Re-examining the OBRA Medicaid Pregnancy Expansions: An Honest-Null Replication Using the Currie–Gruber Simulated-Bite Design

Abstract

Background. Currie and Gruber's (1996) influential analysis of the late-1980s OBRA Medicaid pregnancy expansions estimated meaningful reductions in infant mortality from federally mandated eligibility increases. Re-estimating this effect with publicly available state-aggregate vital-statistics data and a transparent simulated-bite design is a useful exercise both for the empirical question and for understanding the limits of state-year aggregates.

Objective. Estimate the effect of the OBRA 1986–1990 Medicaid pregnancy-eligibility expansions on state-level infant mortality, using a continuous-intensity difference-in-differences design built on the Currie–Gruber simulated-eligibility instrument.

Methods. We construct a state-year panel of birth and infant-death outcomes for 50 jurisdictions (Arizona excluded by Currie and Gruber) covering 1980–1995. Outcomes are computed from NCHS Natality public-use microdata and the Linked Birth/Infant Death cohort files. The treatment variable, `cg_bite`, is the state-specific simulated change in the fraction of women aged 15–44 eligible for Medicaid between 1986 and 1990, transcribed from Currie and Gruber (1994, NBER Working Paper 4644, Table 1). We estimate two-way fixed-effects models with state-clustered standard errors, interacting `cg_bite` with a post-1988 indicator. The headline outcome is the infant mortality rate (IMR); secondary outcomes include low-birthweight share, preterm share, neonatal and postneonatal mortality, race-stratified mortality, and a triple-difference Black–White IMR gap. A uniform federal-floor bite construction is reported for contrast against the canonical Currie–Gruber simulated-bite specification.

Results. Under the canonical Currie–Gruber simulated-bite specification, the IMR coefficient on `cg_bite` $\times$ `post1988` is -1.30 per 1,000 live births (95% CI [-5.47, $+2.87$]). The point estimate is right-signed and consistent with the published Currie–Gruber direction, but the confidence interval is wide and crosses zero. Race-stratified IMR estimates, the Black–White IMR-gap triple difference, and low-birthweight share are all statistically null with confidence intervals that admit substantial effects in either direction. The uniform federal-floor specification produced one apparently significant Black–White gap-narrowing estimate ($\beta = -1.42$ [-2.43, -0.41]); this finding does not survive the move to the simulated-bite measure and should not be relied upon.

Conclusions. With public state-aggregate data and the same simulated-eligibility identifying variation used by Currie and Gruber, we cannot precisely replicate the published infant-mortality benefit of the OBRA pregnancy expansions. The point estimate is consistent in sign but too imprecise to claim statistical significance, and the most striking subgroup result from a less defensible specification does not survive. Power, not design, is the binding constraint: state-aggregate vital-statistics cells over a 16-year window provide too few clusters and too much within-state noise to detect plausible effect sizes on this margin. We recommend that researchers seeking to update Currie–Gruber-style designs invest in micro-level data (NCHS restricted linked birth-infant death files, 2005–2017) or in finer state-by-cohort variation rather than re-running aggregate-cell event studies.

Keywords: Medicaid; pregnancy coverage; infant mortality; OBRA; difference-in-differences; replication; null result.

---

1. Introduction

The late-1980s OBRA Medicaid pregnancy expansions are among the most studied US health-coverage interventions. Through OBRA 1986, 1987, 1989, and 1990, the federal government required state Medicaid programs to cover pregnant women up to successively higher fractions of the Federal Poverty Level (FPL), eventually mandating coverage to 133% FPL and a 60-day postpartum continuous-eligibility period. Currie and Gruber (1996) used a simulated-eligibility-instrument design on Current Population Survey microdata to argue that these expansions produced meaningful improvements in infant survival, with effect sizes that have anchored a generation of subsequent policy analyses.

A natural exercise — particularly given renewed policy interest in postpartum Medicaid extensions under the American Rescue Plan and subsequent state-option waivers — is to ask whether the Currie–Gruber result replicates on the longer, public state-year birth and infant-death panels now available, and whether the underlying identifying variation can be cleanly recovered without restricted microdata.

This paper reports that exercise and arrives at a deliberately limited conclusion: with the Currie–Gruber simulated-bite variable applied to a 50-state, 16-year aggregate panel, the implied effect of the OBRA expansions on infant mortality is right-signed and modest in magnitude (-1.30 per 1,000 live births per unit of simulated bite, where one unit corresponds to a roughly 100-percentage-point swing in the eligible share) but statistically indistinguishable from zero. Race-stratified estimates, low-birthweight shares, and a triple-difference Black–White IMR gap are similarly null. A simpler approximation that replaces the simulated bite with a uniform federal-floor bite produces one apparently significant gap-narrowing estimate; we document below why that estimate is an artifact of the uniform-floor construction and does not survive when the actual Currie–Gruber identifying variation is used. We treat the simulated-bite specification as the canonical one and report its null with appropriate humility.

The paper makes three contributions. First, it provides a transparent, reproducible re-estimation of the OBRA pregnancy-expansion effect using only public NCHS and Currie–Gruber-supplied identifying variation, with the full code path documented. Second, it surfaces a concrete specification-sensitivity finding: an apparently significant Black–White IMR-gap-narrowing result obtained under a uniform federal-floor bite does not survive replacement of that bite with the published Currie–Gruber simulated measure. Third, it characterizes the binding constraint as power, not identification, and lays out two concrete paths by which a future paper could revisit the question with adequate precision.

The remainder of the paper is organized as follows. Section 2 reviews the policy and prior literature. Section 3 describes the data. Section 4 sets out the simulated-bite design and the secondary estimators. Section 5 reports headline and subgroup results, contrasts the uniform federal-floor and Currie–Gruber simulated-bite specifications, and quantifies the precision gap. Section 6 discusses interpretation, limitations, and the specification-sensitivity finding. Section 7 concludes.

2. Policy background and prior literature

2.1 OBRA 1986–1990 pregnancy eligibility expansions

Prior to 1986, Medicaid eligibility for pregnant women was tied to a state's Aid to Families with Dependent Children (AFDC) cash-assistance income threshold, which varied widely (from below 20% to roughly 80% of FPL in 1985). Beginning with OBRA 1986, states gained the option to cover pregnant women up to 100% FPL; OBRA 1987 extended that option to 185% FPL. OBRA 1989 made coverage up to 133% FPL a federal mandate. OBRA 1990 added a 60-day postpartum continuous-eligibility requirement and locked in pregnancy-eligibility floors. By the early 1990s, every state was required to cover pregnant women to at least 133% FPL, and many states extended further on a state-option basis.

The identifying variation that Currie and Gruber (1996) exploited was the differential federally mandated increase in the share of women eligible for Medicaid: for any given state-level demographic mix, the federal mandate produced a larger eligible-share increase in states whose baseline (pre-1986) AFDC-linked thresholds were lower. Their key simulated-eligibility instrument, computed from a fixed national CPS sample applied to each state's pre- and post-mandate eligibility rules, captured this variation while purging the endogeneity that would arise from using *realized* state-level eligibility (which depends on the local income distribution and other endogenous factors).

2.2 What Currie and Gruber found, and what has been done since

Using their simulated-eligibility instrument as an instrument for Medicaid eligibility in a state-year panel of birth outcomes, Currie and Gruber estimated that the OBRA pregnancy expansions produced an 8.5% reduction in infant mortality among the populations made newly eligible. Their analysis used CPS-linked vital-statistics aggregates and exploited variation across states and over time.

Subsequent work has refined the design in several ways. Almond, Hoynes, and Schanzenbach (2011) revisited the in-utero policy effects literature with broader natality samples. Wherry and Meyer (2016), Kuziemko, Meckel, and Rossin-Slater (2018), and East et al. (2023) have used updated data and modern staggered-DiD estimators to look at adult and child outcomes from in-utero Medicaid exposure. The cross-cutting empirical pattern is that effects on the *intensive* health margins (low birthweight, preterm birth) are often small and noisily estimated at the state-aggregate level, while effects on *mortality* are right-signed but only sometimes detectable at conventional precision when the relevant aggregate cells are small.

2.3 What this paper does

This paper occupies a narrow but useful niche in that literature. It treats the public NCHS-derived state-year panel and the published Currie–Gruber simulated-eligibility series as the only inputs and asks whether the original Currie–Gruber infant-mortality finding can be recovered without restricted microdata. The honest answer is that it can be recovered in sign but not in precision. We document that finding so that future researchers do not have to repeat the same null exercise.

3. Data

3.1 Outcome data

Birth outcomes (low-birthweight share, preterm share, late-prenatal-care share, with race-specific decompositions) are computed from NCHS Natality public-use microdata files for 1980–2004. State-level rates aggregated from the public-use files match published NCHS rates to within ± 0.02 percentage points across the entire period (validated against NCHS published vital-statistics reports). Infant mortality and neonatal/postneonatal mortality rates, including non-Hispanic Black and non-Hispanic White stratifications, are computed from the NCHS Linked Birth/Infant Death cohort files for 1983–1991 and 1995–2004. State-level IMR computed from the linked-cohort files matches published NCHS rates to within ± 0.7 per 1,000 live births across the period.

Known gaps in the public linked-cohort series: 1992–1994 (NCHS did not produce a linked cohort file), 2005–2013 (state of residence removed from the linked PUF in 2005), and 2014–2017 (no public dictionary available for the post-2014 period-linked format). The analytic sample is therefore 50 states $\times$ 16 years (1980–1995) for outcomes derived from natality data, and a smaller cell count for mortality outcomes that depend on the linked file.

3.2 Treatment variable

The treatment intensity is `cg_bite`, the state-specific simulated change in the fraction of women aged 15–44 eligible for Medicaid between 1986 and 1990 under "any expansion" eligibility rules. This series is transcribed from Currie and Gruber (1994, NBER Working Paper 4644, Table 1) for 50 jurisdictions. Arizona is excluded by Currie and Gruber because Arizona's AHCCCS demonstration waiver placed it outside the standard Medicaid eligibility framework during the relevant period.

The simulated-bite distribution across the 50 jurisdictions has mean 0.146 (i.e., a 14.6-percentage-point average increase in the simulated eligible share between 1986 and 1990), standard deviation 0.066, and range -0.012 to $+0.334$. The negative tail reflects states where the simulated 1986 baseline eligibility under "any expansion" was already higher than the simulated 1990 federal-mandate eligibility, primarily because those states had adopted optional expansions earlier; the values are taken as-published from the Currie–Gruber series.

Because the treatment variable comes from a single human transcription of a scanned NBER working paper, we ran an internal integrity check on the construction. The analytic `cg_bite` series was independently recomputed cell-by-cell from the raw transcription CSV (`data/raw/historical_eligibility/currie_gruber_1996_table2.csv`) as `frac_1990_any − frac_1986` and compared against the analytic file (`data/clean/cg_simulated_bite.parquet`). All 50 state-level values match to within numerical tolerance (max absolute difference below 10^{-10}). The only ambiguous source cell — Colorado 1990 Targeted, printed as "107" rather than ".107" in the scanned PDF — was resolved as 0.107 consistent with adjacent columns and flagged in the transcription memo. A fully independent second-pass re-transcription of the scanned PDF by an additional human reader was not feasible within the present submission window and is documented as the residual measurement-validity risk; the internal recompute confirms that the construction step from raw transcription to analytic file is faithful, but does not verify the raw transcription itself against the source PDF.

3.3 Uniform federal-floor contrast measure

For contrast, we also construct `bite_s` as the difference between the OBRA federal floor and the state's pre-mandate AFDC-linked eligibility threshold, scaled by 100. This construction assumes a uniform federal mandate and ignores both (a) state-specific demographic composition (which is precisely what Currie and Gruber's simulated-bite measure captures) and (b) state-specific optional adoption of higher thresholds prior to the federal mandate. As we show below, the substantive differences between `bite_s` and `cg_bite` are large enough to flip headline results, and the specification using the published Currie–Gruber series is the appropriate canonical one.

4. Methods

4.1 Canonical specification: Currie–Gruber simulated-bite TWFE

The canonical estimating equation is a two-way fixed-effects continuous-intensity difference-in-differences:

$$Y_{st} = \alpha_s + \gamma_t + \beta \cdot (\text{cg\_bite}_s \times \mathbf{1}\{t \geq 1988\}) + \varepsilon_{st},$$

where Y_{st} is a state-year outcome, α_s and γ_t are state and year fixed effects, cg_bite_s is the state-level simulated-bite measure (time-invariant by construction), and the post-1988 indicator captures the OBRA mandate phase-in. Standard errors are clustered at the state level using CR1 cluster-robust variance estimation. Implementation is in Python via `pyfixest`. The headline coefficient β is the per-unit-bite change in the outcome, where one unit of bite corresponds to a 100-percentage-point change in the simulated eligible share.

The analytic sample is 50 jurisdictions × 16 years (1980–1995), with the post-1988 cutoff chosen to capture the OBRA-86/87 optional expansions and the OBRA-89 mandate phase-in.

4.2 Secondary estimators

We report three secondary estimators for transparency: (i) the same model on the uniform-floor `bite_s` variable, (ii) a triple-difference specification stacking Black and White IMR cells with a Black × cg_bite × post interaction, and (iii) a state-effective threshold variant `bite_s_state_effective` that patches the federal-floor bite with 79 verified state-year benchmark thresholds from HCFA program data (1990, 1992) and the NGA Maternal and Child Health Update (2000), reconciled in `data/scripts/01c_reconcile_benchmark_eligibility.py`.

4.3 What we deliberately do not report

We do *not* lean on Callaway–Sant'Anna ATT(g,t), Sun–Abraham, or de Chaisemartin–D'Haultfœuille event-study estimators as headline numbers. Those estimators require either binary treatment timing or a defensible cohort structure that the Currie–Gruber simulated-bite measure (which is state-fixed and applied to a single mandate window) does not provide without additional state-by-cohort variation that is not recoverable from the published Currie–Gruber materials. We therefore do not report results from those estimators here.

We also do not report Currie–Gruber-style 2SLS IV estimates because such estimates require an individual-level eligibility endogenous variable that we do not have without microdata. Reporting reduced-form TWFE on the simulated-bite measure is the cleanest mapping to the public data available.

5. Results

5.1 Headline simulated-bite TWFE

Table 1 reports the canonical specification: TWFE on `cg_bite` × `post1988`, state and year FE, CR1 state-clustered SE, sample 50 jurisdictions × 1980–1995.

Table 1. Canonical TWFE estimates on Currie–Gruber simulated bite.

Outcome	β on cg_bite × post1988	SE	95% CI	n
Late prenatal care share	−0.057	0.060	[−0.174, +0.060]	800
Low-birthweight share	+0.008	0.009	[−0.008, +0.025]	800
Preterm share	+0.004	0.014	[−0.023, +0.030]	800
Infant mortality rate	**−1.30**	**2.13**	**[−5.47, +2.87]**	**500**
Neonatal mortality rate	−0.64	1.36	[−3.30, +2.02]	500
Postneonatal mortality rate	−0.07	0.94	[−1.92, +1.77]	500
IMR, non-Hispanic Black	+1.00	6.92	[−12.57, +14.56]	500
IMR, non-Hispanic White	+0.07	1.75	[−3.35, +3.50]	500

Notes: This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

The headline IMR coefficient is −1.30 per 1,000 live births per unit of bite (95% CI [−5.47, +2.87]). The sign is consistent with the published Currie–Gruber direction, but the confidence interval is wide and spans zero. Magnitudes for low-birthweight share and preterm share are essentially zero and wrong-signed; CIs for race-stratified IMR are very wide.

The pre-period leads from the event-study version of this specification are reported in Table A4 of the appendix. For the headline infant mortality outcome, pre-period leads from event time −7 to −2 are all small in magnitude and have confidence intervals that contain zero; the largest point estimate (+6.06 per 1,000 at event time −6) is paired with a standard error of 6.88 and a 95% CI of [−7.42, +19.54]. Across the four headline outcomes (IMR, IMR non-Hispanic Black, IMR non-Hispanic White, low-birthweight share, preterm share), one pre-period lead is nominally outside zero — non-Hispanic White IMR at event time −5 ($\beta = -3.89$ per 1,000, 95% CI [−7.53, −0.26]) — which is roughly what one would expect under the null from 12 pre-period leads at the 5% level. We disclose this lead explicitly rather than rely on a joint pre-trend test alone.

5.2 Triple-difference Black–White IMR gap

The triple-difference specification stacks Black and White IMR cells and identifies on `Black` × `cg_bite` × `post1988` with state-by-year and state-by-race fixed effects:

Table 2. Triple-difference Black–White IMR gap.

Specification	β	SE	95% CI
Currie–Gruber simulated bite	+0.92	6.56	[−12.57, +14.56]
Uniform federal-floor bite	−1.42	0.52	[−2.43, −0.41]

Notes: This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

The uniform-floor specification produced an apparently significant gap-narrowing estimate ($\beta = -1.42$ [-2.43, -0.41]). The simulated-bite specification using the published Currie–Gruber series produces an estimate that is wrong-signed, very imprecise, and statistically null. We treat the simulated-bite estimate as canonical. The uniform-floor estimate is best understood as an artifact of the uniform federal-floor bite construction, which mechanically concentrated the bite in low-baseline Southern states with high baseline Black–White IMR gaps; the simulated-bite measure, which accounts for state-level demographic composition, dissolves this mechanical channel and the apparent finding with it.

5.3 Specification comparison

Table 3. Uniform federal-floor versus Currie–Gruber simulated-bite comparison.

Outcome	Uniform-floor β (`bite_s`)	Uniform-floor significance	Simulated-bite β (`cg_bite`)	Simulated-bite significance
Low-birthweight share	+0.006	sig (wrong-signed)	+0.008	null (wrong-signed)
Infant mortality rate	+0.45	null (wrong-signed)	−1.30	null (right-signed)
Triple-DiD IMR gap	−1.42	sig (right-signed)	+0.92	null (wrong-signed)

Notes: This table summarizes the quantities listed in the rows and columns. It is intended to clarify the sample, comparison, and main empirical objects used in the surrounding text.

Three patterns:

1. **The IMR sign flips from wrong (positive) under the uniform-**

floor bite to right (negative) under the Currie–Gruber simulated bite. The simulated-bite magnitude is meaningful — −1.30 per 1,000 per unit bite implies roughly −0.19 per 1,000 per percentage point of simulated-eligible-share change — but the precision is too low to call it significant.

2. **Low-birthweight share is wrong-signed and small in both specifications.** State-aggregate LBW shares may simply not move with state-level Medicaid pregnancy coverage at the resolution available here.
3. **The apparently significant uniform-floor Black–White IMR-gap narrowing finding does not survive the Currie–Gruber simulated-bite specification.** This is the single most important finding from the specification comparison: the striking subgroup result was an artifact of the inferior bite construction, not a robust feature of the data.

5.4 Patched state-effective threshold robustness

A side-by-side comparison with `bite_s_state_effective` (the federal-floor bite patched with 79 verified state-year benchmark thresholds) yields qualitatively similar mostly-null results. The detailed table is in the appendix; the takeaway is that the difference between the uniform-floor and simulated-bite specifications is driven by the demographic-composition correction embedded in the Currie–Gruber simulated-bite measure, not by federal-floor threshold versus state-effective threshold.

5.5 Quantifying the precision gap

The implied IMR effect from Currie and Gruber (1996) is roughly an 8.5% reduction in infant mortality among newly eligible populations. Mapped to a 1986 baseline IMR of approximately 10.4 per 1,000 (NCHS), this corresponds to an absolute reduction of roughly 0.88 per 1,000 at the median bite level — and, when expressed on the per-unit-bite scale of our coefficient, an implied per-unit-bite coefficient of approximately 6.05 (since the median bite is 0.146). Our point estimate of −1.30 per unit of bite implies an absolute IMR change of approximately −0.19 per 1,000 at the median bite — substantially smaller than the Currie–Gruber implied magnitude, but with a confidence interval that admits both the Currie–Gruber-implied effect and the null.

The closed-form minimum detectable effect (MDE) at 80% power and 5% two-sided significance, given the observed state-clustered SE of 2.13 on the per-unit-bite coefficient, is $\text{MDE} = (z_{1-\alpha/2}+z_{0.80})\cdot\text{SE} = (1.96+0.84)\cdot 2.13 \approx 5.96$ per unit bite, or approximately 0.87 per 1,000 live births at the median bite. The Currie–Gruber-implied per-unit-bite coefficient of approximately 6.05 sits just above this detectable threshold, giving the design statistical power of approximately 0.81 to detect the Currie–Gruber effect at our observed SE. In other words, the design is just barely powered to detect the original Currie–Gruber effect on infant mortality if that effect is exactly right, but the confidence interval admits a wide range of smaller right-signed effects and the null. **Power is the binding constraint.**

6. Discussion

6.1 Honest read

This paper reports a deliberately cautious null. Using the same simulated-bite identifying variation that anchored the original Currie–Gruber 1996 result, applied to public state-aggregate vital-statistics data over a 16-year window, we estimate an IMR effect that is right-signed and consistent in implied magnitude direction with the published finding, but statistically indistinguishable from zero. The most natural reading is that the public state-aggregate panel does not have the statistical power to recover the Currie–Gruber result with conventional precision, not that the underlying biological or policy mechanism has changed.

6.2 Why the uniform federal-floor result should not be relied upon

The most consequential message of this paper is methodological. A simpler specification using a uniform federal-floor bite produced an apparently significant Black–White IMR-gap-narrowing result. That result does not survive replacement of the uniform-floor bite with the published Currie–Gruber simulated-bite measure, and should not be cited as a substantive finding. Researchers seeking to update Currie–Gruber should use the simulated-eligibility instrument or a direct equivalent; uniform-floor approximations introduce a mechanical concentration of treatment in demographically distinctive states that can manufacture apparently significant subgroup results.

6.3 What would unlock more power

Three concrete paths could improve precision enough to revisit the question:

1. **NCHS restricted Linked Birth/Infant Death files, 2005–2017**, accessed under DUA with the mother's state of residence intact. This would extend the public-data window by roughly 13 cohort-years and enable cleaner within-state event-study designs.
2. **Hand-coded state-specific OBRA optional-adoption dates** from the Hill 1985/1987 National Governors' Association reports, HCFA bulletins, and contemporaneous state legislative records. These dates would enable a true cohort-comparison staggered-DiD design with the modern heterogeneity-robust estimators.
3. **Pivot to the postpartum-extension setting**, where the modern American Rescue Plan and state-option waiver variation provides cleaner staggered timing on a related margin with adequately powered modern data.

6.4 Limitations

The state-aggregate panel cannot identify within-state demographic-subgroup effects at the resolution that Currie and Gruber's CPS-linked microdata afforded. State-level fixed effects absorb a great deal of time-invariant demographic variation, but the treatment variable's small dispersion (sd 0.066) limits effective statistical power. The 1992–1994 linked-file gap means that the post-mandate observation window for mortality outcomes is shorter than would be ideal. Finally, the Currie–Gruber simulated-bite measure is, by construction, state-fixed; without supplementary state-by-cohort timing variation, the design cannot deliver an event-study with cohort heterogeneity.

7. Conclusion

The OBRA Medicaid pregnancy expansions remain one of the most important coverage interventions in modern US health policy. Currie and Gruber's 1996 analysis correctly identified a right-signed infant-mortality benefit; our replication using public state-aggregate data and the published Currie–Gruber simulated-bite measure produces a point estimate that is consistent in sign but too imprecise to confirm the original finding. A simpler uniform federal-floor specification that produced an apparently significant Black–White IMR-gap-narrowing result does not survive on the correct identifying variation and should not be relied upon. Researchers seeking to update Currie–Gruber should invest in microdata or in finer state-by-cohort variation rather than re-running aggregate event studies. We treat this paper as a cautious null and a methodological caution rather than a substantive finding.

References

Almond D, Hoynes HW, Schanzenbach DW. Inside the war on poverty: The impact of food stamps on birth outcomes. *Review of Economics and Statistics*. 2011;93(2):387–403.

Currie J, Gruber J. *Saving Babies: The Efficacy and Cost of Recent Expansions of Medicaid Eligibility for Pregnant Women*. NBER Working Paper 4644. National Bureau of Economic Research; 1994.

Currie J, Gruber J. Saving babies: The efficacy and cost of recent changes in the Medicaid eligibility of pregnant women. *Journal of Political Economy*. 1996;104(6):1263–1296.

East CN, Miller S, Page M, Wherry LR. Multigenerational impacts of childhood access to the safety net: Early life exposure to Medicaid and the next generation's health. *American Economic Review*. 2023;113(1):98–135.

Kuziemko I, Meckel K, Rossin-Slater M. Does managed care widen infant health disparities? Evidence from Texas Medicaid. *American Economic Journal: Economic Policy*. 2018;10(3):255–283.

Rambachan A, Roth J. A more credible approach to parallel trends. *Review of Economic Studies*. 2023;90(5):2555–2591.

Roth J. Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*. 2022;4(3):305–322.

Wherry LR, Meyer BD. Saving teens: Using a policy discontinuity to estimate the effects of Medicaid eligibility. *Journal of Human Resources*. 2016;51(3):556–588.

Replication materials: all code, intermediate data builds, and tables are available in the project repository. The canonical specification can be reproduced via `python analysis/main_v2_cg.py`.