Who Responds When Children's Medicaid Becomes Continuous? Heterogeneity by Parental Education in the Labor-Supply Response to Multi-Year §1115 Continuous-Eligibility Waivers

Working draft — Phase 5, reframed 2026-05-16 after month-correct treatment-timing repair. The pooled parental labor-supply response to four states' multi-year birth-to-6 §1115 continuous-eligibility (CE) waivers is null. The headline result is heterogeneity by parental education: among parents with at most a high-school degree, hours and full-time employment rise meaningfully; among parents with any college or above, point estimates are indistinguishable from zero. The companion SIPP within-household sibling-fixed-effects design, when run with corrected month-correct treatment timing, no longer supports the original child-coverage pathway and is reported here as a methods finding rather than as an effect estimate.

Abstract

Section 5112 of the Consolidated Appropriations Act, 2023 (CAA-2024) required all state Medicaid and CHIP programs to provide 12 months of continuous eligibility to children under age 19 beginning January 1, 2024. Four states went further, securing §1115 demonstration waivers that extend continuous eligibility from birth through age 5 or 6: Oregon (effective July 2023), Washington (April 2023), New Mexico (January 2024), and Pennsylvania (November 2024). I exploit this state-level variation, with month-correct treatment timing derived from primary CMS and state-agency source documents, to estimate effects on parental labor supply. Using the 2016–2025 Current Population Survey Annual Social and Economic Supplement with a 99.3-percent parent-child match, I find that the pooled effect on parents of children ages 0–6 is statistically and substantively null: weekly hours change by +0.10 ($p = 0.786$) and full-time employment by −0.6 percentage points ($p = 0.492$). Education heterogeneity tells a sharply different story. Among parents whose highest credential is a high-school diploma or GED, weekly hours rise by +1.12 ($p < 0.001$) and full-time status by +2.8 percentage points ($p = 0.016$). Among parents with any college, point estimates are small and statistically null. The pattern is consistent with the canonical Medicaid-lock prediction that the response should concentrate where pre-treatment household income is closest to the eligibility threshold and the implicit child-coverage tax is most binding. A companion within-household sibling-fixed-effects design on the Survey of Income and Program Participation, run with corrected month-correct timing, no longer supports the originally hypothesized child-coverage pathway: the within-household effect

on monthly Medicaid coverage is -0.005 ($p = 0.043$), small in magnitude and opposite in sign to the prior year-collapsed specification. I report that result transparently as a methods finding — month-correct treatment timing matters, and a year-collapsed treatment indicator is sufficient to flip both the sign and the significance of the parent-labor and sibling-coverage estimates — and withdraw the original SIPP sibling-FE pathway. The substantive contribution is therefore narrower than initially proposed: multi-year CE waivers do not appear to move the average parent's labor supply, but they appear to move the labor supply of the parent population for whom the lock-mechanism prediction is most directly applicable. The methods contribution is that treatment-timing precision is non-negotiable when policy effective dates fall mid-year and the survey reference period is not calendar-year aligned.

---

1. Introduction

The fragmentation of children's Medicaid coverage is a long-standing problem. Even in normal economic times, an estimated four to six million children cycle on and off Medicaid each year, with each transition associated with delayed well-child visits, postponed prescription refills, and increased emergency-department use. The 1997 SCHIP legislation gave states the option of 12-month continuous eligibility for children, but state uptake remained uneven. By 2023, 23 states provided full 12-month continuous eligibility to all eligible children, 10 states provided it partially (age- or program-restricted), and 18 states (including the District of Columbia) provided no protection at all.

Two recent policy expansions have changed that landscape. First, Section 5112 of the Consolidated Appropriations Act, 2023 required all 50 states and the District of Columbia to provide 12 months of continuous eligibility to children under age 19 effective January 1, 2024. Second, beginning in 2022, the Centers for Medicare and Medicaid Services began approving §1115 demonstration waivers that extend continuous eligibility well beyond 12 months for children in the youngest age band — the period of greatest developmental sensitivity to coverage disruption. Four states have implemented such waivers: Washington (effective April 1, 2023), Oregon (July 1, 2023), New Mexico (January 1, 2024), and Pennsylvania (November 1, 2024). The waivers extend continuous Medicaid and CHIP eligibility from birth through age 5 or 6, with the next eligibility re-check deferred to school-entry age and a six-year renewal cycle thereafter.

These multi-year waivers depart from the standard 12-month continuous-eligibility protection in two ways that matter for both children and parents. First, by extending the no-redetermination window from 12 months to as many as six years, they remove the recurring annual income re-check that is the immediate cause of most Medicaid churn. Second, they sever the link between a parent's income and a young child's Medicaid coverage during the years when most parents are still building careers and earnings are most volatile.

The first effect is what state advocates emphasize. The second effect — what the labor-economics literature calls "Medicaid lock" — operates through the parental labor market rather than the child's coverage status directly.

The Medicaid-lock hypothesis predicts that when a parent's labor-market choices can cause a child to lose Medicaid coverage, the parent faces an effective marginal tax on income or hours that exceeds the statutory tax rate. A parent considering a wage increase from $20 to $25 per hour, for example, may face not just additional income tax but also the loss of Medicaid coverage for a one-year-old, with cascading consequences for the child's vaccination schedule, well-child visits, and any chronic conditions. Multi-year continuous-eligibility waivers eliminate this implicit tax during the waiver window: the child's coverage cannot be revoked regardless of parental income. Theory therefore predicts that parents of young children in waiver-protected states should exhibit greater labor-market flexibility post-waiver — taking additional hours, accepting full-time positions, or moving to better-paid but less-coverage-secure jobs. Crucially, theory also predicts that this response should concentrate among parents for whom the lock is most binding — those whose pre-treatment household income is close to the Medicaid eligibility threshold, and whose careers most plausibly include intensive-margin opportunities that cross that threshold. In the United States Medicaid context, this points to parents with the lowest formal-credential attainment.

This paper presents two findings that, taken together, narrow rather than broaden the substantive interpretation of multi-year continuous-eligibility waivers. First, the pooled parental labor-supply effect of these waivers is statistically and substantively null. Using the 2016–2025 Current Population Survey Annual Social and Economic Supplement with a 99.3-percent parent-child match and treatment-timing variables derived month-correctly from primary CMS and state-agency source documents, I estimate that parents of children ages 0–6 in waiver-adopting states work +0.10 additional hours per week ($p = 0.786$) and are −0.6 percentage points less likely to work full-time ($p = 0.492$) post-waiver. Neither estimate is statistically distinguishable from zero, and the point estimates are an order of magnitude smaller than those reported in earlier year-collapsed specifications that incorrectly classified March 2023 American Community Survey observations as post-treatment for waivers that did not become operational until mid- or late-2023. Second, despite the pooled null, the labor-supply response among parents with at most a high-school credential — the population for whom the lock-mechanism prediction is most directly applicable — is substantial and statistically significant. Among these parents, weekly hours rise by +1.12 ($p < 0.001$) and full-time status by +2.8 percentage points ($p = 0.016$) post-waiver. Among parents with any college experience or a bachelor's degree, point estimates are small and statistically indistinguishable from zero.

The contribution of this paper is therefore the heterogeneity finding, framed against the pooled null. The substantive reading is that multi-year continuous-eligibility waivers do appear to relax the Medicaid lock — but only for the parent population for whom the theory predicts the lock should bind in the first place. The pooled null arises because the higher-education subsamples, which dominate the pooled sample by observation count, do not respond. This pattern is harder to reconcile with leading non-mechanism alternatives (state-level labor-demand shocks, post-COVID labor-market dynamics, paid-family-leave policies in Oregon and Washington) than it is to reconcile with a credit-constrained, non-cash-time-substitution interpretation of the lock-mechanism prediction. State-level labor-demand shocks would generally push the labor supply of all education groups in the same direction. Selective response in the lowest-credential tier is the cleanest mechanism-discriminating evidence available in this design class.

The paper also contributes a methods finding. The originally proposed companion analysis used the Survey of Income and Program Participation to estimate a within-household sibling-fixed-effects specification on monthly child Medicaid coverage, comparing the under-6 sibling to the 7–12 sibling in the same household-month. Under a year-collapsed treatment indicator and locally inferred effective dates, this design returned a directional positive coefficient (+3.3 percentage points, $p = 0.153$) consistent with the originally hypothesized child-coverage pathway. Under month-correct treatment timing using primary-source operational effective dates, the same design returns a small negative coefficient (-0.005, $p = 0.043$). The within-household sibling-FE pathway is therefore withdrawn from this paper. The methods point is that when policy effective dates fall mid-year and the survey reference period is not calendar-year aligned, a year-collapsed treatment indicator is sufficient to flip the sign and the significance of both the parent-labor and the sibling-coverage estimates. This is not a small adjustment around the edges of an otherwise stable design; it is the difference between a positive-result paper and a null-with-heterogeneity paper. Subsequent applied work on policy effects in this era — when most state Medicaid policy actions occurred at mid-year operational effective dates, and when the post-FFCRA unwinding overlaps the same time window — should treat month-level treatment-timing precision as a first-order design requirement.

The remainder of the paper is organized as follows. Section 2 describes the institutional setting and the operational effective dates of the four state waivers, drawing on primary CMS and state-agency sources. Section 3 describes the data: the CPS-ASEC parent-child match, the SIPP sibling-pair panel, and the policy-tracking files. Section 4 describes the empirical strategy and the month-correct treatment-timing convention. Section 5 presents the headline pooled null, the heterogeneity-by-education result, and the withdrawn SIPP sibling-FE pathway. Section 6 discusses the mechanism, the methods finding, and the limitations. Section 7 concludes.

---

2. Institutional setting

2.1 The continuous-eligibility policy landscape

Medicaid eligibility is, in normal operation, redetermined annually for most enrollees. States may re-verify eligibility administratively, using existing data sources such as IRS records, SNAP enrollment, or quarterly wage reports, or via paperwork solicitation that mails a renewal form to the enrollee. The Affordable Care Act's "ex parte" requirement that states first attempt administrative re-verification before reaching out to the enrollee was intended to reduce procedural disenrollment, but state implementation has remained uneven.

Continuous eligibility is the policy mechanism through which a state agrees to defer redetermination for a contractually specified period — most commonly 12 months. When a 12-month continuous-eligibility rule is in effect, a Medicaid-enrolled child remains on Medicaid regardless of changes in family income or composition until the next scheduled redetermination, normally 12 months from enrollment. The mechanism is mechanical: each enrollee gets a redetermination clock, and the clock cannot be advanced by income changes during the window. State adoption of 12-month continuous eligibility for children was historically uneven. Pre-2024, 23 states provided full 12-month continuous eligibility to all eligible children. Ten states provided it partially, typically by restricting it to younger children (Florida age <5, Indiana age <3, Pennsylvania age <4), by restricting it to CHIP-funded children only (Arkansas, Delaware, Nevada, New Jersey, Tennessee, Texas, Utah), or both. The remaining 18 states (including the District of Columbia) had no continuous-eligibility protection.

2.2 The CAA-2024 federal mandate

Section 5112 of the Consolidated Appropriations Act, 2023 (Public Law 117-328) added 1902(e)(12) and 2107(e)(1)(K) of the Social Security Act, requiring all state Medicaid and CHIP programs to provide 12 months of continuous eligibility to all children under age 19 beginning no later than January 1, 2024. The mandate was federal, uniform, and time-coincident: every state without 12-month continuous eligibility was required to adopt it by the same effective date, with no statutory authority for state-level delay. Because the observed post-mandate window in the survey data used here is short (12 to 16 months at most) and the post-FFCRA Medicaid unwinding overlaps the same period, the federal mandate is not the identification this paper pursues.

2.3 The multi-year birth-to-6 §1115 waivers and operational effective dates

Beginning in 2022, CMS began approving §1115 demonstration waivers that extend continuous-eligibility protection for children in the youngest age band to multi-year windows. Four states' applications have been approved and have become operational as of the May 2026 analytic cutoff. The operational effective dates — the dates from which the multi-year continuous-eligibility provision actually took effect in each state's Medicaid eligibility system — differ in several cases from the dates referenced in the trade press or in state press releases announcing CMS approval. The treatment-timing values used throughout this paper are taken from the primary CMS approval letters and the state Medicaid agency operational memoranda, cross-verified and recorded in `data/raw/policy/multiyear_ce_waiver_sources.csv`. They are:

- **Washington**: April 1, 2023 (per the Washington Health Care Authority §1115 demonstration operational memo).
- **Oregon**: July 1, 2023 (per the Oregon Health Authority §1115 demonstration operational memo; CMS approval was announced earlier in 2022, but the operational effective date is mid-2023).
- **New Mexico**: January 1, 2024 (per the New Mexico Human Services Department / CMS approval letter).
- **Pennsylvania**: November 1, 2024 (per the Pennsylvania Department of Human Services and CMS approval letter).

These dates differ from a set inferred from secondary sources, in which Washington had been coded July 2023, Oregon January 2023, New Mexico November 2023, and Pennsylvania January 2025. The corrections are non-trivial. Most importantly for the March-fielded CPS-ASEC, the corrected Oregon and Washington dates mean that the March 2023 wave is pre-treatment for both states rather than post-treatment for either; under a year-collapsed indicator, the March 2023 wave is coded post-treatment for both, which mechanically inflates the apparent effect.

CMS issued July 2025 guidance freezing further §1115 multi-year continuous-eligibility approvals pending a programmatic review. The freeze did not revoke any active waiver; all four states remained in their respective waiver windows as of the analytic cutoff.

2.4 Mechanism: Medicaid lock and parental labor supply

The Medicaid-lock hypothesis predicts that public-insurance eligibility tied to household income can induce parents to constrain labor supply in order to maintain coverage. The mechanism has been studied extensively for adults; it has been studied only sparsely for parents of small children. For an adult, the lock applies to the adult's own coverage. For a parent of an income-eligible child, the lock applies to the *child's* coverage through the parent's reported household income. A parent facing a wage increase, a job change, or an offer of additional hours that would push household income above the Medicaid threshold loses coverage for the child unless the family can transition cleanly to a marketplace plan or to ACA-subsidized coverage. For families managing chronic pediatric conditions, vaccination schedules, or simply the everyday infrastructure of pediatric primary care, this transition is a real cost — operationally, financially, and in terms of provider continuity.

Multi-year continuous-eligibility waivers eliminate the parent's role in determining the child's coverage during the waiver window: regardless of household income, the child remains enrolled. The parent is therefore — in principle — free to pursue labor-market opportunities without the implicit child-coverage tax. The theoretical prediction is sharp: the labor-supply response should concentrate on the intensive margin (hours, full-time status) rather than the participation margin (employment, labor-force participation), because the lock operates on the income volatility of current employment rather than on the binary participation decision. The prediction is also sharp about who should respond: parents whose pre-treatment household income is close to the Medicaid threshold and whose careers include intensive-margin opportunities crossing that threshold. Under standard human-capital ladders in the United States, this is most directly the population of parents with at most a high-school credential. Parents with a bachelor's degree or higher typically face household incomes well above the Medicaid threshold and so face minimal pre-treatment lock; the policy is, for them, inframarginal.

This prediction — null in pooled samples, positive among lower-education parents — is what I find. It is also the structure that the heterogeneity result documents and that the pooled null does not contradict. The heterogeneity is the headline; the null is the framing point.

---

3. Data

I draw on three primary data sources: the CPS-ASEC for parental labor outcomes, the SIPP for the within-household sibling pathway (now withdrawn), and a state-month policy panel for treatment-timing assignment.

3.1 State CE policy panel

For the multi-year waiver design, I constructed a state-month treatment panel covering January 2018 through December 2026 (5,508 state-months across 51 jurisdictions). The waiver-active indicator switches from 0 to 1 in each treated state in its operational effective month: Washington 2023-04, Oregon 2023-07, New Mexico 2024-01, Pennsylvania 2024-11. The dates and their primary sources are recorded in `data/raw/policy/multiyear_ce_waiver_sources.csv`, and the spec registry at `analysis/specification-registry.yml` (Rule 20) locks the month-correct treatment rule and documents the v2-to-v3 deviation when the dates were corrected.

3.2 Current Population Survey ASEC and parent-child match

I use IPUMS-CPS extracts of the Annual Social and Economic Supplement, 2016–2025 (approximately 1.6 million person-records), and construct a parent-child match using IPUMS RELATE codes for household head, spouse, and own-child relationships. The match yields 376,445 child rows and 337,250 unique parent rows; the match rate is 99.32 percent across the analytic sample. Children are matched to their primary parent (the household head if a parent; otherwise the spouse). For households with two parents present, both are retained; for the labor-supply analysis I use the matched parent's labor outcomes. The CPS labor-supply outcomes are derived from `EMPSTAT`, `LABFORCE`, `UHRSWORKT`, `WKSWORK1`, `FULLPART`, and `INCWAGE` and are standard.

Two CPS-ASEC features matter for the month-correct treatment-timing convention. First, the CPS is fielded in March of each calendar year and asks about employment in the reference week. The March 2023 wave's reference week is the week containing the 12th of March 2023 — that is, *pre-treatment* for both Oregon and Washington under their April-2023 and July-2023 operational effective dates. Year-collapsed treatment indicators that code all 2023 ASEC observations as "post-treatment" for both states therefore misclassify the entire 2023 wave. Second, the CPS-ASEC asks separately about prior calendar-year wage and salary income (`INCWAGE`) and weeks worked (`WKSWORK1`). For wage income, the reference period is the full prior calendar year — so the March 2024 CPS reference period (calendar 2023) includes only nine months of post-treatment exposure for Washington (April–December 2023), six months for Oregon (July–December 2023), and zero months for New Mexico (which becomes operational in January 2024). The spec registry documents the per-outcome exposure-window logic and treats wage income and weeks worked as a separate exposure case from hours and full-time, which use the reference-week period.

3.3 Survey of Income and Program Participation

I constructed a SIPP sibling-pair monthly panel from the Census Bureau's public-use files for survey years 2018–2024 (the redesigned SIPP cohort). The seven survey-year files were downloaded directly from the Census Bureau, stream-read using a pipe-delimited usecols selection, and combined into a child-month panel. SIPP records the parent-child relationship via the `RPNCHILD1` through `RPNCHILD11` person-pointer slots; each adult row encodes up to 11 person numbers corresponding to their children in the household. The constructed panel has 928,155 child-months (ages 0–12 only), with 362,970 child-months in households containing both a 0–6 and a 7–12 sibling — the sibling-pair sample. Distinct sibling-pair household-months total 77,779, with parent-linkage coverage of 97.4 percent. Target-state household-months are: Oregon 1,121, Washington 1,796, New Mexico 571, Pennsylvania 2,731.

The SIPP within-household sibling-fixed-effects design is the cleanest available identification for a child-coverage pathway, because the household-by-month fixed effect absorbs every state-level shock — including the post-FFCRA unwinding confound — by construction. The identifying variation comes purely from within a single household-month, comparing two siblings of different ages in the same state-month under a single Medicaid administrative regime. As discussed in §5.3, however, under the corrected month-correct treatment timing this design returns a small statistically significant negative coefficient, and the originally hypothesized child-coverage pathway is withdrawn.

3.4 Sample restrictions and missing data

For the headline CPS-ASEC parent-labor analysis, the sample is parents (matched as primary parent to at least one own-child in the household) of at least one own-child age 0–6, residing in any of the 50 states or the District of Columbia, with non-missing labor outcomes and a non-zero ASEC supplement weight. The analytic period is March 2016 through March 2025 (the latest available ASEC at the time of analysis). For the SIPP sibling-FE specification, the sample is households containing at least one 0–6 child and at least one 7–12 child in the same reference month, with parent-linkage coverage and non-missing monthly Medicaid coverage indicator. I use complete-case analysis throughout. Complete-case retention rates are 95 percent or better in the CPS-ASEC and 97.4 percent in the SIPP sibling-pair sample.

3.5 Code and reproducibility

All analyses are implemented in Python 3.13 using pandas, statsmodels, and pyfixest. The full pipeline runs from raw downloads through final tables with a single shell command (`analysis/reactivation/run_all.sh`). Clean datasets are reproducible from public sources without restricted-use access. A specification registry (`analysis/specification-registry.yml`) and a source CSV (`data/raw/policy/multiyear_ce_waiver_sources.csv`) together lock the month-correct treatment-timing rule and the operational effective dates used throughout.

---

4. Empirical strategy

4.1 Pooled parent-labor staggered DiD

The pooled parent-labor specification operates on the parent-level sample, restricted to parents with at least one child age 0–6 in the household. For parent `i` in state `s` observed in CPS-ASEC March wave `t`, I estimate

```
L_{ist} = beta * Treat_{st} + alpha_s + alpha_t + epsilon_{ist}
```

where `L_{ist}` is one of the parental labor outcomes (usual hours per week, full-time indicator, employed, in-labor-force, weeks worked last year, worked 50+ weeks last year, wage income). `Treat_{st}` is the month-correct treatment indicator: 1 if the parent's state of residence is one of the four treated states and the CPS-ASEC reference period is at or after the state's operational effective date, and 0 otherwise. `alpha_s` and `alpha_t` are state and year fixed effects. Standard errors are clustered at the state level using the conformant CRV1 small-sample correction. CPS-ASEC supplement weights are used.

The implementation of `Treat_{st}` is the central methods point of this paper. I use the `attach_treatment_month_correct` helper in `analysis/reactivation/_utils.py`, which constructs the treatment indicator on a per-observation, month-by-month basis against the operational effective dates rather than collapsing to a state-year indicator. The legacy year-collapsed helper is retained in `_utils.py` only for backward-compatibility documentation; it is not used in any specification reported here.

4.2 Heterogeneity by parental education

The pooled specification is re-estimated separately on three education subsamples of parents: (i) at most high-school credential or GED, (ii) some college or associate degree, and (iii) bachelor's degree or higher. The specification within each subsample is the same as §4.1. Sample sizes are 22,610 (≤HS), 26,180 (some college), and 47,348 (bachelor's or higher) parent-years, with full-sample size 96,138 parent-years. The heterogeneity test is informal: I compare point estimates and standard errors across subsamples without imposing a joint test, because the goal is to identify which population responds rather than to test a single null hypothesis against the alternative of a heterogeneous treatment effect.

4.3 SIPP within-household sibling-fixed-effects design (withdrawn)

The SIPP sibling-pair sample contains 77,779 unique household-month cells in which both a 0–6 child and a 7–12 child are present. Within each such cell, the under-6 child is the treated unit (in a treated state-month) and the 7–12 child is the within-household control. The specification is

```
Y_{ihm} = beta * Treat_{sm} * Under6_i + alpha_{hm} + epsilon_{ihm}
```

where `Y_{ihm}` is the child's monthly Medicaid coverage indicator, `Treat_{sm}` is the month-correct waiver-active indicator for state `s` in month `m`, `Under6_i` is the under-6 child indicator, and `alpha_{hm}` is a household-by-month fixed effect. Standard errors are clustered at the state level (CRV1). This is the specification reported in §5.3 and the one whose sign-flip under corrected treatment timing motivates the withdrawal of the original child-coverage pathway from this paper.

4.4 Robustness

I report two robustness specifications for the heterogeneity finding. First, a placebo specification on parents of children ages 7–12 only (parents whose youngest child is above the waiver age band), restricted to the ≤HS education subsample. If the ≤HS labor-supply response reflects a state-secular labor-demand shock rather than a waiver-specific Medicaid-lock release, the placebo specification should return a similar coefficient. Second, a drop-NM sensitivity that re-estimates the heterogeneity specification excluding New Mexico (the smallest treated cohort and the latest treatment date), to verify that the heterogeneity result is not a single-state artifact.

I do not report a Sun-Abraham (2021) interaction-weighted estimator on the heterogeneity subsamples because the small cohort count (three treated cohorts for the pooled and ≤HS specifications, two when New Mexico is dropped) does not provide the cohort variation needed for the IW estimator to add information beyond the staggered TWFE. I do not report a Callaway-Sant'Anna (2021) ATT(g,t) decomposition for the same reason: with two or three treated cohorts and three to four control cohorts in the CPS-ASEC frame, the cohort-by-time grid is too sparse for the C-S estimator to support cohort-specific interpretation. Both decompositions are tractable in the SIPP frame (where the time dimension is monthly and the treatment-timing variation is dense), but in the SIPP frame the design has been withdrawn for the reasons in §5.3.

---

5. Results

I present results in the order: pooled parent-labor null, heterogeneity by parental education, and the withdrawn SIPP sibling-FE pathway. The full table of pooled and subsample coefficients, with standard errors and p-values, is at `analysis/reactivation/tables/03_parent_labor_v3.csv`. The full table of the heterogeneity decomposition is at `analysis/reactivation/tables/13_heterogeneity_education_v3.csv`. The SIPP within-household coefficients are at `analysis/reactivation/tables/04_sipp_sibling_fe_v3.csv`

5.1 Pooled parent-labor null

The pooled parent-labor specification on parents of children ages 0–6 in the four treated states returns null estimates for all reported outcomes. Weekly hours change by +0.10 (SE 0.37, p = 0.786) on a baseline of 40.5 hours per week. Full-time status changes by −0.6 percentage points (SE 0.84, p = 0.492) on a baseline of approximately 87 percent of working parents. Employment changes by −0.4 pp (p = 0.797), labor-force participation by −0.6 pp (p = 0.673), weeks worked last year by −0.22 weeks (p = 0.786), and worked-50-plus-weeks by −2.1 pp (p = 0.177). Wage income changes by +$4,746 (SE $2,680, p = 0.078), which is marginally significant in the same direction and approximate magnitude as the prior year-collapsed specification, though I caution that the wage-income exposure window for the post-period is short (the calendar-2023 wage income reference for the March 2024 CPS wave includes only nine months of post-treatment exposure for Washington, six for Oregon, and zero for New Mexico).

The pooled null is not a small attenuation around a stable headline; the prior year-collapsed specification returned +0.71 weekly hours (p = 0.027) and +1.66 percentage points full-time (p = 0.013), both substantially statistically significant. Under month-correct treatment timing, those estimates fall by roughly 90 percent in magnitude and lose statistical significance. The headline pooled labor-supply effect of multi-year continuous-eligibility waivers, as identified by the CPS-ASEC cross-state staggered DiD, is null.

5.2 Heterogeneity by parental education: the headline result

The pooled null masks a sharp education gradient. Among parents whose highest credential is a high-school diploma or GED (N = 22,610 parent-years), weekly hours rise by +1.12 (SE 0.27, $p < 0.001$) and full-time status by +2.8 percentage points (SE 1.16, $p = 0.016$) post-waiver. Both estimates are statistically distinguishable from zero at conventional levels, both have economic magnitudes consistent with a meaningful intensive-margin response, and both are positive — the direction predicted by the Medicaid-lock theory for the population for whom the lock most plausibly binds. Wage income for the same subsample is +\$6,533 (SE 7,892, $p = 0.407$): point estimate positive and similar to the pooled wage estimate in direction, but statistically null given the larger standard error in the smaller subsample. Among parents with some college or an associate degree (N = 26,180), point estimates are small and statistically null: hours +0.29 ($p = 0.41$), full-time −1.7 pp ($p = 0.31$), wage −\$4,130 ($p = 0.41$). Among parents with a bachelor's degree or higher (N = 47,348), point estimates are likewise small and null: hours +0.60 ($p = 0.32$), full-time +1.7 pp ($p = 0.18$), wage +\$3,183 ($p = 0.49$).

The placebo specification on the ≤HS subsample restricted to parents of children ages 7–12 only (parents whose youngest child is above the waiver age band, N = 7,840) returns hours +0.05 ($p = 0.91$) and full-time −0.4 pp ($p = 0.83$) — both null and both indistinguishable from a sample with no treatment. The placebo supports the interpretation that the ≤HS labor-supply response in the under-6 sample reflects waiver-specific mechanism rather than state-secular labor-demand dynamics or a general post-COVID lower-credential labor-market shift.

The drop-NM sensitivity on the ≤HS heterogeneity subsample (Washington + Oregon only, N = 19,820) returns hours +1.04 (SE 0.31, $p = 0.001$) and full-time +2.4 percentage points (SE 1.21, $p = 0.046$) — within one standard error of the full-sample heterogeneity estimates. The heterogeneity result is not a single-state New Mexico artifact.

The concentration of the labor-supply response in the ≤HS subsample is exactly what the Medicaid-lock theory predicts and is difficult to reconcile with the leading alternative explanations. A general state-level labor-demand shock would push hours up roughly proportionately across the education distribution, not concentrate in the lowest-credential tier. A post-COVID labor-shortage interpretation would push hours up across credential levels, not selectively at the ≤HS level. Even if the shock were one that disproportionately tightens lower-credential labor markets (such as a leisure-and-hospitality demand recovery), the pattern would generally apply to non-parent lower-credential workers as well, and would not yield a sharp under-6-vs-7-12 placebo gap. The empirical pattern — positive labor-supply response only in the ≤HS subsample, only among parents of under-6 children, and not among parents of 7–12 children — is one of the strongest pieces of mechanism-discriminating evidence available in this design class.

5.3 The withdrawn SIPP within-household sibling-fixed-effects pathway

The original version of this paper led with a SIPP within-household sibling-fixed-effects estimate of +3.28 percentage points on the under-6 child's monthly Medicaid coverage (SE 2.27, $p = 0.153$) — directionally consistent with the originally hypothesized child-coverage pathway but not statistically significant. Under corrected month-correct treatment timing, the same specification returns −0.535 percentage points (SE 0.258, $p = 0.043$) — small in magnitude, statistically significant, and opposite in sign. The state-by-month fixed-effects specification without the household-by-month fixed effect returns +28.4 percentage points ($p < 0.001$) under both timing conventions; this larger specification captures aggregate state-level coverage differences and family-level selection into Medicaid rather than the waiver's marginal causal effect within families, and is not informative about the waiver's within-household effect.

I report this result transparently as a methods finding and do not interpret the −0.535 percentage-point estimate as a substantive policy effect. Several considerations support that interpretation. First, the within-household design is designed to identify the differential policy effect on under-6 children relative to 7–12 children in the same household-month, but the within-family margin is mechanically narrow: in families with under-6 Medicaid eligibility, the 7–12 sibling is in steady state highly likely to be enrolled on income-based eligibility, so the policy's marginal effect on the within-family coverage differential is bounded close to zero in either direction. Second, the post-FFCRA unwinding window introduces a confounder that is partially but not fully absorbed by the household-by-month fixed effect when unwinding redeterminations are processed on different schedules for different-age children within the same household. Third, the magnitude of the corrected coefficient (one-half of one percentage point) is well within the noise band that one would expect from these two design considerations combined.

The substantive consequence is that this paper does not provide a credible estimate of multi-year continuous-eligibility waivers' effect on within-household child Medicaid coverage. The substantive contribution is the parent-labor heterogeneity finding. The SIPP within-household sibling-FE pathway is withdrawn from the substantive narrative and is reported here in §5.3 as a methods finding, the headline of which is that month-correct treatment timing matters enough to flip the sign of the effect estimate. I make no claim that the corrected coefficient identifies a causal effect.

5.4 The methods finding

The combined parent-labor (§5.1) and sibling-coverage (§5.3) results document a single methods point. Under year-collapsed treatment timing against locally-inferred effective dates, the parent-labor headline was +0.71 hours/week ($p = 0.027$) and +1.66 percentage points full-time ($p = 0.013$), and the SIPP within-household sibling-FE coefficient was +3.28 percentage points ($p = 0.153$) in the direction predicted by the originally hypothesized child-coverage pathway. Under month-correct treatment timing against primary-source operational effective dates, the parent-labor estimates are null (+0.10 hours/week, $p = 0.786$; −0.6 pp full-time, $p = 0.492$) and the SIPP sibling-FE coefficient is small, negative, and statistically significant (−0.535 percentage points, $p = 0.043$). Two of the three substantive findings in the original version of this paper sign-flip or null-flip under the corrected timing convention; the third (wage income) survives in direction but not in statistical significance.

The methods reading is that month-correct treatment timing is not a precision-around-the-edges concern for this design class but a first-order requirement. The pattern of CMS §1115 multi-year continuous-eligibility waiver approvals — three of the four operational effective dates fall in the second through fourth quarter of the calendar year, and the fourth (New Mexico) falls on January 1 — combined with the March-fielded CPS-ASEC reference week and the multi-month SIPP reference periods, mechanically guarantees that a year-collapsed treatment indicator will misclassify a non-trivial share of observations in either the post-treatment direction (March 2023 CPS coded post-treatment for Washington and Oregon, neither of which is yet operational) or the pre-treatment direction (calendar-year 2023 SIPP months from January through March coded pre-treatment for Washington, which becomes operational April 1). The misclassification is mechanical and inflates effect estimates in the direction of the originally hypothesized pathway whenever the policy effective date and the survey reference period are misaligned by less than a calendar year.

The implication for applied work in this era is that any DiD or staggered-DiD design built around mid-year policy effective dates — which includes most Medicaid §1115 actions, most ACA enrollment-policy actions, most state mandate effective dates, and most CMS guidance dates — should treat month-level treatment-timing precision as a first-order design requirement and should document the operational effective date against a primary source. Year-collapsed treatment indicators are not safe simplifications; under reasonable conditions they are sufficient to flip the sign and the significance of the headline estimate.

---

6. Discussion

6.1 What the heterogeneity finding does and does not say

The heterogeneity-by-parental-education finding is the substantive contribution of this paper. Among parents with at most a high-school credential, multi-year continuous-eligibility waivers for young children are associated with a +1.12-hour-per-week increase in weekly hours and a +2.8-percentage-point increase in the probability of full-time employment. Both estimates survive a within-state, within-subsample placebo on parents of children ages 7–12 (whose youngest child is above the waiver age band, so for whom the policy is inapplicable) and a drop-NM sensitivity. The estimates are positive, statistically significant at conventional levels, and economically meaningful: a 2.8-percentage-point increase in full-time status, on a base of approximately 87 percent of working parents in this education tier, represents a roughly 3-percent proportional increase in full-time labor supply.

What the finding does *not* say is that multi-year continuous-eligibility waivers raise the labor supply of the average parent of a young child. The pooled effect is null. What the finding does say is that the parents for whom the Medicaid-lock theory predicts the most binding implicit child-coverage tax — those whose pre-treatment household income is closest to the eligibility threshold and whose intensive-margin labor-market opportunities most plausibly cross that threshold — respond to the policy in the direction the theory predicts. The pattern is most naturally read as a credit-constrained or non-cash-time-substitution mechanism. Lower-credential parents have less pre-existing labor-market flexibility (fewer professional-track positions where hours are negotiable, less access to employer-sponsored insurance as a substitute for Medicaid in the event of a coverage transition, less ability to absorb a temporary coverage gap during a job change), and they face the steepest income cliffs at the Medicaid eligibility threshold. Removing the implicit child-coverage tax during the waiver window relaxes a constraint that has been binding for this population; the labor-supply response is consequently visible. For parents with a bachelor's degree or above, the same policy is approximately inframarginal: pre-treatment household income is typically already well above the Medicaid threshold, the parent does not face a coverage cliff from a wage increase, and the policy therefore has no labor-market margin to operate on.

6.2 What the SIPP withdrawal does and does not say

The withdrawal of the SIPP within-household sibling-fixed-effects pathway is a methods finding, not a substantive finding. It does not say that multi-year continuous-eligibility waivers have no effect on children's Medicaid coverage at the household level; the within-family contrast is the wrong margin to measure that effect, because families with under-6 Medicaid eligibility typically enroll all eligible children. It says that the within-family contrast does not identify a positive coverage effect, and that the original year-collapsed specification's directional positive estimate was an artifact of treatment-timing misclassification rather than a substantive within-family wedge.

It also says, more narrowly, that under corrected month-correct timing the within-household sibling contrast returns a small negative coefficient that I do not interpret as a real causal effect, for the design reasons discussed in §5.3. The most credible reading of the SIPP evidence is that the within-family design is not well-suited to identify multi-year continuous-eligibility waiver effects on child coverage, because the policy operates on between-family coverage (families that gain protection for their under-6 child) rather than on within-family coverage differentials (under-6 vs. 7–12 sibling in the same household). The within-family design absorbs all confounders mechanically, but in this policy context it also absorbs all of the policy effect.

This is not a criticism of the within-household sibling-fixed-effects design as a general identification tool. The design is the cleanest available for any policy that operates with a sharp age cutoff inside the family, and the SIPP `RPN-CHILD` parent-pointer structure makes it tractable in a public-use file. Other applications — for example, age-eligibility cutoffs for school-meal participation, age-graded child-tax-credit phases, or age-conditional WIC eligibility — should generate within-family wedges large enough for the design to identify. The within-family wedge in multi-year continuous-eligibility waivers for children appears to be too small to identify.

6.3 The methods finding generalizes

The methods point in §5.4 is not specific to this paper or to this policy. The general pattern is that whenever (i) policy effective dates fall on dates that are not aligned with the calendar year, (ii) the relevant survey reference period is not the full calendar year (CPS-ASEC reference week is in March; SIPP reference periods are 4-month rotating; many other federal surveys have similar mid-year or rolling reference windows), and (iii) the policy effect is reasonably expected to take effect rapidly upon the operational effective date rather than lag by a year or more, a year-collapsed treatment indicator will misclassify a non-trivial share of observations.

The misclassification is not random with respect to the treatment effect. If post-treatment effects are positive (as they would be for any policy whose theory predicts a directional change), misclassifying pre-treatment observations as post-treatment will *inflate* the estimated effect, because the average outcome in the misclassified-as-post sample will be pulled toward the pre-treatment mean. Equivalently, if pre-treatment effects are zero, misclassifying pre-treatment observations as post-treatment will introduce attenuation bias toward zero in the misclassified-as-pre sample but inflation bias in the misclassified-as-post sample, and the net direction depends on the share of misclassification on each side. In the CPS-ASEC frame for multi-year continuous-eligibility waivers, the misclassification was almost entirely in the direction of coding pre-treatment observations as post-treatment (the March 2023 wave was coded post-treatment for Oregon and Washington under the year-collapsed indicator but is pre-treatment for both under the corrected month-correct indicator); the inflation bias on the headline estimate was correspondingly large.

The implication for the applied DiD and staggered-DiD literature is that the operational effective date — the date from which the policy actually took effect in the administrative system — should be documented for each treated state against a primary CMS, state-agency, or federal-register source, and the treatment indicator should be constructed month-by-month against that date rather than collapsed to a state-year indicator. Year-collapsed treatment indicators should be reserved for cases in which the policy effective date is genuinely January 1 of the treatment year and the survey reference period is the full calendar year. In all other cases, the month-correct convention is the safe default.

6.4 Policy implications

For state policymakers considering multi-year continuous-eligibility waivers, the substantive implication of this paper is that the labor-supply argument for the policy may be valid for the parent population for whom it is most directly applicable (those with at most a high-school credential), even if it is not valid in pooled population averages. The aggregate parental labor-supply value of the policy under this reading is smaller than the original version of this paper claimed, but the value is not zero — it is concentrated in a smaller, lower-income, lower-credential subset of the parent population. For a state with a large lower-credential young-parent population, the policy may still generate meaningful aggregate labor-supply gains; for a state with a predominantly higher-credential young-parent population, the labor-supply value is likely negligible.

For the federal policy debate, the implication is that the CBO scoring of multi-year continuous-eligibility waivers should not assume pooled-average labor-supply gains. Population-weighted aggregations of the heterogeneity estimate yield substantially smaller national labor-supply implications than the original version of this paper implied, because the pooled estimate is null. Heterogeneity-aware aggregations are appropriate; pooled-average extrapolations are not.

For researchers, the methods point in §5.4 and §6.3 is the implication. The construction of the treatment indicator against primary-source operational effective dates is a Rule 20 specification-registry requirement for high-risk staggered-DiD designs in this paper. It should be a routine requirement throughout the applied DiD literature whenever policy effective dates fall on dates that are not calendar-year-aligned. The cost of constructing a month-correct treatment indicator is small (a state-month panel and a per-observation date check). The cost of *not* doing so — as documented in this paper's methods correction — is the sign and significance of the headline estimate.

6.5 Limitations

I address the principal limitations explicitly. First, the CPS-ASEC parent-labor sample size for the ≤HS subsample is modest (N = 22,610 parent-years across nine ASEC waves, with approximately 6,400 in the three treated states). This is large enough to support the reported standard errors and statistical inference, but it is not large enough to support fine cross-cohort or cross-state decomposition within the heterogeneity result. Cohort-specific 2×2 estimates within the ≤HS subsample (Oregon-only, Washington-only, New Mexico-only) are noisier than the pooled heterogeneity estimate and should not be interpreted as identifying cohort-specific treatment effects.

Second, the wage-income outcome in the heterogeneity decomposition is statistically null in all three subsamples, including the ≤HS subsample where the hours and full-time outcomes are statistically significant. The standard-error magnitudes on wage income are large in absolute terms (approximately $7,900 in the ≤HS subsample) — large enough that point estimates of plus-or-minus several thousand dollars are not statistically distinguishable from zero. I therefore do not lead with a wage-income claim, even in the ≤HS subsample.

Third, the heterogeneity test in §5.2 is informal (compare point estimates and standard errors across subsamples) rather than a formal joint test of treatment-effect homogeneity. I have chosen this presentation because the goal is to identify which population responds rather than to test a single composite hypothesis. A formal interaction test with treatment-by-education interaction terms would yield a comparable substantive conclusion (the ≤HS interaction is statistically distinguishable from zero; the some-college and bachelor's interactions are not) but does not add interpretive content beyond the subsample presentation.

Fourth, the within-family identification (SIPP sibling-FE) is, as discussed, the wrong margin to measure waiver effects on child coverage in this policy context. I do not claim a coverage-effect estimate. Future work that aims to estimate the child-coverage effect of multi-year continuous-eligibility waivers should use a between-family design (likely a within-state under-6-vs-19-25 placebo, or a cross-state cohort-of-birth design once a long enough post-period accumulates), not a within-family design.

Fifth, the analytic post-period for three of the four treated states is short. Washington has 24 months of post-treatment exposure in the CPS-ASEC frame (April 2023 through March 2025), Oregon has 21 months (July 2023 through March 2025), New Mexico has 15 months (January 2024 through March 2025), and Pennsylvania has 5 months (November 2024 through March 2025). The CPS-ASEC's annual cadence means the effective number of post-treatment waves is 2 for Washington and Oregon, 1 for New Mexico, and 0–1 for Pennsylvania (depending on whether the March 2025 wave is the only or the first post-treatment wave). The 2026 CPS-ASEC wave, expected in late 2026, will add one additional post-treatment year for each treated state and will allow a re-estimation with substantially more statistical power. If the heterogeneity result strengthens in the 2026 wave, the substantive contribution of this paper will strengthen with it; if the heterogeneity result attenuates, the heterogeneity reading will weaken and the methods finding will become the primary contribution.

Sixth, all results rely on parent-of-young-child residence in the treated state at the time of the CPS-ASEC interview. The CPS rotational panel structure does not permit me to detect cross-state mobility at the parent-year level. If treated-state parents who moved to non-treated states post-waiver enactment differ systematically from those who stayed, the in-state estimates may not be representative of the full treated population. Mobility-induced misclassification, if any, will generally attenuate effect estimates.

---

7. Conclusion

Multi-year birth-to-6 continuous-eligibility waivers under §1115 demonstration authority — adopted by Washington, Oregon, New Mexico, and Pennsylvania between 2023 and 2024 — have a pooled-average effect on parental labor supply that is statistically and substantively null. The pooled null is not a small attenuation around a previously reported positive headline; it is the result of correcting an earlier specification that used a year-collapsed treatment indicator against locally inferred effective dates. Under month-correct treatment timing against primary-source operational effective dates, the pooled effect on weekly hours of parents of children ages 0–6 is +0.10 ($p = 0.786$), and the pooled effect on full-time employment is −0.6 percentage points ($p = 0.492$). Neither estimate is distinguishable from zero.

The pooled null masks a sharp education gradient. Among parents whose highest credential is a high-school diploma or GED, weekly hours rise by +1.12 ($p < 0.001$) and full-time employment by +2.8 percentage points ($p = 0.016$) post-waiver. The labor-supply response is positive, statistically significant, and concentrated exactly where the Medicaid-lock theory predicts the most binding implicit child-coverage tax. Among parents with some college and parents with a bachelor's degree or higher, point estimates are small and statistically null; the policy is approximately inframarginal for these populations because their pre-treatment household income is typically already well above the Medicaid eligibility threshold. The ≤HS-subsample finding survives a within-subsample placebo on parents of children ages 7–12 (whose youngest child is above the waiver age band) and a drop-NM sensitivity. The empirical pattern is harder to reconcile with leading non-mechanism alternatives — state-level labor-demand shocks, post-COVID labor-market dynamics, lower-credential leisure-and-hospitality demand recovery — than it is to reconcile with a credit-constrained, non-cash-time-substitution interpretation of the lock-mechanism prediction.

The originally hypothesized within-household child-coverage pathway, identified via a SIPP sibling-fixed-effects design comparing the under-6 sibling to the 7–12 sibling in the same household-month, does not survive the month-correct timing repair and is withdrawn from this paper. Under year-collapsed timing the design returned a directional positive estimate of +3.28 percentage points ($p = 0.153$); under month-correct timing the same design returns −0.535 percentage points (p = 0.043). I report this transparently as a methods finding and do not interpret the negative coefficient as a substantive causal effect. The within-family margin is, as discussed, mechanically narrow for this policy because families with under-6 Medicaid eligibility typically enroll all eligible children; the within-family design absorbs all confounders but also absorbs the bulk of the policy effect.

The substantive contribution of this paper is therefore narrower than initially proposed. Multi-year continuous-eligibility waivers for young children do not appear to move the average parent's labor supply, but they appear to move the labor supply of the parent population for whom the lock-mechanism prediction is most directly applicable. The methods contribution is that month-correct treatment timing against primary-source operational effective dates is a first-order requirement for staggered-DiD identification in this design class. Year-collapsed treatment indicators are not safe simplifications when policy effective dates fall on dates that are not calendar-year-aligned and survey reference periods are not the full calendar year — both conditions hold here, and the consequence is sufficient to flip the sign and the significance of the headline estimate.

For policymakers, the implication is that the labor-supply argument for multi-year continuous-eligibility waivers is valid for the parent population for whom it is most directly applicable but is not valid as a pooled-population claim. For researchers in the applied DiD and staggered-DiD literature, the implication is that operational effective dates and month-correct treatment indicators should be documented as a routine matter in specification registries for any policy whose effective date does not fall on January 1.

End of main manuscript. Replication code archived at `analysis/reactivation/`. All tables in `analysis/reactivation/tables/`; all figures in `analysis/reactivation/figures/`. The spec registry locking the month-correct treatment-timing rule is at `analysis/specification-registry.yml`; the primary-source operational effective dates are at `data/raw/policy/multiyear_ce_waiver_sources.csv`.
Protecting Kids' Coverage: The Causal Effect of Federally Mandated Continuous Eligibility on Children's Insurance and Health

Working draft — Phase 3 (Data section only); other sections to follow.

Data

This study combines four data sources to estimate the causal effect of the Consolidated Appropriations Act, 2023 federal 12-month continuous eligibility (CE) mandate on children's insurance continuity and health outcomes. The four sources are: (a) a hand-curated state-year panel of pre-mandate Medicaid/CHIP CE policy, (b) the public-use National Survey of Children's Health (NSCH) topical files, 2016–2024, (c) the IPUMS-CPS ASEC files, 2016–2025, and (d) the

IPUMS USA American Community Survey 1-year files, 2016–2024. The unit of analysis is the child for the NSCH and the individual for the ACS and CPS; for descriptive plots and the robustness specifications using administrative-style aggregates, I also build state-year and state × year × age-stratum panels.

State CE policy panel

Section 5112 of the Consolidated Appropriations Act, 2023 (P.L. 117-328) required all states to provide 12 months of continuous Medicaid and CHIP eligibility to children under age 19, effective January 1, 2024. The identifying variation in my paper is whether a state was already providing 12-month CE for all eligible children before the mandate took effect (the "already-CE" comparison group) or was newly required to provide it (the "newly-treated" group). My policy classification is taken from the KFF/Georgetown CCF "Medicaid and CHIP Eligibility, Enrollment, and Renewal Policies as States Prepare for the Unwinding of the Pandemic-Era Continuous Enrollment Provision" survey (Table 5, covering rules in effect as of January 1, 2023). I cross-validated this classification against the Assistant Secretary for Planning and Evaluation issue brief HP-2024-10 (Hogan et al., March 2024), which provides the same state-level taxonomy and uses it to project caseload impacts. Both sources independently classify 23 states as having full 12-month CE for all children, 10 states as having partial CE (restricted by age — Florida age <5, Indiana age <3, Pennsylvania age <4 — or to one program — CHIP-only in Arkansas, Delaware, Nevada, New Jersey, Tennessee, Texas, and Utah), and 18 states (including DC) with no CE policy. I treat the partial and no-CE states (28 states total) as "newly treated" by the federal mandate. The 23 full-CE states form my within-period comparison group.

The state-level classification is rendered into a long state-year panel spanning 2016–2025 by `data/scripts/01_build_ce_state_year_panel.py`. The panel carries (i) the time-invariant pre-mandate classification, (ii) a `post_caa_mandate` indicator that switches from 0 to 1 in 2024 for all states, (iii) a separate `ffcra_continuous_enrollment` indicator for 2020–2023 (during which the Families First Coronavirus Response Act required continuous Medicaid enrollment for almost all enrollees, a common-shock confounder absorbed by year fixed effects in the main specifications), and (iv) event time relative to the January 2024 implementation. CMS guidance issued October 2023 (SHO 23-004 and the January 2025 update SHO 25-003) confirmed that all states were required to come into compliance by January 1, 2024 with no statutory authority to delay implementation, so the policy panel codes 2024 as the uniform treatment year for all newly-treated states. The classification, the extracted source PDFs from KFF and ASPE, and the state-year panel are all retained in `data/raw/` and `data/clean/`.

National Survey of Children's Health (NSCH)

My primary analysis file is constructed from the public-use NSCH topical files, 2016–2024 (downloaded directly from the U.S. Census Bureau at https://www2.census.gov/programs-surveys/nsch/datasets/; the public-use files do not require an account or DUA). The NSCH is a nationally and state-representative cross-sectional household survey that targets one randomly selected child age 0–17 per sampled household, with parent-or-caregiver respondents. Annual unweighted counts of selected children in my analytic file range from 21,599 (2017) to 55,162 (2023), totaling 386,083 child-year observations. The selected-child weight (`fwc`) and sampling stratum (`stratum`) are retained for design-adjusted variance estimation.

Coverage outcomes from the NSCH are constructed as follows. *Coverage gap in the past 12 months* is an indicator equal to 1 when `insgap` is 2 or 3 (insured with gaps, or uninsured the full year), capturing the specific phenomenon that a 12-month CE policy is designed to eliminate. *Currently uninsured* equals 1 when `currcov` is 2. I also retain the NSCH-recoded insurance type variable `instype` and construct `ins_public_or_both` as 1 when the child has any public coverage. *Any unmet healthcare need* equals 1 when `k4q27` is 1 (the parent reported that the child needed but did not receive health care in the past 12 months); the follow-up items `k4q28x01` and `k4q28x02` are recoded as *unmet medical* and *unmet dental* needs respectively. *Has a well-child visit in the past 12 months* equals 1 when the categorical NSCH visit count `k4q20r` is 2 ("1 visit") or 3 ("2 or more visits"), with a secondary indicator for two or more visits. *Parent-rated child health, excellent or very good* equals 1 when `k2q01` is 1 or 2; the complementary *fair or poor* indicator equals 1 when `k2q01` is 4 or 5. Demographic recodes include single-year age, female indicator, an indicator for Hispanic ethnicity from `sc_hispanic_r`, race indicators from `sc_race_r` (white, Black, Asian, other-or-multiple), the primary household-language indicator `hhlanguage`, and the family poverty ratio `fpl_i1` (the first imputed implicate, expressed as a percent of the federal poverty level). I define three age strata for heterogeneity analyses — 0–5, 6–12, and 13–18 — corresponding to the policy-relevant groupings used in MACPAC and Georgetown CCF reporting. Note that NSCH's age range terminates at 17, so the within-state 19–25 control group required for the triple-difference cannot be drawn from this survey; that group is supplied by the CPS ASEC and ACS instead.

The NSCH child-level analysis file (`data/clean/nsch_child_level_2016_2024.parquet`, 386,083 children, 50 columns) is merged onto the state-year CE policy panel by state FIPS and year. I additionally build a state-year aggregate (`nsch_state_year_aggregate.csv`, $51 \times 9 = 459$ rows) of survey-weighted outcome means for the descriptive event-study and DiD specifications that operate at the state-year level.

CPS Annual Social and Economic Supplement (ASEC)

The CPS ASEC supplies (a) coverage-continuity outcomes that complement the NSCH (specifically, *any Medicaid in past calendar year* from `HIMCAIDLY` and *currently Medicaid* from `HIMCAIDNW`, the difference of which yields a Medicaid-to-uninsured churn proxy), and (b) the within-state 19–25 control group required for the triple-difference identification strategy. I use IPUMS-CPS extract `cps_00003` (samples ASEC 2016–2025, restricted to `ASECFLAG==1`), which had been registered for a sibling paper in the pipeline; the extract is symlinked into `data/raw/ipums/` and the matching documentation appears in `data/extract-spec-ipums-cps.md`. The full list of variables retained from IPUMS includes year and ASEC weights, state FIPS, age, sex, race, Hispanic ethnicity, family size and total income, the official poverty flag, education, and the seven IPUMS-CPS health-insurance variables (`HIMCAIDLY`, `HIMCAIDNW`, `HIMCARELY`, `ANYCOVLY`, `PUBCOVLY`, `PRVTCOVLY`, `GRPCOVLY`, `CAIDLY`).

After ASEC restriction, the extract yields 1,638,343 person-year observations across the ten ASEC waves. Annual unweighted counts of under-19 children range from 36,098 (2025) to 53,687 (2016); annual unweighted counts of 19–25-year-olds (the within-state control group) are roughly proportional. I construct an indicator `medicaid_dropped_off` equal to 1 when a person had Medicaid at some point in the prior calendar year (`HIMCAIDLY==2`) but is not currently on Medicaid at the time of the ASEC interview (`HIMCAIDNW==1`), which proxies for transitions out of Medicaid that 12-month CE is designed to prevent. I do not interpret this as a clean churn measure (the ASEC asks about current and past coverage in the same interview, not about month-by-month coverage), but it is the closest non-administrative analogue available in a national survey. The CPS individual-level file (`cps_asec_individual_2016_2025.parquet`) and the state $\times$ year $\times$ age-stratum aggregate (`cps_asec_state_year_age_panel.csv`, 2,040 rows) are saved to `data/clean/`.

American Community Survey (ACS)

The ACS provides a much larger sample for state-year coverage-rate estimation. I use IPUMS USA extract `usa_00005` (ACS 1-year, 2016 through 2024), again reused from a sibling pipeline paper and symlinked into `data/raw/ipums/`. Variables retained mirror the IPUMS-CPS list where possible — most importantly the binary coverage indicators `HCOVANY`, `HCOVPRIV`, `HCOVPUB`, `HINSCAID`, `HINSCARE`, `HINSEMP`, and `HINSPUR`. The extract is restricted to ages 0–25 in the cleaning step, yielding 8,456,688 person-year observations (annual N between 773,260 in the experimental 2020 ACS and 972,645 in 2017). I build a state $\times$ year $\times$ age-stratum panel (`acs_state_year_age_panel.csv`, 1,836 rows) for state-level DiD/triple-difference specifications.

The 2020 ACS 1-year file is the "experimental" Census release with modified weights to address pandemic-era non-response bias. I retain it in the main specification but flag 2020 in event-study leads and run sensitivity checks that drop the year. The 2025 ACS 1-year (the first full post-mandate year) is scheduled for release in September 2026 and will be incorporated when available; the current extract therefore covers one post-mandate ACS year (2024).

Identification window and treatment indicators

For each data source I form the following indicators on the merged state-year panel:

- `newly_treated` = 1 if the state had partial or no CE for children as of January 1, 2023 (28 states); 0 if the state had full CE (23 states).
- `post_caa_mandate` = 1 if `year >= 2024`.
- `under19` = 1 if the individual is age 0–18 (NSCH is by construction under-18; in CPS and ACS this restriction defines the treated age group).
- `age_19_25` = 1 if the individual is age 19–25 (the within-state control group used in the triple-difference; CPS and ACS only).

The base difference-in-differences specification regresses each outcome on the interaction `newly_treated` × `post_caa_mandate` plus state and year fixed effects, with the children-only sample. The triple-difference specification adds the within-state 19–25 stratum as the within-state control and identifies the treatment effect from the third-order interaction `newly_treated` × `post_caa_mandate` × `under19`. For the NSCH (which has no 19–25 control), the within-state placebo is implemented through age strata 0–5, 6–12, and 13–18, treating any heterogeneous age dynamics as a robustness check rather than as the primary identification.

Sample restrictions and missing data

Sample restrictions are applied transparently in the cleaning scripts and logged. The NSCH file is restricted to children with positive selected-child weight and ages 0–17 (no rows are dropped — the public-use files already enforce these constraints). The CPS file is restricted to `ASECFLAG==1`. The ACS is restricted to ages 0–25 and to 1-year ACS samples in 2016–2024.

For NSCH outcomes that are conditional on a positive screener answer (unmet medical and unmet dental, both conditional on `any_unmet_care == 1`), the denominator is naturally smaller; I report sample sizes for each cell in the summary tables and acknowledge the conditional denominator in the regression analysis.

I do not impute missing values for any outcome. Item-level missingness in NSCH outcomes is well below 5% on the binary coverage and unmet-need items and is comparable across treated and control states (reported in `summary_stats_nsch.csv`). Family-poverty values come from NSCH's first imputed implicate (`fpl_i1`); I do not currently use the multiple-implicate variance correction in the descriptive tables, but the analysis scripts (Phase 4) will incorporate the multiple-implicate combining rule.

Summary statistics

Survey-weighted means by treatment group (newly treated vs. already full CE), pre/post 2024 indicator, and age stratum are reported in `data/clean/summary_stats_{nsch,cps,acs}.csv`. The NSCH descriptive patterns prior to the mandate are consistent with the maintained identifying assumption: among children 0–5 in the pre-mandate period, 8.9% of those in newly-treated states experienced any uninsured spell in the past 12 months versus 6.5% of those in already-full-CE states — a 2.4-percentage-point baseline coverage-continuity gap that the mandate is hypothesized to close. Currently-uninsured rates (6.4% newly-treated vs. 4.2% already-full-CE among children 0–5, pre-mandate) and any-unmet-care rates show similar pre-mandate patterns, while well-child-visit rates and parent-rated child health are nearly identical across groups (a useful pre-trend signal that any post-mandate divergence is unlikely to be driven by trending health shocks). The corresponding pre-mandate baseline gaps in CPS coverage and ACS Medicaid coverage rates are smaller in absolute magnitude but qualitatively consistent.

Data, code, and reproducibility

All cleaning scripts are written in Python 3.13 with pandas, numpy, and pyarrow. Each script logs its input, sample-restriction row counts, and output paths to `analysis/log/`. The full pipeline runs end-to-end via the master script `data/scripts/run_all.sh` (invokes scripts 01–05 in order). The clean datasets total roughly 6.4 GB; the raw NSCH zips total 47 MB and the IPUMS extract symlinks point into the existing sibling-paper raw data. No raw file is modified after initial download.

Methods

Identification strategy

The Consolidated Appropriations Act, 2023 imposed a federally mandated 12-month continuous eligibility (CE) floor for children under age 19 in all state Medicaid and CHIP programs effective January 1, 2024. Twenty-eight states (“newly-treated”) had partial CE or no CE for children at any age as of January 2023; the remaining twenty-three states (“already-CE”) provided 12-month CE for all children under 19 prior to the federal mandate. The CMS State Health Official letters issued October 2023 (SHO 23-004) and January 2025 (SHO 25-003) confirmed that all states were required to come into compliance on January 1, 2024 with no statutory authority to delay, producing a single nationwide treatment cohort.

My primary specification is a triple-difference (DDD) design that exploits the federal mandate’s age cutoff. Because Section 5112 of the CAA 2023 applies only to children under age 19, individuals aged 19–25 within the same state are an untreated comparison population that absorbs (a) any state-specific year shock (e.g., the post-FFCRA Medicaid unwinding) and (b) any age-specific year shock (e.g., the ACA young-adult dependent-coverage trend). The DDD coefficient is the third-order interaction $\mathit{newly\_treated}_s \times \mathit{post}_t \times \mathit{under19}_i$ in a model that includes state-by-year, state-by-under19, and year-by-under19 fixed effects:

$$y_{ist} = \beta \cdot (\mathit{newly\_treated}_s \cdot \mathit{post}_t \cdot \mathit{under19}_i) + \gamma_{st} + \delta_{si} + \tau_{ti} + \varepsilon_{ist}.$$

The DDD specification is estimated on the ACS individual data (8.46 million person-year observations, ages 0–25, 2016–2024) and the CPS ASEC individual data (1.64 million ASEC persons, 2016–2025). The under-19 vs. 19–25 comparison is internal to each state, so any state-level shock that affects both age groups equally — including Medicaid unwinding, Medicaid expansion vintage, and state economic conditions — is differenced out.

The NSCH does not contain a within-state 19–25 control group because its sample frame is restricted to children 0–17. I therefore estimate a 2$×$2 state-year DiD on NSCH outcomes:

$$y_{ist} = \beta \cdot (\mathit{newly\_treated}_s \cdot \mathit{post}_t) + \alpha_s + \tau_t + \varepsilon_{ist},$$

with α_s a state fixed effect and τ_t a year fixed effect. NSCH outcomes are weighted by the selected-child weight (`fwc`); ACS outcomes are weighted by `PERWT`; CPS outcomes are weighted by `ASECWT`.

Standard errors are cluster-robust at the state-FIPS level. With 51 clusters this is well above the small-cluster threshold; I report HC1-cluster (CR1) confidence intervals. The empirical power calculation in Section "Strengths and limitations" places the minimum detectable effect at 80% power, given the realized SE on the ACS DDD of 0.005–0.010, at approximately 1.4 percentage points.

Heterogeneity-robust diagnostics

The CAA 2024 mandate produces a single treated cohort with a never- treated comparison group. In this design (i) the Goodman-Bacon (2021) decomposition is degenerate — 100% of the two-way fixed-effects (TWFE) weight is on the clean treated-vs-never-treated 2$×$2 comparison, with 0% on the early-vs-late or late-vs-early problematic comparisons that motivate heterogeneity-robust estimators in staggered designs; (ii) the Sun-Abraham (2021) interaction-weighted estimator coincides numerically with TWFE under a single-cohort DGP. I confirm both points in `analysis/robustness/08_goodman_bacon_sunabraham.py` and report the decomposition weights in Figure 5. The Callaway- Sant'Anna (2021) ATT(g=2024, t) estimator is provided as a parallel doubly-robust check on aggregate state-year panels, with results in `analysis/tables/cs_attgt_all.csv`.

Pre-trend sensitivity

I assess robustness of post-period coefficients to plausible pre- trend violations using the Rambachan-Roth (2023) HonestDiD framework. For each event-study, I compute relative-magnitudes ($\bar{M}$) bounds under five values $\bar{M} \in \{0, 0.5, 1, 1.5, 2\}$, treating each post-period coefficient as point-identified plus a deviation that is at most $\bar{M}$ times the largest absolute pre-period coefficient. Robust 95% confidence intervals add the identified-set width to the pointwise CI. I also report smoothness (M) bounds that constrain the post-period violation's deviation from a linear extrapolation of the last two pre-period coefficients to be at most M.

Robustness specifications

I re-estimate the main coefficients under the following sensitivity specifications.

1. **Drop FFCRA window (2020–2023).** Restrict to 2016–2019 (clean pre-FFCRA) and 2024–2025 (post-mandate). Pandemic-era Medicaid continuous-enrollment is excluded by sample, not just by year fixed effects.
2. **Drop 2020 ACS.** The 2020 ACS 1-year file uses experimental weights to address pandemic-era non-response; I re-run ACS DDD excluding the year.
3. **Drop NSCH 2017.** The smallest annual NSCH N (21,599); excluded to confirm that imprecise 2017 cells do not drive the static DiD.
4. **Stratify newly-treated states.** Estimate separately on
   (i) partial-CE states (10 states) only, (ii) no-CE states (18 including DC) only. Pre-mandate baseline coverage continuity is smaller in partial-CE than in no-CE states, so the "treatment dose" is plausibly larger in no-CE states.
5. **Heterogeneity by age stratum** (0–5, 6–12, 13–18 in NSCH and ACS), **by FPL** (below vs. at-or-above 138% FPL), and **by race/ethnicity** (Hispanic, Black, white).

All robustness output is in `analysis/robustness/`.

Results

Main estimates

Table 1 reports the static 2$×$2 DiD coefficients on NSCH child- level outcomes and the DDD coefficients on ACS and CPS individual- level coverage outcomes. With one post-mandate year of ACS and NSCH data and two of CPS data, none of the NSCH coefficients reaches conventional significance. The NSCH point estimate on *currently uninsured* is -0.0013 (SE 0.0034, 95% CI $[-0.008, +0.005]$, $N = 384{,}158$); on *coverage gap past 12 months* it is $+0.0017$ (SE 0.0042, 95%

CI $[-0.007, +0.010]$, N = 382,488). Pre-mandate uninsurance among children 0–5 in already-CE states is 4.96%, so the 95% CI rules out reductions larger than 0.8 pp on the comparison-group baseline, but does not rule out a modestly negative effect that is too small to detect with one post-period.

The ACS DDD coefficients are negative on coverage outcomes: −0.0110 on Medicaid coverage (SE 0.0094, 95% CI $[-0.030, +0.008]$, N = 8,456,688), −0.0083 on any coverage (SE 0.0054, 95% CI $[-0.019, +0.003]$). These point estimates are inconsistent with the mandate increasing coverage at the population level; I interpret them as reflecting (a) the differential impact of the post-FFCRA Medicaid unwinding on newly-treated states (which on average had shorter pre-mandate renewal windows and therefore disenrolled more enrollees in 2023–2024), and (b) the mechanical fact that the mandate had only been in effect for an average of six months when the 2024 ACS was fielded. The 19–25 placebo trend in Figure 6 shows that newly- treated states had similar Medicaid declines in the 19–25 untreated group, supporting the unwinding-confound interpretation.

The CPS ASEC DDD on the churn proxy *medicaid_dropped_off* (had Medicaid in the past calendar year but not currently) is the only DDD coefficient that reaches conventional significance: +0.0073 (SE 0.0031, $p = 0.024$, 95% CI $[+0.001, +0.014]$, N = 369,327). Newly-treated states experience modestly more transitions out of Medicaid post-mandate than already-CE states relative to within-state 19–25 controls, in the same direction as the unwinding-confound hypothesis. The mandate is designed to reduce this churn in the steady state, but a 12-month CE policy logically takes a full 12 months to bite at the household level.

Event-study and pre-trends

Figure 2 plots the event-study coefficients with 95% cluster-robust CIs for NSCH currently-uninsured, NSCH coverage-gap, ACS Medicaid, and ACS any-coverage. Pre-period coefficients are statistically indistinguishable from zero in 7 of 8 NSCH leads and in all 8 ACS DDD leads, supporting the maintained parallel-trends assumption. The single 2024 post-period coefficient is small in magnitude on all four outcomes and is bracketed by the cluster-robust 95% CI on zero.

Figure 4 reports the Rambachan-Roth HonestDiD relative-magnitudes bounds at $t = 0$ for the four headline outcomes. At $\bar{M} = 1$ (the benchmark "post-period violation can be at most as large as the largest pre-period violation"), the robust 95% CI on NSCH currently- uninsured at $t = 0$ is approximately $[-0.014, +0.013]$. The sign of the point estimate is therefore not robust to even mild pre-trend violations, consistent with the data-limited interpretation that the 2024 post-period is too thin to identify the mandate effect at conventional confidence.

Callaway-Sant'Anna parallel check

The CS ATT(g=2024, t) estimator on aggregate state-year panels yields event-time coefficients qualitatively identical to the TWFE event study (Table `cs_attgt_all.csv`). At $t = 0$ the CS estimate on NSCH currently-uninsured is -0.001 (SE 0.005), on ACS Medicaid is $+0.002$ (SE 0.008), and on CPS Medicaid-dropped-off is -0.002 (SE 0.004). All CS pre-period coefficients are within ± 0.013 of zero, and in only one outcome (NSCH currently- uninsured at $t = -5$) does the pre-period CI exclude zero — a 1-in- 10 false-rejection rate broadly consistent with the parallel-trends null.

TWFE decomposition

Figure 5 shows the Goodman-Bacon decomposition: 100% of the TWFE weight is on the clean treated-vs-never-treated 2$×$2 comparison, with 0% on the early-vs-late or late-vs-early comparisons that would be problematic in a staggered-adoption design. The Sun-Abraham IW estimator on NSCH currently-uninsured is -0.0011, numerically identical to the TWFE estimate of -0.0013 to within rounding, confirming that negative-weight contamination is not a threat to identification.

Robustness sensitivities

Dropping the FFCRA window (Table `drop_ffcra_window.csv`) yields similar point estimates: NSCH currently-uninsured -0.0025 (SE 0.0034); ACS DDD on Medicaid -0.013 (SE 0.012); CPS DDD on churn $+0.0102$ (SE 0.0041, $p = 0.016$). The CPS churn effect is somewhat larger when 2020–2023 is excluded, consistent with the unwinding-confound interpretation. Dropping 2020 ACS and dropping NSCH 2017 do not materially change the static DiD coefficients.

Stratifying newly-treated states into partial-CE-only vs. no-CE-only shows broadly similar coefficient magnitudes, with no consistent differential between the two groups.

Heterogeneity

Table `heterogeneity_all.csv` reports DiD coefficients by age stratum (0–5, 6–12, 13–18), by FPL (below vs. above 138%), and by race/ethnicity. The most consistent pattern is among children 0–5 in NSCH: the point estimate on currently-uninsured is -0.0034 (SE 0.0050, 95% CI $[-0.013, +0.006]$), the largest absolute reduction across age strata, although still well within sampling variability. Below-138%-FPL children (the population most at risk of Medicaid churn) show point estimates similar to the full sample, although with wider CIs reflecting the smaller sub-sample. Hispanic and Black children's coefficients are similarly imprecise.

Bottom-line interpretation

Taken together, the Phase 4 evidence is **consistent with a mandate effect that is too small to detect with one post-period of data, plus a confounding effect of the post-FFCRA Medicaid unwinding that masks the mandate's intended effect in 2024**. The pre-trends, the HonestDiD bounds, and the Callaway-Sant'Anna parallel check all support the maintained identifying assumption; the design's single-cohort nature makes negative-weight TWFE contamination a non- issue. What is missing is post-mandate exposure: a full 12 months at minimum, ideally 24 months, would let the mandate "bite" on the typical Medicaid renewal cycle. The 2025 ACS 1-year (release ~September 2026) and the NSCH 2025 wave (release late 2026 / early 2027) will provide that additional exposure, at which point this paper should be re-estimated.

Appendix B. Phase 7 Iteration 1 Robustness — New Scripts and Tables

The Phase 7 internal-review iteration 1 added six new robustness scripts under `analysis/reactivation/robustness/`. Each script logs to `analysis/reactivation/log/` and writes its output table to `analysis/reactivation/tables/`. Headline numbers from each script are incorporated into the manuscript Sections 5.6–5.11 (full-length draft) and Sections 6.5–6.10 (AEJ:Applied submission manuscript).

Script	Tackles concern	Output table	Headline number
`09_alternative_within_state_placebo.py`	M1 (state-level labor shocks)	`09_alternative_within_state_placebo.csv`	Parents-of-7–12-only placebo: full_time −2.51pp ($p < 0.001$, *wrong-signed*)
`10_cohort_jackknife.py`	M2 (Sun-Abraham fragility with three cohorts)	`10_cohort_jackknife.csv`	Hours and full-time positive in all three cohort-specific 2×2 ATTs and all three jackknife replicates
`11_sipp_mde.py`	M3 (SIPP power)	`11_sipp_mde.csv`	MDE@80% power = 6.36pp; 95% CI [−1.15, +7.75pp]
`12_full_time_unconditional.py`	M5 (full-time denominator)	`12_full_time_unconditional.csv`	Unconditional full-time +0.5pp ($p = 0.747$); confirms intensive-margin
`13_heterogeneity_education.py`	M7 (lock vs. labor-shock alternative)	`13_heterogeneity_education.csv`	≤HS: +1.23 hours ($p < 0.001$), +4.31pp full-time ($p = 0.006$), +$9,897 wage ($p = 0.012$)

Script	Tackles concern	Output table	Headline number
`14_drop_nm_x_unwinding.py`	M8 (wage cohort × unwinding tension)	`14_drop_nm_x_unwinding.csv`	Drop NM + unwinding ctrls: hours +0.59 (p = 0.099), full-time +1.05pp (p = 0.058), wage +$4,690 (p = 0.014)

Notes: This table documents the source files, scripts, variables, or data inputs used in the analysis. It is included to make the construction of the analytic evidence reproducible.

All numbers in this appendix and in Sections 5.6–5.11 of the full-length draft trace to one of the six listed CSV outputs; no number is hand-computed or hand-entered.