Emergency Unemployment Compensation 2008). While less of a concern than benefit exhaustion or extension influencing the volume of continued claims, the extended Massenkoff (2021) dataset also shows minimal policy variation in waiting periods between job loss and eligibility for unemployment insurance, which could modesty affect the timing of initial claims; see Appendix B.1 for details.

Like our claims-based unemployment rate, the official IUR similarly omits long-term unemployed workers who have exhausted benefits as well as workers in jobs excluded from UI benefit eligibility and does not take a stance on search requirements. The IUR is measured as the ratio of weekly continued claims divided by covered employment, i.e., workers eligible for state or federal unemployment benefits as reported by employers. At a quarterly frequency, the rate is computed by dividing the average weekly number of continued claims in a given

[15] As an additional robustness check we also compute our claims-based unemployment rate using IC data only, which will not be impacted by changes in maximum duration policies. The IC+CC vs. IC only claims-based unemployment rates track each other very well and are plotted in Appendix Figure B.1. This strong correlation highlights the fact that even after the trough of a business cycle, new separations from employment remain elevated for a significant period of time.

quarter by "covered employment for the first four of the last six completed quarters" (ETA 2022). Our inclusion of IC along with CC will more quickly pick up on inflections in the business cycle, IC being a leading indicator, as will the contemporaneous measures of claims and employment in the denominator of equation (2). But given their conceptual similarities, the official national IUR and our national claims-based unemployment rate are, unsurprisingly, also highly correlated. Figure 3 depicts the IUR (blue) and our national claims-based unemployment rate (red). The two series are a near perfect match both in inflection points and in levels throughout the overlapping sample from over January 1971–May 2022.

Figure 3: Comparison of the Claims-based and Insured Unemployment Rates for the U.S.



Notes: The claims-based unemployment rate is smoothed with a three-month centered moving average. Sample: January 1948–May 2022.

The practical drawback to using the IUR for state-level business cycle data analysis is data availability: While the national IUR is available back to 1971, official state-level IUR data are only available back to January 1986. It would be possible to somewhat backdate the state-level IUR from 1980Q1 through the early 1960s using the primary sources we have collected.[16] But we are not aware of any primary source with monthly state-level data for 1981-85. Even if the 1980s could be backdated as well, such a dataset would, however, truncate the scope of our analysis of unemployment recoveries from ten to seven postwar recessions (to the best of our knowledge no data is consistently reported during the national recessions of 1948-49, 1953-54, or 1957-58), and would thus miss an intriguing deceleration of unemployment recoveries since these earlier recessions, discussed below.

2.4 Fitted CBUR

Given the distinctions between the official unemployment rate and our "raw" claims-based unemployment rate in mind, we also estimate an alternative "fitted" measure of state unemployment rates using a statistical model of the relationship between the two series since January 1976, the observation start date for official state-level unemployment rates. The regression framework captures the idea that a state's unemployment rate is likely higher than the national rate when that state is experiencing a higher claims-based unemployment rate relative to the national claims-based rate, and that the national unemployment rate has predictive power for state unemployment rates, particularly as pertains to the long-term unemployed, exhaustion of state benefits, and workers excluded from UI eligibility. We then use the fitted model to back-cast predicted state unemployment rates prior to January 1976.

To construct our fitted state-level unemployment series, we first estimate the relationship between official unemployment rates for each state with the following specification:

$$UR_{i,t}^{Official} = \beta_{0,i} + \beta_{1,i}(UR_{i,t}^{Claims} - UR_{US,t}^{Claims}) + \beta_{2,i}UR_{US,t}^{Official} + \varepsilon_{i,t} \quad (3)$$

where $UR_{i,t}^{Claims} - UR_{US,t}^{Claims}$ measures the difference between the state and national claims-based unemployment rates (as calculated by (1) and (2) and above) and $UR_{US,t}^{Official}$ is the

[16]The older primary sources have a state-program only breakout, e.g., Table 2 (Average Weekly Insured Unemployment) of the UIS reports has the insured unemployment rate by state only for state programs.

national unemployment rate.[17] Equation (3) is estimated on data from January 1976 through May 2022 for each individual state, and then we use these fitted models to generate unemployment rate predictions for January 1948 through December 1975; the monthly national unemployment rate only dates back to January 1948, limiting our fitted claims-based unemployment rate series by one year relative to our "raw" series.

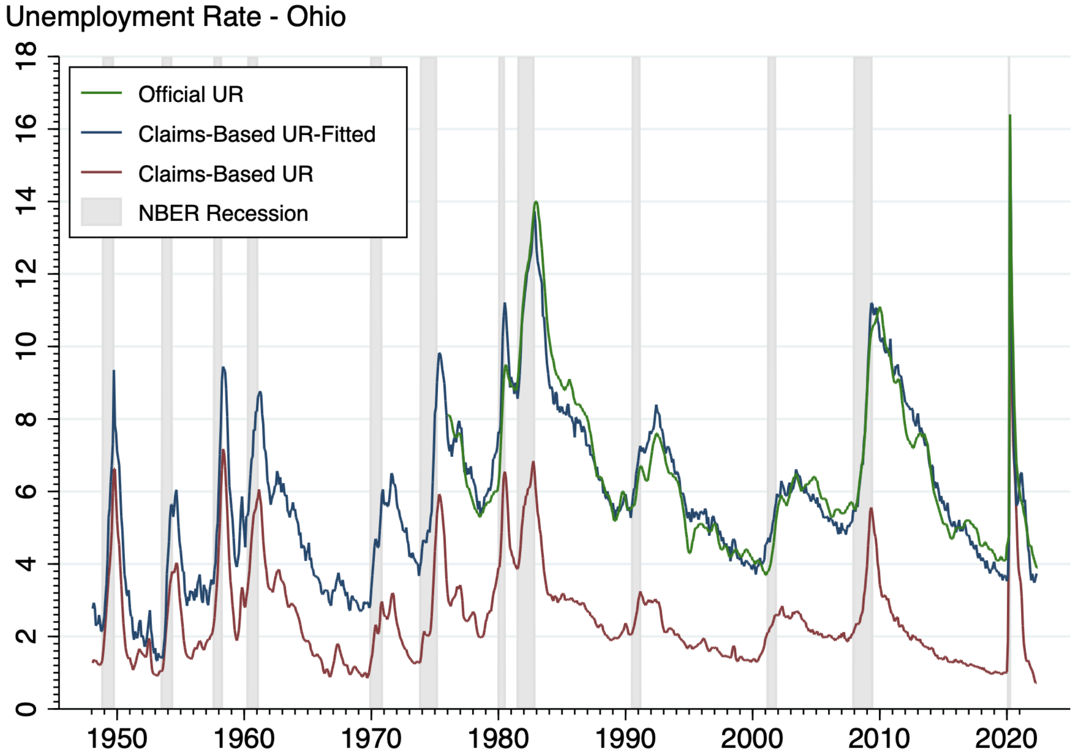
This simple statistical model fits the state-level data extremely well, and the predicted unemployment rates capture features of state business cycles that are entirely consistent with the two unemployment measures from which they are estimated. Both correlates are highly significant predictors of a state's official unemployment rate and the average R^2 is 0.83. The average correlation coefficient of the official and predicted unemployment rates is 0.91, with a maximum of 0.97 (AZ, FL, ID, IN, OH) and a minimum of 0.80 (SD). Revisiting our earlier illustrative example, Figure 4 plots our fitted claims-based unemployment rate (blue) for Ohio along with BLS's official unemployment rate (red) and the "raw" claims-based unemployment rate (green) that were plotted in Figure 1. The fitted claims-based unemployment rate picks up on inflection points in Ohio's business cycle that are nearly identical to those of the official unemployment rate over 1976–2022 and to our raw claims-based unemployment rate over the full 1948-2022 sample.

A simple way to think about this fitting exercise is that we are attempting to match the statistical model output of the BLS's state-level unemployment rate but without use of data from the CPS, although the CES nonfarm employment data is highly correlated with the CPS employment data. Again, our motive for this alternative statistical model of state-level unemployment rates is that the monthly CPS micro-data does not exist pre-1976 and more detailed CES data does not begin until 1990.

Both the raw data and our fitted data have their advantages and drawbacks. One advan-

[17] For both the state- and national-level claims data we use the three-month centered moving averages in this fitting exercise.

Figure 4: Comparison of the Official and Fitted Claims-based Unemployment Rates for Ohio



Notes: The claims-based unemployment rate is smoothed with a three-month centered moving average. Sample: January 1948–May 2022.

tage of these fitted unemployment rates is that they are anchored with the national official unemployment rate. As a result, the level differences disappear but the inflection points are virtually identical. This anchoring to the national unemployment rate also helps to smooth the fitted claims-based unemployment series data, as claims data tend to be noisier than unemployment data. One drawback of the fitted unemployment rates is the fact that out-of-sample observations are constructed on the assumption of a stable empirical relationship. And our fitting exercise modestly truncates claims-based unemployment data availability from starting in January 1947 to January 1948. We include both series in our dataset and let readers determine which is more appropriate for their uses. Encouragingly, both our raw claims-based unemployment rate and fitted unemployment rate series generate similar results when examining the pace and timing of economic recoveries, discussed below.

3 Unemployment Recoveries

With these two measures of state-level unemployment rates constructed back to January 1948, we examine various features of economic recoveries at both the national and state level, over a period spanning twelve national recessions and eleven subsequent recoveries as identified by the NBER Business Cycle Dating Committee (the 1948-49 recession onwards). We also examine state-level unemployment recoveries to date following the Covid-19 pandemic recession.[18] Measuring the pace of economic recoveries and understanding why they differ remains an active research agenda in macroeconomics, particularly as relates to the existence of a natural rate of unemployment (see, e.g., Dupraz et al. (2019) and Hall and Kudlyak (2022)) as well as concerns about so-called "jobless recoveries" from recent recessions (e.g., Galí et al. (2012)). Most of this research agenda is, however, focused on the national business cycle; our new dataset of claims-based unemployment rates allows us to explore the dispersion of recoveries across states over the last 75 years and help shed light on why national unemployment recoveries have decelerated markedly over this time.

3.1 Recession Dating

To analyze the speed and dispersion of unemployment recoveries, one must first must choose a chronology of business cycle inflection points.[19] There are various approaches to identifying peaks and troughs in the business cycle; see Romer and Romer (2019) for a historical overview. We adopt the relatively simple, unemployment-based recession dating algorithm proposed in Dupraz et al. (2019) (DNS, henceforth), which generates a close match to the

[18]The latest available claims data as of this draft is through June 2022, so our three-month centered moving average of the claims-based unemployment rates is limited to May 2022. We study recovery rates from the pandemic trough through May 2022, as there is not yet sufficient data for our recession dating algorithm to identify business cycle peaks following the pandemic recession.

[19]The common alternative to the chronology approach is estimating a Markov regime-switching model, first popularized by Hamilton (1989), in which turning points are unobserved latent variables; the model produces posterior probabilities that a given period is an inflection point, and hence probabilities that the economy is in a recession in any given period. A chronology of inflection points is far more tractable for estimating recovery speeds and comparing results with the recent literature on national unemployment recoveries.

NBER recession dates.[20] Table 1 reports national business cycle peak and trough dates using the DNS algorithm on both our claims-based U.S. unemployment rate and BLS's national unemployment rate, along with the NBER recession dates as a benchmark.

Table 1: Business Cycle Peaks and Troughs

	NBER		DNS Dating Algorithm			
			Claims-based UR		Official UR	
	Peak	Trough	Peak	Trough	Peak	Trough
1	Nov. 1948	Oct. 1949	[Feb. 1948]	Oct. 1949	[Jan. 1948]	Oct. 1949
2	[July 1953]	May 1954	Apr. 1953	Sep. 1954	May 1953	Sep. 1954
3	Aug. 1957	Apr. 1958	Dec. 1955	May 1958	Mar. 1957	July 1958
4	Apr. 1960	Feb. 1961	June 1959	Mar. 1961	Feb. 1960	May 1961
5	Dec. 1969	Nov. 1970	June 1969	Nov. 1970	Sep. 1968	Dec. 1970
6	Nov. 1973	Mar. 1975	Apr. 1973	May 1975	Oct. 1973	May 1975
7a	Jan. 1980	July 1980	Nov. 1978	July 1980	May 1979	
7b	July 1981	Nov. 1982	June 1981	Oct. 1982		Nov. 1982
8	July 1990	Mar. 1991	Nov. 1988	Mar. 1991	Mar. 1989	June 1992
9	Mar. 2001	Nov. 2001	Apr. 2000	Mar. 2002	Apr. 2000	June 2003
10	Dec. 2007	June 2009	Apr. 2006	May 2009	Oct. 2006	Oct. 2009
11	[Feb. 2020]	Apr. 2020	June 2019	May 2020	Sep. 2019	Apr. 2020

Notes: Recession dates for CBUR and UR are generated by applying the DNS algorithm on these two series. For the UR, we use the DNS parameter of 1.5. For CBUR we choose a parameter of 1.0, which is able to capture the NBER recession events. Recession dates in brackets denote some uncertainty about the precise timing of those inflection points. For the NBER recession dates, the peaks in July 1953 and February 2020 are in brackets to note that the identified quarterly peak occurred earlier, in 1953Q2 and 2019Q4, respectively. For the DNS dating algorithm, the peaks in February 1948 and January 1948 are in brackets because the DNS algorithm cannot identify those peaks due to limited data availability, as the official unemployment rate only starts in January 1948. The February 1948 "peak" for the CBUR is a result of matching the January start in the official UR and losing a month to the 3-month centered MA.

Overall, the unemployment-based recession dates generate a relatively consistent match on business cycle dates with those identified by the NBER. One notable difference between

[20]The DNS algorithm identifies peaks and troughs in the U.S. business cycle that are nearly identical to the Hall and Kudlyak (2022) chronology based on observed peaks and troughs in the unemployment rate. As an additional robustness check, we estimate state recession peaks and troughs using the Bry and Boschan (1971) algorithm (B-B, henceforth), another approach to estimating inflection points used in the literature on state and regional business cycles (e.g. Brown (2017)) that is more similar to the DNS algorithm than the Markov regime-switching approach. As with the DNS algorithm, the results of the B-B algorithm are somewhat sensitive to parameter choices, but a reasonable parameterization of the B-B algorithm generates fairly similar state-level recession dates as our preferred parameterization of the DNS algorithm.

the two unemployment-based recession dates is that the claims-based unemployment rate series identifies a double-dip recession in the early 1980s, very much in line with the July 1980–July 1981 recovery identified by NBER, but only a single, longer recession is identified from the official unemployment rate. The lack of a double-dip recession based on the official unemployment rate dates is easily understand by looking at Figure 2, which shows only a modest decline in the official unemployment series in late 1980 and early 1981 but a more pronounced dip in our claims-based unemployment series.[21] It is also generally, though not universally, the case that the claims-based unemployment rate peaks and troughs occur earlier than those generated from the official unemployment rate. One interpretation of this fact is that unemployment claims are faster to pick up changes in the labor market. Appendix Figure B.2 plots cross-correlograms for the official unemployment rate versus either the official insured unemployment rate or our claims-based unemployment rate. These figures highlight that, in addition to being highly correlated with the official unemployment rate, both claims-based indicators tend to lead the official unemployment rate—consistent with peaks and troughs being identified slightly earlier from the claims-based unemployment rate than the official unemployment rate in Table 1.

Relative to the NBER dates, the claims-based unemployment rate generates a closer match on troughs than the official unemployment rate dates. Excluding the 1980 recession (which is not identified from the official unemployment rate), the average absolute difference between the claims-based and NBER dates is 1.4 months, versus 4.3 months for the official unemployment rate dates. Both the claims-based and official unemployment rate series have a worse fit relative to NBER peak dates. The claims-based and official unemployment rate dates have an average absolute discrepancy of 11.7 and 7.8 months, respectively.

The two unemployment-based recession date series generate a relatively consistent match with one another as well. Again, the recession dates align better for troughs than peaks, with

[21] The chronology of recession dates identified by Hall and Kudlyak (2022) from the U.S. unemployment rate similarly does not identify a double-dip recession in the early 1980s.

an average absolute discrepancy of 4.1 months for troughs versus 5.8 months for peaks. Peaks tend to be identified earlier with the claims-based unemployment rate, which is unsurprising, given the use of initial claims—a leading economic indicator—in its construction, whereas the official unemployment rate is a lagging economic indicator. The worst peak match is the 1955/1957 recovery, where the peak dates are 15 months apart. The source of this dating discrepancy can be easily understood by looking at Figure 2. The official unemployment rate (plotted in blue) reaches close to its minimum for this period in 1955, but remains at a similar level and ultimately reaches its minimum in March 1957. The claims-based unemployment rate (red line) hits its minimum for this time period in December 1955, and is on a slight increasing trend into 1957. Thus, the discrepancy in peak dates amounts in part to a disagreement over when the absolute minimum value occurs for a relatively flat "peak" in the business cycle. One of the weaknesses of the DNS algorithm is a sensitivity to such flat "peaks" and "troughs". If we discard this extreme discrepancy of 15 months, the average discrepancy between the peak dates is 4.8 months, roughly in line with the average absolute discrepancy in troughs, implying roughly comparable recession recovery durations on average. In our baseline analysis below we employ the recession dates inferred from the official unemployment rate as a better cross-walk with the existing literature; Appendix B provides additional results using recession dates inferred from our claims-based unemployment rates.

3.2 National and State Economic Recovery Rates

To examine the pace of economic recoveries across at both the national and state level, we follow the general approach in Hall and Kudlyak (2022) and compute the pace of recovery as the mean decline in the log unemployment rate over the recovery period, defined as:

$$\text{Recovery Pace} = -12 \cdot (\log UR_0 - \log UR_T)/T \tag{4}$$

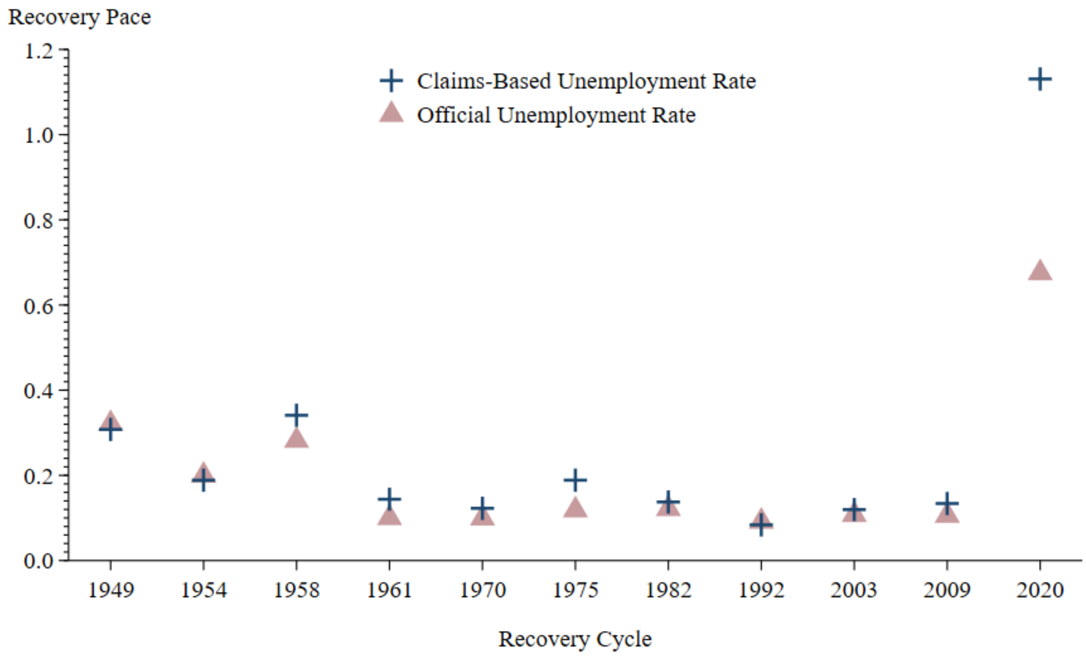
Equation (4) calculates the annualized average log point change in the unemployment rate from the start of recovery (month 0) to the end of recovery (month T). As such, these recovery rates can be interpreted as the annualized percentage decline in the unemployment rate from its peak level at the start of the recession.

Figure 5 depicts these recovery rates for both the official U.S. unemployment rate (red triangles) and our claims-based unemployment rate for the U.S. (blue crosses) across ten post-war economic recoveries. Unsurprisingly, as we employ their methodology for measuring recovery rates, Figure 5 corroborates the finding of Hall and Kudlyak (2022) that U.S. unemployment recoveries have been quite stable since the early 1960s, but the average pace of recovery has decelerated markedly since the 1950s. Encouragingly, our alternative claims-based unemployment rate generates very similar recovery rates as the official unemployment rate following the same ten post-war recessions. It is important to emphasize that the national claims-based unemployment rate is not fitted using the national unemployment rate as in equation (3)—it is simply computed from the raw claims data, see equation (2).

The only major divergence between the two recovery rates comes after the pandemic, when the claims-based unemployment series shows a much faster "recovery rate" than than official unemployment rate, as would be expected based on Figure 2. Both series see a comparable jump in March–May 2020, but during the recovery the claims-based unemployment rate quickly falls well below pre-pandemic levels (to record lows), whereas the official unemployment has yet to recover to its pre-pandemic rate. The differential degrees of recovery are amplified into even greater recovery rates by the historically short time to recovery—which has to be taken with a grain of salt, as we manually set the subsequent trough date to May 2022 based on present data availability, instead of identifying troughs with the DNS algorithm (which would likely identify a later month).

More broadly, the choice of recession dates can influence the calculation of recovery speeds, both in terms of the log point change in the unemployment rate and potentially

Figure 5: National Recovery Rates of Official and Claims-based Unemployment Rates

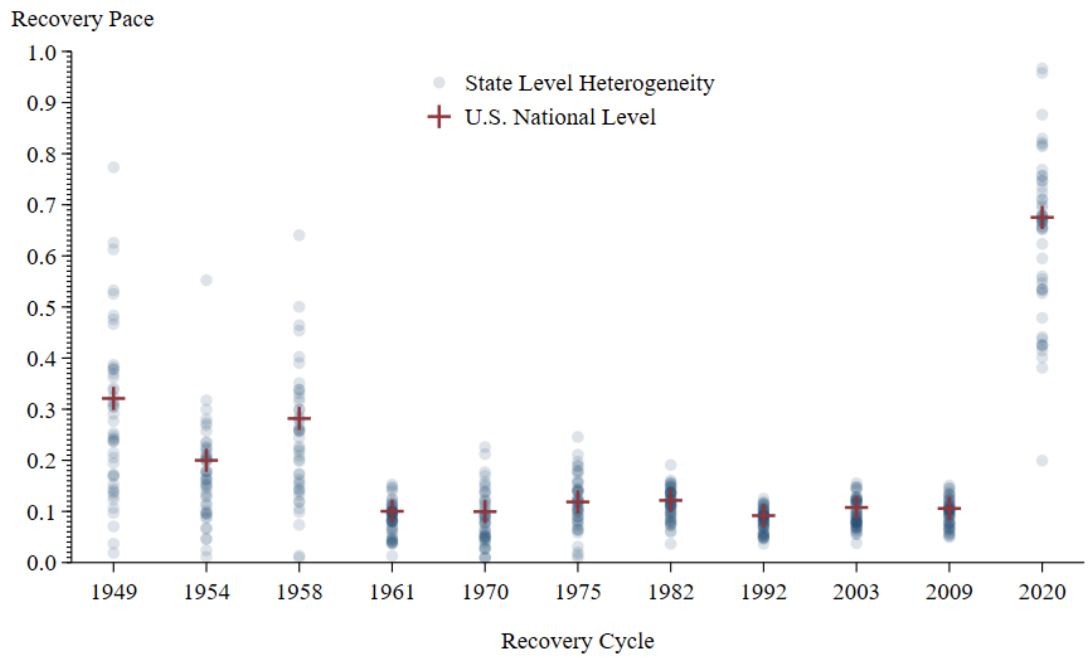


Notes: Recovery dates are estimated from the official unemployment rate using the DNS algorithm, see Table 1 for dates. Recovery from the pandemic recession is dated from trough to May 2022 based on present data availability; subsequent peaks have yet to be identified. Recovery from the 1980 recession is excluded because it is only identified for the claims-based unemployment rate series and recovery is cut short by the more severe 1981-82 double-dip recession.

in the duration of the recession as well. The recovery rates for the official unemployment rate depicted in Figure 5 differ slightly then those depicted in Figure 3 of Hall and Kudlyak (2022) for several recessions because of slight differences in the national recession dates employed. As a robustness check, Appendix Figure B.3 replicates Figure 5 using recovery dates estimated from our claims-based unemployment rates instead of the recession dates derived from the official unemployment rate. The general trends of a marked deceleration in unemployment recoveries since the 1950s and more stable, uniform recovery rates over the last 60 years hold using either set of recession dates; if anything, the deceleration in recovery rates since the 1950s is even more pronounced when using recovery dates estimated from our claims-based unemployment rates.

We next explore the pace of economic recoveries across states for the same ten recessions, using the fitted version of our state claims-based unemployment rates. Figure 6 plots the state-level recovery rates as circles along with red crosses depicting the national claims-based recovery rates (previoulsy plotted in Figure 5). One interesting feature of this data that is immediately apparent is that faster recoveries tend to be associated with much more dispersion in the pace of state-level recoveries: This was true during the faster early post-war recoveries from the recession of 1948-49, 1953-54, and 1957-58, and this dynamic remerged in the rapid recovery from the pandemic, albeit likely for different reasons discussed below.

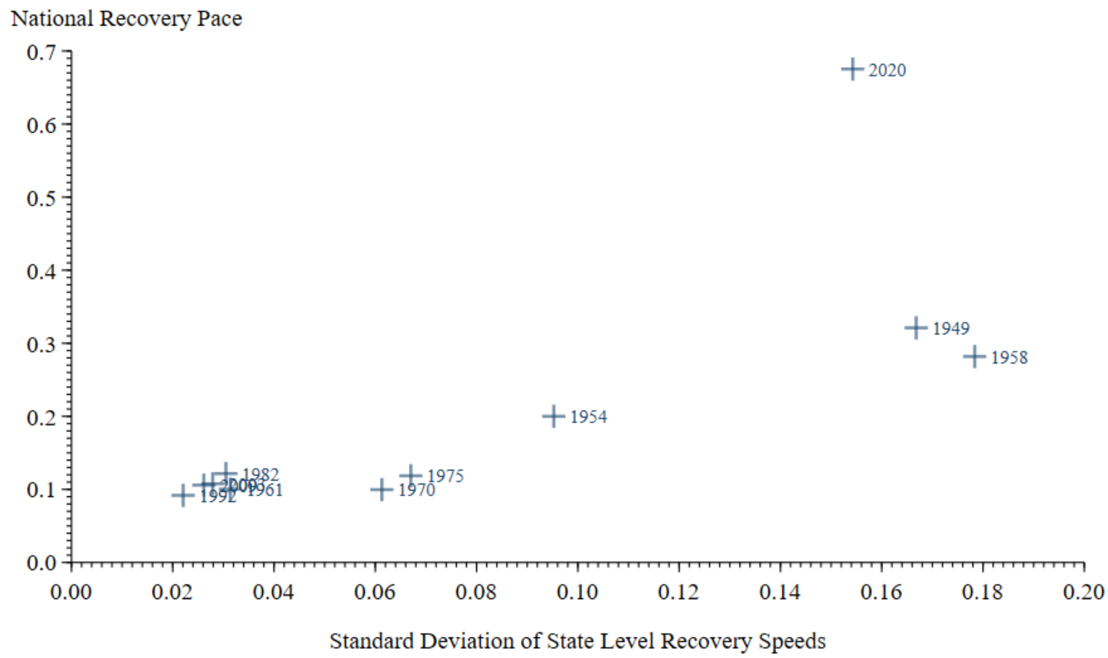
Figure 6: State-level Recovery Rates of Claims-based Unemployment Rates



Notes: Recovery dates are estimated from the official unemployment rate using the DNS algorithm, see Table 1 for dates. Recovery from the pandemic recession is dated from trough to May 2022 based on present data availability. Recovery from the 1980 recession is again excluded, see notes to Figure 5. Recovery rates are negative for a few states, i.e., their unemployment rate rose during the national recovery, but only nonnegative recovery rates are plotted.

To display this association more clearly, we plot the national recovery pace against the standard deviation of state-level recovery paces in Figure 7, which displays a clear increasing relationship. Again, faster national recoveries tend to be ones where states experience very different outcomes, and states experience rather similar outcomes during slower national recoveries throughout the 1960s–2010s.

Figure 7: Dispersion in State-level Recovery Rates and U.S. Unemployment Recovery Rates



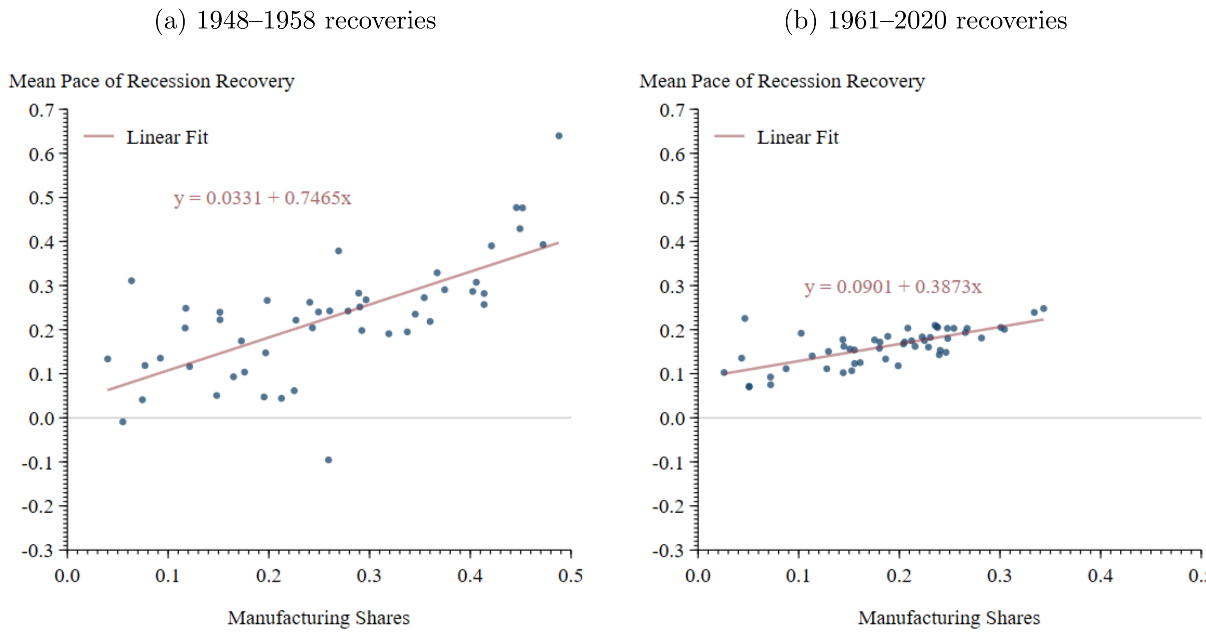
Notes: Recovery dates are estimated from the official unemployment rate using the DNS algorithm, see Table 1 for dates. Recovery from the pandemic recession is dated from trough to May 2022 based on present data availability. Recovery from the 1980 recession is again excluded, see notes to Figure 5.

There has been a significant increase in economic integration of U.S. states over this post-war horizon. The abrupt decrease in the dispersion and average pace of recovery rates depicted in Figure 6 occurs shortly after construction began on the Dwight D. Eisenhower Interstate Highway System, which was enacted by the Federal Aid Highway Act of 1956. Relatedly, our findings are consistent with general features of economic network models: Shocks spread through the system more readily if the nodes of the network have higher connectedness (e.g. Kali and Reyes (2010) and Giroud and Mueller (2019)). And shocks with a more severe impact to more nodes in a network tend to be more severe for the network as a whole (Jackson, 2010).

The more homogenous state recovery rates of the 1960s–2010s suggests the emergence of a more uniform national business cycle, but the earlier heterogeneity of state recovery rates abruptly reemerged following the pandemic, albeit seemingly for very different reasons and with a big caveat. The varied experiences with economic recovery rates in part appears to reflect states' differential exposure to sectors particularly hard-hit but the collapse of in-person services early in the pandemic. With its heavy reliance on gambling, tourism, and leisure and hospitality services, Nevada was inevitably poised for a particularly rough experience with the pandemic, and Las Vegas casinos and non-essential businesses were shut down between March–June 2020; unsurprisingly, Nevada saw the highest spike in claims-based unemployment rates (and the highest official unemployment rate of any state) followed by one of the fastest recovery rates. And some of the heterogeneity appears related to the variable timing of when states experienced waves of Covid cases and when (or if) states introduced lock downs or other public health restrictions. Whereas monetary, fiscal, and oil price shocks should impact most states at roughly the same time, the pandemic spread more slowly and variably throughout the country, partly influenced by prevailing temperatures and the scope for indoor/activity, understandably giving rise to greater heterogeneity in economic fallout from the pandemic. And as noted above, subsequent business cycle peaks are all manually set to May 2022 based on data limitations at present, which may be artificially shortening or lengthening the recovery period in some states, adding more spurious variation in recovery rates. An outlier in so many respects, it is interesting but perhaps unsurprising that data to date show a rather different experience with state unemployment recoveries following the pandemic than the clear trend that had emerged since the 1960s.

Putting the anomalous pandemic recovery aside, data on state recovery rates also allow us to explore other state-level features that might influence the evolving pace of U.S. economic recoveries, particularly the decleration of recoveries after the 1940s–1950s. We investigate one correlate in particular: The share of each state's output stemming from the manufacturing sector. Figure 8 plots states' average manufacturing share against their pace of economic recovery for two time periods: The first three "rapid" recoveries (1948–1958) and the subsequent seven "slower" recoveries (1961–2020) in the left and right panels, respectively. Both time periods show a strong positive correlation: States with larger manufacturing industries tend to experience more rapid recoveries in unemployment.[22] The strength of this relationship diminishes substantially in the latter period, but remains positive and significant.[23]

Figure 8: State-level Recovery Rates by Manufacturing Share of Output



Notes: Recovery dates are estimated from the official unemployment rate using the DNS algorithm, see Table 1 for dates. Recovery from the pandemic recession and the 1980 recession are excluded in the right panel, see notes to Figure 5.

Of course, there could be multiple mechanisms behind these correlations. One could be that manufacturing intensive states are more adversely impacted by recessions, which generates a larger UR_0 in equation (4) and thus faster recoveries. Nearly every state is in a recession during most downturns (discussed below), so for this explanation to hold, it must be that recessions in manufacturing states are more severe. Another possible mechanism is that the pace of recovery is impacted by unique features of the manufacturing industry, for

[22]Linear relationships for both time periods are significant at the 1%-level.

[23]This holds true for the most recent recoveries from the 1990-91, 2001, and 2007-09 recessions, as dated by the NBER. The positive relationship is somewhat weaker than the 1958-2009 time frame, but remains positive and significant at the 1%-level.

example, higher rates of unionization and/or the more intensive use of temporary layoffs.[24] These are questions worthy of further examination using our newly constructed dataset but are beyond the scope of this paper.

3.3 State Recession Dates and Related Robustness Checks

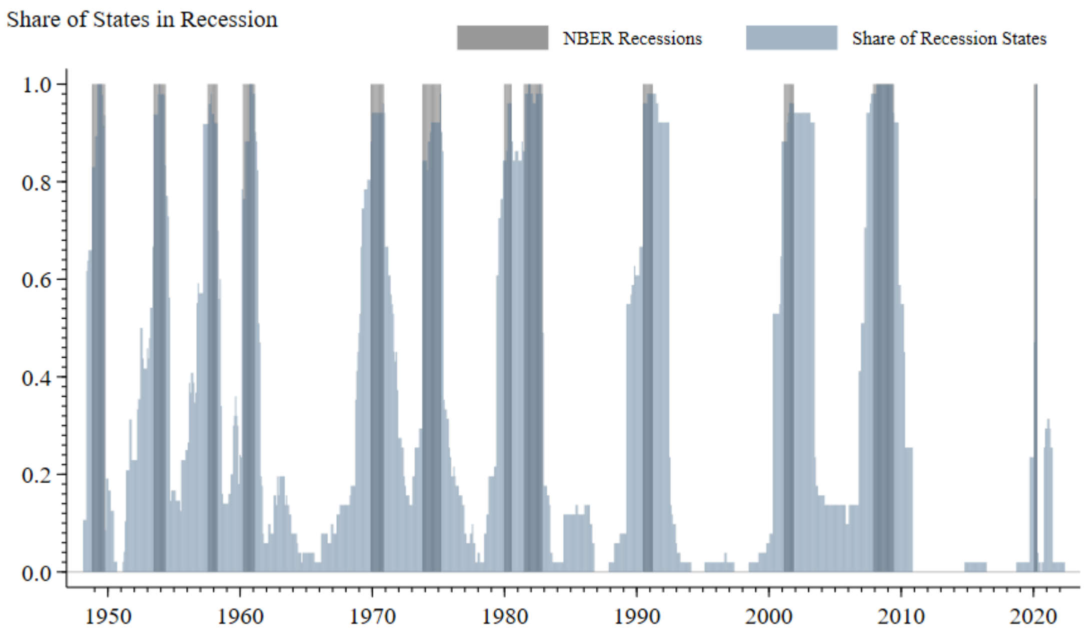
An alternative interpretation of the more disparate pace of recovery across states in the 1940s and 1950s, when the U.S. was less economically integrated, might be that some states never entered a recession and, as a result, their unemployment rates remained relatively flat (or even began rising) during national recoveries. With a flat unemployment rate, equation (4) would estimate a very slow "recovery rate," while a rising unemployment rate would generate a negative "recovery rate." To explore this question we construct recession dates at the state level to identify what share of states enter a recession during a national recession and how this share varies over time. We apply the DNS recession dating algorithm on our fitted claims-based unemployment rate for each state and then compute the share of states that are recorded as being in a recession in each month.[25]

Figure 9 plots the share of states identified as currently in a recession in each month along with national NBER recession dates (gray bars). The first noticeable feature is that the peak share of states experiencing a recession is roughly the same across the three pre-1960s recessions associated with faster recoveries and the seven subsequent recessions. It is

[24] See Lilien (1980) for evidence of high temporary layoff rates in the manufacturing sector, see Nekoei and Weber (2015) for evidence that temporary layoffs experience shorter unemployment spells, and see Gorry et al. (2020) for a theoretical discussion about the importance of permanent displacements for the propagation of unemployment shocks.

[25] DNS set the algorithm parameter "X" to be 1.5, which captures the sufficient increase in the unemployment rate to trigger a recession classification. There is an open question as to what this should be at the state-level. Because some states naturally have lower (higher) unemployment rates on average, 1.5 may under-count (over-count) recessions for these states. We compute the average ratio of a state's unemployment rate to the nation's for the entire dataset and apply this ratio to scale the X parameter for each state, denote these state-level parameters as Y_i. It is also possible that some states' unemployment rates have increased (or declined) relative to the national rate over our data span. Taking the average ratio of state and national unemployment rates over this entire period may result in a state-level DNS parameter that is too coarse to pick up recessions during periods when a state had a low unemployment rate relative to the nation. To be conservative we scale all the Y_i's down by 25% to reduce Type 2 errors in recession dating.

Figure 9: Share of U.S. States in Recession, 1948-2022



Notes: State-level recession dates are estimated from the fitted claims-based unemployment rate rate for each state using the DNS algorithm. The DNS algorithm parameter is adjusted for each state proportionate to its average level of unemployment over the entire time period. Due to data limitations in nonfarm payroll employment, not all states are included early in this sample but are added when feasible: Data for MN begins in January 1950, data for MI begins in January 1956, data for HI begins in January 1958, and data for AK begins in January 1960. Sample: January 1948–May 2022.

certainly the case that the high share of states in a recession is shorter lived in these earlier recessions relative to later ones, but these were also relatively short-lived national recessions, as measured by NBER. These results underscore that the disparate recovery rates in the pre-1960s recessions documented in Figure 6 are not simply a product of many states not being in recession. Another notable feature of Figure 9 is that a number of national "recovery" periods show numerous states remain in recession. In particular, a sizable share of states remains coded as still experiencing a recession throughout the national business cycle expansions of 1954–57, 1958–60, 1970–73, and 2001–07. Though it is hard to see visually in the figure, 100% of states are being coded as "in recession" during the 2020 pandemic recession. The secondary rise in the share of states being coded as "in recession" after the initial onset of the pandemic occurs during the severe third national Covid wave during November 2020–February 2021.

A possible concern with our state-level unemployment rate fitting approach is that the inclusion of the national unemployment rate as a regressor may cause state unemployment rates to follow the national rate too closely in the pre-1976 out-of-sample predictions. As a robustness check, we repeat our state-level recession dating exercise on the raw (unfitted) claims-based unemployment rates and again compute the share of states that are recorded as being in a recession in each month.[26] These alternative recession shares are plotted in Appendix Figure B.4. The state-level recession dating is broadly consistent with those reported in Figure 9. There are two interesting differences worth noting. The first is that the "double-dip" recession in the early 1980s is much more pronounced when using recession dates identified from the raw claims-based unemployment rates: The share of states coded as in recession falls from nearly 90% to roughly 30% between the 1980 and 1981-82 recessions. This may be a product of claims capturing more shorter-duration changes in unemployment and, as depicted in Figure 2, that the short-lived recovery from 1980-81 features prominently in the claims-based unemployment rate but is not the national UR—which is included as a regressor in for the data used to construct Figure 9. The second difference is that the national business cycle expansions of 1954–57 and 1958–60 have slightly fewer states being coded as still being in a recession relative to Figure 9. But in general, both the fitted and raw claims-based unemployment rates appear to capture consistent inflection points and features of both state and national business cycles.

Existing estimates of state recession dates offer a natural validation exercise for the state-

[26]Again, there is a question of how to set the parameter "X" for the DNS algorithm. For the national (unfitted) claims-based unemployment rate we set $X = 1.0$, which generates a good match with NBER recession dates (reported in Figure 1). For states, we compute the ratio of the state-level and national claims-based unemployment rates and scale each state's "X" parameter accordingly, and again scale these down by 25% to be conservative.

level recession dates derived from our claims-based unemployment rates. Owyang et al. (2005) study state and regional business cycles, using a Markov regime-switching model to estimate state recession probabilities from the state coincident indexes of Crone and Clayton-Matthews (2005). Owyang et al. (2005) produce estimates of state recession probabilities for February 1979–June 2002, a sample limited by the availability of state coincident indexes, which in turn are limited by the unavailability of official state unemployment data before January 1976.[27] We use the Owyang et al. (2005) state recession probabilities to benchmark our state recession dates estimated using the DNS algorithm on our fitted claims-based unemployment rates. Appendix Figure B.5 depicts our claims-based unemployment rates (blue lines), state recession dates (gray bars), and the Owyang et al. (2005) state recession probabilities (red lines) for all 50 states.

As a general matter the crosswalk suggests that our claims-based unemployment rates identify similar business cycle dynamics for most states, particularly larger ones, during the overlapping 1979–2002 sample. The similarities and differences between our state-level recession dates and the Owyang et al. (2005) recession probabilities are discussed in more detail in Appendix B.2. Some differences are to be expected. Markov-switching models and the DNS algorithm identify related yet fundamentally different objects, and state-level coincident indexes are a related but broader measure of economic activity than state unemployment rates, our exclusive focus in identifying recession dates.[28] Neither approach is right or wrong per se. But Appendix Figure B.5 underscores a drawback of using the Markov regime-switching approach specifically for studying recovery rates: Our recession dates exhibit fewer erratic, short-lived recessionary spikes or dubiously long recessionary periods, and no judgement is

[27] The Federal Reserve Bank of Philadelphia produces up-to-date monthly state coincident indexes using the model of Crone and Clayton-Matthews (2005), but data are similarly only available starting in January 1979 or later. The coincident indexes are estimated from four state-level variables: Nonfarm payroll employment, average hours worked of production workers in in manufacturing, the official state unemployment rate, and real wage and salary disbursements.

[28] Brown (2017) compares the recession dates generated by a Markov regime-switching model and the B-B algorithm on coincident indexes for states in the Tenth Federal Reserve District, and finds the two models generally identify the same recessions. The regime-switching model tends to identify peaks slightly later.

required regarding a cutoff for recession probabilities to identify recovery dates and durations. The principal advantage to our approach, however, is the ability to identify inflection points in state business cycles for more than 30 additional years when using our claims-based unemployment rate series instead of existing off-the-shelf state coincident indexes. It would be possible to construct backdated coincident indexes using our dataset and estimate state recession probabilities over a longer horizon, but we leave that for future research.

4 Conclusion

In this paper we introduce a new state-level unemployment dataset spanning 1947–2022, a dataset constructed from historical unemployment claims data that we digitized from a series of primary sources and then merged with existing state-level data for 1971 onwards. As BLS's official state unemployment rates only begin in January 1976, our novel dataset represents a sizable expansion of panel data availability for measuring labor market slack, offering practitioners nearly three additional decades of seasonally adjusted monthly state-level data. We construct a "raw" claims-based unemployment rate series going back to January 1947, which is conceptually similar to the official unemployment rate but uses initial and continued unemployment insurance claims as its measure of unemployed workers. We also construct alternative "fitted" unemployment series going back to January 1948, estimating state unemployment rates from a statistical model of the dynamics between our claims-based unemployment and official unemployment rate series. Both claims-based unemployment rate series capture similar state and national business cycle dynamics as omitted by existing data sources for overlapping samples.

Our claims-based unemployment series doubles the number of post-war U.S. recessions that can be studied at the state level. We use our dataset to shed light on the deceleration in unemployment recoveries documented by Hall and Kudlyak (2022) at the national level, by examining state-level labor market recoveries across ten postwar U.S. recessions from 1948 through 2007–09, as well as the recovery to date from the pandemic recession. Consistent with the features of many economic network models, we find that slower national recovery rates since the 1960s—when the U.S. economy was far more interconnected—have been associated with a more uniform pace of recovery across states, whereas faster national labor market recoveries in the 1940s and 1950s were associated with a greater dispersion of recovery rates across states. We also find that states with greater manufacturing shares of economic output have tended to see faster recoveries, particularly in the 1940s and 1950s, and to a smaller but nonetheless significant extent since the 1960s as well. Our results suggest that both the shift from manufacturing to services and increased regional economic integration have contributed to U.S. economic recovery rates slowing and state recovery rates converging over the last 60 years, the rapid and heterogenous rates of recovery across states following the pandemic notwithstanding.

Our dataset would allow for more comprehensive analysis of the evolving network of economic integration across U.S. states or, more broadly, regional business cycles over a longer time horizon than has typically been studied to date—a research agenda previously constrained by the unavailability of state unemployment rate data prior to 1976. One related avenue for future research would be constructing backdated state coincident indexes using our historical claims-based unemployment rates.[29] More broadly, we hope our historical dataset of claims-based unemployment rates, derivative recession dates, and the underlying digitized monthly unemployment claims prove useful controls or outcome variables of interest for empirical macroeconomic studies using U.S. state-level panel data.

[29] Nonfarm payroll employment and wage and salary disbursements, two of the three other inputs used in the latent factor model of Crone and Clayton-Matthews (2005) and the Federal Reserve Bank of Philadelphia, are also available at the state level back to 1948Q1.

References

Bernhardt, R., D. Munro, and E. Wolcott (2021): "How Does the Dramatic Rise of CPS Non-Response Impact Labor Market Indicators?" Tech. rep., GLO Discussion Paper.

Blanchard, O. J. and L. F. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 23, 1–75.

Brown, J. P. (2017): "Identifying State-Level Recessions," *Economic Review*, 102, 85–108.

Bry, G. and C. Boschan (1971): "Programmed Selection of Cyclical Turning Points," Tech. rep., National Bureau of Economic Research.

Chodorow-Reich, G. (2019): "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" *American Economic Journal: Economic Policy*, 11, 1–34.

Crone, T. M. and A. Clayton-Matthews (2005): "Consistent Economic Indexes for the 50 States," *The Review of Economics and Statistics*, 87, 593–603.

Dupraz, S., E. Nakamura, and J. Steinsson (2019): "A plucking model of business cycles," Tech. rep., National Bureau of Economic Research.

Galí, J., F. Smets, and R. Wouters (2012): "Slow Recoveries: A Structural Interpretation," *Journal of Money, Credit and Banking*, 44, 9–30.

Giroud, X. and H. M. Mueller (2019): "Firms' internal networks and local economic shocks," *American Economic Review*, 109, 3617–49.

Glandon, P. J., K. Kuttner, S. Mazumder, and C. Stroup (2022): "Macroeconomic Research, Present and Past," Tech. rep., National Bureau of Economic Research.

Gorry, A., D. Munro, and C. vom Lehn (2020): "Experience, Skill Composition, and the Persistence of Unemployment Fluctuations," *Labour Economics*, 63, 101793.

Hall, R. E. and M. Kudlyak (2022): "The Inexorable Recoveries of US Unemployment," *Journal of Monetary Economics*.

Hamilton, J. D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

Hamilton, J. D. and M. T. Owyang (2012): "The Propagation of Regional Recessions," *The Review of Economics and Statistics*, 94, 935–947.

Hazell, J., J. Herreño, E. Nakamura, and J. Steinsson (2022): "The Slope of the Phillips Curve: Evidence from U.S. States," *The Quarterly Journal of Economics*, 137, 1299–1344.

Jackson, M. O. (2010): *Social and economic networks*, Princeton university press.

Kali, R. and J. Reyes (2010): "Financial contagion on the international trade network," *Economic Inquiry*, 48, 1072–1101.

Kuka, E. and B. A. Stuart (2021): "Racial Inequality in Unemployment Insurance Receipt and Take-Up," Tech. rep., National Bureau of Economic Research.

Lilien, D. M. (1980): "The cyclical pattern of temporary layoffs in United States manufacturing," *The Review of Economics and Statistics*, 24–31.

Massenkoff, M. (2021): "State Unemployment Insurance Laws," http://maximmassenkoff.com/data.html, accessed: 2021-10-25.

Nakamura, E. and J. Steinsson (2014): "Fiscal Stimulus in a Monetary Union: Evidence from US Regions," *American Economic Review*, 104, 753–792.

Nekoei, A. and A. Weber (2015): "Recall expectations and duration dependence," *American Economic Review*, 105, 142–46.

Owyang, M. T., J. Piger, and H. J. Wall (2005): "Business Cycle Phases in the U.S. States," *The Review of Economics and Statistics*, 87, 604–616.

Romer, C. D. and D. H. Romer (2019): "NBER Business Cycle Dating: History and Prospect," *University of California, Berkeley.*

Appendix A Data Appendix

A.1 Data Construction

Monthly state-level unemployment claims are available in digital form through the DOL's website beginning in January 1971. We backdate this dataset from scanned versions of a series of earlier periodical government agency reports, digitizing monthly data for regular state program IC and CC back to December 1946 for all 50 states and Washington, D.C.[30] Our preferred specification for the claims-based unemployment series is a three-month centered moving average, so the start of the sample was chosen to make that series available starting in January 1947. Data for December 1946–October 1949 are digitized from the *Employment Securities Activities* (ESA) report published monthly by the Bureau of Employment Security of the Federal Security Agency Social Security Board.[31] Data for November 1949–October 1963 are digitized from the *Labor Market and Employment Security* (LMES) report published monthly by the Bureau of Employment Security of the U.S. Department of Labor. And data for November 1963 onward are digitized from the *Unemployment Insurance Statistics* (UIS) report published bimonthly by the Bureau of Employment Security of the U.S. Department of Labor Manpower Administration. Digitized scans of all issues of the ESA reports and most issues of the LMES and UIS reports were available through HathiTrust. We supplemented missing monthly reports with Interlibrary Loan Request or Google Book scans and

[30] Fortunately data for Alaska and Hawaii are consistently available before they become states.

[31] ESA reports are available at least as far back as January 1943, so it would be feasible to backdate these historical claims series at least as far March 1943.

data from the Unemployment Insurance Review (UIR) reports published bimonthly by the Bureau of Employment Security of the U.S. Department of Labor Manpower Administration.

Overall, the image quality of the scans we were able to locate was generally quite good and data revisions appeared to be a minimal complication. In some cases we retrieved alternative copies of reports scanned by a different library to resolve uncertainties relating to legibility. We always used reported data on national aggregates to cross-check the sum of state and territory claims against total U.S. claims. The LMES and UIS data tables typically report the percentage change of IC and CC from the prior month alongside the number of claims, and as another cross-check we calculated the corresponding percentage change from the prior month using our digitized data. In cases where image quality presented serious legibility issues or a handful of observations for which data was missing we used reported data on monthly or annual changes in claims to guide our data digitization or impute missing observations.[32] Data was always digitized from the most recently published report available if multiple sources reported claims data for a certain month; data revisions seemed to be more of an issue for the earliest ESA reports than the LMES and UIS reports, but luckily later reports had multiyear tables with revised claims data for most of the observations we digitized from the ESA reports (for September 1947–October 1949).

To construct a complete time series for 1946–2022, the data we digitized from these primary sources were merged with monthly data already digitized and available online from DOL. To be as consistent as possible with data definitions, the more recent data pulled from DOL were always restricted to IC and CC data from regular state programs only, excluding the federal Extended Benefits (EB) program, which was enacted in August 1970;

[32] For instance, claims data are missing for Rhode Island in September 1971, and a footnote in the UIS reports flagged "Data not available" for that state. But RI claims data are reported in the subsequent report for October 1971 along with the percentage change from September 1971, enabling us to impute the missing observation for September 1971 fairly accurately. Similarly, claims data are missing for a handful of states from the ESA reports for 1947, but we were able to fill in all missing observations using data on the year-over-year change in the number of claims reported for 1948 coupled with the actual number of claims reported for 1948.

state-level EB data is only available from DOL for 1986 onwards and almost all of the newly digitized data predates the permanent federal EB program, rendering ours the most consistent data definition. We digitized claims data from the UIS reports through December 1972 to investigate how well our newly digitized data lined up with the existing DOL data, which starts in January 1971. Encouragingly, initial claims data for 1971–72 line up almost perfectly between the DOL data and that of the UIS reports: Only two of the more than 1,200 observations showed any discrepancy, and both were minor.[33] Given the seamless integration of the IC series, we merged our IC data digitized from the various primary sources for December 1946–December 1970 into the DOL data for January 1971–June 2022.

In a potential complication with this merge, the continuing claims data digitized from the UIS reports line up perfectly with the DOL data over 1971-1972 for some states (e.g., CT, DE, MS, MT, PA, and OH) but are significantly higher in the UIS reports than the DOL data for certain other states (e.g., CA, KY, MI, MN, NE, NJ, OK, and VA) and are just slightly (less than 2%) off for many other states (e.g., AL, DE, KS, LA, MA, ME, ND, NH, NM, NV, NY, SD, SC, UT, VT, WI).[34] The state-specific discrepancies between the UIS reports and DOL data could not be explained by certain states triggering EB, geographical regions, or political orientation. But encouragingly, all state-specific discrepancies between the two CC series disappear over a slightly longer horizon, by mid-1977 if not earlier.[35] As such, we extended our digitization of CC from the UIS reports through June 1977 and compared the two series. In almost all cases the two CC series seem to be off by a fairly stable level

[33]For Louisiana in April 1971, the UIS reports reported 17,289 claims whereas the DOL data online showed 17,290 claims. And for Utah in August 1971, the UIS reports reported 6,026 claims whereas the DOL data online showed 6,006 claims. Both data discrepancies were off by less than 0.5%.

[34]Save the following three exceptions, UIS data for CC were consistently greater than or equal to the DOL data available online: The UIS reports showed 3,907 (-3.2%) fewer claims for FL in May 1972, 48 (-0.05%) fewer claims for LA in June 1972, and 33 (-0.03%) fewer claims for LA in July 1972 than the DOL data available online.

[35]There is one later CC discrepancy between the UIS reports and DOL data available online for RI in April-June 1978. Rhode Island exhibited frequent reporting problems in the UIS reports during the 1970s, and the percentage change from June 1978 to July 1978 suggests that the previously reported UIS data is incorrect and the DOL data available online is accurate. Merging the UIS data into the DOL data online in mid-1977 obviates this particular data issue with the UIS reports for RI.

effect—perhaps suggesting a persistent misunderstanding of data reporting requirements at certain state UI offices—but capture similar business cycle fluctuations. For most states with discrepancies between the two continuing claims series, the CC data for 1971–77 digitized from the UIS reports looks less disjoint than the data from DOL (e.g., AZ, CA, CT, DC, FL, KY, MN, NE, NJ, VA, and WA). And in a few states the level of continuing claims in the DOL data seems suspiciously lower than all other observations in surrounding decades (e.g., CA, KY, and WV). Outliers were also a more frequent cause for concern in the existing DOL data than the newly digitized CC data for 1971–77 (discussed below). As such, we take use the CC data digitized from the UIS reports and preceding primary sources for December 1946–June 1977 as our preferred benchmark data specification, which is then merged into the DOL data for July 1977–June 2022.

Neither the newly digitized historical claims data nor the DOL data were seasonally adjusted. We seasonally adjusted the monthly IC and CC data for regular state programs for the full 1946–2022 sample using the U.S. Census Bureau's X-13 ARIMA-SEATS seasonal adjustment software. The unprecedented spike in initial claims starting in March 2020 throws off the seasonal adjustment factors in the lead up to the pandemic. We separately seasonally adjust data for December 1946–February 2020 to avoid this confounding influence, and then splice in data for March 2020–June 2022 from a separate seasonal adjustment of all data for December 1946–June 2022. We also ran tests for outliers using the Win-X13 program, which identified roughly 200 potential additive outliers (AO) and temporary changes (TC) outliers from approximately 91,000 observations (newly digitized historical claims data and existing data combined). These outliers were roughly evenly distributed between our newly digitized data and the existing DOL data. We manually checked each potential outlier to determine if it represented a legitimate change in claims due to plausible or exigent economic circumstances (e.g. a surge in IC in Louisiana and Mississippi in September 2005 as a result of Hurricane Katrina) or a "fat thumb" data coding issue. We used several verification processes. The first was to double check the digitized data against primary source reports when available (to the best of our knowledge, historical claims data is only available through March 1980, when the UIS reports stopped being published). The second was to leverage the relationship between IC and CC, which should move in the same direction contemporaneously or with a one-month lag. For example, a spike in CC, without a concurrent or preceding spike in IC would suggest a data coding issue. And finally, we also examined nonfarm employment data to determine if there was a contemporaneous change in another labor market indicator, reflecting a legitimate change in labor market conditions.

Fat thumb coding issues were relatively rare but can be quite striking and misleading. As an extreme example of a data coding issue identified in the DOL online data, CC in Missouri in June of 1974 surged 4700% from 147,351 to 7,132,843, then collapsed again the following month to 145,365. There is no contemporaneous or lagged surge in IC. And this particular outlier is entirely implausible, as the population of Missouri was less than 5 million in 1974. This is a case in which we believe the first '7' is a typo and the observation should read '132,843,' which is in line with continuing claims data for the prior and subsequent months (147,351 and 145,365, respectively). It is worth noting that the U.S. total for CC in June of 1974 in the DOL data appears to be calculated as the sum of claims for states and territories, and was also flagged as a likely outlier. The U.S. total for CC of 12,910,365 is similarly well above CC data for the prior and subsequent months (roughly 82% higher than 7,110,210 and 7,222,162, respectively), and this is surely a related fat thumb error by aggregation. In the handful of cases thought to reflect "fat thumb" coding errors we replaced this seemingly spurious data with data observations from primary sources, adjusted the first digit when a monthly observation was off by an order of magnitude, or, if necessary, used a linear interpolation between CC data for the prior and subsequent month.[36] The following "fat thumb" outliers were identified and manually adjusted as follows:

[36] Linear interpolation was only needed for adjusting CC in the DOL data when the related UIS primary sources reflected regular state program claims as well as EB, and the UIS dynamics across the current, preceding, and subsequent month were mapped into the DOL data (state programs only, excluding EB) using observations for the preceding and subsequent month.

- DE May 1974 CC: 42,850 to 24,850 (UIS report reads "24,850" not "42,850")
- DE June 1981 CC: 6,433 to 36,433 (off by an order of magnitude)
- FL March 1972 CC: 74,478 to 143,979 (UIS report reads "148,845" not "74,478')
- KY February 1974 CC: 1,218,070 to 121,807 (off by an order of magnitude)
- MA January 1978 IC: 53,954 to 50,829 (UIS report reads "50,829" not "53,954")
- MA February 1978 IC: 90,507 to 86,580 (UIS report reads "86,580" not "90,507")
- MI February 1973 CC: 546,984 to 255,264 (UIS report reads "413,526" not "546,984")
- MO February 1974 CC: 31,088 to 201,743 (off by an order of magnitude)
- MO June 1974 CC: 7,132,843 to 132,843 (off by an order of magnitude)
- NY September 1973 CC: 76,674 to 642,675 (UIS report reads "671,981" not "76,674")
- NY August 1977 CC: 1,762,353 to 1,162,353 UIS report (reads "1,162,353" not "1,762,353")
- RI May 1984 CC: 6,796 to 56,796 (off by an order of magnitude)

In addition to adjusting these fat thumb issues, we use monthly data on average weekly insured unemployment (AWIU) to interpolate data for Illinois in March–April 1977 and for Michigan in April–May 1977, around the merge of the digitized UIS data into the DOL data available online. The UIS data for IL April 1977 was flagged as an outlier, just as the UIS data series lining up with DOL data available online in April–May 1977. The UIS data for IL is consistently higher than the DOL data available online before the merge, but then spike erratically in March 1977 and crater implausibly in April 1977, before aligning at reasonable levels in May 1977. These movements in CC for March–April 1977 do not align with movements in the corresponding IC or AWIU data for IL. The UIS data for MI in April 1977 was also flagged as an outlier before the two series align perfectly starting in

June 1977. Unlike the rest of of the UIS data for MI, which are consistently higher than the DOL data available online, the April and May readings in the UIS reports are much lower. The June 1977 report shows the reading of 498,892 (the same as the DOL data online) having fallen 10.8%, which would put the May reading around 559,296, instead of 317,385, as reported in the UIS reports. Again, these movements in CC for April–May 1977 do not align with movements in the corresponding IC or AWIU data for MI. To interpolate CC in these two cases, we use the UIS reports to calculate the average ratio of CC to AWIU across the two previous and two subsequent months relative to the two months in question, and then multiply the average ratio by AWIU in each of the months in question to back out an estimate of CC. The interpolated CC data for IL and MI are far more consistent with IC and AWIU dynamics throughout 1977.

Our judgement calls about data adjustments will modestly affect the raw claims-based unemployment series. However, data adjustments for 1976 onwards—after BLS's official state unemployment rates are available—will have a negligible effect on our fitted out-of-sample state unemployment rates for January 1948–December 1975. Our fitting exercise will wash these things out in-sample insofar as they are erroneous data. When estimating equation (3) over January 1976–May 2022, erroneous data entering the claims-based unemployment rates on the right-hand side will only show up in the error term. Moreover, the official state unemployment rate being estimated on the left-hand side is also constructed in part from state unemployment insurance claims data subject to similar or identical fat thumb data coding issues.

After manually correcting these handful of fat thumb outliers we re-ran the seasonal adjustment (without hard coding for outliers) for the monthly IC and CC data over December 1946–June 2022, again separately seasonally adjusting data for December 1946–February 2020 to avoid the confounding influence of the pandemic spike in claims on seasonal factors. The seasonally adjusted time series for total U.S. regular state program claims is constructed by summing the seasonally adjusted series for all 50 states plus Washington, D.C., as opposed to seasonally adjusting total U.S. claims.

A.2 Data Validation and Robustness Checks

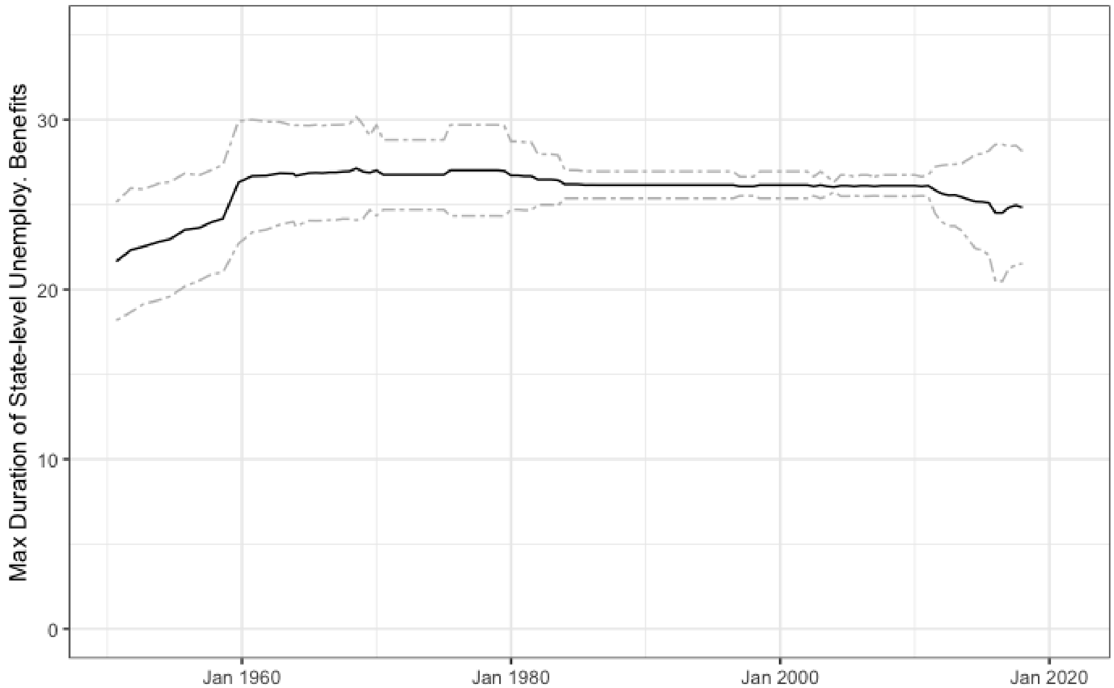
The principle advantage to using unemployment claims as a proxy for the level of unemployment in constructing our claims-based unemployment rates is historical data availability: We exploit a consistently defined monthly measure of unemployment activity at the state level that dates back to the early 1940s. But conceptual differences between the surveyed level of unemployment and the number of unemployment insurance claimants do raise several potential concerns, discussed below.

As noted in Section 2.2, our claims-based unemployment rate omits long-term unemployed workers who have exhausted benefits, just as the official IUR does. Such benefit exhaustion would only pose a serious challenge to our use of claims-based state unemployment rates in studying state business cycles if there was considerable policy variation in maximum benefit durations inducing large changes in the number of unemployed workers receiving benefits, which is not the case. Figure A.1 depicts the mean of the maximum duration of regular state benefits for all 50 states with one standard deviation bands; there is relatively little variation in maximum durations across states in a given year or time.

Similarly, there is very limited policy variation stemming from "waiting periods" or a "waiting week" before benefit eligibility. Since the mid-1950s, all U.S. states have implemented either a one-week waiting period or no waiting period requirement. A handful of states implemented a two-week waiting period at the start of our sample, but these were universally phased out by the late 1940s or early 1950s.[37] Twenty four states never changed their waiting period policies throughout our sample, with a plurality of states consistently

[37] The following states had a two-week waiting periods in the late 1940s: CO, GA, MN, MS, MT, NE, OH, WI, and WY. Colorado and Montana were the last states to still require a two-week waiting period, both of which were reduced to a one-week requirement between 1954 and 1955.

Figure A.1: Maximum Duration of Regular State Unemployment Benefits, 1950–2018



Notes: This figure reports the mean of each state's maximum benefit duration for regular UI programs in the solid black line, along with one-standard-deviation bands reported in the gray dashed lines. January 1950–December 2018. Data source: Massenkoff (2021) and DOL ETA.

imposing a one-week waiting period.[38] Eight states changed their waiting period policy once, eleven states changed their policy twice, and five states changed their policy three times. Only North Carolin and Wisconsin have changed waiting period requirements more than three times over this sample. There were only 58 waiting period policy changes over 1948–2018, just 1.6% of the 3,550 state-year observations.

Figure A.2 depicts the share of unemployed workers who have been out of work for 27 weeks or longer, and would thus have exhausted regular state benefits for most of our data sample. With the notable exception of the Great Recession, the long-term unemployed typically only account for 5% to 25% of unemployed workers. Moreover, excluding the long-term unemployed has very little affect on unemployment dynamics and inflection points at

[38]The following states had a one-week waiting periods throughout the entire 1947–2018 sample: AK, AR, AZ, CA, FL, HI, ID, IL, IN, KS, LA, MO, ND, NM, NY, OK, OR, RI, SD, TN, UT, WA, and WV. Maryland never had a waiting period requirement over this sample.

the national level. Over January 1948–December 2019, the correlation between the log level of unemployed workers and the log level of unemployed workers who have been out of work for 26 weeks or fewer is 0.98.

Figure A.2: Long-term Unemployment as a Share of Total Unemployment

Notes: This figure reports the share of unemployed workers who have been unemployed for 27 weeks or longer relative to all unemployed workers. Sample: January 1948–December 2019. Data source: BLS.

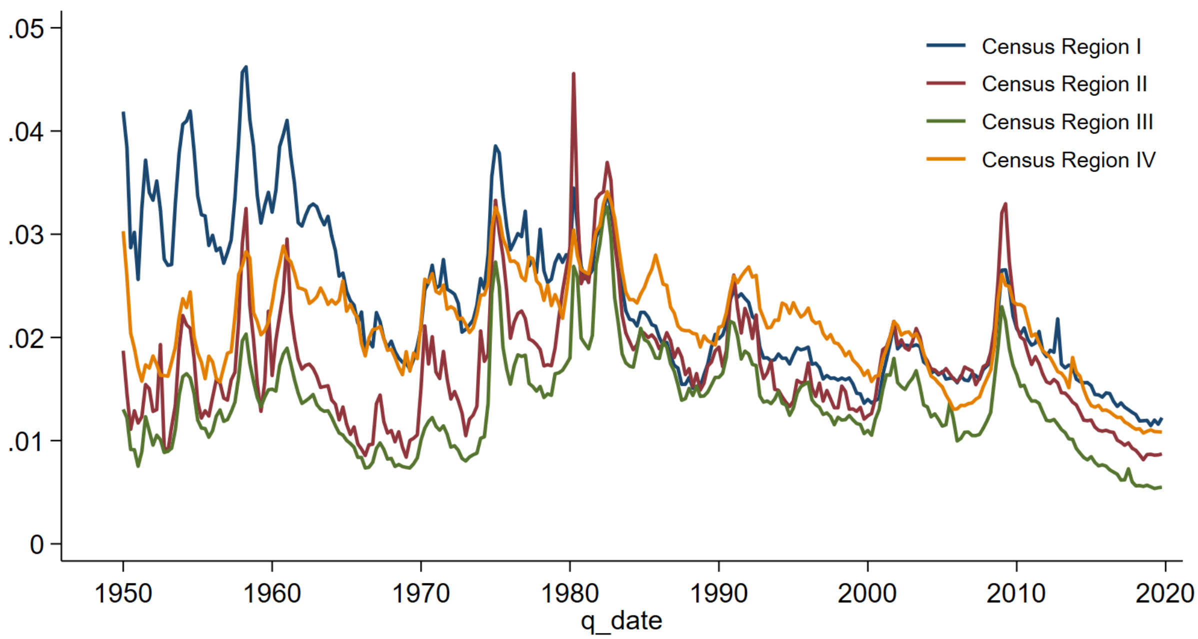
Given the relative stability of the maximum duration of benefits for regular state programs combined with the typically small share of the unemployed who are out of work for 27 weeks or longer, legislative changes to maximum duration should have a limited influence over time variation in continued claims.

Another potential concern with using unemployment claims as a proxy for unemployment relates to time-varying take-up rates in state unemployment programs or denials of unemployment claims. Slow-moving changes in take-up rates and/or denial rates that are uniform across the country pose little threat to our empirical exercise, as they would resemble secular drift in trend unemployment without a first-order effect on unemployment recovery speeds or peaks and troughs identified by the DNS algorithm. In the fitted claims-based unemployment rates, any uniform national effect will be differenced out in the term $(UR_{i,t}^{Claims} - UR_{US,t}^{Claims})$ of equation (3), and any residual level effect would be corrected for with the inclusion of the national unemployment rate on the right-hand side.

More abrupt changes in take-up rates and/or denial rates in only a subset of states would, however, potentially undermine inference from our claims-based unemployment rates. For instance, to the extent that racial discrimination affected take-up rates or denial rates differentially across regions, the Civil Rights Act of 1964 and federally enforced desegregation in the southern United States could have induced divergent trends in state take-up and denial rates across states. Unfortunately, the LMES and UIS reports rarely report claims by race, and even data on claims or denials by race could fail to capture the effects of racial discrimination dissuading applications and take up in the first place. Differential trends in unemployment insurance take-up and denials across race do not, however, appear pronounced in recent decades. Kuka and Stuart (2021) find that racial take-up gaps in unemployment insurance are relatively stable over 1986-2015, which the authors interpret as suggesting that take-up gaps "are explained by persistent economic or social factors." While there is a significant gap between UI take-up and receipt for white and black workers, Kuka and Stuart (2021) find that observed characteristics can explain 66% of the gap in take up and 81% of the gap in benefit receipt. They also find that fixed effects for the South have considerable predictive power for explaining racial UI gaps, whereas other regions don't have much explanatory power; the authors explain that "UI receipt and take-up is lower in the South, where unemployed Black individuals are much more likely to live."

As an additional empirical test predating their sample of study, we examine regular state program initial unemployment insurance claims per capita by Census region, which are plotted in Figure A.3 for 1948–2019. Reassuringly for our claims-based unemployment rates, IC per capita in the South (Region III, depicted in green) behave relatively similarly across the entire sample: They are consistently lower than IC per capita in the other three Census regions, they roughly follow the same inflection points as the other Census regions, and there is no discernible break in these dynamics following the passage of the Civil Rights Act. Interestingly, there is a great deal of co-movement in IC per capita across the four Census regions throughout this entire sample in spite of well documented differences in regional business cycles (Hamilton and Owyang, 2012) and the collapse in heterogeneous state-level recovery rates documented above.

Figure A.3: Initial Claims Per Capita by U.S. Census Region

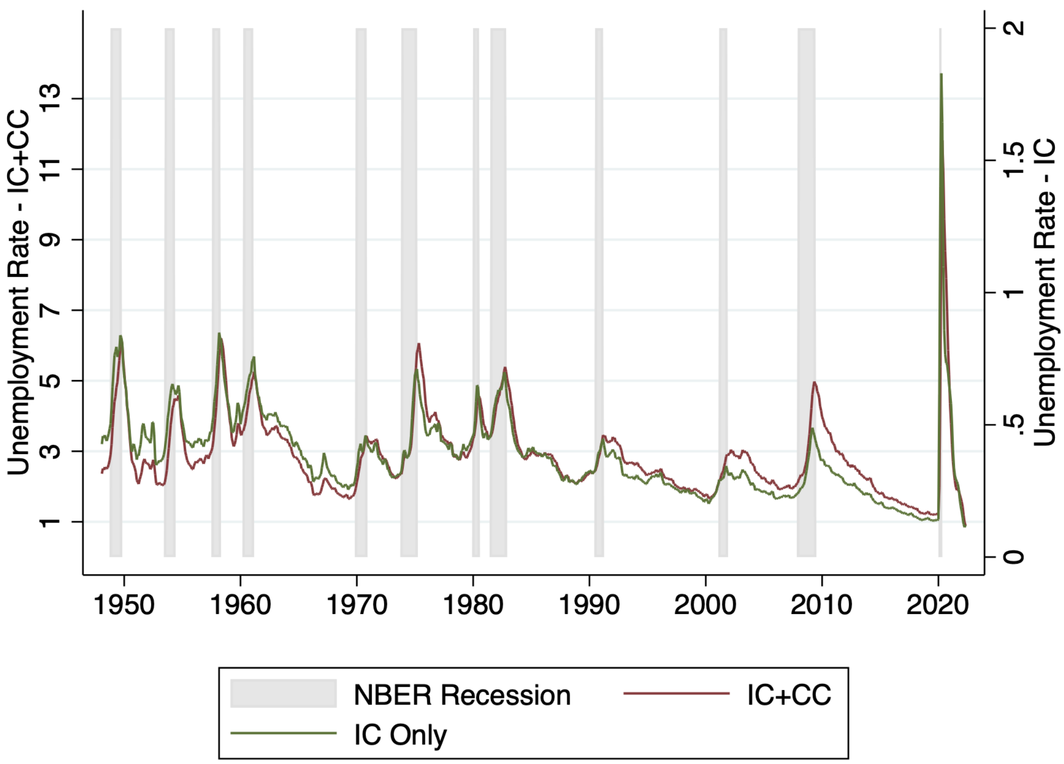


Census Region I: CT, ME, MA, NH, RI, VT, NJ, NY, PA.
Census Region II IN, IL, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD.
Census Region III: DE, DC, FL, GA, MD, NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX.
Census Region IV: AZ, CO, ID, NM, MT, UT, NV, WY, AK, CA, HI, OR, WA.

Appendix B Additional Empirical Results

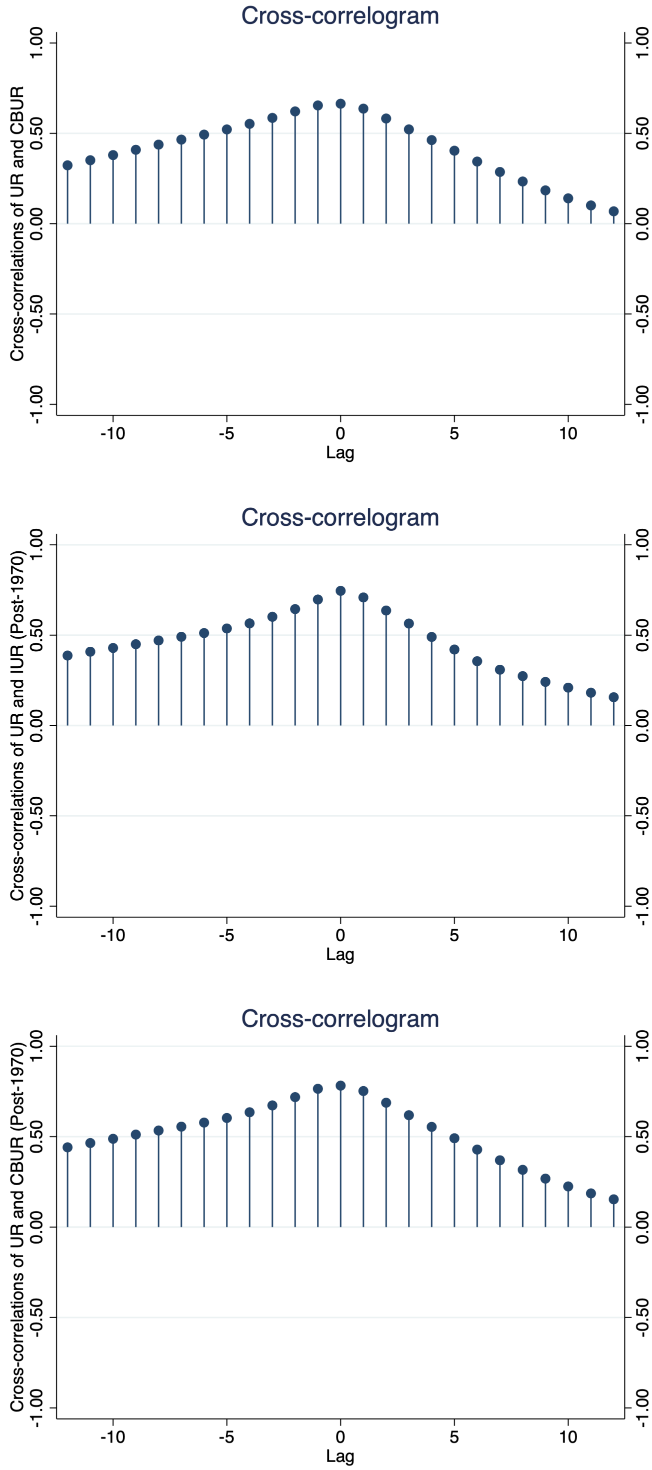
B.1 Robustness Tests

Figure B.1: Comparison of Claims-based Unemployment Rates Using IC+CC Versus IC



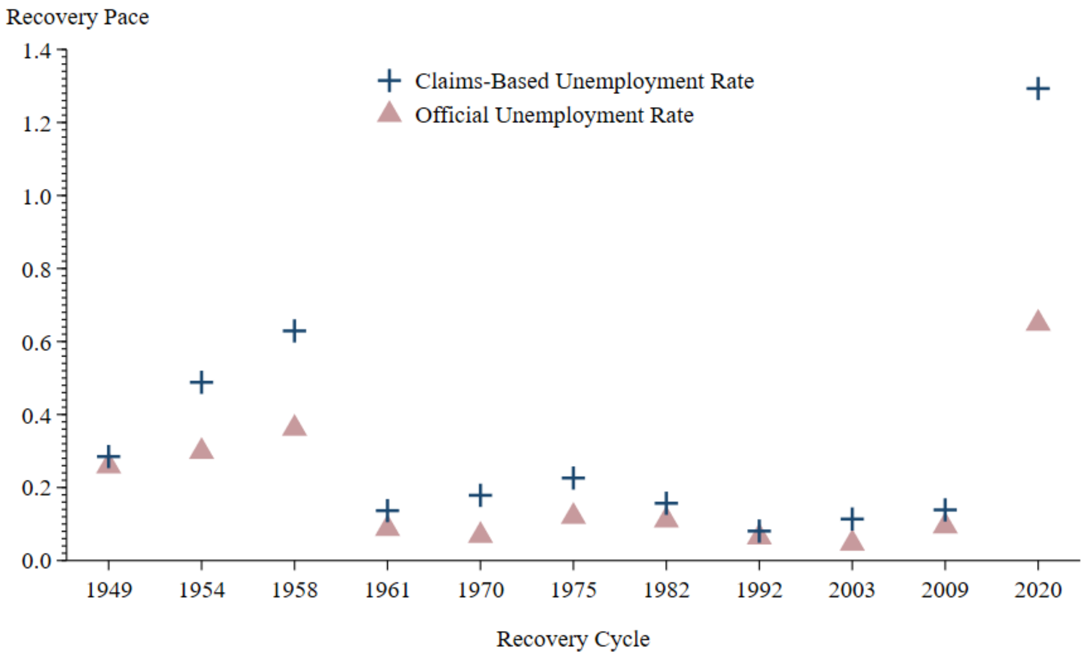
Notes: Claims-based unemployment rates are smoothed with a three-month centered moving average. Claims-based unemployment rates computed from IC+CC data (red line) vs. IC data only (green line) are plotted on the left and right axis, respectively. Sample: January 1948–May 2022.

Figure B.2: Cross Correlations Between the U.S. Unemployment Rate, CBUR, and IUR



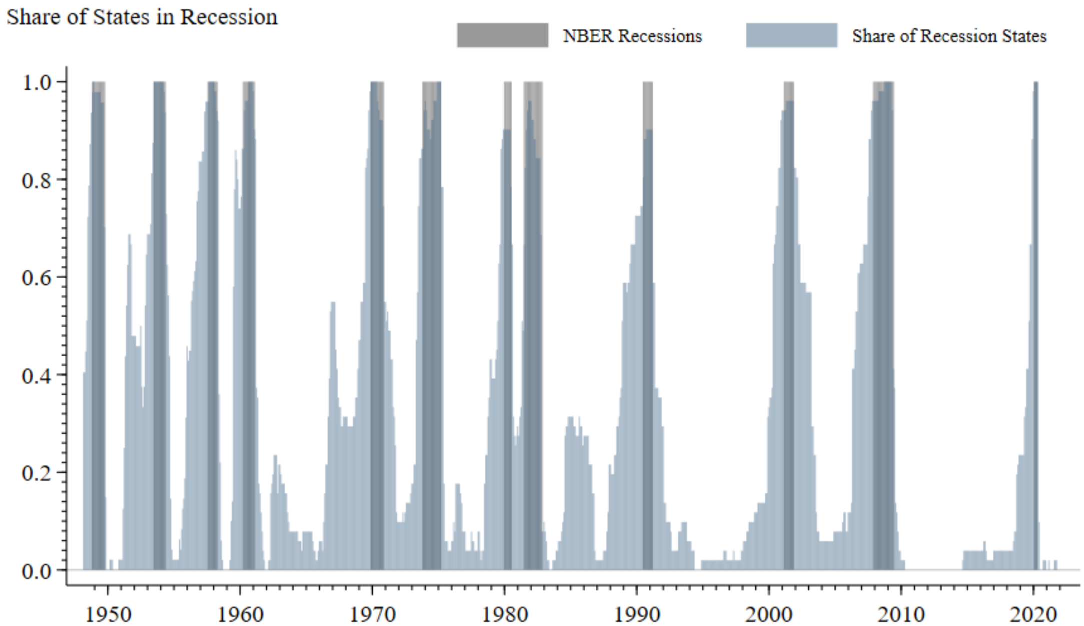
Notes: The top panel plots the cross correlations between the official unemployment rate (UR) versus our claims-based unemployment rate (CBUR) from 1948–2022. The middle panel plots the cross correlations between the official unemployment rate versus the insured unemployment rate (IUR), which is only available from 1971 onwards. The bottom figure re-plots the cross correlations between the official unemployment rate versus our CBUR just over 1971–2022, the same time horizon as the IUR figure in the middle panel.

Figure B.3: National Unemployment Recovery Rates: Recession Dates from CBUR



Notes: Recovery dates from DNS algorithm with recovery dates generated from the claims-based unemployment rate. Recovery from the 1980 recession is again excluded, see notes to Figure 5.

Figure B.4: Share of U.S. States in Recession: Recession Dates from Unfitted CBUR



Notes: State-level recession coding is constructed by applying the DNS algorithm to state-level raw (unfitted) claims-based unemployment rates. The DNS algorithm parameter is adjusted for each state proportionate to its average level of unemployment over the entire time period. Sample: January 1948–May 2022.

B.2 Claims-Based Recession Dates vs. Recession Probabilities

Figure B.5 depicts our claims-based unemployment rates (blue lines), state recession dates (gray bars), and the Owyang et al. (2005) state recession probabilities (red lines) for all 50 states. There are notable similarities for a number of states across to the two datasets when they overlap in the 1979–2002 sample. For many larger states, both our claims-based unemployment recession dates and the Owyang et al. (2005) recession probabilities only identify the same national recessions in the overlapping sample (1980, 1981-82, 1990-91, and 2001), albeit with slightly different peak and trough dates and/or ignoring the distinction between a double-dip recession versus a longer recession in the early 1980s (e.g., AZ, CA, CT, FL, IL, MA, NC, and NY). And in some cases, both datasets identify nearly identically timed recessions that were not experienced on the national level. For instance, we identify

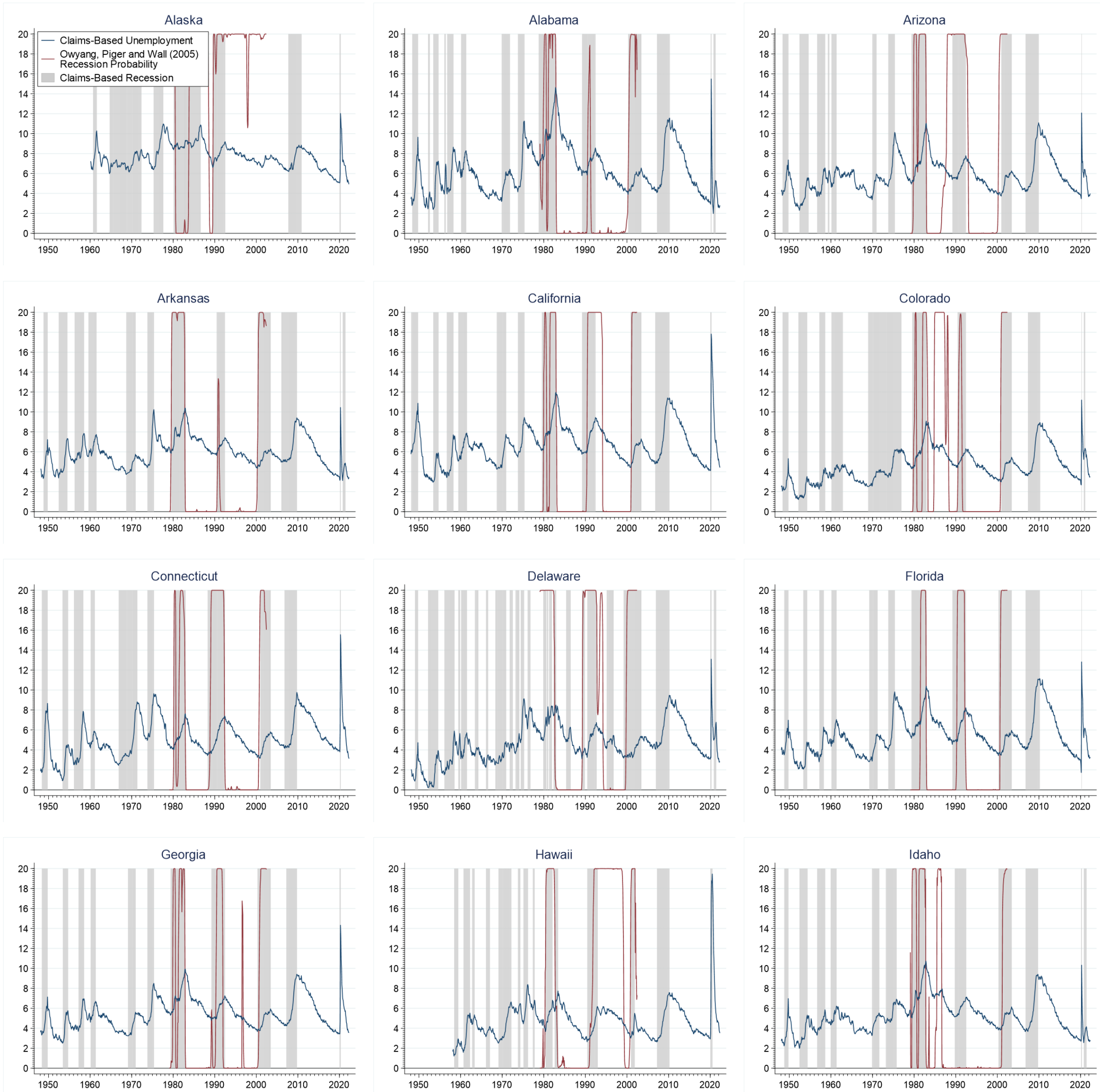
Mississippi as falling into recession over February 1986–June 1986 and Owyang et al. (2005) identify Mississippi as being in recession over February 1986–July 1986 with probabilities exceeding 80% for each of these months. Similarly, Owyang et al. (2005) identify Wyoming as falling into recession over February 1986–March 1987 with probabilities exceeding 80%, while we identify Wyoming as falling into recession over December 1984–October 1986.

There are also striking differences between the two datasets, most notably in smaller states. Out of sync with the national business cycle, Owyang et al. (2005) alone identify short-lived recessions in Idaho, New Mexico, South Dakota, and Utah in the mid-1980s, whereas our dataset alone identifies a short-lived recession in Delaware in the mid-1980s. Owyang et al. (2005) do not identify the 1990-91 recession in a number of states, contrary to our claims-based unemployment recession dates (e.g, IA, ID, LA, ND, OK, SD, TX, UT, and WY). Conversely, Owyang et al. (2005) identify short-lived recessions in Maine, Maryland, New Mexico, and Washington in the mid-1990s, which are not identified in our claims-based unemployment recession dates. And Owyang et al. (2005) do not identify the 2001 recession in Kansas, Oklahoma, or Wyoming, unlike our claims-based unemployment recession dates. And in other states where both datasets identify recessions around 1990-91 and 2001, the Owyang et al. (2005) recession probabilities identify considerably shorter downturns than our claims-based unemployment recession dates (e.g., KY, MN, OR, and WI). In a handful of other states, our claims-based unemployment recession dates show considerably shorter recessions than the Owyang et al. (2005) recession probabilities. At one extreme, the Owyang et al. (2005) recession probabilities show Alaska continuously in a recession from August 1989–June 2002, with recession probabilities averaging 97.5% and never falling below 50% for this sample. Similarly, their recession probabilities show Hawaii in a slump throughout almost all of the 1990s, with recession probabilities averaging 98.3% and never falling below 60% over November 1991–December 1999. In line with a clear, persistent recovery in the claims-based unemployment rates for Alaska in the early 1990s, our state recession dates show Alaska in a much shorter recession, over August 1989–July

1992. And we identify Hawaii as having experienced only a short-lived recession in the early 1990s, followed by a persistent recovery in unemployment.

Neither approach is right or wrong per se, but Figure B.5 underscores that our recession dates exhibit fewer erratic, short-lived recessionary spikes or unusually long recessionary periods, and no judgement is required regarding a cutoff for recession probabilities to identify recovery dates and durations. The principal advantage to our approach, however, is the ability to identify inflection points in state business cycles for more than 30 additional years when using our claims-based unemployment rate series instead of existing off-the-shelf state coincident indexes.

Figure B.5: State Recession Dates and Recession Probabilities



Notes: Our claims-based state unemployment rates (blue) are for January 1948–May 2022, save for the handful of states for which nonfarm payroll employment data is only available starting in the 1950s, see footnote 6 for details. State recession dates (gray bars) are estimated from our claims-based unemployment rates using the Dupraz et al. (2019) algorithm. State recession probabilities (red) for February 1979–June 2002 are from Owyang et al. (2005). The y-axis measures the unemployment rate in percentage points and recession probabilities in five-percentage point increments (20=100%).

Figure B.5: State Recession Dates and Recession Probabilities (Continued...)

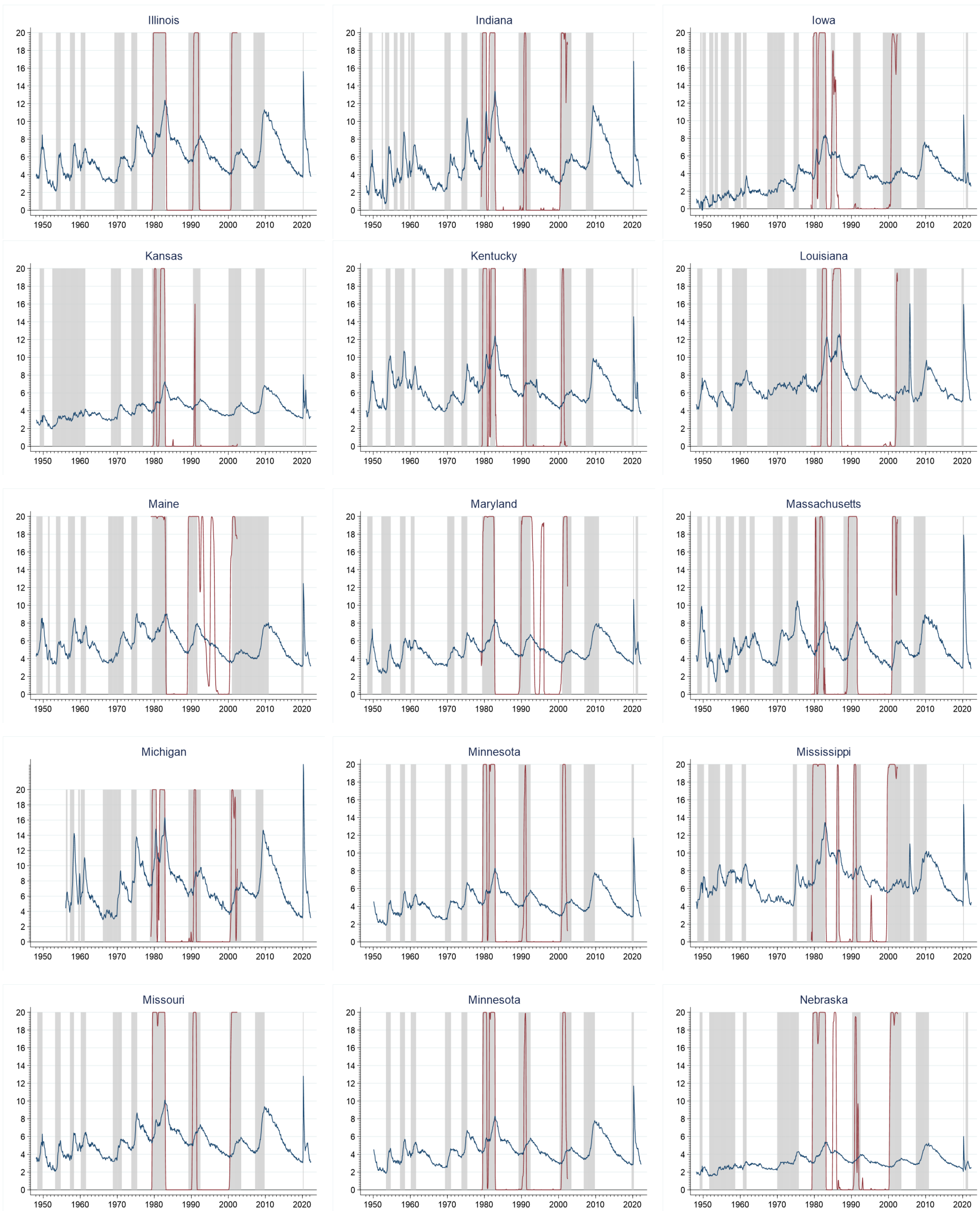


Figure B.5: State Recession Dates and Recession Probabilities (Continued...)

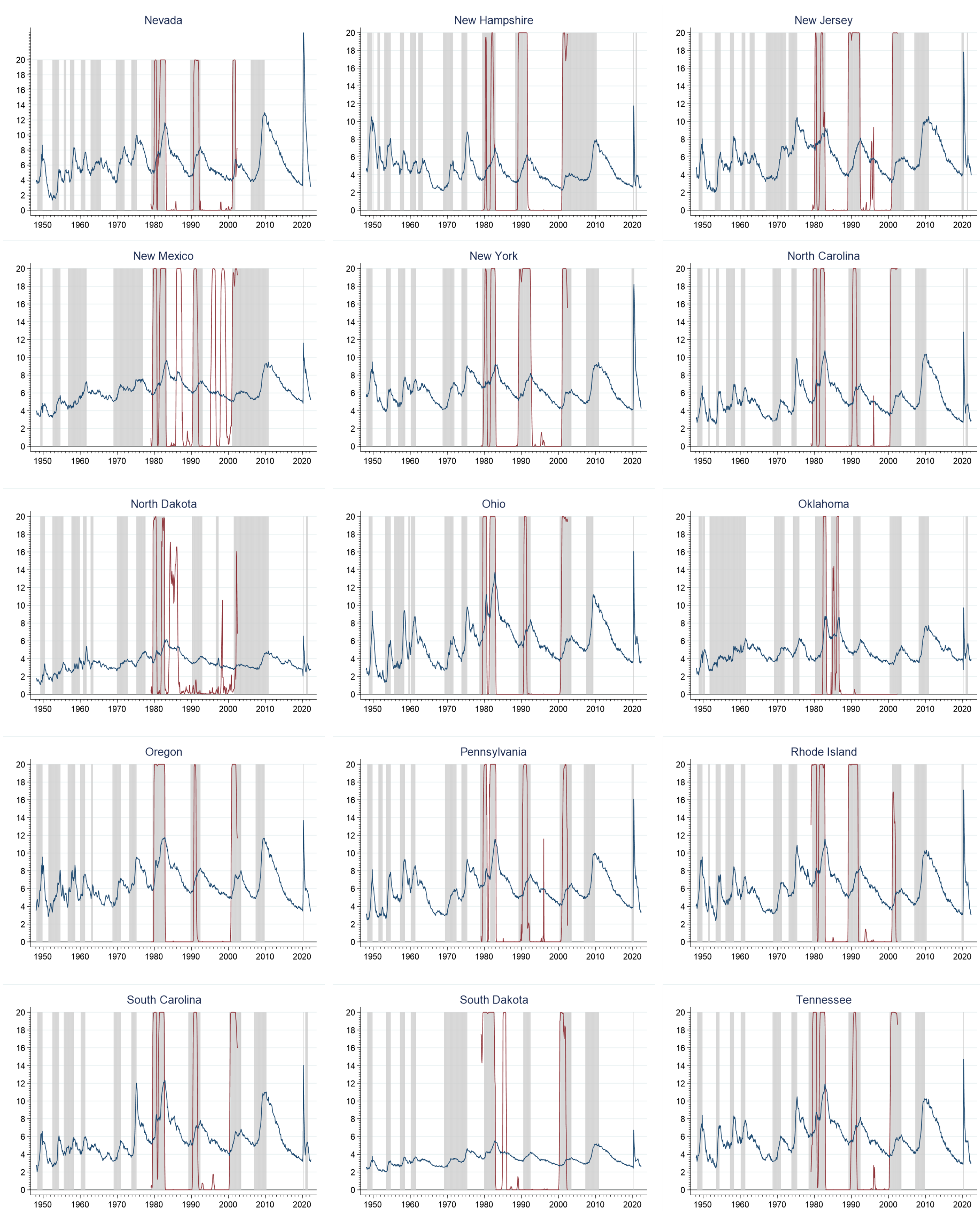


Figure B.5: State Recession Dates and Recession Probabilities (Continued...)