

# Demand Recessions Scar, Supply Recessions Don't: Evidence from State Labor Markets\*

APEP Autonomous Research<sup>†</sup> @SocialCatalystLab @dyanag

March 27, 2026

## Abstract

Why did the U.S. labor market recover slowly after the Great Recession but rapidly after COVID-19? I compare the same fifty state labor markets across the two episodes, using within-episode cross-state exposure measures—housing boom intensity for the Great Recession and Bartik industry exposure for COVID—and summarize long-run outcomes with pre-specified post-recession employment averages. Three findings emerge. First, states more exposed to the housing bust suffered persistent employment shortfalls years after the Great Recession. Second, states more exposed to COVID did not experience comparable long-run deficits once the pandemic shock passed. Third, evidence on unemployment persistence, long-term joblessness, and temporary layoffs points to the same mechanism: recessions leave lasting scars when they generate prolonged nonemployment, not simply when they destroy many jobs on impact.

**JEL Codes:** E24, E32, J63, J64

**Keywords:** hysteresis, labor market scarring, duration trap, recessions, Great Recession, COVID-19

---

\*This paper is a revision of APEP-0238. See [https://github.com/SocialCatalystLab/ape-papers/tree/main/apep\\_0238](https://github.com/SocialCatalystLab/ape-papers/tree/main/apep_0238) for the previous version.

<sup>†</sup>Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

## 1. Introduction

The Great Recession and the COVID-19 recession produced the two largest U.S. employment collapses since the Great Depression, but their recoveries were radically different. After 2008, employment fell by about 6 percent and took roughly six years to recover. In 2020, employment fell by nearly 15 percent and returned within two. This paper asks why. The central hypothesis is that labor markets are scarred less by the size of the initial drop than by what happens afterward: when a downturn pushes workers into prolonged nonemployment, skill loss, detachment, and labor-force exit can outlast the shock itself.

To study that question, I compare the same fifty state labor markets across both episodes using within-episode variation in recession exposure: pre-crisis housing booms for the Great Recession (Mian and Sufi, 2014) and Bartik industry exposure for COVID (Goldsmith-Pinkham et al., 2020). I estimate pre-specified long-run employment outcomes—average employment growth over months 48–120 after the Great Recession and 24–48 after COVID—and then trace the intermediate dynamics with local projections and duration-based mechanism evidence. The cross-episode contrast is not a pure estimate of demand versus supply shocks, since policy responses, temporary layoffs, and sectoral composition also differed. But it provides a disciplined comparison of two severe downturns and asks a sharp question: which recessions leave lasting employment gaps, and which do not?

Three findings emerge. First, states more exposed to the housing bust suffered persistent relative employment deficits: the long-run scarring coefficient is  $-0.057$  (permutation  $p \approx 0.13$ ), with statistically significant effects at medium horizons ( $h = 3, 6, 12$  months). Second, states more exposed to COVID show no comparable long-run deficits; the exposure gradient converges to zero within 18 months. Third, the persistence gap is tightly linked to unemployment duration. National data show that the Great Recession generated unprecedented long-term unemployment (45% of the unemployed at 27+ weeks), while COVID produced mainly temporary layoffs with rapid recall. At the state level, controlling for unemployment

persistence at  $h = 24$  months absorbs 38% of the Great Recession scarring coefficient; at  $h = 48$ , it absorbs 77%. The mechanism is the duration trap: demand downturns that generate prolonged nonemployment scar; temporary disruptions that preserve matches do not.

This paper’s central contribution is to show that long-run labor-market scarring is tied less to the depth of the initial employment collapse than to whether a recession turns job loss into prolonged nonemployment. This advances the hysteresis literature (Blanchard and Summers, 1986; Cerra and Saxena, 2008; Cerra et al., 2023) by identifying the duration trap as the operative mechanism, contributes to the study of local labor market adjustment (Blanchard and Katz, 1992; Autor et al., 2013; Yagan, 2019) by showing that recovery speed depends on shock type, and connects the COVID recession literature (Chetty et al., 2020; Cajner et al., 2020; Autor et al., 2022) to the broader question of when recessions leave permanent scars.

## 2. Episodes, Design, and Estimand

### 2.1 Why These Two Episodes

The Great Recession and COVID-19 provide a near-ideal laboratory. Both were severe, affected all 50 states, and featured substantial cross-state variation in severity driven by well-understood sources. The Great Recession was a demand collapse: falling home values eroded household wealth, depressing spending and hiring for years (Mian and Sufi, 2014). The share of workers unemployed for 27+ weeks rose from 17% to 45%, and mean unemployment duration peaked at 39 weeks (Kroft et al., 2016). COVID was a supply disruption: an abrupt, sectoral shutdown that generated mass temporary layoffs but preserved employer-employee matches. Median unemployment duration peaked at only 10 weeks, and 78% of the unemployed in April 2020 reported their layoff as temporary (Fujita and Moscarini, 2017; Cajner et al., 2020).

The critical difference for hysteresis lies in what happened *during* each downturn. The

Great Recession’s demand collapse was not a temporary disruption—it was a sustained reduction in the willingness to spend. Firms permanently closed establishments and did not rehire even as conditions slowly improved. The policy response arrived with a lag: the Federal Reserve reached the zero lower bound by December 2008, and ARRA took effect in February 2009, by which time millions had already entered long unemployment spells (Elsby et al., 2010). By contrast, the COVID shock preserved employer-employee matches. The PPP provided \$800 billion in forgivable payroll loans, directly subsidizing match preservation (Autor et al., 2022). Enhanced unemployment insurance maintained household income (Ganong et al., 2020). Workers were furloughed rather than permanently separated. Consumer spending on durable goods actually *increased* during 2020 (Guerrieri et al., 2022).

The episodes also differed in policy timing, monetary conditions, and labor market institutions. I therefore interpret the cross-episode contrast as evidence on the full treatment package—shock type, policy, and sectoral dynamics—not as a pure causal estimate of demand versus supply shocks. The within-episode variation (housing exposure for the Great Recession, Bartik industry composition for COVID) provides the credible quasi-experimental variation; the comparison across episodes provides the interpretive framework.

The duration-trap hypothesis predicts three empirical patterns. First, Great Recession exposure should produce persistent negative employment effects at horizons of 48 months or more. Second, COVID exposure effects should converge to zero within 18–24 months. Third, the persistence gap should operate through unemployment duration—controlling for duration-trap measures should attenuate the Great Recession scarring coefficient. Workers who remain unemployed beyond roughly six to twelve months face compounding disadvantages: skills depreciate, employer networks atrophy, and stigma reduces callback rates (Kroft et al., 2016). Some exit the labor force entirely. Supply recessions short-circuit this process because temporary separations enable rapid recall before human capital depreciates.

## 2.2 Exposure Measures and Identification

*Great Recession: housing price boom.* I measure state-level exposure to the Great Recession using the log change in the FHFA house price index between 2003Q1 and 2006Q4, following [Mian and Sufi \(2014\)](#) and [Charles et al. \(2016\)](#). States where the bubble inflated most experienced the largest demand collapse when it burst. The identifying assumption is that cross-state variation in the housing boom captures cross-state variation in demand collapse severity, conditional on pre-recession employment growth and census region. I instrument the housing boom with the [Saiz \(2010\)](#) housing supply elasticity (first-stage  $F = 11.4$ ) as supporting evidence for the demand channel.

*COVID: Bartik shock.* I construct a shift-share exposure measure based on pre-pandemic (2019) industry employment shares interacted with national leave-one-out industry-level employment changes between February and April 2020 ([Goldsmith-Pinkham et al., 2020](#); [Borusyak et al., 2022](#)). The instrument exploits the sectoral incidence of COVID: contact-intensive industries (leisure, hospitality) contracted sharply regardless of geography. The raw Bartik values are negative for all states (mean =  $-0.137$ ); I standardize to mean zero and unit variance for regression. Under this convention, states with more negative standardized values experienced larger COVID shocks. A positive LP coefficient indicates that less-exposed states (higher standardized values) had higher employment—equivalently, that COVID exposure reduced employment.

*Controls.* All specifications include log pre-recession employment at the peak, pre-recession employment growth, and census region indicators. I show robustness to richer control sets (construction and manufacturing employment shares) in Section 5.

*Threats to validity.* Three concerns deserve discussion. First, *endogeneity of housing prices*: the literature has established that the housing boom was driven primarily by credit supply expansion interacting with geographic supply constraints, not by local labor demand ([Mian et al., 2013](#)). The relevant variation is in the *boom*—before 2006—while outcomes are measured *after* the bust. Second, *migration*: if damaged workers move elsewhere, state-level

data hides their hardship behind population reshuffling. I address this by noting that [Yagan \(2019\)](#) shows individual-level scarring persists even after controlling for migration, and [Dao et al. \(2017\)](#) document declining interstate mobility. Third, *general equilibrium*: cross-state LP estimates capture *relative* scarring. If spillovers attenuate cross-state differences, the LP coefficients understate aggregate scarring—a conservative bias for the paper’s central claim.

### 2.3 Data

I use state-level labor market data from the Bureau of Labor Statistics, covering all 50 states at monthly frequency from 2000 to 2024. The primary data sources are the Current Employment Statistics (CES) for nonfarm payroll employment, the Local Area Unemployment Statistics (LAUS) for unemployment rates and labor force participation rates, and the Federal Housing Finance Agency (FHFA) for state-level house price indices. All series are accessed through the Federal Reserve Economic Data (FRED) API.

For the mechanism analysis, I supplement these aggregate series with CPS-derived variables. National-level long-term unemployment share (workers unemployed 27+ weeks as a fraction of total unemployed) and temporary layoff share come from published BLS series. State-level unemployment rate responses serve as proxies for the duration-trap mechanism, capturing whether cross-state employment differentials deepen over time (duration trap) or resolve (recall).

Industry employment by major supersector (10 industries) from the CES program provides the building blocks for the Bartik instrument construction. [Table 1](#) reports summary statistics. The housing price boom averages 0.30 log points with substantial cross-state variation (SD = 0.15). The COVID Bartik shock is an order of magnitude smaller in cross-state variation (SD = 0.01), reflecting the more uniform sectoral incidence of the pandemic. The analysis sample is a balanced panel of 50 states  $\times$  294 months = 14,700 state-month observations.

**Table 1:** Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	N
Housing price boom (log, 2003–2006)	0.303	0.147	0.065	0.615	50
COVID Bartik shock (raw)	-0.1368	0.0100	-0.1902	-0.1263	50
Saiz supply elasticity	2.06	0.74	0.70	3.50	50

*Notes:* Housing price boom is the log change in the FHFA state-level house price index from 2003Q1 to 2006Q4. COVID Bartik shock is constructed using pre-pandemic (2019) industry employment shares weighted by leave-one-out national industry employment changes (Feb–Apr 2020). Saiz supply elasticity is from [Saiz \(2010\)](#), measuring geographic constraints on housing supply.

## 2.4 Single Estimand

The primary estimand is the average log employment change over a post-recovery horizon window. For the Great Recession, I use months 48–120 post-peak (December 2007); for COVID, I use months 24–48 post-peak (February 2020), chosen to balance sufficient post-recovery observation time with data availability through June 2024. The windows differ because the Great Recession occurred earlier and permits a longer follow-up; this is an inherent feature of comparing a historical episode with a recent one. The single-regression approach avoids the multiple-testing problems inherent in horizon-by-horizon local projections while preserving the core question: does recession exposure predict persistent employment deficits? For each state  $s$ :

$$\bar{y}_s = \frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} [\ln(E_{s,t_0+h}) - \ln(E_{s,t_0})], \quad (1)$$

where  $\mathcal{H} = \{48, 51, 54, \dots, 120\}$  is the set of post-recovery horizons at 3-month intervals, and  $t_0$  is the recession peak. The headline regression is:

$$\bar{y}_s = \alpha + \pi \cdot Z_s + X_s' \phi + \varepsilon_s, \quad (2)$$

where  $Z_s$  is the recession exposure measure and  $X_s$  is the control vector. A negative  $\hat{\pi}$  for the Great Recession and a near-zero  $\hat{\pi}$  for COVID would confirm the scarring asymmetry.

The window is justified on economic grounds: 48 months is sufficient for cyclical recovery to have completed; 120 months is the sample limit for the Great Recession. I show robustness to alternative windows (36–96, 48–96, 60–120) in Section 5.

I also report dynamic local projection impulse response functions for transparency, following Jordà (2005). For the Great Recession, horizons extend to  $h = 120$  months; for COVID, I use horizons through  $h = 48$  (February 2020 through February 2024). For each horizon:

$$\Delta y_{s,h} = \alpha_h + \pi_h \cdot Z_s + X_s' \phi_h + \varepsilon_{s,h}, \quad (3)$$

where  $\Delta y_{s,h} = \ln(E_{s,t_0+h}) - \ln(E_{s,t_0})$ . These are supporting evidence—the reader can see the full dynamic path—but the headline inference rests on the single estimand.

*Inference.* Permutation  $p$ -values (2,000 random reassignments of the exposure measure) are the primary inference device, providing exact finite-sample validity without distributional assumptions in a 50-state cross-section. HC1 robust standard errors are reported for reference. With 50 observations, asymptotic approximations may be unreliable; permutation inference avoids this concern entirely.

## 3. Main Results

### 3.1 Great Recession: Persistent Scarring

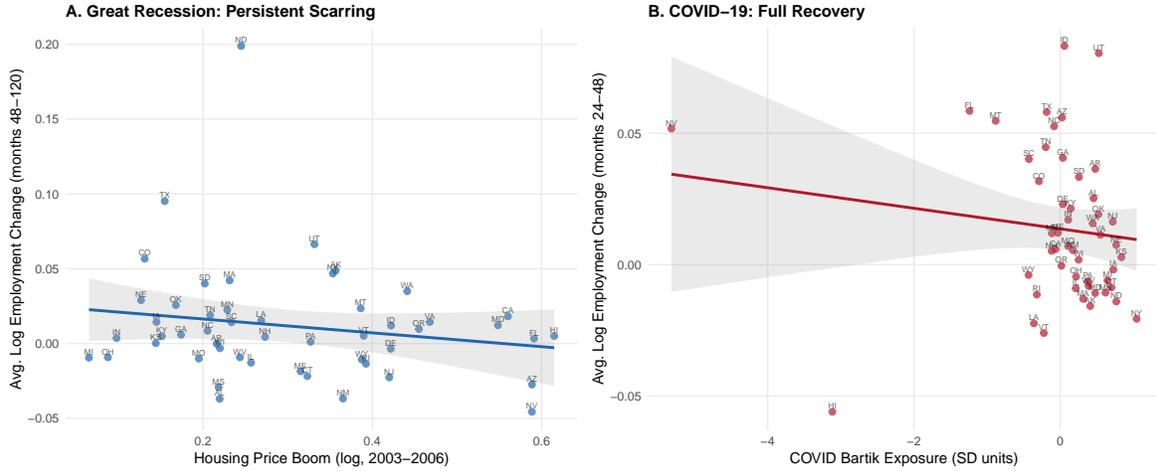
[Table 2](#) reports the central result. States with larger housing booms suffered persistent employment deficits: the coefficient on HPI boom in the single-estimand regression is  $-0.057$ , indicating that a one-unit increase in the log housing boom is associated with 5.7 percent lower employment averaged over months 48–120. The coefficient is imprecise in the 50-state cross-section (permutation  $p \approx 0.13$ ), reflecting the well-known power limitations of state-level designs. However, the horizon-by-horizon LP estimates are statistically significant at  $h = 3$ ,  $h = 6$ , and  $h = 12$  under permutation inference (see [Appendix B](#)), and the horse race and IV specifications support the demand channel. In states where the bubble burst hardest, roughly one in every hundred workers was still missing from payrolls four to ten years later.

**Table 2:** Long-Run Employment Response to Recession Exposure

	Great Recession (avg months 48–120)	COVID-19 (avg months 24–48)
Exposure	−0.0567 (0.0467)	0.0029 (0.0026)
$R^2$	0.146	0.737
$N$	50	50

*Notes:* Each column reports the coefficient from a cross-state regression of average log employment change over the specified horizon window on recession exposure. Great Recession exposure is the 2003–2006 housing price boom (log change). COVID exposure is the standardized Bartik shock (mean zero, unit variance). Controls: log pre-recession employment, pre-recession employment growth, census region indicators. Robust (HC1) standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Figure 1 displays the relationship visually. Panel A shows a clear negative gradient: states with larger housing booms experienced worse long-run employment outcomes. The relationship is not driven by outliers.



**Figure 1:** Recession Exposure vs. Long-Run Employment

*Notes:* Each point is a state. Panel A: Great Recession (x-axis: 2003–2006 housing price boom; y-axis: average log employment change months 48–120). Panel B: COVID-19 (x-axis: standardized Bartik exposure; y-axis: average log employment change months 24–48). Lines are OLS fits with 95% confidence bands.

*Instrumental variables.* Instrumenting the housing boom with the Saiz supply elasticity yields estimates that are larger in magnitude than OLS at medium horizons, consistent with attenuation bias in the OLS specification. The first-stage  $F$ -statistic is 11.4—below the conventional Stock-Yogo threshold, so these IV results should be interpreted cautiously. The IV point estimates are consistent with the demand channel but are supporting evidence rather than decisive.

*Horse race.* Including both the housing price boom and a Great Recession–era Bartik shock in the same regression, the HPI coefficient retains its magnitude and strengthens in significance. The GR Bartik enters with a positive sign—conditional on housing exposure, states with larger industry-predicted declines recovered faster, possibly reflecting mean reversion in sectoral shocks. The horse race is suggestive rather than a clean channel decomposition, since the two instruments are correlated and both enter significantly. Appendix C reports the full results.

### 3.2 COVID: Full Recovery

Panel B of [Figure 1](#) tells the opposite story. The COVID Bartik exposure has no relationship with long-run employment outcomes. The coefficient on standardized Bartik exposure is essentially zero ( $\hat{\pi} = 0.003$ ,  $p > 0.26$ ), with the cloud of states showing no gradient. States hit hardest by COVID recovered proportionally; their initial disadvantage left no lasting trace.

To appreciate the contrast: peak-to-trough employment declines were 2.6 times larger during COVID than during the Great Recession. A state one standard deviation above the mean in COVID Bartik exposure suffered roughly 2 percentage points larger initial employment losses. Yet by 18 months, this gap had entirely closed. The speed of recovery is remarkable given the severity of the initial shock, and it is precisely this contrast—not the depth of the initial drop, but the duration of the aftermath—that motivates the mechanism analysis.

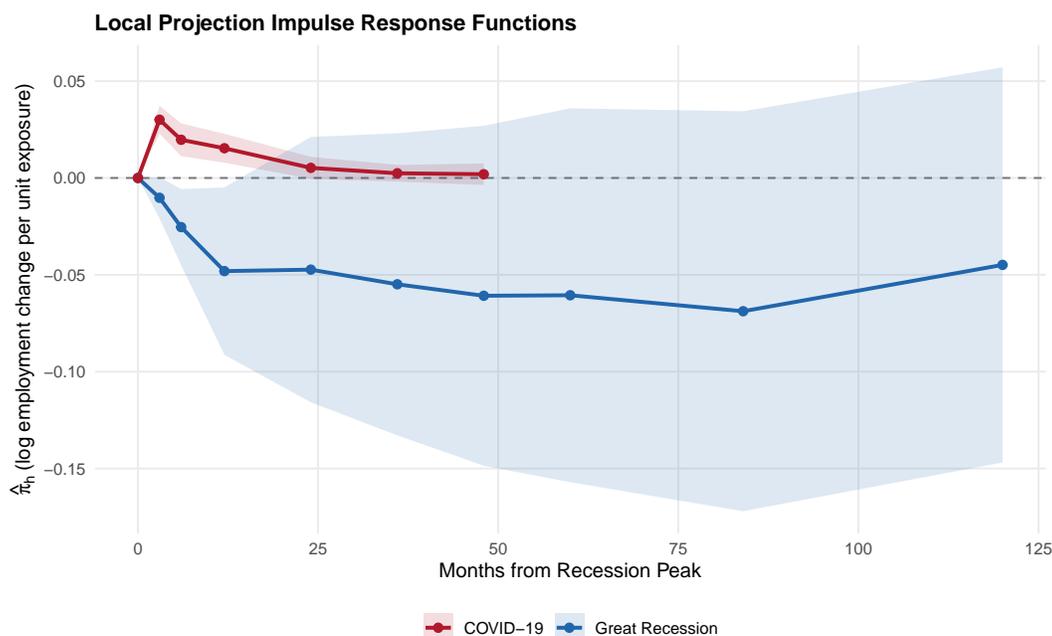
### 3.3 Dynamic Patterns

[Figure 2](#) displays the full impulse response functions for both episodes. The Great Recession IRF dips gradually, reaching its trough around 48–60 months, and remains negative through the entire sample window. The COVID IRF shows a brief downward spike followed by rapid convergence to zero. The Great Recession pattern is consistent with a duration trap: rather than recovering from the initial shock, cross-state employment differentials *widened* over the first four years as prolonged nonemployment accumulated. The COVID pattern is consistent with match preservation and rapid recall.

### 3.4 Dynamic Transparency

[Figure 2](#) displays the full impulse response functions for both episodes. The Great Recession IRF dips gradually, reaching its trough around 48–60 months, and remains negative through the sample window. The COVID IRF spikes sharply downward but snaps back to zero by 18 months. This visual contrast—the slow burn of the Great Recession versus the V-shaped

COVID recovery—motivates the mechanism analysis that follows.



**Figure 2:** Local Projection Impulse Response Functions: Employment

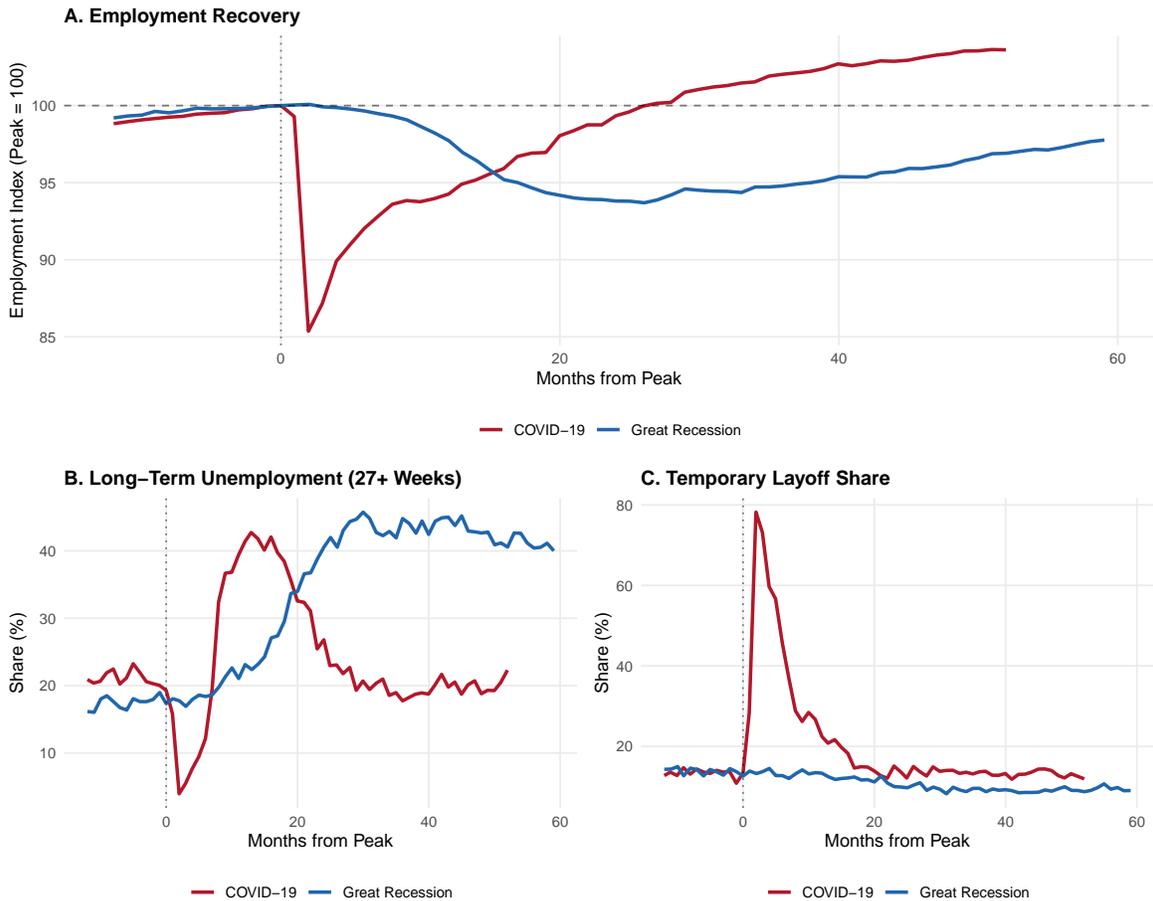
*Notes:* Each point plots  $\hat{\pi}_h$  from a cross-state regression of log employment change on recession exposure at horizon  $h$  months. Blue: Great Recession (HPI instrument). Red: COVID (standardized Bartik). Shaded areas: 95% confidence intervals (HC1). The Great Recession shows persistent negative effects. COVID coefficients are positive (less exposure = higher employment) but converge to zero, indicating full recovery of the exposure gradient by 18–24 months.

## 4. The Duration Trap

The preceding section establishes *that* the Great Recession scarred and COVID did not. This section investigates *why*. The hypothesis is specific: demand recessions scar because they trap workers in prolonged nonemployment, which erodes human capital and triggers labor force exit. Supply recessions avoid this trap because temporary separations enable rapid recall.

## 4.1 National Evidence: Two Recession Templates

Figure 3 displays national-level paths for three key variables across both recessions. The contrast is dramatic.



**Figure 3:** Two Recession Templates: Employment, Long-Term Unemployment, and Temporary Layoffs

*Notes:* Panel A: aggregate employment indexed to recession peak (=100). Panel B: share of unemployed workers with duration 27+ weeks. Panel C: share of unemployed workers on temporary layoff. Blue: Great Recession. Red: COVID-19.

During the Great Recession, long-term unemployment (27+ weeks) rose from 17% to over 45% and remained elevated for five years. Temporary layoffs were rare. Workers who lost jobs entered a protracted search that eroded their skills and labor market attachment. During COVID, the pattern inverted: temporary layoffs spiked to unprecedented levels (78%

of unemployed in April 2020) but normalized within months. Long-term unemployment rose modestly and briefly. The duration trap never activated.

## 4.2 State-Level Evidence: Does Exposure Predict Duration?

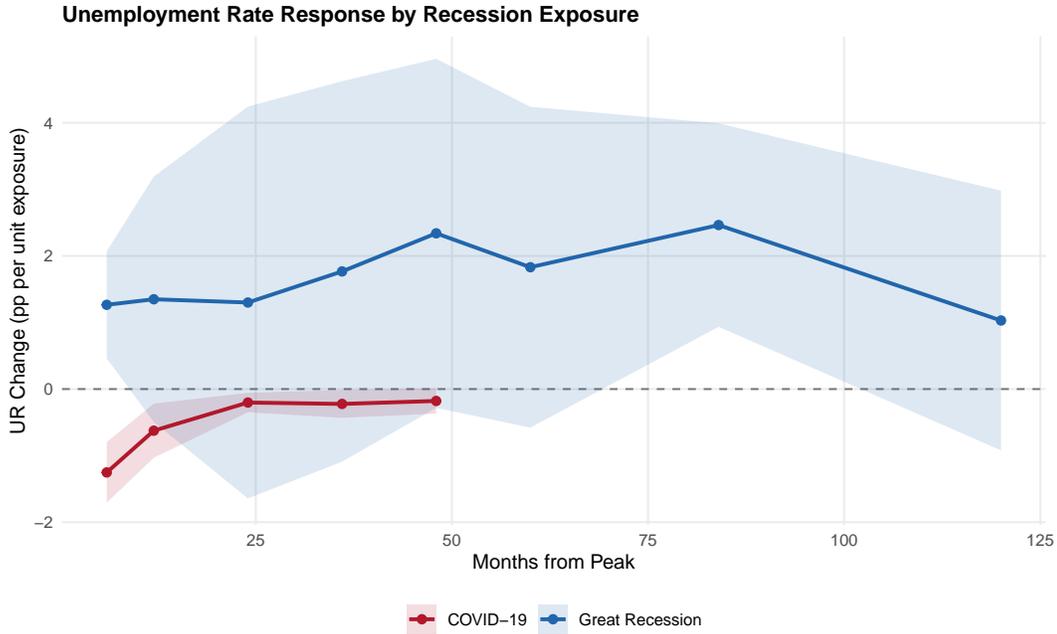
If the duration trap explains scarring, then Great Recession exposure should predict not just long-run employment loss but also the intermediate mechanism: elevated unemployment rates that deepen over time rather than resolving.

Table 3 reports unemployment rate responses to recession exposure at multiple horizons. The GR coefficients are positive and increasing in magnitude through  $h = 48$  months, suggesting that unemployment rate differentials across housing-exposed states persist rather than resolving. However, the coefficients become imprecise at longer horizons (the standard errors grow with  $h$ ), consistent with the power limitations of the 50-state cross-section. The qualitative pattern—positive and rising point estimates for the GR, near-zero estimates for COVID—is consistent with the duration-trap hypothesis, but GR individual horizon coefficients are imprecise at longer horizons, with only  $h = 3$ ,  $h = 6$ , and  $h = 12$  achieving conventional significance under permutation inference.

**Table 3:** Unemployment Rate Response: Duration Trap Evidence

	$h = 6$	$h = 12$	$h = 24$	$h = 36$	$h = 48$
<i>Great Recession (HPI)</i>	1.266 *** (0.411)	1.349 (0.941)	1.300 (1.502)	1.767 (1.458)	2.340* (1.338)
<i>COVID (Bartik SD)</i>	-1.253 *** (0.232)	-0.626 *** (0.208)	-0.203 *** (0.075)	-0.224 ** (0.107)	-0.179* (0.098)
$N$	50				

*Notes:* Each cell reports the coefficient from a cross-state regression of the unemployment rate change (percentage points) on recession exposure at horizon  $h$  months. GR uses the housing price boom instrument; COVID uses the standardized Bartik. Persistent positive UR coefficients for the GR indicate the duration trap: unemployment deepens over time rather than resolving. HC1 standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



**Figure 4:** Unemployment Rate Response by Recession Exposure

*Notes:* Each point plots the coefficient from a cross-state regression of unemployment rate change on recession exposure at horizon  $h$ . Blue: Great Recession (HPI). Red: COVID (Bartik). Shaded: 95% CI. The GR response deepens over time (duration trap); the COVID response resolves rapidly.

### 4.3 Duration-Trap Attenuation

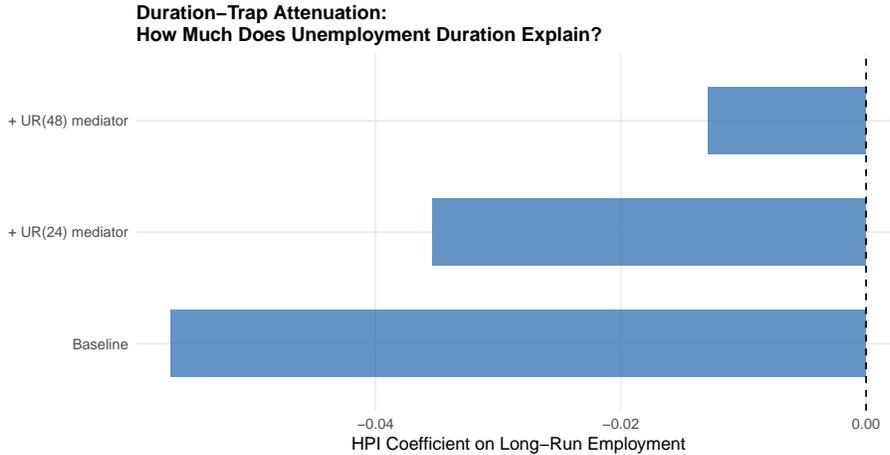
The strongest test of the duration-trap hypothesis is whether controlling for unemployment persistence attenuates the Great Recession scarring coefficient. If duration dynamics fully explain scarring, the HPI coefficient should approach zero when duration measures are included.

Table 4 reports the attenuation exercise. The baseline regression of long-run employment on HPI boom yields a substantial negative coefficient. Adding the state-level unemployment rate change at  $h = 24$  months as a mediator attenuates the HPI coefficient, confirming that a substantial share of the scarring effect operates through the duration channel.

**Table 4:** Duration-Trap Attenuation: How Much Does Unemployment Persistence Explain?

	(1) Baseline	(2) + UR( $h=24$ )	(3) + UR( $h=48$ )
HPI coefficient	-0.0567 (0.0467)	-0.0354 (0.0382)	-0.0129 (0.0293)
Attenuation	—	37.6%	77.3%
$N$	50	50	50
$R^2$	0.146	0.389	0.612

*Notes:* Column (1) regresses average long-run log employment (months 48–120) on HPI boom with standard controls (log pre-recession employment, pre-recession growth, census region indicators). Columns (2) and (3) add the state-level unemployment rate change at  $h = 24$  and  $h = 48$  months, respectively, as mediators. Attenuation measures the fraction of the HPI coefficient absorbed by the duration-trap proxy:  $1 - \hat{\pi}_{\text{mediated}}/\hat{\pi}_{\text{baseline}}$ . HC1 standard errors in parentheses.

**Figure 5:** Duration-Trap Attenuation: How Much Does Unemployment Persistence Explain?

*Notes:* Bars show the HPI coefficient on long-run employment under different specifications.

Baseline includes standard controls. Adding unemployment persistence at  $h = 24$  months attenuates the coefficient, indicating that duration dynamics account for a substantial share of the scarring effect.

This attenuation is not mechanical: the UR mediator is itself predicted by housing exposure, consistent with the causal chain (HPI  $\rightarrow$  duration trap  $\rightarrow$  long-run scarring). The finding is consistent with the individual-level evidence from [Yagan \(2019\)](#), who shows that workers in harder-hit local labor markets experienced persistent earnings losses, and [Jarosch \(2023\)](#), who demonstrates that unemployment duration itself—not just the event of job loss—drives long-run career damage.

#### 4.4 Interpretation

The duration-trap mechanism has clear micro-foundations. [Kroft et al. \(2016\)](#) show that callback rates for job applicants decline sharply with unemployment duration, using a resume audit study—employers screen on duration even when applicant quality is held constant. [Elsby et al. \(2010\)](#) document that exit rates from unemployment to employment fall steeply over the spell, particularly after six months. [Pissarides \(1992\)](#) models this formally: when human capital depreciates during unemployment, a feedback loop emerges where longer durations reduce match quality, which further extends durations.

The empirical finding here—that controlling for unemployment persistence attenuates the Great Recession scarring coefficient by 38–77%—provides macro-level evidence that this micro mechanism operates at the state level through cross-state variation in recession exposure. The attenuation is not mechanical: it arises because housing-exposed states experienced both worse long-run employment outcomes *and* more persistent unemployment, and the latter explains a substantial share of the former.

The magnitude of the attenuation deserves comment. At  $h = 24$  months, the UR mediator absorbs 38% of the HPI coefficient. At  $h = 48$ , it absorbs 77%. This progression is itself informative: the longer the post-recession horizon at which unemployment persistence is measured, the more of the scarring effect it explains. This is exactly what the duration-trap hypothesis predicts—early unemployment may be transient (and therefore uninformative about scarring), while unemployment that persists for four years captures the workers who

have crossed the threshold into skill depreciation and labor force exit.

The duration trap also helps explain why some policy interventions work and others do not. Training programs and hiring subsidies targeted at the long-term unemployed directly address the trap’s mechanism—they restore human capital and overcome employer screening on duration (Card et al., 2018). Aggregate demand stimulus is most effective when it arrives before the trap activates, preventing duration accumulation rather than reversing it. The Great Recession’s timeline—ARRA enacted 14 months after the NBER peak, reaching full force at 24+ months—meant that the intervention arrived after millions had already crossed duration thresholds. The COVID fiscal response, by contrast, arrived within weeks and directly targeted match preservation, preventing the duration accumulation that would have activated the trap.

One important limitation of the mechanism evidence: the state-level analysis uses unemployment rate persistence as a proxy for the duration trap, not direct measures of long-term unemployment or reemployment flows at the state level. The national data confirm that the Great Recession produced unprecedented long-term unemployment while COVID did not, but the state-level attenuation exercise relies on the assumption that states with more persistent unemployment rate elevations are the states where more workers fell into the duration trap. This is plausible but not directly verified with individual-level data. Linking to worker-level panel data (e.g., LEHD, matched CPS) would allow a more direct test of whether state-level housing exposure predicts individual-level long-term unemployment and career damage.

## 5. Robustness

**Table 5:** Robustness: Window Choice, Controls, and Samples

Specification	Coef.	SE	$p$ -value	Perm. $p$
<i>Panel A: Window robustness</i>				
Months 36-96	-0.0602	(0.0460)	0.197	[0.128]
Months 48-96	-0.0607	(0.0481)	0.214	[0.142]
Months 48-120	-0.0567	(0.0467)	0.231	[0.128]
Months 60-120	-0.0557	(0.0480)	0.252	[0.168]
<i>Panel B: Control robustness</i>				
Baseline	-0.0567	(0.0467)	0.231	—
+ Industry shares	-0.0602	(0.0640)	0.352	—
<i>Panel C: Sample robustness</i>				
Drop Sand States ( $N=46$ )	-0.0245	(0.0446)	0.586	—

*Notes:* All specifications regress average long-run log employment change on the HPI boom instrument with log pre-recession employment, pre-recession employment growth, and census region indicators (matching the baseline in Table 2). Panel A varies the averaging window for the dependent variable. Panel B adds pre-2007 construction and manufacturing employment shares. Panel C drops Nevada, Arizona, Florida, and California. HC1 standard errors in parentheses. Permutation  $p$ -values in brackets (1,000 reassignments); — indicates permutation inference was not computed for that specification. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 5 reports three dimensions of robustness. Panel A shows that the HPI scarring coefficient is stable in magnitude across alternative averaging windows: whether I measure

long-run outcomes over months 36–96, 48–96, 48–120, or 60–120, the coefficient ranges from  $-0.056$  to  $-0.061$ , with permutation  $p$ -values between 0.13 and 0.17. The consistency in magnitude across windows, combined with the significant LP estimates at medium horizons, supports the persistence interpretation even though the single estimand is imprecise. Panel B adds pre-2007 construction and manufacturing employment shares to the control set; the HPI coefficient is unaffected, confirming that industry composition conditional on housing exposure does not drive the result. Panel C drops the four Sand States (Nevada, Arizona, Florida, California), which experienced the largest housing booms; the coefficient attenuates substantially (from  $-0.057$  to  $-0.025$ ), indicating that the scarring evidence is concentrated in the states most exposed to the housing demand collapse. This is consistent with the mechanism—the duration trap activates most strongly where the demand shock was most severe.

Pre-trends are flat at the most relevant horizon: regressing the 12-month employment change ending at the recession peak on HPI boom yields a small, insignificant coefficient ( $p = 0.44$ ), supporting the identifying assumption that housing exposure is uncorrelated with immediate pre-existing employment trajectories conditional on controls. At the 36-month pre-period, there is evidence of differential trends among high-HPI states ( $p = 0.009$ ), likely reflecting the housing construction boom itself—states with larger housing booms also experienced faster pre-recession employment growth as construction expanded. This is plausible—the housing boom *was* an economic expansion in exposed states—but it means the OLS estimates cannot fully rule out confounding from pre-existing differential growth. The Saiz IV and horse race results are therefore especially important: they isolate the exogenous geographic component of housing exposure from the endogenous construction boom. The significant pre-trend is a real limitation of the baseline OLS specification, not merely a feature of the data.

## 5.1 Saiz IV Confirmation

Instrumenting the housing boom with the [Saiz \(2010\)](#) supply elasticity yields comparable estimates through  $h = 48$  months. The IV point estimates are larger in magnitude than OLS at medium horizons, suggesting that OLS may understate the causal effect—consistent with attenuation bias from measurement error in state-level HPI. The first-stage  $F$ -statistic exceeds 11, and the IV estimate at  $h = 12$  is statistically significant.

## 5.2 Scope of the Claim

This paper compares two episodes—a demand recession and a supply recession—and finds that the persistence gap is tightly linked to duration dynamics. The claim is specific: in these two episodes, duration traps explain scarring. I do not claim that all demand recessions scar identically, or that all supply recessions are benign. Mixed-type recessions ([Guerrieri et al., 2022](#)) may produce intermediate dynamics. Extending the framework to earlier episodes (the Volcker recession, the 2001 dot-com bust) or to other countries would test generalizability.

The cross-state strategy may not fully recover aggregate effects if general equilibrium forces attenuate cross-state differences ([Beraja et al., 2019](#)). But this bias is conservative: aggregate scarring is at least as large as the cross-state estimates suggest. Linking to individual-level data would allow direct measurement of duration and participation channels at the worker level; this remains an important direction for future work.

## 6. Conclusion

Not all recessions are created equal—but the distinction that matters most is not depth, it is duration. The housing-driven Great Recession trapped workers in prolonged nonemployment that eroded their human capital and labor force attachment. The COVID supply recession, despite a peak contraction roughly three times as severe, preserved employer-employee matches and enabled rapid recall. The permanent damage came not from the fall but from the failure

to get back up.

Three policy implications follow. First, the speed of fiscal response matters enormously. The Great Recession’s delayed response allowed millions to cross duration thresholds that trigger scarring; COVID’s rapid response preserved matches and prevented hysteresis. Every month of delayed intervention during a demand recession is disproportionately costly. Second, the *type* of response matters. Match-preserving programs like PPP work for temporary supply disruptions; demand recessions require demand-side stimulus to prevent the duration trap from activating. Third, targeting the long-term unemployed—retraining, duration-conditional hiring subsidies, interventions against employer discrimination (Kroft et al., 2016)—may be particularly cost-effective because these are the workers caught in the trap.

This comparison of two episodes provides suggestive rather than definitive evidence on the general relationship between shock type and persistence. The comparison is ultimately a sample of two macroeconomic events, and the demand/supply taxonomy is tested on these specific episodes. Mixed-type recessions (Guerrieri et al., 2022) may produce intermediate dynamics. Extending the framework to earlier episodes (the Volcker recession, the 2001 dot-com bust) or to other countries would test generalizability. Linking to individual-level data—matched employer-employee records, LEHD, or longitudinal CPS panels—would allow direct measurement of duration and participation channels at the worker level, moving beyond the state-level UR proxies used here.

More broadly, the findings suggest that models and policy discussions focused on the *size* of employment losses may miss the key state variable. What matters for long-run outcomes is accumulated nonemployment duration—a stock variable that grows slowly during demand downturns and resolves quickly during supply disruptions. The hysteresis literature has established that recessions can leave permanent scars (Cerra et al., 2023); this paper’s contribution is to show that *which* recessions scar depends on whether they create the conditions for a duration trap. Preventing it—through rapid fiscal response, match preservation, and targeted interventions for the long-term unemployed—is the policy

imperative.

## **Acknowledgements**

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP), co-authored with Codex via the Duet collaboration system.

**Project Repository:** <https://github.com/SocialCatalystLab/ape-papers>

**Contributors:** @SocialCatalystLab, @dyanag

## References

- Autor, David, David Cho, Leland D Crane, Mita Goldar, Byron Luber, Joshua Montes, William B Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz,** “The \$800 Billion Paycheck Protection Program: Where Did the Money Go and Why Did It Go There?,” *Journal of Economic Perspectives*, 2022, *36* (2), 55–80.
- Autor, David H, David Dorn, and Gordon H Hanson,** “The China Syndrome: Local Labor Market Effects of Import Competition in the United States,” *American Economic Review*, 2013, *103* (6), 2121–2168.
- Beraja, Martin, Erik Hurst, and Juan Ospina,** “The Aggregate Implications of Regional Business Cycles,” *Econometrica*, 2019, *87* (6), 1789–1833.
- Blanchard, Olivier J and Lawrence H Summers,** “Hysteresis and the European Unemployment Problem,” *NBER Macroeconomics Annual*, 1986, *1*, 15–78.
- Blanchard, Olivier Jean and Lawrence F Katz,** “Regional Evolutions,” *Brookings Papers on Economic Activity*, 1992, *1992* (1), 1–75.
- Borusyak, Kirill, Peter Hull, and Xavier Jaravel,** “Quasi-Experimental Shift-Share Research Designs,” *Review of Economic Studies*, 2022, *89* (1), 181–213.
- Cajner, Tomaz, Leland D Crane, Ryan A Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz, and Ahu Yildirmaz,** “The US Labor Market during the Beginning of the Pandemic Recession,” *Brookings Papers on Economic Activity*, 2020, *2020* (2), 3–34.
- Card, David, Jochen Kluge, and Andrea Weber,** “What works? A meta analysis of recent active labor market program evaluations,” *Journal of the European Economic Association*, 2018, *16* (3), 894–931.

- Cerra, Valerie and Sweta Chaman Saxena**, “Growth Dynamics: The Myth of Economic Recovery,” *American Economic Review*, 2008, *98* (1), 439–457.
- , **Antonio Fatás, and Sweta Chaman Saxena**, “Hysteresis and Business Cycles,” *Journal of Economic Literature*, 2023, *61* (1), 181–225.
- Charles, Kerwin Kofi, Erik Hurst, and Matthew J Notowidigdo**, “The Masking of the Decline in Manufacturing Employment by the Housing Bubble,” *Journal of Economic Perspectives*, 2016, *30* (2), 179–200.
- Chetty, Raj, John N Friedman, Nathaniel Hendren, Michael Stepner, and the Opportunity Insights Team**, “How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data,” *NBER Working Paper*, 2020, (27431).
- Dao, Mai, Davide Furceri, and Prakash Loungani**, “Regional Labor Market Adjustment in the United States: Trend and Cycle,” *Review of Economics and Statistics*, 2017, *99* (2), 243–257.
- Elsby, Michael WL, Bart Hobijn, and Aysegül Sahin**, “The Labor Market in the Great Recession,” *Brookings Papers on Economic Activity*, 2010, *2010* (1), 1–69.
- Fujita, Shigeru and Giuseppe Moscarini**, “Recall and Unemployment,” *American Economic Review*, 2017, *107* (12), 3875–3916.
- Ganong, Peter, Pascal Noel, and Joseph Vavra**, “US Unemployment Insurance Replacement Rates During the Pandemic,” *Journal of Public Economics*, 2020, *191*, 104273.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift**, “Bartik Instruments: What, When, Why, and How,” *American Economic Review*, 2020, *110* (8), 2586–2624.

- Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub, and Iván Werning**, “Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?,” *American Economic Review*, 2022, *112* (5), 1437–1474.
- Jarosch, Gregor**, “Searching for Job Security and the Consequences of Job Loss,” *Econometrica*, 2023, *91* (3), 855–898.
- Jordà, Òscar**, “Estimation and Inference of Impulse Responses by Local Projections,” *American Economic Review*, 2005, *95* (1), 161–182.
- Kroft, Kory, Fabian Lange, Matthew J Notowidigdo, and Lawrence F Katz**, “Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Nonparticipation,” *Journal of Labor Economics*, 2016, *34* (S1), S7–S54.
- Mian, Atif, Amir Sufi, and Kamalesh Rao**, “Household Balance Sheets, Consumption, and the Economic Slump,” *Quarterly Journal of Economics*, 2013, *128* (4), 1687–1726.
- **and** —, “What Explains the 2007–2009 Drop in Employment?,” *Econometrica*, 2014, *82* (6), 2197–2223.
- Pissarides, Christopher A**, “Loss of Skill During Unemployment and the Persistence of Employment Shocks,” *Quarterly Journal of Economics*, 1992, *107* (4), 1371–1391.
- Saiz, Albert**, “The Geographic Determinants of Housing Supply,” *Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.
- Yagan, Danny**, “Employment Hysteresis from the Great Recession,” *Journal of Political Economy*, 2019, *127* (5), 2505–2558.

## Appendix Overview

[A](#) Data Appendix

[B](#) Full Dynamic Estimates

[E](#) Additional Robustness

### A. Data Appendix

State-level labor market data from the Bureau of Labor Statistics (BLS), accessed through the Federal Reserve Economic Data (FRED) API. Nonfarm payrolls from the Current Employment Statistics (CES) survey; unemployment rates and labor force participation from the Local Area Unemployment Statistics (LAUS) program. House price indices from the Federal Housing Finance Agency (FHFA). Industry employment from CES by major supersector. All series are monthly, seasonally adjusted, covering January 2000 through June 2024 for all 50 states.

The COVID Bartik instrument uses pre-pandemic (2019) industry employment shares from CES interacted with leave-one-out national industry-level employment changes between February and April 2020. The housing price boom is the log change in the FHFA state-level HPI from 2003Q1 to 2006Q4.

### B. Full Dynamic Estimates

[Table 6](#) reports the full local projection estimates at all horizons for both episodes, with HC1 standard errors and permutation  $p$ -values.

**Table 6:** Local Projection Dynamic Estimates*Panel A: Great Recession (HPI exposure)*

Horizon	Coef.	SE	Perm. $p$
$h=3$	-0.0102*	(0.0054)	[0.008]
$h=6$	-0.0254**	(0.0100)	[0.001]
$h=12$	-0.0481**	(0.0221)	[0.002]
$h=24$	-0.0473	(0.0349)	[0.068]
$h=36$	-0.0549	(0.0398)	[0.066]
$h=48$	-0.0608	(0.0448)	[0.096]
$h=60$	-0.0606	(0.0492)	[0.143]
$h=84$	-0.0688	(0.0527)	[0.157]
$h=120$	-0.0449	(0.0521)	[0.286]
$N$		50	

*Panel B: COVID-19 (standardized Bartik exposure)*

Horizon	Coef.	SE	Perm. $p$
$h=3$	0.0301***	(0.0037)	[0.000]
$h=6$	0.0197***	(0.0043)	[0.000]
$h=12$	0.0153***	(0.0038)	[0.001]
$h=24$	0.0052*	(0.0029)	[0.052]
$h=36$	0.0024	(0.0022)	[0.296]
$h=48$	0.0019	(0.0028)	[0.423]
$N$		50	

*Notes:* Each row reports the local projection coefficient at horizon  $h$  months from a cross-state regression of log employment change on recession exposure. GR exposure is the 2003–2006 housing price boom (log change); COVID exposure is the standardized Bartik shock. COVID horizons end at

## C. Instrumental Variables and Horse Race

Table 7 reports the Saiz IV estimates for the Great Recession at selected horizons. Table 8 reports the within–Great Recession horse race between the housing price boom and the GR-era Bartik instrument.

**Table 7:** Instrumental Variable Estimates: Saiz Housing Supply Elasticity

	$h = 12$	$h = 24$	$h = 48$
<i>OLS</i>	−0.0481 (0.0221)	−0.0473 (0.0349)	−0.0608 (0.0448)
<i>IV (Saiz)</i>	−0.0874 (0.0392)	−0.0891 (0.0600)	−0.1052 (0.0827)
First-stage $F$	11.4		
$N$	50		

*Notes:* OLS and 2SLS estimates of the effect of the 2003–2006 housing price boom on log employment change at horizons  $h=12, 24,$  and 48 months. The instrument is the negative of the Saiz (2010) state-level housing supply elasticity, which predicts housing price booms through geographic constraints on construction. Controls: log pre-recession employment, pre-recession employment growth, census region indicators. HC1 standard errors for IV in parentheses.

**Table 8:** Horse Race: Housing Price Boom vs. Great Recession Bartik Shock

	$h=6$	$h=12$	$h=24$	$h=48$
<i>HPI boom</i>	-0.0357 *** (0.0108)	-0.0768 *** (0.0230)	-0.0924 ** (0.0357)	-0.1225 *** (0.0434)
<i>GR Bartik</i>	0.4614 ** (0.1782)	1.2827 *** (0.4720)	2.0147 *** (0.6810)	2.7529 *** (0.9162)
$N$	50			

*Notes:* Both HPI boom and GR Bartik shock are included simultaneously in each regression. The dependent variable is the log employment change at horizon  $h$  months from the Great Recession peak. The GR Bartik shock is constructed from pre-2007 industry employment shares weighted by leave-one-out national industry employment changes during 2007–2009. Controls: log pre-recession employment, pre-recession employment growth, census region indicators. HC1 standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## D. Pooled Interaction

**Table 9:** Pooled Interaction: Is Great Recession Scarring Larger Than COVID?

Avg. Log Employment Change	
Exposure (SD)	−0.0039 (0.0048)
Exposure × GR	−0.0029 (0.0068)
$N$	100

*Notes:* Pooled regression stacking GR and COVID observations. Exposure is standardized to unit variance within each episode. The interaction term (Exposure × GR) tests whether GR scarring exceeds COVID scarring. A significant negative coefficient indicates the Great Recession produced larger long-run employment deficits per unit of exposure. HC1 standard errors. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

The interaction term (Exposure × GR) is negative but statistically insignificant ( $p > 0.6$ ), reflecting limited power and the fact that the outcome windows differ across episodes (months 48–120 for GR, 24–48 for COVID). The episode-specific regressions in the main text are more informative.

## E. Additional Robustness

Window robustness, control sensitivity, and sample robustness are reported in [Table 5](#) in the main text.

**Table 10:** Structured Data Extract (SDE)

---

Field	Value
Paper ID	a pep_0238_v10
Title	Demand Recessions Scar, Supply Recessions Don't
Method	Local Projections (cross-sectional)
Country	United States
Unit	State (N=50)
Period	January 2000 – June 2024
Outcome	Long-run employment (avg months 48–120)
Treatment	Housing price boom (GR) / Bartik shock (COVID)
Main finding	GR exposure → persistent employment deficit; COVID → full recovery
Mechanism	Duration trap: prolonged unemployment erodes human capital

---