

Supplementary Materials

Psychiatric Medication Use at the Medicaid Eligibility Threshold in States with Separate CHIP

Zihang Li

^a*Department of Economics, Auburn University, Auburn, AL, 36849, USA*

Contents

S1. General-Utilization Diagnostics	3
S2. Psychiatric-Medication RD Plots by Diagnosis Status	5
S3. Heterogeneity, Transition-Year Diagnostics, and Evidence Map	6
S3.1. Clinical Rationale and Diagnosis-Group Heterogeneity	6
S3.2. Evidence Map	9
S4. Variable Construction with NSCH Item Mappings	11
S5. Year-by-Year Psychiatric-Medication Estimates	17
S6. Sample-Definition, Specification, and Income-Measurement Sensitivity	20
S6.1. State-Program Sample Sensitivity	20
S6.2. Specification Sensitivity for the Primary All-Child Outcome	22
S6.3. State-by-Year Fixed-Effect Sensitivity Across Populations	24
S6.4. Survey-Year versus Prior-Income-Year Policy Thresholds	25
S6.5. Family-Income Imputation Sensitivity	26
S6.6. Medication-Measure Sensitivity Across Populations	28
S7. Design Validation and Diagnosis-Composition Diagnostics	31
S7.1. Density and Measured-Composition Continuity	31
S7.2. Shifted-Cutoff Diagnostics	34
S7.3. Donut-Width Sensitivity	35
S7.4. Severity and Diagnosis Composition Within Current MEDB	36
S8. Local-Polynomial and Kernel Sensitivity	37

S9. Multiple-Hypothesis Correction	39
S10. Additional Supporting Tables	41
S10.1. Sample Sizes by Period, CE Group, and Current-MEDB Status	41
S10.2. Descriptive Statistics Around the Threshold	41
S10.3. Outcomes Among Children Without a Current MEDB Diagnosis	43
S10.4. Identification Limits and Diagnostic Evidence	45
S10.5. Medicaid and Separate-CHIP Coverage Terms	46
S11. Continuous-Eligibility Policy Documentation and Cross-Checks	47
S11.1. Year-by-year CE verification and strict-sample roster	47
S11.2. Non-KFF documentary context for selected threshold values	50
S11.3. Institutional Medicaid-to-separate-CHIP gap distribution	52
S11.4. Sensitivity of the strict estimand to the 50-FPL gap filter	52
S12. Diagnosis-Composition and Survey-Weight Sensitivity	55
S12.1. All-Child Diagnosis-Entry and Composition Checks	55
S12.2. Survey-Weight Leverage Sensitivity	57
S13. Discreteness-Robust Inference and Running-Variable Diagnostics	60
S13.1. Honest confidence intervals	60
S13.2. Shifted-cutoff diagnostics	61
S13.3. Cross-period distribution checks for the observed running variable	61
S13.4. Alternative inference across the medication population hierarchy	61
S14. Exploratory Coverage Check in the American Community Survey	64
AI-Assistance Disclosure	67

This supplement collects tables and figures referenced in the main paper as supplementary materials. Unless explicitly labeled otherwise, specifications, samples, and variable definitions follow the main text; sensitivity samples and bandwidths are identified in captions and notes.

Abbreviations. RD = regression discontinuity; CE = children’s Medicaid continuous eligibility, with non-CE denoting states without a January 2016 12-month Medicaid CE policy covering ages 6–17; CHIP = Children’s Health Insurance Program; MEDB = mental, emotional, developmental, or behavioral; NSCH = National Survey of Children’s Health; KFF = Kaiser Family Foundation; FPL = federal poverty level; COVID = coronavirus disease 2019; ADHD = attention-deficit/hyperactivity disorder; SSI = Supplemental Security Income; MAGI = modified adjusted gross income; FFCRA = Families First Coronavirus Response Act; CAA = Consolidated Appropriations Act; MSE = mean squared error; ACS = American Community Survey; IPUMS = Integrated Public Use Microdata Series; SE = standard error; CI = confidence interval; BH = Benjamini–Hochberg; FDR = false discovery rate; FWER = family-wise error rate. Table and figure notes repeat key abbreviations where needed.

Map to the main paper. The main paper leads with the reduced-form discontinuity in psychiatric medication use, treats switching costs and therapy disruption as candidate interpretations rather than identified mechanisms, distinguishes the stronger COVID-period inference from the more inference-sensitive pre-period estimate, and treats the 2024 estimate as exploratory. Psychiatric-medication use among all children is the sole population endpoint. Diagnosed-subgroup estimates are heterogeneity results; the COVID-period cell is the strongest, but it coincides with diagnosis-composition changes and its fixed-bandwidth honest confidence interval includes zero. The institutionally clean equal-bandwidth specification ($h = b = 50$ FPL) is primary; results using the former $h = 50$, $b = 70$ specification are explicitly labeled sensitivity analyses. The supplement reports general-utilization diagnostics (§S1); psychiatric-medication plots by diagnosis status (§S2); diagnosis-group heterogeneity, transition-year diagnostics, and an evidence map (§S3); outcome construction (§S4); year-by-year estimates (§S5); sample-definition, state-program, specification, income-timing, income-imputation, and medication-measure checks (§S6); design validation and diagnosis-composition diagnostics (§S7); local-polynomial and kernel sensitivity (§S8); multiple-hypothesis correction (§S9); additional supporting tables (§S10); continuous-eligibility documentation and threshold cross-checks (§S11); diagnosis-composition and survey-weight sensitivity (§S12); discreteness-robust inference and running-variable diagnostics (§S13); and an exploratory American Community Survey coverage check (§S14). The full organization appears in the table of contents.

S1. General-Utilization Diagnostics

Table S1 applies the main paper’s all-child design to four additional measures of health-care use. These exploratory outcomes help assess whether the medication result is part of a broader change in health-care utilization at the threshold.

Table S1: All-Child General-Utilization Diagnostics, Primary Specification

Outcome	Pre-pandemic (2016–2019)	Federal freeze (2021–2022)	2024
Emergency-room visit	−0.003 (0.066) [3,517]	0.176* (0.097) [2,056]	0.182** (0.080) [1,036]
Hospitalization	0.098* (0.051) [1,642]	0.046* (0.026) [2,054]	0.012 (0.049) [1,036]
Preventive care	0.013 (0.066) [3,503]	0.065 (0.094) [2,041]	0.062 (0.144) [1,026]
Specialist visit	0.062 (0.059) [3,508]	0.069* (0.040) [2,051]	−0.086 (0.070) [1,036]

Notes. Each cell reports a robust bias-corrected local-linear RD estimate, its robust standard error in parentheses, and N_h in brackets. The sample includes all children ages 6–17 in strict separate-CHIP state-years that lacked a January 2016 12-month Medicaid continuous-eligibility policy covering these ages. Children are not excluded based on SSI receipt or chronic functional limitation. Estimates use the primary $h = b = 50$ FPL specification, survey weights, a triangular kernel, demographic covariates plus state and period-year indicators, and standard errors clustered on running-variable support points. The hospitalization item was not fielded in 2016–2017, so its pre-pandemic estimate uses 2018–2019. These four outcomes are exploratory and are not part of the five-outcome multiple-testing family in the main analysis; stars use unadjusted robust-bias-corrected significance levels. RD = regression discontinuity; FPL = federal poverty level; SSI = Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The estimates do not show a consistent general-utilization pattern. Before the pandemic, the hospitalization estimate is 9.8 percentage points and rejects zero at the 10 percent level but not at 5 percent; the other three estimates do not reject zero at the 10 percent level. During the federal freeze, the emergency-room, hospitalization, and specialist estimates are positive and reject zero at the 10 percent level but not at 5 percent, while the preventive-care estimate does not reject zero. In 2024, the emergency-room estimate is 18.2 percentage points and rejects zero at the 5 percent level; the other three estimates do not reject zero at the 10 percent level. Because these are 12 exploratory tests without a multiple-testing adjustment, the isolated 5 percent finding does not by itself establish a broader utilization change. Taken together, the results do not support a common increase in all forms of health-care use at the threshold, although they leave open outcome- and period-specific changes, particularly in emergency-room use during 2024.

S2. Psychiatric-Medication RD Plots by Diagnosis Status

Figure S1 complements the all-child RD plots in the main paper by showing psychiatric-medication use separately for children with and without a current MEDB diagnosis. The upper row shows children with a current MEDB diagnosis, and the lower row shows children without one. These plots describe where the estimated discontinuity is concentrated, but they do not formally test whether the two diagnosis-group estimates differ. Because diagnosis status can itself change at the threshold, the plots also do not identify effects within compositionally fixed diagnosis groups.

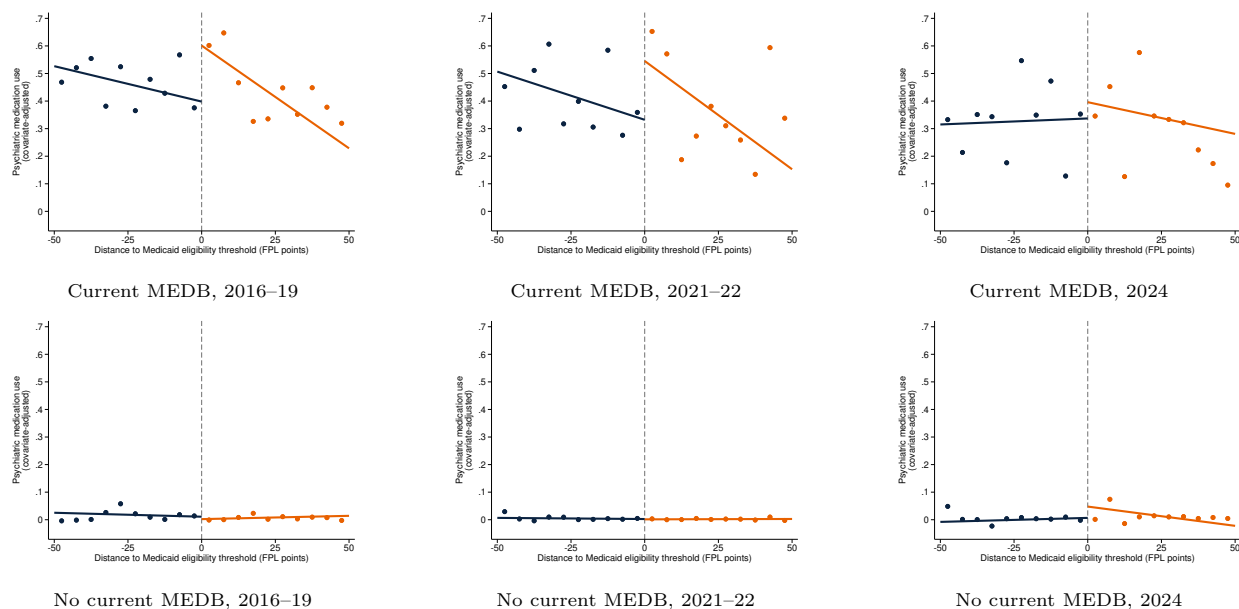


Figure S1: Psychiatric-Medication Use at the Medicaid Eligibility Threshold by Current MEDB Status. Columns show 2016–2019, 2021–2022, and 2024. The sample includes children ages 6–17 in strict separate-CHIP state-years among historically non-CE states. Children are not excluded based on SSI receipt or chronic functional limitation. The outcome is the primary past-12-month K4Q23 psychiatric-medication measure. Each panel plots survey-weighted 5-FPL-bin means of the outcome residualized on the main covariates and state and period-year indicators and recentered at the within-bandwidth weighted mean. Lines are conventional survey-weighted triangular-kernel local-linear fits on each side using an effect bandwidth of $h = 50$ FPL points. All six panels use the same vertical scale. The plots are descriptive and do not report robust bias-corrected estimates or confidence intervals. For children with a current MEDB diagnosis, the visible jumps in the conventional fitted lines are smaller than the robust bias-corrected estimates reported below. The corresponding robust bias-corrected estimates and alternative inference appear in the main paper and Supplementary Section S13.4. RD = regression discontinuity; CHIP = Children’s Health Insurance Program; CE = continuous eligibility; non-CE = no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17; MEDB = mental, emotional, developmental, or behavioral; SSI = Supplemental Security Income; FPL = federal poverty level.

The robust bias-corrected medication estimates among children with a current MEDB diagnosis are 32.8 percentage points before the pandemic, 43.8 points during the federal freeze, and 12.6 points in 2024. The corresponding estimates among children without a current MEDB diagnosis are -0.8 , -0.2 , and 6.9 points. The first two current-MEDB estimates reject zero at the 5 percent level, whereas neither 2024 estimate nor any no-current-MEDB estimate does. Together, the plots and estimates show that the pre-pandemic and federal-freeze medication discontinuities are concentrated among children with a current diagnosis. They do not establish that diagnosis status caused the all-child pattern because the probability and composition of a reported diagnosis can also change at the threshold.

S3. Heterogeneity, Transition-Year Diagnostics, and Evidence Map

This section reports three evidence blocks that connect directly to the main paper. Section S3.1 explains the clinical basis for the diagnosis-group comparison and reports the complete primary and wider-bias-bandwidth estimates. It also reports the excluded 2020 and 2023 waves as transition-year diagnostics. Section S3.2 maps the paper’s three hypotheses, descriptive period patterns, and principal validity checks to the corresponding evidence and limitations.

S3.1. Clinical Rationale and Diagnosis-Group Heterogeneity

The diagnosis-group comparison asks whether medication discontinuities are more consistently positive for conditions whose guideline-supported care often includes psychotherapy than for ADHD, for which medication is a guideline-supported first-line treatment. American Academy of Pediatrics guidance recommends approved medication together with parent-training and/or classroom behavioral intervention for school-aged children and recommends medication, with assent, for adolescents (Wolraich et al., 2019). Guidance for anxiety and depression includes psychotherapy, medication, or combined care, with treatment depending on diagnosis, severity, response, and preferences (Walter et al., 2020; Higa-McMillan et al., 2016; Walter et al., 2023). These recommendations motivate the group comparison but do not reveal what treatment an individual child received. Three other diagnoses could have expanded the ADHD group: behavior or conduct problems (k2q34b), autism spectrum disorder (k2q35b), and Tourette syndrome (k2q38b). Their guidance includes behavioral and pharmacologic options that depend on the condition, symptom burden, and patient preferences (Steiner and Remsing, 2007; Riley et al., 2016; Hyman et al., 2020; Murphy et al., 2013; Pringsheim et al., 2019). Adding them would broaden the clinical mix rather than simply increase the sample size. The analysis therefore retains the prespecified ADHD definitions and treats the resulting small-cell imprecision as an important limitation.

Table S2: Condition Heterogeneity in Psychiatric-Medication Use

Subgroup	2016–2019	2021–2022	2024
<i>Panel A. Primary $h = b = 50$ specification</i>			
Anxiety or depression	0.366*** (0.107)	0.401*** (0.097)	0.312* (0.172)
N_h	498	363	214
ADHD without anxiety or depression	0.002 (0.114)	0.040 (0.293)	−0.172 (0.115)
N_h	275	168	84
ADHD as the only MEDB diagnosis	−0.121 (0.147)	−0.294 (0.383)	−0.002 (0.257)
N_h	114	70	33
<i>Panel B. $h = 50, b = 70$ sensitivity</i>			
Anxiety or depression	0.336*** (0.100)	0.441*** (0.096)	0.269* (0.144)
N_h	498	363	214
ADHD without anxiety or depression	−0.069 (0.101)	0.230 (0.315)	−0.595*** (0.164)
N_h	275	168	84
ADHD as the only MEDB diagnosis	0.141 (0.157)	0.489 (0.379)	−0.352 (0.251)
N_h	114	70	33

Notes. Entries are robust bias-corrected local-linear RD estimates for the primary past-12-month K4Q23 psychiatric-medication outcome, with robust standard errors in parentheses; N_h is the unweighted complete-case count inside the $h = 50$ -FPL effect bandwidth. The sample includes children ages 6–17 with a current MEDB diagnosis in strict separate-CHIP state-years among historically non-CE states. Children are not excluded based on SSI receipt or chronic functional limitation. Estimates use survey weights, a triangular kernel, demographic covariates and additive state and within-period year indicators, and standard errors clustered on running-variable support points. Panel A uses the institutionally clean $h = b = 50$ primary specification. Panel B retains the former $h = 50, b = 70$ specification as a sensitivity because its bias-correction window crosses the next eligibility cutoff in some state-years. ADHD is reported both without anxiety or depression and as the child’s only MEDB diagnosis; the latter is nested within the former. The diagnosis labels do not identify treatment pathways. The 18 reported tests are not adjusted for multiple testing. Because the subgroups share state clusters, covariance-aware between-group contrasts are not estimated; significance in one subgroup and not another does not establish a subgroup difference. RD denotes regression discontinuity, ADHD denotes attention-deficit/hyperactivity disorder, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Under the primary specification, the medication estimates are positive in all three periods for children with anxiety or depression. The 2016–2019 and 2021–2022 estimates reject zero at the 1 percent level, while the 2024 estimate rejects at the 10 percent level but not at the 5 percent level. Neither ADHD definition

produces a consistently positive pattern, and none of the primary ADHD estimates rejects zero at the 10 percent level. This sign pattern is consistent with H2, but it is not evidence that the subgroup effects differ. The 2024 ADHD cells contain only 84 children without anxiety or depression and 33 children with ADHD as their only current MEDB diagnosis, making those estimates especially imprecise.

Transition-Year Diagnostics. The primary regime-specific analyses exclude 2020 and 2023 because neither year represents a clean enrollment regime. The 2020 past-12-month medication window can straddle the March 18 FFCRA implementation and pandemic onset, while the prior-calendar-year income measure generally refers to 2019. The 2023 survey overlaps the Medicaid unwinding. The following tables report those waves for transparency and compare them with the retained periods, but the transition-year estimates are not interpreted as effects of the freeze or unwinding.

Table S3: Psychiatric-Medication RD Estimates in the Transition Years, with Regime-Period Benchmarks

Period	Context	Pooled sample (A)	Strict sample (B)	$A - B$
2016–2019	No freeze or mandate	0.284*** (0.098) [1,471]	0.328*** (0.113) [998]	−0.044
2020	FFCRA/pandemic transition	0.151 (0.191) [498]	0.224 (0.206) [321]	−0.073
2021–2022	Federal enrollment freeze	0.415*** (0.123) [943]	0.438*** (0.138) [653]	−0.022
2023	Medicaid unwinding	−0.265** (0.127) [570]	−0.347** (0.137) [339]	0.082
2024	Federal CE mandate	0.009 (0.089) [626]	0.126 (0.090) [354]	−0.118

Notes. Entries are robust bias-corrected local-linear RD estimates for the primary past-12-month K4Q23 psychiatric-medication outcome, with robust standard errors in parentheses and N_h in brackets. $A - B$ is the difference between the two point estimates; no covariance-aware comparison is estimated. Sample A combines separate-CHIP and CHIP-funded Medicaid-expansion state-years. Sample B retains only strict separate-CHIP state-years. Both samples include children ages 6–17 with a current MEDB diagnosis in historically non-CE states and do not exclude children based on SSI receipt or chronic functional limitation. Estimates use the primary $h = b = 50$ FPL specification, survey weights, a triangular kernel, demographic covariates and additive state and within-period year indicators, and standard errors clustered on running-variable support points. The 2020 and 2023 rows each contain one survey wave and are excluded from the three-regime design. FFCRA denotes Families First Coronavirus Response Act, RD denotes regression discontinuity, CE denotes continuous eligibility, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table S4: Insurance RD Estimates in the Transition Years, Pooled Current-MEDB Sample

Year	Medicaid/CHIP coverage	Private-only insurance
2020	0.292*** (0.100) [501]	−0.271*** (0.072) [495]
2023	−0.068 (0.083) [572]	−0.026 (0.093) [568]

Notes. Entries are robust bias-corrected local-linear RD estimates, with robust standard errors in parentheses and N_h in brackets. The pooled state-program sample combines separate-CHIP and CHIP-funded Medicaid-expansion state-years and includes children ages 6–17 with a current MEDB diagnosis in historically non-CE states. Children are not excluded based on SSI receipt or chronic functional limitation. Estimates use the primary $h = b = 50$ FPL specification, survey weights, a triangular kernel, demographic covariates and additive state and within-period year indicators, and standard errors clustered on running-variable support points. FFCRA denotes Families First Coronavirus Response Act, RD denotes regression discontinuity, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Neither 2020 medication estimate rejects zero at the 5 percent level. Both 2023 medication estimates are negative and reject zero at that level. In the pooled current-MEDB sample, Medicaid/CHIP coverage rises by 29.2 percentage points and private-only coverage falls by 27.1 points in 2020; both estimates reject zero at the 1 percent level. The corresponding 2023 insurance estimates are close to zero and do not reject at the 10 percent level. These results show that the two excluded waves contain distinct threshold patterns. They do not identify effects of the FFCRA freeze, the pandemic, or the unwinding because each wave combines policy timing with other contemporaneous changes and possible sample-composition differences.

S3.2. Evidence Map

Table S5 maps the paper’s three framework-derived hypotheses, two descriptive period patterns, and three principal validity checks to their empirical evidence and limitations.

Table S5: Evidence Map: Hypotheses, Descriptive Patterns, and Validity Checks

Question	Empirical pattern	Evidence	Reading
<i>Framework-derived hypotheses</i>			
H1. Pre-period medication discontinuity concentrated among children with a current diagnosis	The all-child and current-MEDB estimates are positive; the no-current-MEDB estimate is near zero	The primary robust bias-corrected estimates are 9.7 percentage points among all children, 32.8 points among children with a current MEDB diagnosis, and -0.8 points among children without one. Support-point and state-clustered tests reject zero for the all-child and current-MEDB estimates, while state wild-bootstrap tests for the distinct conventional estimates do not.	Directionally consistent; inference-sensitive and not a causal decomposition
H2. More consistently positive medication discontinuity for anxiety or depression	Positive estimates in all three periods for anxiety or depression; neither ADHD definition is consistently positive	Table S2. The primary anxiety-or-depression estimates are 36.6, 40.1, and 31.2 percentage points. The modality labels are heuristic, the subgroup estimates share state clusters, and no formal between-group test is estimated.	Suggestive sign pattern, not a demonstrated subgroup difference
H3. No detectable uninsurance discontinuity	No detectable discontinuity before the pandemic or during the freeze; negative 2024 estimate before multiplicity adjustment	The all-child estimates are -7.6 , -0.6 , and -22.1 percentage points. The first two do not reject zero. The 2024 estimate rejects before adjustment but not after Benjamini–Hochberg or Holm adjustment.	Consistent before and during the freeze; 2024 does not survive adjustment
<i>Descriptive period patterns, not causal policy tests</i>			
COVID-period pattern	The all-child medication estimate is largest during the federal enrollment freeze	The estimate is 16.9 percentage points. The corresponding estimates are 9.7 points before the pandemic and 7.5 points in 2024. No covariance-aware cross-period comparison is estimated.	Descriptive; freeze effects are not separated from other pandemic changes
2024 pattern	The all-child medication estimate is smaller and imprecise in the first mandate year	The estimate is 7.5 percentage points and does not reject zero. The 2024 outcome window can include pre-mandate months and unwinding activity, and the 2024 sample is smaller.	Exploratory; attenuation is not identified
<i>Validity and specificity checks</i>			
Diagnosis composition	Rated severity does not change detectably, but ADHD prevalence and the number of current diagnoses rise at the threshold in earlier periods	ADHD prevalence rises by 7.5 points before the pandemic and 17.1 points during the freeze. Under the $h = 50, b = 70$ sensitivity, the five-condition diagnosis count rises by 0.70 during the freeze. Adding that count reduces the current-MEDB medication estimate from 36.7 to 22.9 points.	Diagnosis composition remains a plausible contributor to the current-MEDB estimates
Timing	The all-child estimates are positive in every retained year; the current-MEDB estimates show no monotonic pre-period trend but differ sharply between 2021 and 2022	Section S5 and the transition-year tables above. The annual and transition-year estimates are descriptive, and each annual cell contains only one survey wave.	The all-child pattern spans the retained waves; not a formal pre-trend or policy-effect test
Measurement and inference	The COVID medication sign remains positive across outcome definitions and missing-response codings, but inference varies across methods	Sections S6.6, S9, and S13. The pre-period estimate is sensitive to missing-response coding and the medication-question response rate; fixed-bandwidth honest intervals include zero for the current-MEDB pre and COVID estimates; state wild-bootstrap inference is weaker than support-point inference.	COVID has the most stable positive sign; magnitude and statistical conclusions remain method-sensitive

Notes. The table maps the paper’s hypotheses and descriptive patterns to the evidence used to assess them. Cross-period patterns do not identify effects of the federal enrollment freeze or the 2024 continuous-eligibility mandate, and none of the validity checks proves the RD identifying assumptions. Unless a sensitivity is named explicitly, point estimates use local-linear `rdrobust`, a triangular kernel, $h = b = 50$ FPL, demographic covariates plus additive state and within-period year indicators, survey weights, and standard errors clustered on running-variable support points. RD denotes regression discontinuity, CHIP denotes Children’s Health Insurance Program, CE denotes continuous eligibility, MEDB denotes mental, emotional, developmental, or behavioral, ADHD denotes attention-deficit/hyperactivity disorder, and FPL denotes federal poverty level.

S4. Variable Construction with NSCH Item Mappings

This section documents the analysis populations, periods, running variable, strict-sample restriction, and construction of all binary outcomes. The all-child population is primary. Medication and mental-health-treatment estimates among children with a current mental, emotional, developmental, or behavioral (MEDB) diagnosis are secondary, and medication estimates among children without a current MEDB diagnosis provide a specificity check. The item-level construction rules appear first; Table S6 then summarizes the empirical role of each outcome.

Analysis Populations and Periods. The analysis includes children ages 6–17. The pre-period pools 2016–2019, the COVID period pools 2021–2022, and the post period is 2024. The primary population includes all children who satisfy the institutional sample restrictions described below. The current-MEDB and no-current-MEDB analyses partition that population using the diagnosis classification in the next paragraph. The main period-specific analyses exclude the transition years 2020 and 2023.

Current-MEDB Classification. Children are classified as having a current MEDB diagnosis if a parent reports that the child currently has any of the following: ADHD (**k2q31b**), depression (**k2q32b**), anxiety (**k2q33b**), behavioral or conduct problems (**k2q34b**), autism spectrum disorder (**k2q35b**), developmental delay (**k2q36b**), speech disorder (**k2q37b**), or Tourette syndrome (**k2q38b**). The variable `has_medb_condition` equals one if any of these eight items is affirmative. It equals zero otherwise, including when a current-condition item is skipped because the child was never reported to have that diagnosis. Thus, the no-current-MEDB group means that none of the eight current-diagnosis indicators is affirmative; it does not establish that the child has never experienced a behavioral-health condition.

Running Variable. The running variable `distance_M` is family income, measured as a percentage of FPL, minus the state-year Medicaid income-eligibility threshold for children ages 6–18.¹ Nonnegative values are classified as above and negative values as below the threshold under the study’s KFF/NSCH measure. The data do not reveal a child’s actual Medicaid eligibility, program enrollment, or movement between Medicaid and separate CHIP. Family income comes from the harmonized NSCH FPL measure, and annual state Medicaid and separate-CHIP thresholds come from KFF eligibility surveys cited in the main paper.

Continuous-Eligibility and Strict-Sample Classifications. States are classified as historically CE or non-CE according to whether their January 2016 Medicaid policy provided 12-month continuous eligibility to children. The primary analyses use historically non-CE states; historically CE states provide descriptive context. The *strict separate-CHIP sample* retains state-years with a nonmissing separate-CHIP threshold only when the absolute difference between that threshold and the school-age Medicaid threshold is at least 50 FPL points.

¹KFF’s published Table 1 reports this school-age threshold for a 6–18 age bracket; the analytic population is restricted to ages 6–17, the NSCH age range used throughout. References to the “ages 6–18” threshold in this supplement follow KFF’s bracket labeling.

This restriction isolates observations for which the primary $h = 50$ Medicaid-threshold window does not reach the next observed public-eligibility cutoff. The analysis retains all otherwise eligible children and does not exclude children based on SSI receipt or chronic functional limitation.

Medication Measures. The primary medication outcome, `util_psych_meds`, is an exact alias for `util_psych_meds_k4q23`. It equals one when `k4q23==1` and zero when `k4q23==2`. K4Q23 asks whether, during the past 12 months, the child took medication because of difficulties with emotions, concentration, or behavior. Other K4Q23 responses remain missing and are omitted from the primary complete-case analysis.

The measurement sensitivity, `util_psych_meds_expanded`, equals one if K4Q23 is affirmative, if the child is currently taking ADHD medication (`k2q31d==1`), or if the child is currently taking autism medication (`autismmed==1`). It equals zero when K4Q23 is negative and neither current-medication item is affirmative. It remains missing when K4Q23 has no substantive yes/no response and neither current-medication item is affirmative. This expanded construction is not a second primary outcome. It is a sensitivity check that mixes a past-12-month item with two current-use items. The ADHD and autism medication items are also asked only for children with the corresponding reported diagnosis, so the information used by the expanded measure differs across children.

Other Outcome Variables. All remaining outcomes are binary. Each entry below lists the variable name, the NSCH item or items used to construct it, the exact coding rule, the concept it measures, and its empirical role in the paper.

Coverage measures (4 outcomes). NSCH derives the composite coverage-type variable `instype`, which classifies each child into five mutually exclusive categories: 1 = public only (Medicaid and/or CHIP, with no private coverage), 2 = private only (any private plan, including marketplace and TRICARE, with no public coverage), 3 = public and private together, 4 = other coverage that NSCH could not classify, and 5 = uninsured. A child with public coverage only is classified in `instype = 1` whether that coverage is Medicaid or separate CHIP; the umbrella `ins_medchip` indicator remains one under either program. These measures can test for changes in umbrella coverage and other coverage categories, but they cannot reveal an internal Medicaid-to-separate-CHIP transition. For the private-only and uninsured indicators, `instype==4` and ordinary nonresponse remain missing.

- `ins_medchip` (Medicaid/CHIP coverage, `k12q12`). One if `k12q12==1`; zero if `k12q12==2` or the current-insurance item reports that the child is uninsured (`currins==2`). The latter rule codes the questionnaire’s uninsured logical skip as zero; ordinary item nonresponse remains missing. `K12Q12` asks whether the child is covered by Medicaid, Medical Assistance, or another government-assistance plan for low-income children or children with disabilities. It does not separate Medicaid from CHIP, so it captures traditional Medicaid, CHIP-funded Medicaid expansion, and separately administered CHIP together, with some parent-reporting measurement error at the separate-CHIP margin. *Role.* primary all-child coverage diagnostic.
- `ins_private` (private-only insurance, `instype==2`). One if `instype==2`; zero if `instype ∈ {1,3,5}`. The private-only category includes privately purchased coverage, marketplace plans, employer-sponsored plans, and TRICARE. Children with both private and public coverage (`instype==3`) are coded as zero. A positive RD estimate therefore indicates a higher probability of private-only coverage relative to the pooled public, dual-coverage, and uninsured alternatives; it does not identify which alternative changed. *Role.* primary all-child coverage diagnostic.
- `ins_employer` (current or former employer- or union-sponsored insurance, `k12q03`). One if `k12q03==1`; zero if `k12q03==2` or the current-insurance item reports that the child is uninsured (`currins==2`). The latter rule codes the questionnaire’s uninsured logical skip as zero; ordinary item nonresponse remains missing. `K12Q03` asks whether the child has coverage through a current or former employer or union. *Role.* other exploratory.
- `ins_uninsured` (uninsured, `instype==5`). One if `instype==5`; zero if `instype ∈ {1,2,3}`. `instype==5` is the NSCH “not insured” category. *Role.* primary all-child outcome and H3 coverage-continuity diagnostic.

Behavioral-health utilization (3 outcomes).

- `util_psych_meds` (past-12-month psychiatric medication, `k4q23`). The exact construction appears under *Medication Measures* above. The item does not report the prescribing clinician, program type,

or whether therapy was displaced. *Role.* primary all-child population endpoint and current-MEDB secondary outcome.

- **util_mh_treatment** (mental-health treatment or counseling, `k4q22_r`). One if `k4q22_r==1`; zero if `k4q22_r ∈ {2,3}`. K4Q22_R asks whether, during the past 12 months, the child received any treatment or counseling from a mental-health professional such as a psychiatrist, psychologist, psychiatric nurse, or clinical social worker. The indicator does not separate psychotherapy from medication-management visits. A positive sign here paired with a positive sign on **util_psych_meds** can therefore reflect more therapy, more prescriber visits, or both. *Role.* primary all-child behavioral-health companion and current-MEDB secondary outcome.
- **util_mh_unmet** (unmet mental-health need, alternative coding of `k4q22_r`). One if `k4q22_r==2`; zero if `k4q22_r ∈ {1,3}`. Code 2 indicates that the parent reported the child needed but did not receive mental-health treatment or counseling. *Role.* other exploratory.

Access and unmet need (3 outcomes).

- **access_unmet_any** (any unmet need, `k4q27`). One if `k4q27==1`; zero if `k4q27==2`. K4Q27 asks whether the child needed but did not receive health care during the past 12 months. *Role.* other exploratory.
- **access_unmet_medical** (unmet medical need, `k4q28x01` and `k4q27`). One if `k4q28x01==1`; zero if `k4q28x01==2` or if `k4q27==2`, which denotes no unmet health-care need of any type. K4Q28X01 identifies medical care among reported types of unmet need and is a logical skip when K4Q27 indicates no unmet need; coding those skips as zero makes this a population-level indicator. Other item nonresponse remains missing. *Role.* other exploratory.
- **access_medical_home** (personal doctor or nurse proxy, `k4q04_r`). One if `k4q04_r ∈ {1,2}`; zero if `k4q04_r==3`. Despite the legacy variable name, this indicator records whether the child has one or more people the parent thinks of as the child’s personal doctor or nurse. It is not the NSCH-published patient-centered medical-home composite. *Role.* other exploratory.

Perceived insurance adequacy and cost (3 outcomes).

- **insqual_covers** (insurance covers needed services, `k3q20`). One if `k3q20 ∈ {1,2}` (“always” or “usually”); zero if `k3q20 ∈ {3,4}` (“sometimes” or “never”). Questionnaire skips, including those for uninsured children, remain missing. The resulting estimand is conditional on a valid item response and typically on current coverage, rather than an unconditional diagnosed-subgroup population outcome. It measures the parent’s perceived adequacy of coverage; it does not directly identify benefit design, formulary breadth, or covered-service rules. *Role.* exploratory current-MEDB adequacy outcome.
- **insqual_providers** (insurance allows seeing needed providers, `k3q22`). One if `k3q22 ∈ {1,2}`; zero if `k3q22 ∈ {3,4}`. Questionnaire skips, including those for uninsured children, remain missing. The

resulting estimand is conditional on a valid item response and typically on current coverage, rather than an unconditional diagnosed-subgroup population outcome. It measures parent-reported provider-access adequacy; it does not directly observe a plan’s formal network breadth. *Role.* exploratory current-MEDB adequacy outcome.

- `insqual_paybills` (problems paying medical or health-care bills, `k3q25` and `howmuch`). One if `k3q25==1`; zero if `k3q25==2` or `howmuch==1`, which reports \$0 in medical or health-care spending for the child. The latter rule codes the corresponding logical skip as zero; ordinary item nonresponse remains missing. *Role.* other exploratory.

General-utilization placebos (4 outcomes, supplement only).

- `util_er` (emergency-room visit, `hospitaler`). One if `hospitaler` $\in \{2, 3, 4\}$, indicating any ER use; zero if `hospitaler==1`.
- `util_hospital` (hospitalization, `hospitalstay`). One if `hospitalstay==1`; zero if `hospitalstay==2`. The item is unavailable in 2016–2017, so pre-period hospitalization estimates use only 2018–2019.
- `util_preventive` (preventive care, `k4q20r` and `s4q01`). One if `k4q20r` $\in \{2, 3\}$, indicating any preventive visit; zero if `k4q20r==1` or the parent reports no doctor or health-care-provider visit in the past 12 months (`s4q01==2`). The latter rule codes the corresponding logical skip as zero; ordinary item nonresponse remains missing.
- `util_specialist` (specialist visit, `k4q24_r`). One if `k4q24_r==1`; zero if `k4q24_r` $\in \{2, 3\}$.

Table S6: Outcome Menu and Empirical Role

Outcome	Empirical role	Why this is measured
<i>Primary all-child outcome family (5).</i>		
Past-12-month psychiatric medication (util_psych_meds)	Primary population endpoint; current-MEDB secondary outcome	Parent-reported medication use in the past 12 months. The outcome does not identify program enrollment, switching, or substitution away from psychotherapy.
Mental-health treatment or counseling (util_mh_treatment)	Behavioral-health companion; current-MEDB secondary outcome	Captures treatment or counseling by a mental-health professional but cannot separate psychotherapy from medication-management visits, visit counts, or continuity.
Medicaid/CHIP coverage (ins_medchip)	Umbrella coverage diagnostic	Tests for a discontinuity in umbrella public coverage at the threshold; cannot distinguish Medicaid from separate CHIP within the umbrella. The Medicaid-to-separate-CHIP sort is inferred from state eligibility rules and the strict-sample construction, not directly observed in this indicator.
Private-only insurance (ins_private)	Coverage diagnostic	Parent-reported private-only coverage; does not observe a transition into or out of a specific public program.
Uninsured (ins_uninsured)	H3 coverage-continuity diagnostic	Tests for a detectable uninsurance discontinuity. A null estimate does not prove that coverage or coverage type is fixed.
<i>Medication measurement sensitivity (1 alternative construction).</i>		
Expanded psychiatric medication (util_psych_meds_expanded)	Measurement sensitivity; not part of the primary family	Adds current ADHD or autism medication to the past-12-month K4Q23 outcome. Because the construction mixes reference periods and diagnosis-contingent items, it tests measurement sensitivity rather than replacing the primary outcome.
<i>General-utilization placebos (4 outcomes, supplement only).</i>		
ER visits (util_er), hospitalization (util_hospital), preventive care (util_preventive), specialist visit (util_specialist)	Broad-access diagnostics	Mixed and period-unstable estimates provide no consistent evidence of a common access discontinuity, although the tests do not exclude outcome-specific access changes.
<i>Other exploratory outcomes (8 outcomes, domain-level role).</i>		
Employer-sponsored insurance (ins_employer)	Coverage-source check	Additional coverage-source indicator alongside the mutually exclusive coverage-type outcomes.
Insurance covers needed services (insqual_covers), insurance allows seeing needed providers (insqual_providers)	Perceived insurance adequacy	Parent-reported coverage and provider adequacy in the current-MEDB analysis; neither measure identifies formal benefit design or network breadth.
Any unmet need (access_unmet_any), unmet medical need (access_unmet_medical), personal doctor or nurse (access_medical_home)	Access-domain checks	Exploratory context for access and unmet need.
Problems paying medical or health-care bills (insqual_paybills), unmet mental-health need (util_mh_unmet)	Cost and behavioral-health context	Additional exploratory context.

Notes. Seventeen substantive outcomes are analyzed across the paper and supplement: thirteen non-general-utilization outcomes and four general-utilization placebos. The expanded medication measure is an alternative construction of the primary medication concept and is not counted as an eighteenth outcome. The five all-child outcomes in the first panel form the multiple-testing family reported in Section S9. Current-MEDB medication and counseling estimates are secondary. Other population-outcome combinations are exploratory unless noted. NSCH question-text crosswalks are documented in this section. NSCH denotes National Survey of Children's Health, CHIP denotes Children's Health Insurance Program, ER denotes emergency room, and MH denotes mental health.

S5. Year-by-Year Psychiatric-Medication Estimates

This section estimates the psychiatric-medication RD separately for each retained survey year. It reports the primary all-child population first and the current-MEDB secondary population second. Both use the primary $h = b = 50$ specification in the strict separate-CHIP sample. Each annual cell contains one survey cross-section, so the estimates describe the annual pattern but do not form an event study or a formal pre-trend test.

In the all-child population, the estimates are positive in all seven retained years, ranging from 6.0 to 16.8 percentage points. None rejects zero at the 5 percent level; the 2021 and 2022 estimates reject at the 10 percent level but not at the 5 percent level. Their similar magnitudes—16.8 points in 2021 and 13.9 points in 2022—show that the positive pooled COVID-period estimate appears in both retained waves rather than in only one. The four pre-period estimates also show no monotonic trend toward zero.

The current-MEDB estimates are less stable across years. The pre-period estimates are positive and reject zero at the 5 percent level in 2016, 2018, and 2019, but the 2017 estimate is near zero. During the federal enrollment freeze, the 2021 estimate is 60.2 percentage points and rejects zero at the 1 percent level, while the 2022 estimate is 6.2 points and does not reject at the 10 percent level. The pooled current-MEDB COVID estimate therefore averages two different annual patterns. The 2024 estimate is 12.6 points and does not reject at the 10 percent level. Because the annual samples are smaller than the pooled-period samples and no covariance-aware comparison across years or populations is estimated, these patterns do not establish that the discontinuities changed over time or differ between populations.

A natural follow-up question is whether the smaller 2024 medication estimate is timed with the January 1, 2024 mandate or reflects a more gradual post-COVID process. Table S3 reports the excluded 2020 and 2023 waves under the primary $h = b = 50$ specification. The 2020 survey straddles implementation of the Families First Coronavirus Response Act and the pandemic onset, while the 2023 field period overlaps the Medicaid unwinding. The 2020 medication estimates are positive but imprecise in both the pooled and strict samples. The 2023 estimates are negative and reject zero at the 5 percent level in both samples, but they may reflect redetermination timing, sample composition, other contemporaneous changes, or sampling variation (Bensken et al., 2025; Chua et al., 2025). Table S4 reports the companion insurance estimates. In the pooled current-MEDB sample, Medicaid/CHIP coverage rises and private-only coverage falls in 2020, while both 2023 insurance estimates are close to zero. These transition-year results are reported transparently but are not clean pre-trend or policy-effect tests.

Table S7: Year-by-Year Psychiatric-Medication RD Estimates by Analysis Population

Year	All children (primary)			Current MEDB (secondary)		
	Estimate	Robust SE	N_h	Estimate	Robust SE	N_h
2016	0.060	0.062	1,316	0.336***	0.121	351
2017	0.120	0.089	545	0.019	0.159	158
2018	0.078	0.059	895	0.398**	0.177	249
2019	0.118	0.089	726	0.550***	0.168	240
2021	0.168*	0.096	1,052	0.602***	0.216	319
2022	0.139*	0.077	970	0.062	0.162	334
2024	0.075	0.063	1,026	0.126	0.090	354

Notes. Entries are robust bias-corrected local-linear RD estimates, with robust standard errors and the unweighted complete-case count inside the effect bandwidth. The primary specification sets the effect and bias-correction bandwidths to $h = b = 50$ FPL points and uses a triangular kernel, demographic covariates plus additive state indicators, survey weights, and standard errors clustered on running-variable support points. Both populations include children ages 6–17 in strict separate-CHIP state-years among historically non-CE states and do not exclude children based on SSI receipt or chronic functional limitation. The current-MEDB population is a subset of the all-child population; the table does not report covariance-aware comparisons across years or populations. The 2020 and 2023 transition years are excluded and reported separately in Table S3. RD denotes regression discontinuity, FPL denotes federal poverty level, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

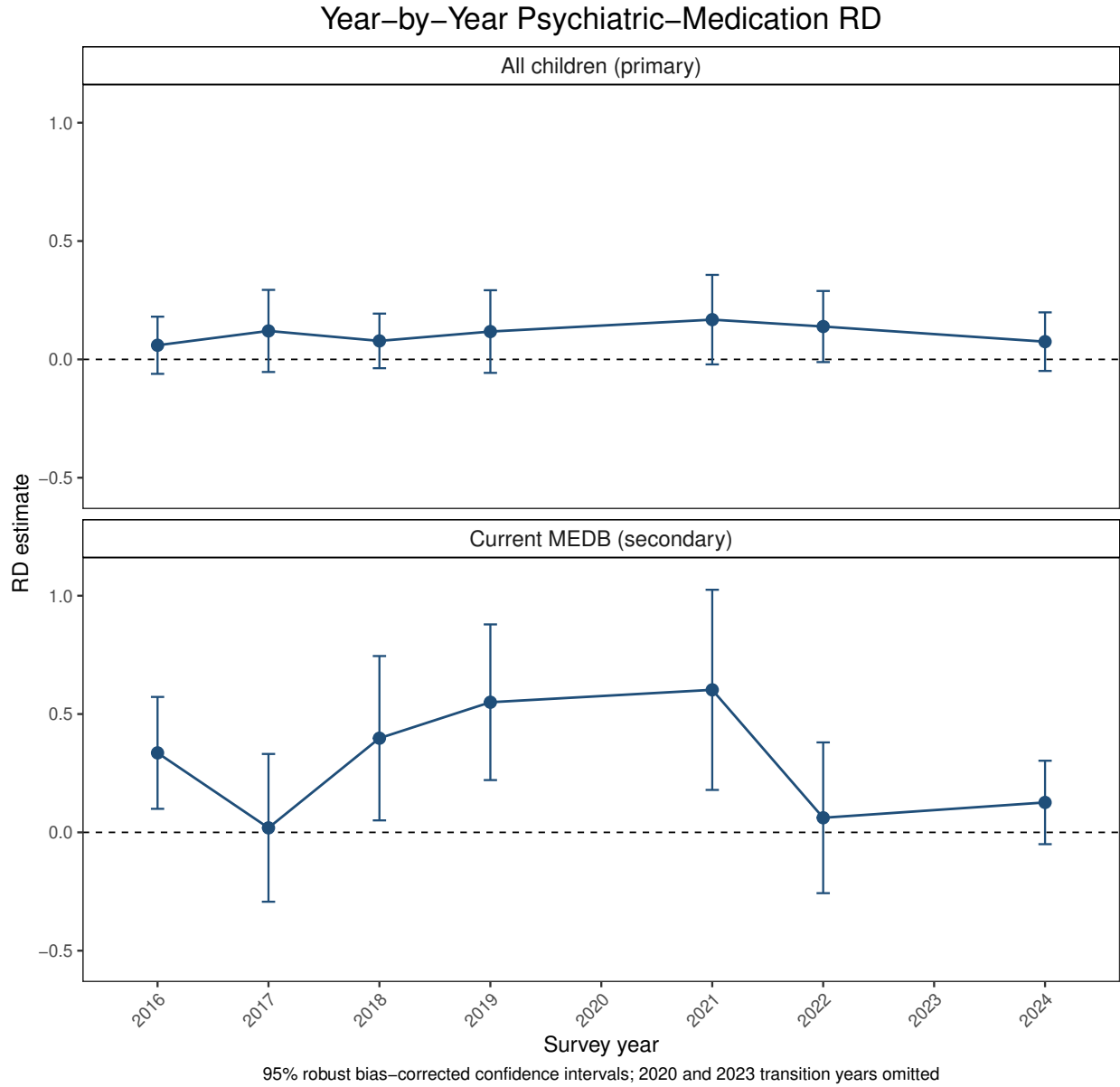


Figure S2: Year-by-Year Psychiatric-Medication RD Estimates by Analysis Population. Points are robust bias-corrected local-linear estimates and bars are 95% confidence intervals from the primary $h = b = 50$ -FPL specification. Panels share the same vertical scale. Lines connect the retained survey years only; 2020 and 2023 are omitted, and the connecting segments do not represent a continuously observed event-study path. The current-MEDB population is a subset of the all-child population, and no covariance-aware comparison between panels is estimated. RD denotes regression discontinuity, FPL denotes federal poverty level, and MEDB denotes mental, emotional, developmental, or behavioral.

S6. Sample-Definition, Specification, and Income-Measurement Sensitivity

This section collects five sensitivity analyses tied to the main paper’s sample and measurement choices. Section S6.1 compares the strict separate-CHIP sample with a pooled state-program sample. Sections S6.2 and S6.3 vary the estimator and adjustment structure. Sections S6.4 and S6.5 examine the timing and imputation of reported family income. Section S6.6, which follows this candidate block unchanged, compares the primary and expanded psychiatric-medication measures.

S6.1. State-Program Sample Sensitivity

The main paper uses the strict separate-CHIP sample because its Medicaid threshold is followed by a separately administered CHIP program at least 50 FPL points above. The pooled sensitivity adds state-years in which CHIP is implemented through a Medicaid expansion. In those state-years, the Medicaid threshold does not mark a Medicaid-to-separate-CHIP boundary. Actual enrollment and individual transitions are unobserved in both samples, so their comparison tests sensitivity to the state-program restriction; it does not test whether children switch between programs.

Table S8: State-Program Sample Definitions

Sample	Definition
Pooled sensitivity	Retains historically non-CE state-years after dropping any observed Medicaid-to-separate-CHIP gap below 50 FPL points. State-years without a separate-CHIP threshold remain in the sample, so CHIP-funded Medicaid-expansion states are included.
Strict separate-CHIP sample	Applies the pooled rule and additionally requires a nonmissing separate-CHIP threshold. This is the main paper’s primary institutional sample.

Notes. Both samples include children ages 6–17 in historically non-CE states and retain all otherwise eligible children; neither excludes children based on SSI receipt or chronic functional limitation. The 50-FPL gap is evaluated at the state-year level. Section S11.1 reports the contributing state-program classifications. Non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, CHIP denotes Children’s Health Insurance Program, FPL denotes federal poverty level, and SSI denotes Supplemental Security Income.

Table S9: Primary $h = b = 50$ RD Estimates in Pooled and Strict State-Program Samples

Population	Outcome	2016–2019		2021–2022		2024	
		Pooled	Strict	Pooled	Strict	Pooled	Strict
All children	Psychiatric medication	0.088** (0.035) [5,398]	0.097*** (0.037) [3,482]	0.155** (0.063) [3,070]	0.169** (0.069) [2,022]	0.015 (0.061) [1,776]	0.075 (0.063) [1,026]
	Mental-health treatment or counseling	0.003 (0.041) [5,436]	−0.009 (0.043) [3,512]	0.089* (0.053) [3,115]	0.104* (0.055) [2,054]	−0.186 (0.126) [1,796]	−0.164 (0.136) [1,036]
	Medicaid/CHIP coverage	0.094 (0.074) [5,195]	0.107 (0.083) [3,395]	−0.117 (0.072) [3,107]	−0.126 (0.078) [2,047]	0.256** (0.128) [1,791]	0.275** (0.140) [1,033]
	Private-only insurance	−0.033 (0.066) [5,378]	−0.022 (0.071) [3,466]	0.134* (0.074) [3,072]	0.137* (0.077) [2,020]	−0.105 (0.100) [1,770]	−0.104 (0.120) [1,022]
	Uninsured	−0.066 (0.046) [5,378]	−0.076 (0.051) [3,466]	−0.007 (0.039) [3,072]	−0.006 (0.045) [2,020]	−0.182* (0.093) [1,770]	−0.221** (0.107) [1,022]
Current MEDB	Psychiatric medication	0.284*** (0.098) [1,471]	0.328*** (0.113) [998]	0.415*** (0.123) [943]	0.438*** (0.138) [653]	0.009 (0.089) [626]	0.126 (0.090) [354]

Notes. Entries are robust bias-corrected local-linear RD estimates, with robust standard errors in parentheses and N_h in brackets. All estimates use the primary $h = b = 50$ FPL specification, a triangular kernel, survey weights, demographic covariates plus additive state and within-period year indicators, and standard errors clustered on running-variable support points. The current-MEDB population is a subset of the all-child population. Neither sample excludes children based on SSI receipt or chronic functional limitation. Stars use unadjusted robust p -values; the comparison does not report covariance-aware tests between samples. RD denotes regression discontinuity, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, FPL denotes federal poverty level, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The all-child estimates are similar across the pooled and strict samples before the pandemic and during the federal enrollment freeze. In 2024, restricting to separate-CHIP state-years raises the medication estimate from 1.5 to 7.5 percentage points, although neither estimate rejects zero at the 10 percent level. The current-MEDB medication estimate shows a larger 2024 difference, rising from 0.9 to 12.6 points, and neither estimate rejects at that level. The all-child coverage and counseling estimates retain the same signs across the two samples in every period. These comparisons show that the central pre-period and COVID medication patterns do not depend on excluding CHIP-funded Medicaid-expansion state-years. They also show greater sample sensitivity in 2024. Because the samples represent different statutory margins and actual program enrollment is unobserved, the comparison does not identify a switching mechanism.

S6.2. Specification Sensitivity for the Primary All-Child Outcome

Table S10 reports seven specifications for the primary all-child psychiatric-medication outcome in the strict separate-CHIP sample. The primary specification sets both bandwidths to 50 FPL points. The remaining columns vary the bias-correction bandwidth, covariate adjustment, effect bandwidth, treatment of observations nearest the cutoff, or fixed-effect structure.

Table S10: All-Child Psychiatric-Medication RD Across Seven Specifications

Period	Primary	Wider bias	No covariates	Narrow	Narrow, no covariates	Donut ± 3	State-by-year FE
2016–2019	0.097*** (0.037) [3,482]	0.087*** (0.034) [3,482]	0.108** (0.044) [3,556]	0.099** (0.040) [2,112]	0.105** (0.045) [2,162]	0.103 (0.066) [3,308]	0.097*** (0.035) [3,482]
2021–2022	0.169** (0.069) [2,022]	0.143** (0.063) [2,022]	0.154** (0.068) [2,040]	0.178** (0.070) [1,208]	0.153** (0.069) [1,221]	0.277*** (0.093) [1,904]	0.159** (0.067) [2,022]
2024	0.075 (0.063) [1,026]	0.069 (0.058) [1,026]	0.048 (0.060) [1,034]	0.100 (0.070) [607]	0.052 (0.064) [612]	0.125 (0.087) [976]	0.075 (0.063) [1,026]

Notes. Entries are robust bias-corrected local-linear RD estimates, with robust standard errors in parentheses and N_h in brackets. Primary uses $h = b = 50$, demographic covariates, and additive state and within-period year indicators. Wider bias changes only the bias-correction bandwidth to $b = 70$. No covariates uses $h = b = 50$ and retains only within-period year indicators. Narrow uses $h = 30, b = 50$; its no-covariate version again retains only within-period year indicators. Donut ± 3 uses $h = b = 50$ after dropping observations with $|\text{distance}_M| < 3$. State-by-year FE replaces additive state and year indicators with state-by-year cutoff-cell indicators while retaining demographics. All rows use the strict separate-CHIP historically non-CE all-child sample, a triangular kernel, survey weights, and standard errors clustered on running-variable support points. Children are not excluded based on SSI receipt or chronic functional limitation. The $b = 70$ sensitivity can cross the next institutional cutoff in some state-years; the primary $h = b = 50$ specification cannot. FE denotes fixed effects, RD denotes regression discontinuity, FPL denotes federal poverty level, CHIP denotes Children’s Health Insurance Program, and SSI denotes Supplemental Security Income. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

The all-child medication estimate is positive in every specification and period. Before the pandemic, six of the seven estimates reject zero at the 5 percent level; the donut estimate does not. During the federal enrollment freeze, all seven reject zero at that level. None of the 2024 estimates rejects zero at the 10 percent level. The sign pattern is therefore stable, while the exact magnitude and precision depend on the specification. The larger COVID donut estimate also shows that dropping observations nearest the cutoff is not an innocuous precision check.

S6.3. State-by-Year Fixed-Effect Sensitivity Across Populations

The state-by-year specification replaces additive state and within-period year indicators with a separate intercept for each contributing state-year cutoff cell. This provides a flexible check on the normalized multi-cutoff pooling restriction while retaining the primary bandwidths, kernel, weights, demographic covariates, and support-point clustering.

Table S11: Psychiatric-Medication RD with Additive and State-by-Year Fixed Effects

Population	Adjustment	2016–2019	2021–2022	2024
All children	Additive state and year FE	0.097*** (0.037) [3,482]	0.169** (0.069) [2,022]	0.075 (0.063) [1,026]
All children	State-by-year FE	0.097*** (0.035) [3,482]	0.159** (0.067) [2,022]	0.075 (0.063) [1,026]
Current MEDB	Additive state and year FE	0.328*** (0.113) [998]	0.438*** (0.138) [653]	0.126 (0.090) [354]
Current MEDB	State-by-year FE	0.327*** (0.075) [998]	0.386*** (0.137) [653]	0.126 (0.090) [354]

Notes. Entries are robust bias-corrected local-linear RD estimates, with robust standard errors in parentheses and N_h in brackets. Every row uses the strict separate-CHIP historically non-CE sample, ages 6–17, $h = b = 50$ FPL, a triangular kernel, survey weights, demographic covariates, and standard errors clustered on running-variable support points. The additive specification includes state and within-period year indicators. The state-by-year specification includes 64 pre-period, 32 COVID-period, and 15 post-period cutoff-cell indicators. Children are not excluded based on SSI receipt or chronic functional limitation. FE denotes fixed effects, RD denotes regression discontinuity, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, FPL denotes federal poverty level, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Replacing additive indicators with state-by-year indicators changes the all-child estimates by at most 1.0 percentage point. The current-MEDB estimate changes little before the pandemic and in 2024, while the COVID estimate falls from 43.8 to 38.6 points. The positive pre-period and COVID estimates still reject zero at the 5 percent level in both populations. This check shows that those signs are not generated by the

additive state/year structure, but it does not resolve sorting, diagnosis selection, or other threats to the RD design.

S6.4. Survey-Year versus Prior-Income-Year Policy Thresholds

NSCH family income refers to the prior calendar year, whereas the primary running variable subtracts the KFF threshold indexed to the survey year. Table S12 rebuilds the Medicaid distance and strict separate-CHIP gap with prior-calendar-year thresholds, including the 2015 threshold for the 2016 survey. Each period uses the intersection of observations eligible under both definitions so that the comparison isolates threshold timing.

Table S12: Psychiatric-Medication RD Using Survey-Year and Prior-Income-Year Thresholds

Population	Period	Survey-year threshold	Prior-income-year threshold
All children	2016–2019	0.096**	0.094**
		(0.038) [3,387]	(0.038) [3,389]
	2021–2022	0.169**	0.161**
		(0.069) [2,022]	(0.070) [2,020]
	2024	0.075	0.075
		(0.063) [1,026]	(0.063) [1,026]
Current MEDB	2016–2019	0.324***	0.312***
		(0.113) [972]	(0.113) [972]
	2021–2022	0.438***	0.418***
		(0.138) [653]	(0.147) [652]
	2024	0.126	0.126
		(0.090) [354]	(0.090) [354]

Notes. Entries are robust bias-corrected local-linear RD estimates, with robust standard errors in parentheses and N_h in brackets. Specifications use the strict separate-CHIP historically non-CE sample, survey weights, a triangular kernel, demographic covariates plus additive state and within-period year indicators, $h = b = 50$ FPL, and standard errors clustered on support points defined by the indicated running variable. The common sample requires an institutional gap of at least 50 FPL under both threshold definitions. Within that pool, the Medicaid threshold changes in one state-year cell per period before 2024 and by at most 2 FPL; separate-CHIP thresholds change in two pre-period cells, one COVID cell, and one 2024 cell and by at most 25 FPL. The latter changes can affect the common-sample screen without changing the Medicaid-distance value. Children are not excluded based on SSI receipt or chronic functional limitation. KFF denotes KFF’s Medicaid and CHIP eligibility survey, RD denotes regression discontinuity, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, FPL denotes federal poverty level, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The two threshold definitions produce nearly identical all-child estimates: the difference is 0.2 percentage points before the pandemic, 0.8 points during the federal enrollment freeze, and zero in 2024. The current-MEDB differences are 1.3, 1.9, and zero points. The statistical conclusions are unchanged in both

populations. This result shows that the observed threshold-timing mismatch has little effect in this panel, principally because school-age Medicaid thresholds changed little across adjacent years. It does not make reported NSCH income equivalent to Medicaid modified adjusted gross income, recover within-year policy timing, or address income measurement error.

S6.5. Family-Income Imputation Sensitivity

The primary analysis uses the first NSCH family-poverty-ratio implicate and drops records explicitly flagged as income-imputed. The flag is unavailable for 2016, so the exclusion does not apply to that wave. This sensitivity retains flagged records, reconstructs the running variable separately with each of the six public-use poverty-ratio implicates, estimates the otherwise unchanged primary specification six times, and applies Rubin’s rules to the six bias-corrected estimates and squared robust standard errors.

Table S13: Psychiatric-Medication RD Across Six Family-Income Implicates

Analysis	2016–2019	2021–2022	2024
<i>Panel A. All children (primary)</i>			
Primary: implicate 1, flagged imputations excluded	0.097*** (0.037) [3,482]	0.169** (0.069) [2,022]	0.075 (0.063) [1,026]
Retained, implicate 1	0.082** (0.034) [4,476]	0.114** (0.051) [2,748]	0.037 (0.065) [1,440]
Retained, implicate 2	0.116*** (0.037) [4,609]	0.132** (0.054) [2,759]	0.055 (0.055) [1,453]
Retained, implicate 3	0.082** (0.032) [4,593]	0.086 (0.057) [2,742]	0.115* (0.062) [1,469]
Retained, implicate 4	0.084** (0.037) [4,550]	0.156*** (0.041) [2,755]	0.054 (0.058) [1,445]
Retained, implicate 5	0.069** (0.031) [4,528]	0.178*** (0.064) [2,724]	0.081 (0.057) [1,425]
Retained, implicate 6	0.083** (0.037) [4,630]	0.095* (0.056) [2,746]	0.154** (0.064) [1,467]
Rubin combined ($M = 6$)	0.086** (0.039) [4,476–4,630]	0.127* (0.066) [2,724–2,759]	0.083 (0.077) [1,425–1,469]
<i>Panel B. Current MEDB summary</i>			
Primary: implicate 1, flagged imputations excluded	0.328*** (0.113) [998]	0.438*** (0.138) [653]	0.126 (0.090) [354]
Rubin combined ($M = 6$)	0.333*** (0.104) [1,226–1,279]	0.331** (0.157) [818–839]	0.137 (0.191) [451–489]

Notes. Each implicate-specific cell reports a robust bias-corrected local-linear RD estimate, its robust standard error in parentheses, and N_h in brackets. Rubin rows report the combined estimate, total standard error, and range of N_h across implicates; there is no single combined count because implicate-specific income changes bandwidth membership. Sensitivity rows retain records flagged as income-imputed and use the strict separate-CHIP historically non-CE sample, a triangular kernel, survey weights, demographic covariates plus additive state and within-period year indicators, $h = b = 50$ FPL, and clustering on each implicate-specific integer running-variable support. Rubin total variance equals the mean squared robust standard error plus $(1 + 1/M)$ times the between-implicate variance; t reference degrees of freedom are used when that variance is positive. Children are not excluded based on SSI receipt or chronic functional limitation. RD denotes regression discontinuity, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, FPL denotes federal poverty level, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In the all-child population, the Rubin-combined estimates are 8.6 percentage points before the pandemic, 12.7 points during the federal enrollment freeze, and 8.3 points in 2024. The pre-period estimate rejects zero at the 5 percent level. The COVID estimate rejects at the 10 percent level but not at the 5 percent level,

and the 2024 estimate does not reject at the 10 percent level. The positive pre-period and COVID signs therefore remain when flagged records are restored and all six implicates are used, but the COVID inference is weaker than in the complete-case primary analysis.

The current-MEDB Rubin-combined estimates are 33.3, 33.1, and 13.7 percentage points. The first rejects zero at the 1 percent level, the second at the 5 percent level, and the third does not reject at the 10 percent level. These results preserve the positive earlier-period pattern but do not eliminate sensitivity to income measurement. Across implicates, the all-child 2024 estimates range from 3.7 to 15.4 points, while the current-MEDB estimates range from 0.9 to 32.6 points.

This diagnostic propagates the six public-use income implicates and tests the exclusion of flagged records. It does not resolve the difference between reported survey income and Medicaid MAGI, threshold measurement error, outcome nonresponse, or diagnosis selection. Rubin’s rules combine nonlinear local estimates whose samples change across implicates, so the combined values are sensitivity summaries rather than substitutes for identification.

S6.6. Medication-Measure Sensitivity Across Populations

The primary psychiatric-medication outcome is the past-12-month K4Q23 item. The expanded measure classifies a child as using psychiatric medication if K4Q23, current ADHD medication, or current autism medication is affirmative. As Section S4 explains, the expanded measure mixes reference periods and adds diagnosis-contingent items. Table S14 compares the two measures under the primary $h = b = 50$ specification for all children, children with a current MEDB diagnosis, and children without a current MEDB diagnosis.

Table S14: Primary and Expanded Psychiatric-Medication Measures Across Analysis Populations

Population	Period	Measure	$\hat{\tau}^{bc}$	SE	p (RV)	p (state)	$\hat{\tau}^{conv}$	Wild p	N_h
All children	Pre	K4Q23	0.097	0.037	0.009	0.032	0.056	0.174	3,482
		Expanded	0.080	0.038	0.034	0.107	0.044	0.230	3,483
	COVID	K4Q23	0.169	0.069	0.014	0.063	0.101	0.154	2,022
		Expanded	0.194	0.072	0.007	0.030	0.115	0.081	2,025
	2024	K4Q23	0.075	0.063	0.236	0.388	0.066	0.408	1,026
		Expanded	0.111	0.065	0.088	0.230	0.083	0.386	1,028
Current MEDB	Pre	K4Q23	0.328	0.113	0.004	0.010	0.198	0.172	998
		Expanded	0.258	0.140	0.065	0.036	0.148	0.326	999
	COVID	K4Q23	0.438	0.138	0.001	< 0.001	0.249	0.021	653
		Expanded	0.469	0.137	0.001	< 0.001	0.279	0.005	656
	2024	K4Q23	0.126	0.090	0.160	0.442	0.133	0.732	354
		Expanded	0.240	0.102	0.018	0.135	0.206	0.435	356
No current MEDB	Pre	K4Q23	−0.008	0.013	0.544	0.620	−0.011	0.255	2,484
		Expanded	−0.008	0.013	0.544	0.620	−0.011	0.255	2,484
	COVID	K4Q23	−0.002	0.007	0.781	0.791	−0.001	0.822	1,369
		Expanded	−0.002	0.007	0.781	0.791	−0.001	0.822	1,369
	2024	K4Q23	0.069	0.060	0.249	0.133	0.049	0.792	672
		Expanded	0.069	0.060	0.249	0.133	0.049	0.792	672

Notes. K4Q23 is the primary past-12-month psychiatric-medication outcome. The expanded measure equals one if K4Q23, current ADHD medication, or current autism medication is affirmative. It mixes a past-12-month item with two current-use items that are asked according to the corresponding reported diagnosis, so it is a measurement sensitivity rather than a second primary outcome. $\hat{\tau}^{bc}$ is the robust bias-corrected local-linear RD estimate; SE and p (RV) use clustering on running-variable support points; and p (state) clusters on states. Wild p is the state-level wild-cluster-bootstrap p -value for the conventional local-linear estimate $\hat{\tau}^{conv}$. The primary specification uses $h = b = 50$ FPL points, a triangular kernel, survey weights, covariates and state and period-year indicators, and strict separate-CHIP state-years among historically non-CE states. The sample includes children ages 6–17 and does not exclude children based on SSI receipt or chronic functional limitation. N_h is the unweighted complete-case count inside the effect bandwidth. The expanded measure adds only 1, 3, and 2 complete local observations in the pre, COVID, and 2024 all-child samples, respectively; the same additions occur in the current-MEDB samples, and none occurs in the no-current-MEDB samples. Reported p -values are unadjusted. Pre and COVID denote 2016–2019 and 2021–2022. RV denotes running variable, RD denotes regression discontinuity, MEDB denotes mental, emotional, developmental, or behavioral, ADHD denotes attention-deficit/hyperactivity disorder, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level.

For all children, replacing K4Q23 with the expanded measure changes the estimates from 9.7 to 8.0 percentage points before the pandemic, from 16.9 to 19.4 points during COVID, and from 7.5 to 11.1 points in 2024. Both measures give positive estimates in all three periods. Before the pandemic and during COVID, both reject zero at the 5 percent level with support-point clustering. State-clustered inference is less uniform. The K4Q23 estimate rejects at the 5 percent level before the pandemic, whereas the expanded-measure estimate rejects at that level during COVID. Neither conventional all-child estimate rejects under the state wild bootstrap at the 5 percent level. In 2024, neither measure rejects zero at the 5 percent level under any of the three approaches.

Among children with a current MEDB diagnosis, the expanded measure changes the estimates from 32.8 to 25.8 percentage points before the pandemic, from 43.8 to 46.9 points during COVID, and from 12.6 to 24.0 points in 2024. During COVID, both measures reject zero at the 5 percent level under support-point clustering, state clustering, and the wild bootstrap. Before the pandemic, the expanded measure does not reject at the 5 percent level with support-point clustering or the wild bootstrap, although it does with state clustering. In 2024, the expanded measure rejects at the 5 percent level only with support-point clustering. The primary K4Q23 measure does not reject at that level under any of the three approaches in 2024.

For children without a current MEDB diagnosis, the two measures give identical local estimates and inference in every period. The estimates are -0.8 percentage points before the pandemic, -0.2 points during COVID, and 6.9 points in 2024, and none rejects zero at the 5 percent level under any of the three approaches. Adding the current ADHD and autism medication items therefore does not create a detectable local medication discontinuity in the specificity population.

Across the two medication definitions, the all-child and current-MEDB COVID estimates remain positive, and both current-MEDB COVID estimates reject at the 5 percent level under all three inference approaches. The pre-period significance pattern changes with the population, medication measure, and clustering level. The expanded measure increases the 2024 point estimates, but state-clustered and wild-bootstrap inference still do not reject zero at the 5 percent level. These comparisons support using K4Q23 as the primary outcome and the expanded measure as a measurement sensitivity rather than as a second preferred outcome.

S7. Design Validation and Diagnosis-Composition Diagnostics

This section reports four checks on the primary design and its diagnosed-subgroup interpretation. Section S7.1 reports running-variable density and measured-composition results for the primary all-child population. Sections S7.2 and S7.3 examine all-child psychiatric-medication estimates at shifted cutoffs and after excluding observations nearest the threshold. Section S7.4 reports severity and diagnosis-count results within the current-MEDB secondary population.

S7.1. Density and Measured-Composition Continuity

The primary RD design requires potential outcomes to vary smoothly through the Medicaid threshold. Density and measured-composition checks can reveal irregular sorting or observed composition changes that would make this requirement less plausible, but they cannot rule out discontinuities in unmeasured characteristics. Table S15 reports the density tests underlying the main paper’s all-child density figure. Table S16 reports the 18 all-child balance tests summarized in the main paper.

Table S15: Running-Variable Density Tests in the Primary All-Child Population

Period	T -statistic	p -value	Observations
2016–2019	0.774	0.439	24,487
2021–2022	−0.955	0.339	14,361
2024	0.272	0.786	8,646

Notes. The table reports the robust `rddensity` statistic and two-sided p -value (Cattaneo et al., 2020). The running variable is family income minus the state-year Medicaid threshold, both measured in FPL points. Tests use the full available running-variable support in the strict separate-CHIP historically non-CE all-child sample, ages 6–17. Children are not excluded based on SSI receipt or chronic functional limitation. The density test is unweighted and does not condition on the medication outcome being observed. RD denotes regression discontinuity, FPL denotes federal poverty level, CHIP denotes Children’s Health Insurance Program, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, and SSI denotes Supplemental Security Income.

None of the three density tests rejects continuity at the 5 percent level. The measured-composition results show the same pattern before the pandemic and during 2021–2022: none of the 18 tests rejects at that level. In 2024, the residual household-language category rises by 15.2 percentage points and rejects at the 5 percent level. One rejection among 18 unadjusted tests can occur by chance, so it is a caution rather than evidence that the two sides differ systematically. Because the balance measures are observed rather than exhaustive, these results do not establish balance in unmeasured determinants.

Table S16: Measured-Composition Balance in the Primary All-Child Population

Measured characteristic	2016–2019	2021–2022	2024
Family structure (harmonized code)	−0.036 (0.256) [3,457]	−0.100 (0.205) [2,039]	−0.253 (0.329) [1,032]
Race: Black or African American alone	−0.029 (0.045) [3,601]	−0.077 (0.087) [2,079]	−0.239* (0.128) [1,048]
Race: American Indian or Alaska Native alone	0.021 (0.020) [3,601]	0.033* (0.020) [2,079]	0.019 (0.021) [1,048]
Race: Asian alone	0.030* (0.016) [3,601]	−0.024 (0.024) [2,079]	0.086* (0.048) [1,048]
Race: Native Hawaiian or Other Pacific Islander alone	0.002 (0.005) [3,601]	0.003 (0.008) [2,079]	0.006 (0.007) [1,048]
Race: two or more races	0.013 (0.016) [3,601]	−0.068* (0.041) [2,079]	−0.021 (0.056) [1,048]
Female	−0.078 (0.072) [3,601]	−0.082 (0.158) [2,079]	−0.162 (0.129) [1,048]
Household language: Spanish	−0.117 (0.075) [3,561]	0.068 (0.091) [2,067]	−0.112 (0.151) [1,043]
Household language: other	0.011 (0.025) [3,561]	−0.018 (0.029) [2,067]	0.152*** (0.059) [1,043]
Adult education: high school/vocational	−0.055 (0.091) [3,586]	−0.075 (0.087) [2,079]	−0.102 (0.191) [1,048]
Adult education: more than high school	0.044 (0.094) [3,586]	0.123 (0.083) [2,079]	−0.132 (0.183) [1,048]
Two children	−0.045 (0.095) [3,601]	0.016 (0.120) [2,079]	0.114 (0.187) [1,048]
Three children	−0.030 (0.103) [3,601]	−0.052 (0.117) [2,079]	0.382* (0.218) [1,048]
Four or more children	−0.056 (0.069) [3,601]	0.027 (0.079) [2,079]	−0.295 (0.196) [1,048]
Adult 1 mental health: very good	−0.078 (0.083) [3,580]	−0.144 (0.090) [2,073]	−0.016 (0.186) [1,045]
Adult 1 mental health: good	0.038 (0.078) [3,580]	−0.058 (0.076) [2,073]	−0.080 (0.144) [1,045]
Adult 1 mental health: fair	0.044 (0.028) [3,580]	0.046 (0.028) [2,073]	0.125* (0.065) [1,045]
Adult 1 mental health: poor	−0.015 (0.013) [3,580]	0.000 (0.016) [2,073]	0.001 (0.002) [1,045]

Notes. Each cell reports a robust bias-corrected local-linear RD estimate, its robust standard error in parentheses, and N_h in brackets. Models use survey weights, a triangular kernel, $h = b = 50$ FPL, additive state and within-period year indicators, and standard errors clustered on running-variable support points. They do not adjust for the other measured characteristics. The sample is the strict separate-CHIP historically non-CE all-child population, ages 6–17, and retains children regardless of SSI receipt or chronic functional limitation. Adult 1 mental health is contemporaneous, and family structure can also change over time; the table therefore describes measured composition rather than a set of exclusively predetermined covariates. Reference groups are White alone for race, male for sex, English for household language, less than high school for adult education, one child for household size, and excellent for Adult 1 mental health. No multiplicity adjustment is applied. RD denotes regression discontinuity, FPL denotes federal poverty level, CHIP denotes Children’s Health Insurance Program, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

S7.2. Shifted-Cutoff Diagnostics

Figure S3 re-estimates the primary all-child medication model at cutoffs from 60 FPL points below to 60 points above the Medicaid threshold. For every nonzero shift, the sample is restricted to the corresponding side of the actual threshold so that neither 50-FPL estimation window includes it. The shifted estimates are diagnostics rather than placebo treatment effects because income may be related to medication use away from the eligibility boundary.

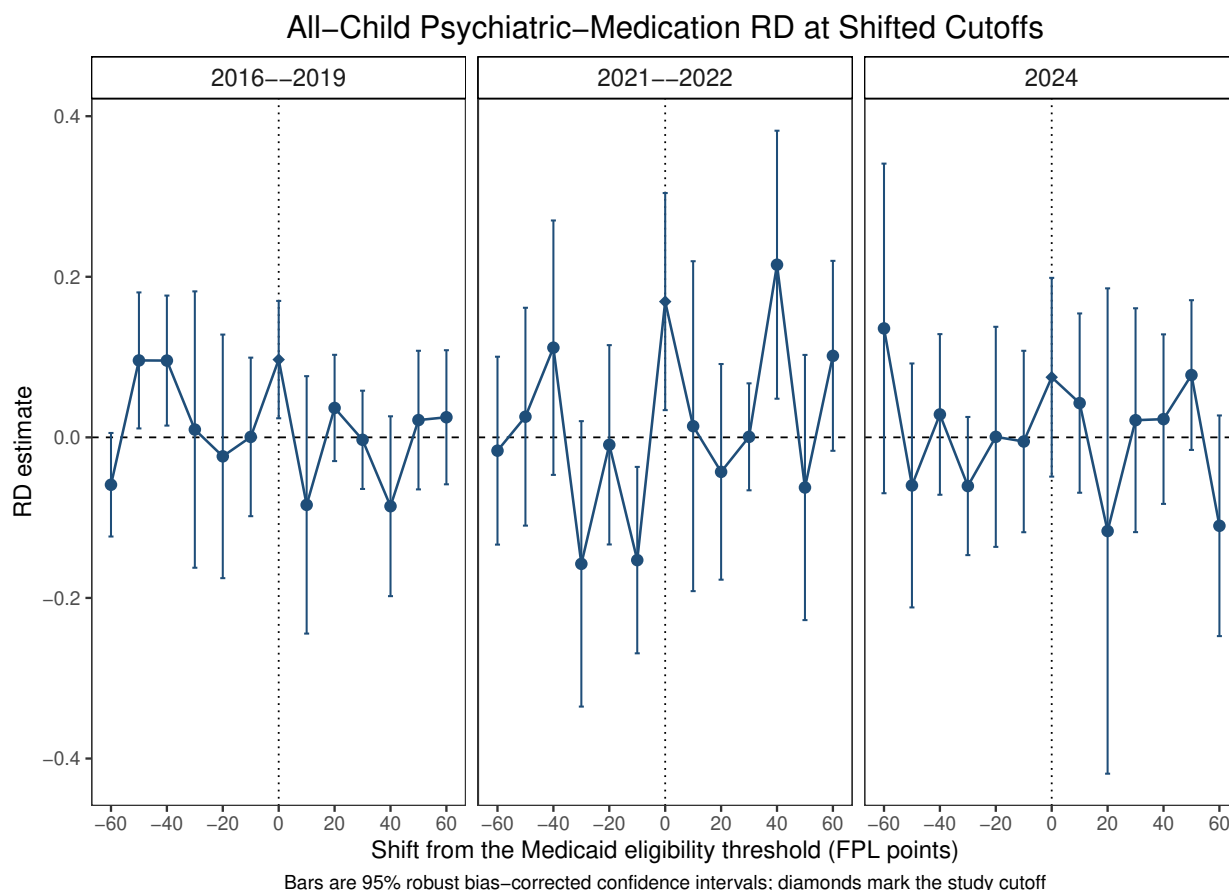


Figure S3: All-Child Psychiatric-Medication RD at Shifted Cutoffs

Notes. Points are robust bias-corrected local-linear estimates and bars are 95 percent confidence intervals. Diamonds mark the study cutoff. Negative shifts use only observations below the study cutoff; positive shifts use only observations above it. All models use the strict separate-CHIP historically non-CE all-child sample, ages 6–17, the primary $h = b = 50$ FPL bandwidths, a triangular kernel, survey weights, demographic covariates plus additive state and within-period year indicators, and standard errors clustered on shifted running-variable support points. Children are not excluded based on SSI receipt or chronic functional limitation. No multiplicity adjustment is applied across shifted cutoffs. RD denotes regression discontinuity, FPL denotes federal poverty level, CHIP denotes Children’s Health Insurance Program, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, and SSI denotes Supplemental Security Income.

The estimates at the study cutoff reproduce the 9.7-point pre-period, 16.9-point 2021–2022, and 7.5-point 2024 primary results. Among the 36 nonzero shifted cutoffs, two pre-period estimates and two 2021–2022

estimates reject zero at the unadjusted 5 percent level; none does so in 2024. The nonzero signs alternate, and the tests use overlapping data and are not independent. Even so, the actual threshold is not the only location with a raw 5-percent rejection. The grid therefore does not provide a clean falsification pass and leaves open the possibility of other local irregularities in the income–medication relationship.

S7.3. Donut-Width Sensitivity

Figure S4 drops observations increasingly close to the Medicaid threshold while keeping the primary estimator and 50-FPL bandwidths fixed. Radius zero is the primary model, and radius three is the donut specification reported in Section S6.2.

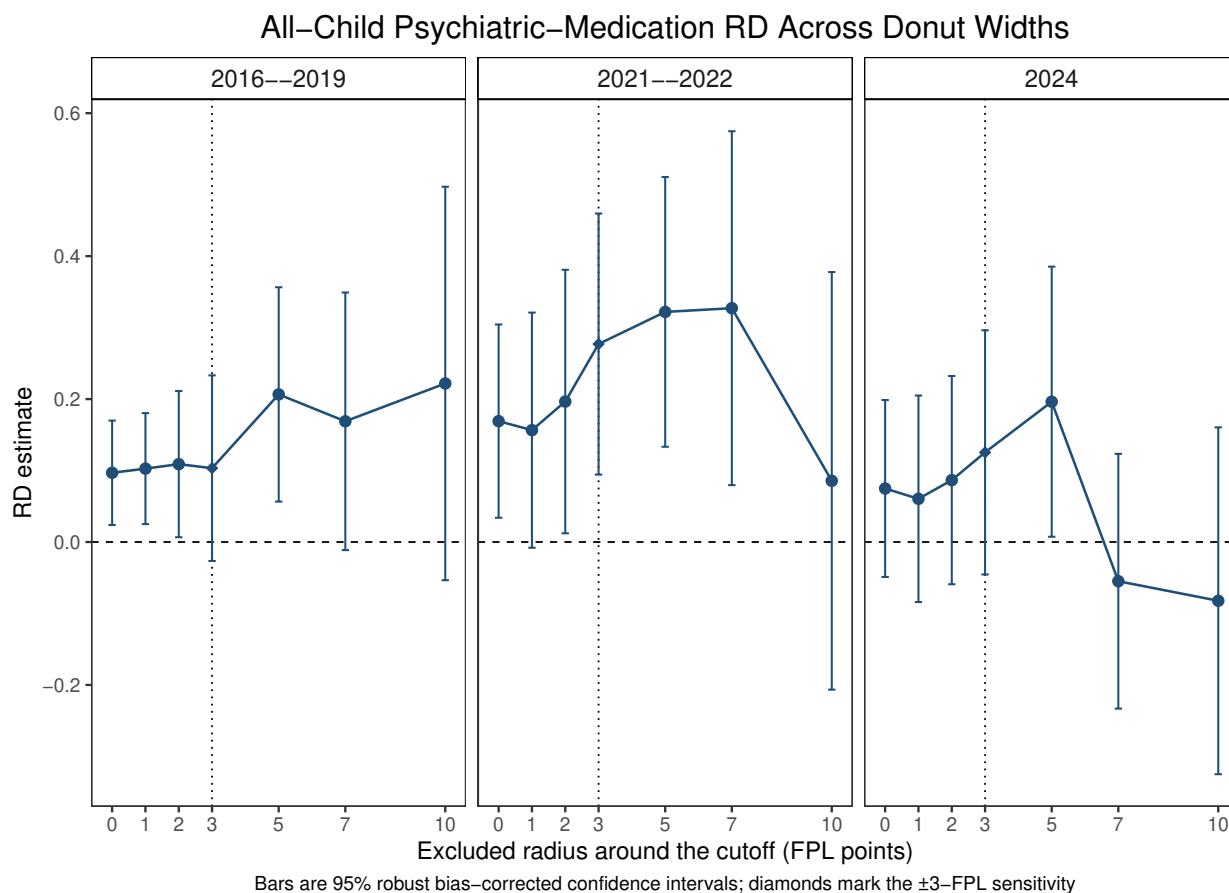


Figure S4: All–Child Psychiatric–Medication RD Across Donut Widths

Notes. Points are robust bias-corrected local-linear estimates and bars are 95 percent confidence intervals. A radius of r drops observations with $|\text{distance}_M| < r$ while retaining observations exactly at $\pm r$; diamonds mark the ± 3 -FPL sensitivity. All models use the strict separate-CHIP historically non-CE all-child sample, ages 6–17, $h = b = 50$ FPL, a triangular kernel, survey weights, demographic covariates plus additive state and within-period year indicators, and standard errors clustered on running-variable support points. Children are not excluded based on SSI receipt or chronic functional limitation. RD denotes regression discontinuity, FPL denotes federal poverty level, CHIP denotes Children’s Health Insurance Program, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, and SSI denotes Supplemental Security Income.

The pre-period and 2021–2022 estimates remain positive at every retained radius, although their magnitudes and precision change. At the ± 3 -FPL radius, the estimates are 10.3 and 27.7 percentage points; the former does not reject zero at the 10 percent level, while the latter rejects at the 1 percent level. At the widest ± 10 -FPL radius, the 2021–2022 estimate falls to 8.6 points and does not reject at the 10 percent level. The 2024 estimate remains positive through ± 5 FPL but becomes negative at the two widest radii. These results support earlier-period sign stability over the displayed radii, not a stable magnitude, and they reinforce the fragility of the 2024 estimate.

S7.4. Severity and Diagnosis Composition Within Current MEDB

The all-child diagnosis-selection results in Section S12.1 test whether entry into the current-MEDB population changes at the threshold. Table S17 instead asks whether reported severity and the number of diagnoses differ within the diagnosed population. These measures cannot restore a common subgroup composition because diagnosis status and diagnosis mix may themselves change at the threshold.

Table S17: Severity and Five-Condition Diagnosis-Count Diagnostics Within Current MEDB

Outcome	2016–2019	2021–2022	2024
Any condition rated moderate or severe	0.052 (0.110) [1,007]	0.169 (0.133) [661]	0.149 (0.114) [358]
Any condition rated severe	0.053 (0.079) [1,007]	−0.018 (0.108) [661]	−0.131 (0.142) [358]
Five-condition diagnosis count	0.146 (0.215) [1,007]	0.695** (0.275) [661]	0.151 (0.180) [358]

Notes. Each cell reports a robust bias-corrected local-linear RD estimate, its robust standard error in parentheses, and N_h in brackets. The sensitivity specification uses the strict separate-CHIP historically non-CE current-MEDB sample, ages 6–17, $h = 50$, $b = 70$ FPL, a triangular kernel, survey weights, demographic covariates plus additive state and within-period year indicators, and standard errors clustered on running-variable support points. Children are retained regardless of SSI receipt or chronic functional limitation. The five harmonized conditions are ADHD, depression, anxiety, behavior or conduct problems, and autism. Developmental delay, speech disorder, and Tourette syndrome enter the broader MEDB definition but not these three diagnostics. The 70-FPL bias-correction window can cross the next institutional cutoff in some state-years, so the table is a sensitivity rather than a primary specification. RD denotes regression discontinuity, ADHD denotes attention-deficit/hyperactivity disorder, MEDB denotes mental, emotional, developmental, or behavioral, FPL denotes federal poverty level, CHIP denotes Children’s Health Insurance Program, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Neither reported-severity measure rejects continuity at the 10 percent level in any period. During 2021–2022, however, diagnosed children just above the threshold have an estimated 0.70 more of the five counted diagnoses, and that difference rejects zero at the 5 percent level. Adding the diagnosis count to the medication model lowers the corresponding current-MEDB estimate from 36.7 to 22.9 percentage points; the adjusted estimate does not reject zero at the 5 percent level, although it is close to doing so. This reduction suggests that diagnosis composition may account for part of the diagnosed-subgroup medication estimate. It does not eliminate selection concerns because diagnosis count is itself measured after the threshold assignment and may change at the threshold.

S8. Local-Polynomial and Kernel Sensitivity

The primary model uses a local-linear polynomial and triangular kernel. Section S6.2 varies bandwidths, covariate adjustment, observations nearest the cutoff, and fixed effects while retaining that estimator. Table S18 instead varies the local-polynomial order and kernel while holding the all-child population, strict state sample, covariates, weights, clustering level, and $h = b = 50$ bandwidths fixed. The local-quadratic rows use a cubic polynomial for bias correction within the same 50-FPL window.

Table S18: All-Child Psychiatric-Medication RD Across Local-Polynomial Orders and Kernels

Local polynomial	Kernel	2016–2019	2021–2022	2024
Linear	Triangular	0.097*** (0.037) [3,482]	0.169** (0.069) [2,022]	0.075 (0.063) [1,026]
Linear	Uniform	0.091** (0.040) [3,553]	0.163*** (0.063) [2,065]	0.059 (0.059) [1,033]
Linear	Epanechnikov	0.103*** (0.040) [3,482]	0.173** (0.069) [2,022]	0.068 (0.063) [1,026]
Quadratic	Triangular	0.061 (0.041) [3,482]	0.142* (0.085) [2,022]	0.103 (0.073) [1,026]
Quadratic	Uniform	0.119*** (0.044) [3,553]	0.160** (0.078) [2,065]	0.064 (0.076) [1,033]
Quadratic	Epanechnikov	0.070* (0.041) [3,482]	0.140* (0.085) [2,022]	0.098 (0.076) [1,026]

Notes. Each cell reports a robust bias-corrected RD estimate, its robust standard error in parentheses, and N_h in brackets. All models use the strict separate-CHIP historically non-CE all-child sample, ages 6–17, $h = b = 50$ FPL, survey weights, demographic covariates plus additive state and within-period year indicators, and standard errors clustered on running-variable support points. Children are retained regardless of SSI receipt or chronic functional limitation. The linear models use a quadratic polynomial for bias correction, and the quadratic models use a cubic polynomial. Uniform-kernel rows have slightly larger N_h because observations exactly 50 FPL points from the cutoff receive positive weight; the triangular and Epanechnikov kernels assign zero weight at that boundary. RD denotes regression discontinuity, FPL denotes federal poverty level, CHIP denotes Children’s Health Insurance Program, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17, and SSI denotes Supplemental Security Income. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The estimated medication discontinuity is positive in all six model combinations in every period. Before the pandemic, four estimates reject zero at the 5 percent level and one additional estimate rejects at the 10 percent level. During 2021–2022, four reject at the 5 percent level and the other two reject at the 10 percent level. The local-quadratic results therefore retain the earlier-period positive signs but show that inference

depends on the polynomial and kernel. In 2024, the six estimates range from 5.9 to 10.3 percentage points, and none rejects zero at the 10 percent level. The 2024 conclusion remains imprecise even though all six point estimates are positive.

S9. Multiple-Hypothesis Correction

The main paper reports five outcomes for the primary all-child population: past-12-month psychiatric medication, mental-health treatment or counseling, Medicaid/CHIP coverage, private-only coverage, and uninsured status. This section shows whether the period-specific findings remain statistically significant after accounting for those five tests. Benjamini–Hochberg and Holm adjustments are applied separately within each period under the primary $h = b = 50$ specification.

Method choice. Benjamini–Hochberg controls the expected false-discovery proportion among findings declared statistically significant, whereas Holm controls the probability of at least one false rejection within the five-outcome family (Benjamini and Hochberg, 1995; Holm, 1979). Holm is therefore the stricter check in this application. These analytic adjustments are not a substitute for a Romano–Wolf resampling procedure.

Reading the table. The pre-period psychiatric-medication estimate remains statistically significant at the 5 percent level under both adjustments. The COVID medication estimate remains significant at the 10 percent level but not at the 5 percent level. No 2024 outcome remains significant at the 10 percent level after either adjustment. Thus, among the five outcomes tested together, only the pre-period medication estimate remains statistically significant at the 5 percent level under both methods.

Table S19: Multiple-Hypothesis Correction Across the Five Primary All-Child Outcomes, Primary $h = b = 50$ Specification

Period	Outcome	$\hat{\tau}$	SE	p (unadj.)	p (BH)	p (Holm)	N_h
Pre (2016–2019)	Psychiatric medication	0.097	0.037	0.009***	0.047**	0.047**	3,482
	Mental-health treatment or counseling	−0.009	0.043	0.842	0.842	1.000	3,512
	Medicaid/CHIP coverage	0.107	0.083	0.196	0.326	0.587	3,395
	Private-only insurance	−0.022	0.071	0.756	0.842	1.000	3,466
	Uninsured	−0.076	0.051	0.138	0.326	0.553	3,466
COVID (2021–2022)	Psychiatric medication	0.169	0.069	0.014**	0.071*	0.071*	2,022
	Mental-health treatment or counseling	0.104	0.055	0.058*	0.126	0.233	2,054
	Medicaid/CHIP coverage	−0.126	0.078	0.105	0.131	0.233	2,047
	Private-only insurance	0.137	0.077	0.076*	0.126	0.233	2,020
	Uninsured	−0.006	0.045	0.894	0.894	0.894	2,020
Post (2024)	Psychiatric medication	0.075	0.063	0.236	0.295	0.684	1,026
	Mental-health treatment or counseling	−0.164	0.136	0.228	0.295	0.684	1,036
	Medicaid/CHIP coverage	0.275	0.140	0.049**	0.123	0.197	1,033
	Private-only insurance	−0.104	0.120	0.383	0.383	0.684	1,022
	Uninsured	−0.221	0.107	0.039**	0.123	0.193	1,022

Notes. $\hat{\tau}$ is the robust bias-corrected local-linear RD estimate; SE is its robust standard error; p (unadj.) is the unadjusted robust p -value; and N_h is the unweighted complete-case count inside the effect bandwidth. BH and Holm adjustments are applied within each period’s family of five all-child outcomes. The primary specification sets effect and bias-correction bandwidths to $h = b = 50$ FPL points and uses a triangular kernel, survey weights, covariates and state and period-year indicators, and standard errors clustered on running-variable support points. The sample includes all children ages 6–17 in strict separate-CHIP state-years among historically non-CE states. It does not exclude children based on SSI receipt or chronic functional limitation. BH denotes Benjamini–Hochberg, FDR denotes false discovery rate, FWER denotes family-wise error rate, RD denotes regression discontinuity, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

S10. Additional Supporting Tables

This section reports sample sizes, threshold summary statistics, the full non-MEDB falsification table, identification limits, and institutional coverage terms referenced in the journal manuscript. The main paper carries a compact summary of the non-MEDB falsification that condenses the full table reported here.

S10.1. Sample Sizes by Period, CE Group, and Current-MEDB Status

Table S20 reports counts before applying an outcome-specific complete-case restriction or the RD effect bandwidth. The primary all-child historically non-CE population is the sum of the current-MEDB and non-current-MEDB rows: 24,487 children before the pandemic, 14,361 during COVID, and 8,646 in 2024. The corresponding current-MEDB secondary populations contain 6,147, 4,219, and 2,823 children. The smaller N_h values in the result tables arise because those models additionally restrict to the effect bandwidth and require nonmissing values for the outcome, survey weight, and model covariates.

Table S20: Strict Separate-CHIP Sample Sizes by Period, Historical Continuous Eligibility, and Current-MEDB Status

Period	Historical CE Group	Current MEDB	N	N_{below}	N_{above}
Pre (2016–2019)	Non-CE	Yes	6,147	1,188	4,959
Pre (2016–2019)	Non-CE	No	18,340	2,696	15,644
Pre (2016–2019)	CE	Yes	6,119	1,265	4,854
Pre (2016–2019)	CE	No	18,413	2,675	15,738
COVID (2021–2022)	Non-CE	Yes	4,219	874	3,345
COVID (2021–2022)	Non-CE	No	10,142	1,764	8,378
COVID (2021–2022)	CE	Yes	5,454	1,110	4,344
COVID (2021–2022)	CE	No	12,353	2,039	10,314
Post (2024)	Non-CE	Yes	2,823	441	2,382
Post (2024)	Non-CE	No	5,823	783	5,040
Post (2024)	CE	Yes	1,984	358	1,626
Post (2024)	CE	No	3,923	582	3,341

Notes. N is the unweighted count on the full observed running-variable support; it is not an effect-bandwidth or outcome-specific count. N_{below} and N_{above} partition that count at the state-year Medicaid eligibility threshold. The strict sample includes children ages 6–17, requires a nonmissing separate-CHIP threshold, and retains state-years only when the absolute difference between the separate-CHIP and Medicaid thresholds is at least 50 FPL points. It does not exclude children based on SSI receipt or chronic functional limitation. Pre, COVID, and post denote 2016–2019, 2021–2022, and 2024. Historical CE status records whether the state’s January 2016 Medicaid policy provided 12-month continuous eligibility to children; 24 jurisdictions are classified as CE from KFF Table 14. The 2024 federal mandate applied 12-month continuous eligibility nationwide, so “Non-CE” in 2024 means historically non-CE. CE denotes continuous eligibility, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level.

S10.2. Descriptive Statistics Around the Threshold

Table S21 describes the pre-period current-MEDB secondary population within 50 FPL points of the threshold. It reports unweighted levels and does not estimate the RD. The table is retained to show the observed prevalence of the main coverage and behavioral-health measures in the diagnosed subgroup that motivated the original analysis.

Table S21: Unweighted Descriptive Statistics within 50 FPL Points of the Medicaid Threshold (Pre-Period, Historically Non-CE States, Current-MEDB Children)

Variable	Below Threshold			Above Threshold			Raw Diff.
	Mean	SD	<i>N</i>	Mean	SD	<i>N</i>	
<i>Coverage Type</i>							
Medicaid/CHIP	0.766	0.424	474	0.555	0.497	544	−0.211
Private-only insurance	0.198	0.399	479	0.404	0.491	557	0.206
Uninsured	0.040	0.195	479	0.050	0.219	557	0.011
<i>Behavioral-Health Utilization</i>							
Past-12-month psychiatric medication	0.501	0.501	481	0.432	0.496	560	−0.069
Mental-health treatment or counseling	0.445	0.498	485	0.435	0.496	559	−0.011
<i>Access and Unmet Need</i>							
Any unmet need	0.117	0.322	486	0.129	0.336	565	0.012
<i>Perceived Insurance Adequacy</i>							
Insurance covers needed services	0.903	0.297	463	0.898	0.303	531	−0.005

Notes. Entries are unweighted descriptive means and standard deviations; *N* is the unweighted nonmissing count. “Below” is `distance_M` $\in [-50, 0)$ and “Above” is `distance_M` $\in [0, 50]$, where `distance_M` is family income in FPL percentage points minus the Medicaid threshold. Raw Diff. is the above-threshold mean minus the below-threshold mean. The table does not use survey weights, fit side-specific income slopes, adjust for covariates, or estimate fitted limits at the threshold. It therefore should not be interpreted as an RD effect or used for statistical inference. The sample is strict separate-CHIP historically non-CE current-MEDB children ages 6–17 in 2016–2019. SD denotes standard deviation, CHIP denotes Children’s Health Insurance Program, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy, MEDB denotes mental, emotional, developmental, or behavioral, and FPL denotes federal poverty level.

Across the full 50-point intervals, the raw coverage means show less Medicaid/CHIP coverage and more private-only coverage above the threshold, while uninsurance is similar on the two sides. Those patterns describe coverage composition but cannot establish that children move from Medicaid to separate CHIP because the NSCH public-coverage item combines the two programs. The raw psychiatric-medication mean is 6.9 percentage points lower above the threshold, whereas the adjusted local-linear RD estimate for the current-MEDB subgroup is positive. The difference in sign is not a numerical inconsistency: the table averages all observations within each 50-point interval, while the RD compares fitted limits at the threshold after weighting and adjustment. The table should therefore be read as context for outcome levels, not as an alternative estimate of the discontinuity.

S10.3. Outcomes Among Children Without a Current MEDB Diagnosis

Children without a current MEDB diagnosis provide the medication-specificity check for H1. Table S22 reports psychiatric medication and ancillary outcomes in both historical CE groups. The historically non-CE columns correspond to the paper’s primary institutional comparison; the historically CE columns provide descriptive context. The medication row asks whether the all-child medication discontinuity is also present among children for whom none of the eight current-diagnosis indicators is affirmative. The remaining rows show whether the same threshold is associated with broader coverage, counseling, access, or perceived insurance-adequacy changes in that population.

Table S22: RD Estimates for Children Without a Current MEDB Diagnosis Across Reported Specificity Outcomes

Outcome	Historically Non-CE States			Historically CE States		
	Pre	COVID	Post	Pre	COVID	Post
<i>Coverage Type</i>						
Medicaid/CHIP	0.165** (0.084)	−0.148 (0.105)	0.200 (0.166)	0.005 (0.059)	0.024 (0.126)	−0.365*** (0.117)
N_h	2,418	1,389	677	2,511	1,612	447
Private-only insurance	−0.078 (0.087)	0.146 (0.103)	−0.051 (0.149)	0.059 (0.056)	−0.137 (0.146)	0.315*** (0.119)
N_h	2,472	1,370	668	2,573	1,590	444
Uninsured	−0.075 (0.074)	0.016 (0.065)	−0.327** (0.155)	−0.024 (0.034)	0.060 (0.059)	0.056 (0.034)
N_h	2,472	1,370	668	2,573	1,590	444
<i>Behavioral-Health Utilization</i>						
Psychiatric medication	−0.008 (0.013)	−0.002 (0.007)	0.069 (0.060)	−0.010* (0.006)	−0.008* (0.005)	−0.047 (0.033)
N_h	2,484	1,369	672	2,580	1,579	440
Mental-health treatment or counseling	0.014 (0.030)	0.017 (0.035)	0.001 (0.072)	−0.037 (0.023)	0.065** (0.027)	0.073 (0.055)
N_h	2,510	1,397	681	2,597	1,603	447
<i>Access and Unmet Need</i>						
Any unmet need	−0.040 (0.059)	0.024 (0.027)	0.029 (0.051)	0.016 (0.014)	0.034 (0.041)	0.024 (0.037)
N_h	2,515	1,392	681	2,598	1,614	448
<i>Perceived Insurance Adequacy</i>						
Insurance covers needed services	0.111* (0.065)	−0.015 (0.050)	0.031 (0.046)	0.005 (0.056)	0.045 (0.035)	−0.053 (0.053)
N_h	2,269	1,284	590	2,449	1,507	399

Notes. Each cell is a robust bias-corrected local-linear RD estimate, with its robust standard error in parentheses; N_h is the unweighted complete-case count inside the effect bandwidth. The strict separate-CHIP sample includes children ages 6–17 without a current MEDB diagnosis and does not exclude children based on SSI receipt or chronic functional limitation. Estimates use the primary $h = b = 50$ FPL-point specification, a triangular kernel, survey weights, covariates and state and period-year indicators, and standard errors clustered on running-variable support points. Pre, COVID, and post denote 2016–2019, 2021–2022, and 2024. RD denotes regression discontinuity, CE denotes continuous eligibility, MEDB denotes mental, emotional, developmental, or behavioral, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level. Stars use unadjusted robust p -values; these ancillary outcomes are not adjusted as a common multiple-testing family. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In historically non-CE states, the medication estimates are −0.8 percentage points before the pandemic, −0.2 points during COVID, and 6.9 points in 2024. None is statistically significant at the 10 percent level. The near-zero pre-period and COVID estimates support H1’s prediction that the medication discontinuity is concentrated among children with a current diagnosis. They do not rule out a smaller change because medication use is uncommon among children without a current diagnosis. The positive but imprecise 2024

estimate does not reproduce the earlier near-zero pattern.

The ancillary outcomes do not show a common pattern across periods. In the primary historically non-CE comparison, Medicaid/CHIP coverage is higher above the threshold before the pandemic and uninsurance is lower above the threshold in 2024 under unadjusted 5 percent inference. The other coverage, counseling, access, and adequacy estimates do not reject zero at that level. These scattered results do not indicate a general increase or decrease across all outcomes, and they do not change the medication-specific interpretation.

S10.4. Identification Limits and Diagnostic Evidence

Table S23 links five identification limits discussed in the main paper to the corresponding diagnostic evidence and to the information that remains unavailable.

Table S23: Identification Limits and Diagnostic Evidence

#	Limit	Diagnostic evidence and remaining information need
1	The running variable is a prior-calendar-year family-income measure, not Medicaid MAGI	Section S6.4 aligns the KFF threshold with the income reference year on a common sample. The all-child estimates change from 0.096 to 0.094 before the pandemic, from 0.169 to 0.161 during 2021–2022, and not at all in 2024. Section S6.5 retains records flagged as income-imputed and combines all six NSCH income implicates; the all-child combined estimates are 0.086, 0.127, and 0.083. These checks do not reconstruct MAGI, recover within-year income, or remove income measurement error
2	The state-year cutoff is a KFF-published threshold, not an eligibility value reconstructed from each state’s operational Medicaid rules	The analysis checks the archived threshold inputs against the corresponding annual KFF tables. Section S11.2 also reports a 31-row non-KFF documentary ledger for selected states and years. That ledger provides institutional context rather than a complete year-specific validation audit: several documents are later or current, some displayed comparisons use a MAGI-disregard conversion, and two Florida rows retain a 5-FPL difference. Administrative eligibility files or contemporaneous state manuals for every state-year would be needed to validate the operational cutoff directly
3	Current-MEDB analyses condition on a diagnosis that may differ across the threshold	Section S12.1 estimates diagnosis prevalence in the all-child population, and Section S7.4 examines severity and diagnosis count within the current-MEDB population. Diagnosis prevalence changes in some cells, and the five-condition diagnosis count rises during 2021–2022. The current-MEDB estimates therefore do not compare a diagnosed subgroup with fixed composition on both sides. A design using diagnosis measured before the income threshold or before any coverage change would be needed to hold diagnosis composition fixed
4	The psychiatric-medication outcome is missing for a small share of children, and item response is not smooth before the pandemic	Within the estimation window, the medication item is observed for more than 98 percent of children in each period. The response discontinuity is 0.060 before the pandemic and rejects continuity at the 5 percent level; the 0.025 and 0.037 discontinuities in the later periods do not. Coding every missing response as medication reduces the medication estimates to 0.040, 0.145, and 0.046, and the pre-period estimate no longer rejects zero at the 5 percent level. Administrative medication fills or complete outcome follow-up would be needed to remove this response concern
5	NSCH is a repeated cross-section and does not observe program enrollment, switching, prescribers, or treatment continuity	The all-child design avoids conditioning the primary estimate on diagnosis, and the population-specific results show where the medication pattern is concentrated. These comparisons do not identify the pathway producing the discontinuity, and the current-MEDB and no-current-MEDB estimates are not compared with a covariance-aware test. Longitudinal records linking income and eligibility histories to Medicaid or separate-CHIP enrollment, plan rules, providers, psychotherapy, and medication fills would be needed to distinguish program switching, prescribing-rule changes, provider disruption, and other mechanisms

Notes. In the primary all-child population, none of the 18 measured-composition tests rejects at the 5 percent level before the pandemic or during 2021–2022; one rejects in 2024 (Table S16). None of the three density tests rejects continuity at that level (Table S15). Among 36 nonzero shifted-cutoff estimates, four reject at the unadjusted 5 percent level. Those tests use overlapping data and are dependent, but the shifted-cutoff grid is not a clean falsification pass (Section S7.2). The balance and shifted-cutoff tests are not adjusted for multiple comparisons, four caregiver-mental-health measures are contemporaneous, and measured balance cannot establish balance in unobserved characteristics. The COVID period also coincides with changes in therapy supply, telehealth, severity, and survey behavior. H2 is the psychotherapy-first versus medication-first subgroup comparison; those categories are relative clinical heuristics. MAGI denotes modified adjusted gross income, NSCH denotes National Survey of Children’s Health, RD denotes regression discontinuity, MEDB denotes mental, emotional, developmental, or behavioral, FPL denotes federal poverty level, KFF denotes KFF’s Medicaid and CHIP eligibility survey, and CHIP denotes Children’s Health Insurance Program.

S10.5. Medicaid and Separate-CHIP Coverage Terms

Table S24 summarizes selected federal baselines and areas of state or plan variation. It does not establish that either program is systematically more restrictive.

Table S24: Selected Medicaid and Separate-CHIP Coverage Terms for Children

	Medicaid	Separate CHIP
Federal benefit baseline	EPSDT provides comprehensive and preventive services to Medicaid-enrolled children under age 21 (42 U.S.C. § 1396d(r)), implemented through state Medicaid programs	Separate CHIP follows Title XXI benefit rules rather than the Medicaid EPSDT entitlement. The state plan defines the covered package subject to federal requirements
Cost sharing for children	Most children are exempt from premiums and most service-level cost sharing, subject to limited statutory and regulatory exceptions (Centers for Medicare & Medicaid Services, 2026b)	States may impose enrollment fees, premiums, deductibles, coinsurance, and copayments (Centers for Medicare & Medicaid Services, 2026a). For children at or below 150% FPL, premiums and service-level charges are subject to specific federal caps. Total family cost sharing may not exceed 5% of family income (U.S. Code of Federal Regulations, 2026)
Behavioral-health delivery and network	Fee-for-service and managed-care arrangements vary by state, eligibility group, and contract	State-administered and contracted arrangements vary by state and contract
Prior authorization	Varies by state and plan. Many Medicaid programs use prior authorization for pediatric antipsychotics, as documented by Schmid et al. (2015) and Blake et al. (2019)	Prior-authorization practices vary across state-contracted plans. Published evidence directly comparing prior-authorization burden between separate CHIP and state Medicaid is limited
Continuous eligibility	State option before 2024. Since January 1, 2024, federal law has required 12 months of continuous eligibility for enrolled children under age 19, subject to limited exceptions (CAA 2023 § 5112)	The same federal 12-month requirement applies. One exception permits separate-CHIP coverage to end if the child becomes eligible for Medicaid

Notes. The table summarizes federal baselines, not the terms of a particular child's coverage. Benefits, cost sharing, delivery systems, networks, and authorization rules vary across states, eligibility groups, plans, and contracts. Direct comparisons of children's behavioral-health coverage are limited. See Bisgaier and Rhodes (2011), Hu et al. (2023), and Zhu et al. (2022). CHIP denotes Children's Health Insurance Program, CE denotes continuous eligibility, EPSDT denotes Early and Periodic Screening, Diagnostic, and Treatment, CAA denotes Consolidated Appropriations Act, and FPL denotes federal poverty level. NSCH does not distinguish Medicaid from separate-CHIP enrollment.

S11. Continuous-Eligibility Policy Documentation and Cross-Checks

The paper’s `continuous_eligibility` indicator is fixed at whether the state’s January 2016 policy provided 12-month Medicaid CE to children ages 6–17 (24 jurisdictions classified as CE from KFF Table 14 and the remaining 27 as historically non-CE). This section documents that baseline classification across the study window, identifies the state-years retained in the strict sample, checks the archived threshold panel against the annual KFF source tables, reports non-KFF documentary context for selected rules, and describes the institutional gap that defines the strict sample.

S11.1. Year-by-year CE verification and strict-sample roster

A natural concern is whether within-period changes in state CE policy or in separate-CHIP program structure could contaminate the paper’s three-regime design. For the threshold panel, I reviewed all nine annual KFF Table 1 reports from January 2016 through May 2024, including January 2021; the 2024 report’s threshold table is dated May 2024. For the CE policy cross-walk, I reviewed the seven reports with comparable children’s Medicaid CE tables (January 2016, 2017, 2018, 2019, 2020, 2022, and 2023). The January 2021 Table 1 extract used for threshold coding does not contain a comparable CE table. The review identifies no state-policy change that requires recoding the school-aged strict-sample observations under the paper’s pre-mandate-baseline CE definition (Kaiser Family Foundation, 2016–2024).

Children’s Medicaid CE state changes, January 2016 through January 2023. I cross-walked children’s Medicaid CE status across KFF’s annual 50-state surveys for January 2016 (Table 14), January 2017 (Table 14), January 2018 (Table 13), January 2019 (Table 13), January 2020 (Table 11), January 2022 (Table 11), and January 2023 (Table 5). The January 2021 survey was fielded, and its Table 1 supplies that year’s threshold values; that Table 1 extract does not include the comparable children’s Medicaid CE table. This is a one-year gap in the CE-table cross-walk, not a gap in the threshold panel. The FFCRA Medicaid enrollment freeze covered every state from March 2020 through March 2023 regardless of state CE policy, limiting the operational consequence of that documentation gap. Across the reviewed CE tables, only two material changes occurred in the children’s Medicaid CE column. First, Pennsylvania adopted limited-group continuous eligibility for children under age 4 by January 2020. Pennsylvania is the only non-CE-as-of-2016 state in the paper’s 27-jurisdiction non-CE universe to adopt any state-level children’s Medicaid CE before the federal mandate. Pennsylvania’s pre-mandate CE applies to ages 0–3 only, outside the paper’s 6–17 MEDB analysis sample, so it does not alter the paper’s school-aged sample classification. Second, New Jersey’s January 2023 row in KFF Table 5 carries fewer Y entries than its January 2022 row, suggesting that New Jersey may have dropped children’s Medicaid CE between the two surveys. New Jersey remains in the historical-CE comparison group by construction; 2023 is excluded from the three-regime analysis, and the federal CE mandate covers 2024. Among the paper’s 27 non-CE-as-of-2016 jurisdictions, the reviewed surveys identify no adoption of full state-level children’s Medicaid CE applicable to school-aged children before January 1, 2024; the paper therefore retains its baseline classification throughout the study window.

Table S25: Children’s Medicaid Continuous-Eligibility State Counts, January 2016 through January 2023 (KFF 50-State Surveys)

KFF survey	Full-Y count	Limited-Group	Net change vs. prior year
January 2016	24	FL (under 5), IN (under 3)	baseline (paper’s coding)
January 2017	24	FL, IN	none
January 2018	24	FL, IN	none
January 2019	24	FL, IN	none
January 2020	23	FL, IN, PA (under 4)	PA adopts limited-group CE; SC reports “Not Reported”
January 2022	24	FL, IN, PA	SC’s data gap filled
January 2023	23	FL, IN, PA	NJ row carries fewer Y entries; historical classification retained

Notes. Each row reports the number of states with full children’s Medicaid continuous eligibility (“Y” in KFF’s annual 50-state survey) and age-restricted limited-group CE. Full-Y remains 23–24. The figures 24 and 32 refer to different same-year definitions in the January 2016 survey: 24 states reported full children’s Medicaid CE, whereas 32 had CE in Medicaid and/or separate CHIP; they do not describe a 24-to-32 increase over time. The January 2021 survey was fielded and supplies the threshold table, but the Table 1 extract used for threshold coding does not include the comparable CE table. FL = Florida; IN = Indiana; PA = Pennsylvania; SC = South Carolina; NJ = New Jersey; KFF = KFF’s annual 50-state Medicaid and CHIP eligibility survey; CE = continuous eligibility; CHIP = Children’s Health Insurance Program. Source: KFF survey Tables 5 and 11–14 (numbering varies by year).

There is therefore no documented 24-to-32 temporal expansion in this cross-walk. The number 32 is the broader January 2016 count of states with CE in Medicaid and/or separate CHIP, while 24 is the same survey’s Medicaid-only count. Because the analysis classification is based on children’s Medicaid CE, a state can remain in the historically non-CE sample even if its separate-CHIP program offers CE. The direction of any resulting bias is not identified: separate-CHIP retention could affect enrollment, treatment continuity, or sample composition in several ways. A joint Medicaid-and-CHIP CE sensitivity would therefore be useful but is not available in the current panel.

Separate-CHIP to CHIP-funded Medicaid-expansion transitions and CHIP eligibility income limits. The 15 always-CHIP-funded Medicaid-expansion jurisdictions across 2016–2024 are Alaska, California, DC, Hawaii, Maryland, Michigan, Minnesota, Nebraska, New Hampshire, New Mexico, Ohio, Oklahoma, Rhode Island, South Carolina, and Vermont. The six transition states are North Dakota in 2020, Wyoming in 2022, Illinois in 2023, and Kentucky, Maine, and North Carolina in 2024. The strict-sample filter drops state-years where `abs(separate_chip - medicaid_elig_6_18) < 50` or `separate_chip` is missing. Each transition state enters the strict sample only when it operates a separate-CHIP program whose upper limit is at least 50 FPL points above its Medicaid threshold. The KFF panel applies that classification separately to every state-year.

The KFF source-table review identified two transcription discrepancies in legacy project files. Kansas’s January 2021 Medicaid-funded limit for children ages 1–5 had been recorded as 149 rather than the Table 1 value of 154 FPL. The school-aged limit used by this analysis remains 138 FPL, so the correction affects neither the sample nor the results. For Kentucky in 2024, Table 1 reports Title XIX Medicaid limits of

142 FPL for ages 1–5 and 109 FPL for ages 6–18, with CHIP-funded Medicaid expansion extending both age groups to 218 FPL. The eligibility panel and analysis-ready inputs used the combined Medicaid/CHIP-funded Medicaid-expansion upper limit of 218, while one legacy pre-setup intermediate retained 164 and has now been synchronized to 218. Kentucky 2024 has no separate-CHIP threshold and is therefore excluded from the strict estimator; the pooled state-program sensitivity retains it at the combined 218-FPL limit. Because every analysis-ready input already used 218, synchronizing the legacy intermediate changes no output reported in the paper or supplement.

Strict-sample state-year roster. Table S26 lists every historically non-CE state-year retained by the strict rule. Fifteen states enter in all seven analysis years. Kentucky enters in 2016–2019 and 2021–2022 but not in 2024, when it no longer operates a separate-CHIP program. The resulting sample contains 64 state-years in 2016–2019, 32 in 2021–2022, and 15 in 2024, for 111 state-years in total. This roster documents the state-year application of the strict rule referenced in the main paper.

Table S26: Historically Non-CE State-Years Retained by the Strict Institutional Rule

State	2016–2019	2021–2022	2024
Arizona	Yes	Yes	Yes
Arkansas	Yes	Yes	Yes
Connecticut	Yes	Yes	Yes
Delaware	Yes	Yes	Yes
Florida	Yes	Yes	Yes
Georgia	Yes	Yes	Yes
Indiana	Yes	Yes	Yes
Kentucky	Yes	Yes	No
Massachusetts	Yes	Yes	Yes
Missouri	Yes	Yes	Yes
Nevada	Yes	Yes	Yes
Pennsylvania	Yes	Yes	Yes
Texas	Yes	Yes	Yes
Utah	Yes	Yes	Yes
Virginia	Yes	Yes	Yes
Wisconsin	Yes	Yes	Yes
Retained states	16	16	15
Retained state-years	64	32	15

Notes. A “Yes” means that the state satisfies the strict institutional rule in every year covered by that column. The rule retains historically non-CE state-years that operate a separate-CHIP program with an upper limit at least 50 FPL points above the Medicaid income threshold for ages 6–18. The grouped columns contain four, two, and one analysis years, respectively. The total is 111 retained state-years. Kentucky is excluded in 2024 because it transitioned from separate CHIP to CHIP-funded Medicaid expansion. CE = continuous eligibility; CHIP = Children’s Health Insurance Program; FPL = federal poverty level.

S11.2. Non-KFF documentary context for selected threshold values

A separate human-coded ledger pairs selected KFF rows with non-KFF documentary materials. It is supplied for source transparency and program-structure context, not as an independent year-specific validation audit. The annual KFF reports remain the contemporaneous state-year source for the threshold panel used in the analysis.

Scope. The ledger contains 31 rows, not a full 51-jurisdiction sweep: five focal strict-sample separate-CHIP non-CE states (Florida, Texas, Indiana, Pennsylvania, and Georgia) at five KFF anchor years spanning the three policy regimes (2016, 2019, 2021, 2022, and 2024), plus six state-years in which the KFF panel records the end of separate CHIP (North Dakota 2020, Wyoming 2022, Illinois 2023, Kentucky 2024, Maine 2024, and North Carolina 2024). The year identifies the paired KFF row, not necessarily the vintage of the non-KFF document.

Ledger fields. For each row the ledger preserves source metadata, working threshold transcriptions, the KFF values, mechanical differences between the two sets of values, working program-structure classifications, and a flag indicating that a non-KFF link was located. Some working transcriptions incorporate a MAGI-disregard conversion. Neither a zero difference nor the presence of a link establishes that a document is contemporaneous with the paired KFF year.

Document timing and interpretation. The linked materials vary in timing and authority. For example, the document paired with Florida 2021 is a 2024 eligibility-extension request; the Indiana 2021 and 2022 rows use a document dated January 2023 and hosted by a primary-care association; the Pennsylvania 2022 and 2024 rows link charts effective in 2025 and 2026, respectively; the current North Dakota page describes a limit effective in 2026 rather than the 2020 row; and the Wyoming link is a Cornell Legal Information Institute reproduction rather than a state-hosted source. Current handbook or agency pages are also paired with several earlier Texas, Pennsylvania, and Georgia rows. These documents may illuminate program structure or later continuity, but they cannot uniformly validate the corresponding historical numeric values.

Mechanically, the ledger’s working numeric fields differ from KFF by no more than 2 FPL points in 29 rows and by 5 points in the two Florida rows. That count is descriptive, not a validation rate, because document vintages differ and some working transcriptions include conversions. No analytic threshold was replaced from this exercise: the annual KFF tables remain the inputs to the state-year panel. A compact context table appears below; the full 31-row ledger, including document descriptions and limitations, is shipped as `data/kff/state_manual_validation.csv`.

Table S27: Non-KFF Documentary Context Paired with Selected KFF Threshold Rows

State	KFF year	Non-KFF document	Section/Page	KFF Medicaid (% FPL)	KFF separate CHIP (% FPL)	Link located
<i>Panel A. Five focal states × five anchor years</i>						
Florida	2016	DCF appendix (date not established)	Appx. A-7	138	215	✓
Florida	2019	AHCA Overview	Slide 6	138	215	✓
Florida	2021	AHCA 2024 eligibility request	Section II	138	215	✓
Florida	2022	AHCA current KidCare page	Eligibility	138	215	✓
Florida	2024	FL HB 121 (2023)	pp. 4–5	138	215	✓
Texas	2016	TX Works Hbk. (current page)	C-130	138	206	✓
Texas	2019	TX Works Hbk. (current page)	C-130	138	206	✓
Texas	2021	TX Works Hbk. (current page)	C-130	138	206	✓
Texas	2022	TX Works Hbk. (current page)	C-130	138	206	✓
Texas	2024	SPA TX-22-0001	§4.1	138	206	✓
Indiana	2016	CMS factsheet (undated file)	Eligibility	165	262	✓
Indiana	2019	CMS factsheet (undated file)	Eligibility	165	262	✓
Indiana	2021	Indiana Medicaid hierarchy (2023)	MA categories	163	255	✓
Indiana	2022	Indiana Medicaid hierarchy (2023)	MA categories	163	255	✓
Indiana	2024	FSSA OMPP aid category (2025)	MA categories	163	255	✓
Pennsylvania	2016	PA OIM Ch. 312 Appx. A	MAGI Income	138	319	✓
Pennsylvania	2019	PA OIM Ch. 312 Appx. A	MAGI Income	138	319	✓
Pennsylvania	2021	PA DHS current CHIP page	Elig. chart	138	319	✓
Pennsylvania	2022	PA DHS FPIG chart (2025)	FPIG table	138	319	✓
Pennsylvania	2024	PA DHS FPIG chart (2026)	FPIG table	138	319	✓
Georgia	2016	GA DFCS PAMMS	§2194	138	252	✓
Georgia	2019	GA DFCS PAMMS	§2194	138	252	✓
Georgia	2021	GA DCH current PeachCare page	Elig. table	138	252	✓
Georgia	2022	GA DCH current PeachCare page	Elig. table	138	252	✓
Georgia	2024	MACPAC Exhibit 35	Exhibit 35	138	252	✓
<i>Panel B. Program-structure transition cells</i>						
North Dakota	2020	ND HHS current page (2026 rule)	Prog. Overview	175	—	✓
Wyoming	2022	Cornell LII code reproduction	Kid Care CHIP	205	—	✓
Illinois	2023	IDHS PM 15-06-01-d	All Kids Assist	318	—	✓
Kentucky	2024	KCHIP State Plan	Eligibility	218	—	✓
Maine	2024	MaineCare Rule 303A	Part 4 §4(B/D/E)	305	—	✓
North Carolina	2024	NC Medicaid Transition	Policy Memos	216	—	✓

Notes. Panel A pairs five focal strict separate-CHIP non-CE states with five KFF anchor years (2016, 2019, 2021, 2022, 2024). The year identifies the paired KFF row, not necessarily the vintage of the non-KFF document. A checkmark means only that the listed link was located; it does not establish contemporaneous authority or numeric validation. The displayed thresholds are the KFF inputs used in the analysis. The ledger preserves separate working transcriptions and mechanical differences: 29 rows differ from KFF by no more than 2 FPL points and the two Florida rows differ by 5 points, but this is not a validation rate. Panel B lists KFF transition rows and paired program-structure context. KFF = KFF’s Medicaid and CHIP eligibility survey; FPL = federal poverty level; MAGI = modified adjusted gross income; DCF = Department of Children and Families; AHCA = Agency for Health Care Administration; SPA = State Plan Amendment; CMS = Centers for Medicare & Medicaid Services; FSSA = Indiana Family and Social Services Administration; OMPP = Office of Medicaid Policy and Planning; DHS = Department of Human Services; FPIG = Federal Poverty Income Guidelines; DFCS = Division of Family and Children Services; PAMMS = Policy and Manual Management System; DCH = Georgia Department of Community Health; MACPAC = Medicaid and CHIP Payment and Access Commission; HHS = Health and Human Services; IDHS = Illinois Department of Human Services; KCHIP = Kentucky Children’s Health Insurance Program; non-CE = no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17.

S11.3. Institutional Medicaid-to-separate-CHIP gap distribution

The primary specification fixes both the effect and bias-correction bandwidths at $h = b = 50$ FPL on institutional grounds. The strict-sample filter excludes any state-year where the gap between separate CHIP and Medicaid is below 50 FPL points, so neither primary window crosses the next institutional cutoff. The full gap distribution from the KFF panel also explains why the former $h = 50, b = 70$ specification is retained only as a sensitivity.

Table S28: Distribution of the Institutional Medicaid-to-Separate-CHIP Gap (FPL Points)

Statistic	Value	Description
Nonmissing separate-CHIP state-years	311	2016–2024; full policy-panel universe
Mean gap among nonmissing	105.34	FPL points
State-years with gap ≥ 50	280	Strict-eligible universe
Mean gap among strict-eligible	113.29	FPL points
Min gap among strict-eligible	51	FPL points
Max gap among strict-eligible	251	FPL points
Share of strict-eligible with gap $\in [50, 70)$	0.300	84 of 280 state-years
Analysis-year strict-eligible state-years	218	2016–2019, 2021–2022, and 2024
Share of analysis-year strict-eligible with gap $\in [50, 70)$	0.298	65 of 218 state-years
Historically non-CE analysis-year strict-eligible state-years	111	Headline state-year universe
Share of historically non-CE analysis-year cells with gap $\in [50, 70)$	0.423	47 of 111 state-years

Notes. The gap is the difference between the separate-CHIP income threshold and the Medicaid eligibility threshold for ages 6–18, in FPL percentage points, from the KFF panel. The strict-sample filter requires a gap of at least 50 FPL. The primary $h = b = 50$ windows do not cross the next institutional cutoff in any strict-eligible state-year. Among analysis years (2016–2019, 2021–2022, and 2024), the former $b = 70$ sensitivity window crosses it in 29.8% of all strict-eligible state-years and 42.3% of the historically non-CE state-years used for the headline estimates, so it can affect the bias correction. The first four rows summarize the full 2016–2024 policy panel; the final four restrict to analysis years. CHIP = Children’s Health Insurance Program; CE = continuous eligibility; FPL = federal poverty level; KFF = KFF’s Medicaid and CHIP eligibility survey.

Among analysis-year state-years, 29.8 percent of all strict-eligible cells and 42.3 percent of the historically non-CE headline cells sit in the 50–69 FPL range. A bias-correction bandwidth of $b = 70$ therefore crosses the next cutoff for some retained state-years and can influence the robust bias-corrected estimate. For that reason, $h = b = 50$ is primary and $h = 50, b = 70$ is a sensitivity rather than an institutional safeguard.

S11.4. Sensitivity of the strict estimand to the 50-FPL gap filter

The strict sample requires the separate-CHIP upper limit to lie at least 50 FPL points above the Medicaid threshold, matching the primary $h = b = 50$ windows. Table S29 reruns that primary specification after lowering the minimum gap to 40 FPL points or raising it to 60. The 40- and 50-FPL samples are identical because no historically non-CE analysis-year cell has a gap from 40 through 49 FPL points. Raising the minimum to 60 drops two states in the pre and COVID periods and one state in 2024.

Among all children, increasing the minimum gap from 50 to 60 FPL points changes the estimate from 0.097 to 0.106 before the pandemic, from 0.169 to 0.173 during the COVID period, and from 0.075 to 0.090 in 2024. The pre and COVID estimates remain positive and statistically significant at the 5 percent level, while the 2024 estimate remains positive but imprecise. The main all-child conclusions therefore do not depend on including state-years with gaps from 50 through 59 FPL points.

Table S29: Psychiatric-Medication RD Estimates under Alternative Institutional-Gap Filters

	Gap ≥ 40	Gap ≥ 50 (primary)	Gap ≥ 60
<i>Panel A. All children</i>			
Pre (2016–19)	0.097*** (0.037)	0.097*** (0.037)	0.106*** (0.041)
N_h	3,482	3,482	3,050
COVID (2021–22)	0.169** (0.069)	0.169** (0.069)	0.173** (0.077)
N_h	2,022	2,022	1,790
2024	0.075 (0.063)	0.075 (0.063)	0.090 (0.062)
N_h	1,026	1,026	995
<i>Panel B. Children with a current MEDB diagnosis</i>			
Pre (2016–19)	0.328*** (0.113)	0.328*** (0.113)	0.366*** (0.130)
N_h	998	998	882
COVID (2021–22)	0.438*** (0.138)	0.438*** (0.138)	0.460*** (0.152)
N_h	653	653	568
2024	0.126 (0.090)	0.126 (0.090)	0.154* (0.092)
N_h	354	354	340

Notes. Entries are robust bias-corrected local-linear RD estimates, with standard errors in parentheses. All models use $h = b = 50$ FPL points, a triangular kernel, survey weights, the demographic adjustment set, additive state and within-period year indicators, and standard errors clustered on running-variable support points. The sample includes children ages 6–17 in historically non-CE state-years that operate separate CHIP and satisfy the indicated minimum Medicaid-to-separate-CHIP gap. Panel A includes all children and is the primary population analysis. Panel B restricts to children with a current mental, emotional, developmental, or behavioral diagnosis and is a secondary heterogeneity analysis. For the 40- and 50-FPL filters, the pre, COVID, and 2024 models use 16, 16, and 15 states (64, 32, and 15 state-years). The 60-FPL filter uses 14 states (56, 28, and 14 state-years). N_h is the unweighted number of observations with complete outcome and covariate data inside the effect bandwidth. RD = regression discontinuity; non-CE = no January 2016 12-month Medicaid continuous-eligibility policy covering ages 6–17; MEDB = mental, emotional, developmental, or behavioral; CHIP = Children’s Health Insurance Program; FPL = federal poverty level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The current-MEDB estimates are larger than the all-child estimates under every filter. They remain positive and statistically significant at the 1 percent level in the pre and COVID periods. In 2024, the estimate is not statistically significant at the 5 percent level under any filter. These results show that the discontinuity is larger among children with a current MEDB diagnosis, although this subgroup comparison does not determine how much of the all-child change it accounts for. Together, the two panels show that tightening the institutional-gap rule leaves both the primary population result and the secondary subgroup pattern substantively unchanged.

S12. Diagnosis-Composition and Survey-Weight Sensitivity

S12.1. All-Child Diagnosis-Entry and Composition Checks

The all-child medication estimate does not condition inclusion on diagnosis. The secondary current-MEDB estimate does, so the children observed in that subgroup can differ across the threshold if the probability or mix of recorded diagnoses changes. Table S30 reports the primary all-child medication estimate alongside the all-child discontinuity in any current MEDB diagnosis and each diagnosis component. These are diagnostic checks, not a way to remove diagnosis selection: diagnoses are parent-reported, the component indicators overlap, and the component tests are exploratory and unadjusted for multiplicity.

Table S30: All-Child Psychiatric-Medication and Diagnosis-Entry RD Estimates

Outcome	Pre (2016–2019)	COVID (2021–2022)	Post (2024)
<i>Primary all-child medication outcome</i>			
Psychiatric medication use	0.097*** (0.037)	0.169** (0.069)	0.075 (0.063)
N_h	3,482	2,022	1,026
<i>Diagnosis-entry and composition checks</i>			
Any current MEDB diagnosis	0.059 (0.054)	0.140 (0.088)	−0.143 (0.191)
ADHD	0.075** (0.031)	0.171** (0.084)	−0.011 (0.090)
Depression	0.002 (0.015)	0.055 (0.052)	−0.271** (0.116)
Anxiety	−0.038 (0.052)	0.081 (0.051)	−0.292** (0.137)
Behavior or conduct problems	0.025 (0.041)	0.094 (0.066)	0.079 (0.065)
Autism spectrum disorder	0.024 (0.024)	0.016 (0.037)	0.138 (0.153)
Developmental delay	0.041 (0.032)	0.048 (0.052)	−0.232 (0.153)
Speech disorder	0.023 (0.020)	−0.003 (0.044)	−0.045 (0.044)
Tourette syndrome	−0.005 (0.003)	−0.009 (0.008)	0.022 (0.017)
N_h	3,525	2,061	1,040

Notes. Each cell is a robust bias-corrected local-linear RD estimate with its robust standard error in parentheses; N_h is the unweighted complete-case count inside the $h = 50$ -FPL effect bandwidth. The sample includes all children ages 6–17 in strict separate-CHIP state-years among historically non-CE states and is not conditioned on MEDB status. It does not exclude children based on SSI receipt or chronic functional limitation. Specifications use a triangular kernel, covariates and state and period-year indicators, survey weights, $h = b = 50$ FPL, and standard errors clustered on running-variable support points. The medication outcome has slightly fewer complete observations than the diagnosis outcomes. MEDB denotes mental, emotional, developmental, or behavioral, ADHD denotes attention-deficit/hyperactivity disorder, RD denotes regression discontinuity, non-CE denotes no January 2016 12-month Medicaid continuous-eligibility policy, CHIP denotes Children’s Health Insurance Program, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level. Diagnosis-component tests are exploratory and not adjusted for multiple testing. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The all-child medication estimate is positive and statistically significant at the 5 percent level before the pandemic and during COVID, but not in 2024. The probability of any current MEDB diagnosis does not change detectably at the 5 percent level in any period. The overall diagnosis indicator therefore provides no

evidence of a clear threshold discontinuity in entry into the current-MEDB group.

The diagnosis components tell a more qualified story. ADHD prevalence is 7.5 percentage points higher above the threshold before the pandemic and 17.1 points higher during COVID; both estimates are statistically significant at the 5 percent level. In 2024, depression and anxiety prevalence are 27.1 and 29.2 points lower above the threshold, and both estimates are statistically significant at the 5 percent level. Because children can have more than one diagnosis, increases in some components can be offset by decreases in others when the broad any-MEDB indicator is formed. The component estimates are also exploratory and unadjusted for multiple testing. They show that the diagnostic composition of the current-MEDB subgroup may differ across the threshold even when the aggregate diagnosis indicator does not. The all-child estimates remain primary because they do not condition on that potentially changing subgroup.

S12.2. Survey-Weight Leverage Sensitivity

The NSCH survey weights define the target-population estimand and are used throughout the paper. A small number of large weights can nevertheless exert substantial influence in a local sample. Table S31 therefore holds the RD estimator, bandwidths, kernel, covariates, and clustering rule fixed while limiting the largest survey weights. For each population and period, one sensitivity winsorizes weights at the 1st and 99th percentiles, and another removes records above the 99th percentile. These are leverage checks, not alternative preferred estimands.

Table S31: Psychiatric-Medication RD Sensitivity to Large Survey Weights

Weight rule	Pre (2016–2019)	COVID (2021–2022)	2024
<i>Panel A. All children</i>			
Survey weights	0.097*** (0.037)	0.169** (0.069)	0.075 (0.063)
Winsorized at p1/p99	0.093** (0.036)	0.169** (0.068)	0.073 (0.064)
Drop above p99	0.078** (0.037)	0.183*** (0.066)	0.034 (0.067)
<i>Panel B. Children with a current MEDB diagnosis</i>			
Survey weights	0.328*** (0.113)	0.438*** (0.138)	0.126 (0.090)
Winsorized at p1/p99	0.321*** (0.110)	0.443*** (0.139)	0.126 (0.091)
Drop above p99	0.294** (0.115)	0.443*** (0.141)	0.066 (0.120)

Notes. Each entry is a robust bias-corrected local-linear RD estimate with its robust standard error below it. All specifications use $h = b = 50$ FPL points, a triangular kernel, the demographic adjustment set, additive state and within-period year indicators, and standard errors clustered on running-variable support points. The sample includes children ages 6–17 in strict separate-CHIP state-years among historically non-CE states and does not exclude children based on SSI receipt or chronic functional limitation. Panel A includes all children and is the primary population analysis. Panel B restricts to children with a current mental, emotional, developmental, or behavioral diagnosis and is a secondary heterogeneity analysis. Percentile cutoffs are calculated separately by population and period from the original NSCH survey weight among complete cases inside the $h = 50$ effect bandwidth. The winsorized model replaces weights below the 1st percentile or above the 99th percentile with the corresponding cutoff. The trimmed model removes records above the 99th percentile. RD = regression discontinuity; NSCH = National Survey of Children’s Health; CE = continuous eligibility; CHIP = Children’s Health Insurance Program; SSI = Supplemental Security Income; FPL = federal poverty level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table S32: Survey-Weight Leverage in the Local Psychiatric-Medication Samples

Population	Period	N_h	Kish N_{eff}	Design effect	Top-1% share	Largest-case share
All children	Pre	3,482	718.0	4.85	11.7%	1.1%
	COVID	2,022	465.8	4.34	12.0%	1.0%
	2024	1,026	191.6	5.35	11.3%	2.6%
Current MEDB	Pre	998	246.1	4.05	8.7%	1.9%
	COVID	653	173.3	3.77	8.0%	2.3%
	2024	354	64.7	5.47	11.5%	6.4%

Notes. Diagnostics use the same complete cases as Table S31 inside the $h = 50$ -FPL effect bandwidth. N_h is the unweighted number of complete observations. Kish $N_{\text{eff}} = (\sum_i w_i)^2 / \sum_i w_i^2$ uses the product of the survey and triangular-kernel weights. The design effect is N_h / N_{eff} . “Top-1% share” is the combined-weight share of records whose original survey weight exceeds the population- and period-specific 99th percentile. “Largest-case share” is the largest record’s share of the combined local weight. MEDB = mental, emotional, developmental, or behavioral; FPL = federal poverty level.

Removing records above the 99th percentile of survey weight changes the all-child estimate from 9.7 to 7.8 percentage points before the pandemic and from 16.9 to 18.3 points during COVID. Both estimates remain positive and statistically significant at the 5 percent level. The corresponding current-MEDB estimates change from 32.8 to 29.4 points and from 43.8 to 44.3 points, and both also remain statistically significant at the 5 percent level. The positive estimates in these two periods are therefore not produced by the records with the largest 1 percent of survey weights.

In 2024, removing those records changes the all-child estimate from 7.5 to 3.4 percentage points and the current-MEDB estimate from 12.6 to 6.6 points; neither estimate is statistically significant at the 5 percent level under any reported weight rule. The leverage diagnostics also show that the current-MEDB local sample has an effective size of about 65, and its largest record carries 6.4 percent of the combined local weight. These results reinforce the paper’s treatment of the 2024 estimates as imprecise and particularly sensitive to a small local diagnosed sample.

S13. Discreteness-Robust Inference and Running-Variable Diagnostics

Purpose. The running variable is recorded in integer FPL percentage points and has mass points and limited support, while each period contains only 15 or 16 states (Kolesár and Rothe, 2018). Section S13.1 reports smoothness-bounded honest confidence intervals for the all-child and current-MEDB medication estimates. Section S13.2 points to the consolidated all-child shifted-cutoff diagnostic in Section S7.2. Section S13.3 describes the observed running-variable distribution across periods, and Section S13.4 reports state-clustered and wild-cluster-bootstrap inference across the medication population hierarchy. The estimation checks use survey weights; the distribution and density diagnostics are descriptive and unweighted.

S13.1. Honest confidence intervals

Table S33 compares the primary robust bias-corrected estimates with **rdhonest** estimates and smoothness-bounded intervals using the Armstrong–Kolesár rule-of-thumb second-derivative bound (Armstrong and Kolesár, 2020). The honest specifications use survey weights, demographic covariates, and within-period year indicators but omit the state indicators in the primary specification. They are therefore separate estimators, not alternative intervals around the primary point estimates. I report both the **rdhonest** MSE-optimal bandwidth and a fixed 50-FPL bandwidth matching the primary effect window.

Table S33: Smoothness-Bounded Honest Confidence Intervals by Medication Population

Period	Primary $h = b = 50$ $\hat{\tau}$ [robust 95% CI]	Honest, MSE-optimal $\hat{\tau}$ [honest 95% CI]	MSE-optimal h (FPL points)	Honest, $h = 50$ $\hat{\tau}$ [honest 95% CI]
<i>Panel A. All children</i>				
Pre	0.097 [0.024, 0.170]	0.112 [0.005, 0.220]	29.5	0.069 [−0.072, 0.210]
COVID	0.169 [0.034, 0.304]	0.156 [0.006, 0.307]	29.8	0.109 [−0.079, 0.298]
2024	0.075 [−0.049, 0.199]	0.050 [−0.127, 0.228]	41.1	0.058 [−0.131, 0.246]
<i>Panel B. Children with a current MEDB diagnosis</i>				
Pre	0.328 [0.106, 0.550]	0.324 [0.056, 0.591]	31.7	0.193 [−0.133, 0.519]
COVID	0.438 [0.168, 0.708]	0.391 [0.041, 0.741]	30.3	0.305 [−0.134, 0.745]
2024	0.126 [−0.050, 0.303]	0.108 [−0.243, 0.460]	62.0	0.097 [−0.251, 0.444]

Notes. The primary column reports robust bias-corrected **rdrobust** estimates and running-variable-clustered intervals using $h = b = 50$ FPL points, a triangular kernel, survey weights, the demographic adjustment set, and additive state and within-period year indicators. Honest columns use **rdhonest**, survey weights, demographic covariates, within-period year indicators, and the Armstrong–Kolesár rule-of-thumb second-derivative bound (Armstrong and Kolesár, 2020); they omit state indicators. “MSE-optimal” denotes the **rdhonest** mean-squared-error-optimal bandwidth. The package issued high-leverage warnings in these local samples, so the honest results also require caution. Both panels use children ages 6–17 in strict separate-CHIP state-years among historically non-CE states and do not exclude children based on SSI receipt or chronic functional limitation. Panel A is the primary population analysis; Panel B is a secondary heterogeneity analysis. CI = confidence interval; MEDB = mental, emotional, developmental, or behavioral; FPL = federal poverty level; CHIP = Children’s Health Insurance Program; CE = continuous eligibility; SSI = Supplemental Security Income.

At the MSE-optimal bandwidth, the honest intervals exclude zero before the pandemic and during COVID in both populations, although the all-child lower bounds are only 0.005 and 0.006. At the fixed 50-FPL bandwidth, every honest interval includes zero. The point estimates remain positive, but the statistical conclusions depend on the bandwidth, smoothness restriction, covariate specification, and estimator. The

fixed-bandwidth honest results therefore do not confirm the primary conventional inference, and the package’s high-leverage warnings add a further reason for caution.

S13.2. Shifted-cutoff diagnostics

Section S7.2 reports the shifted-cutoff analysis once, using the primary all-child population, $h = b = 50$ FPL points, and samples restricted so that nonzero shifted windows do not include the study threshold. I do not repeat the Base supplement’s second rank table, which used the secondary current-MEDB population and the former $h = 50, b = 70$ sensitivity specification. Its bias-correction windows overlapped the study threshold for some nominally clean shifts. Consolidating the evidence in S7.2 makes the less favorable result clear: four of 36 nonzero shifted-cutoff estimates reject zero at the unadjusted 5 percent level, so the grid is not a clean falsification pass.

S13.3. Cross-period distribution checks for the observed running variable

Table S34 describes the all-child running variable inside the primary 50-FPL window. Its standard deviation is between 28.7 and 29.1 FPL points in all three periods. Section S7.1 separately reports the full-support density tests, none of which rejects continuity at the 5 percent level. These are descriptive distribution checks only. They do not measure error in the NSCH income proxy, establish classical measurement error, or determine the magnitude or direction of resulting RD bias.

Table S34: Cross-Period Running-Variable Distribution in the Primary All-Child Population

Period	N within ± 50 FPL	Mean	Standard deviation	Variance
Pre (2016–2019)	3,673	2.31	28.85	832.12
COVID (2021–2022)	2,122	2.17	29.05	844.12
2024	1,056	2.63	28.67	821.72

Notes. The running variable is reported family income in FPL percentage points minus the state-year Medicaid threshold. N is the unweighted number of children with a nonmissing running variable inside the inclusive window $|\text{distance_M}| \leq 50$; unlike an outcome-specific N_h , it does not require the medication outcome or model covariates to be complete. The sample includes all children ages 6–17 in strict separate-CHIP state-years among historically non-CE states and does not exclude children based on SSI receipt or chronic functional limitation. The unweighted full-support density tests appear in Section S7.1. FPL = federal poverty level; CHIP = Children’s Health Insurance Program; CE = continuous eligibility; SSI = Supplemental Security Income.

S13.4. Alternative inference across the medication population hierarchy

Table S35 compares inference for psychiatric medication among all children, children with a current MEDB diagnosis, and children without a current MEDB diagnosis under the primary $h = b = 50$ specification. Support-point clustering addresses dependence created by the integer-valued running variable. State clustering addresses within-state dependence. The state wild bootstrap provides an additional small-cluster check for the distinct conventional local-linear estimate.

For all children, the support-point and state-clustered tests are statistically significant at the 5 percent level before the pandemic. During COVID, the support-point test is significant at the 5 percent level,

whereas the state-clustered test is significant at the 10 percent level but not at the 5 percent level. The wild-bootstrap tests for the conventional all-child estimates are not statistically significant at the 5 percent level in either period. For children with a current MEDB diagnosis, the support-point and state-clustered tests are significant at the 5 percent level before the pandemic and during COVID, while the wild bootstrap is significant at that level only during COVID. For children without a current MEDB diagnosis, none of the three inference approaches finds a statistically significant medication discontinuity at the 5 percent level in any period. In 2024, no population has a medication estimate that is statistically significant at the 5 percent level under any of the three approaches.

Only the current-MEDB COVID estimate is statistically significant at the 5 percent level under support-point clustering, state clustering, and the wild bootstrap. The disagreements elsewhere show that the strength of the evidence depends on how uncertainty is measured and, for the wild bootstrap, on using the conventional rather than the bias-corrected estimator. The point estimates remain informative about direction and magnitude, but the alternative procedures limit how confidently the pre-period and all-child COVID estimates can be distinguished from zero.

Table S35: Inference under Alternative Clustering for Psychiatric Medication by Population, Primary $h = b = 50$ Specification

Population	Period	N_h	States	$\hat{\tau}^{bc}$	p (RV)	p (state)	$\hat{\tau}^{conv}$	Wild p
All children	Pre	3,482	16	0.097	0.009	0.032	0.056	0.174
	COVID	2,022	16	0.169	0.014	0.063	0.101	0.154
	2024	1,026	15	0.075	0.236	0.388	0.066	0.408
Current MEDB	Pre	998	16	0.328	0.004	0.010	0.198	0.172
	COVID	653	16	0.438	0.001	0.001	0.249	0.021
	2024	354	15	0.126	0.160	0.442	0.133	0.732
No current MEDB	Pre	2,484	16	-0.008	0.544	0.620	-0.011	0.255
	COVID	1,369	16	-0.002	0.781	0.791	-0.001	0.822
	2024	672	15	0.069	0.249	0.133	0.049	0.792

Notes. The primary specification sets the effect and bias-correction bandwidths to $h = b = 50$ FPL points and uses a triangular kernel, covariates and state and period-year indicators, and survey weights. The strict separate-CHIP sample includes children ages 6–17 in historically non-CE states and does not exclude children based on SSI receipt or chronic functional limitation. N_h is the unweighted complete-case count inside the effect bandwidth. $\hat{\tau}^{bc}$ is the robust bias-corrected RD estimate; p (RV) clusters on running-variable support points and p (state) on states. Wild p is a state-level wild-cluster-bootstrap p -value for the conventional local-linear estimate $\hat{\tau}^{conv}$ because `rdrobust` does not implement the bootstrap for the bias-corrected estimator. Wild-bootstrap 95 percent intervals for pre, COVID, and 2024 are $[-0.038, 0.132]$, $[-0.024, 0.219]$, and $[-0.174, 0.210]$ for all children; $[-0.125, 0.449]$, $[0.075, 0.391]$, and $[-0.362, 0.771]$ for children with a current MEDB diagnosis; and $[-0.032, 0.008]$, $[-0.012, 0.008]$, and $[-0.049, 0.155]$ for children without a current MEDB diagnosis. Pre, COVID, and 2024 denote 2016–2019, 2021–2022, and 2024. RV denotes running variable, RD denotes regression discontinuity, MEDB denotes mental, emotional, developmental, or behavioral, SSI denotes Supplemental Security Income, and FPL denotes federal poverty level.

S14. Exploratory Coverage Check in the American Community Survey

The NSCH cannot show whether children move from Medicaid to separate CHIP at the statutory threshold because its Medicaid/CHIP measure pools the two programs. This section uses the American Community Survey (ACS) to ask a narrower question: whether public coverage, any private coverage, or uninsurance changes detectably at the same constructed Medicaid threshold (Ruggles et al., 2025). The exercise is not an external validation of the psychiatric-medication result or a first-stage test of Medicaid-to-CHIP switching. The ACS has no treatment measures and also does not distinguish Medicaid from separate CHIP.

I use ACS person records for children ages 6–17 in the same survey years as the NSCH analysis: 2016–2019, 2021–2022, and 2024. I merge the same KFF state-year Medicaid eligibility limits used in the main analysis and define the ACS running variable as the child’s family income-to-poverty ratio minus the applicable Medicaid limit. The sample applies the same strict separate-CHIP, historically non-CE state-year filter. Within the inclusive descriptive window $|distance| \leq 50$, the sample contains 16 states and 246,495 children: 148,183 before the pandemic, 67,814 during COVID, and 30,498 in 2024. Because the triangular kernel assigns zero weight to observations exactly 50 FPL points from the threshold, the corresponding estimation counts are 145,433, 66,541, and 29,910.

The ACS running variable is only an approximation to the income measure used for program eligibility. It uses the Census family income-to-poverty ratio rather than Medicaid modified adjusted gross income. The resulting measurement difference can blur or otherwise distort a discontinuity. The ACS coverage measures also differ from the NSCH measures: ACS any private coverage can include children who also report public coverage, whereas the NSCH private-only outcome excludes dual coverage. For these reasons, the ACS levels should not be compared directly with the NSCH levels.

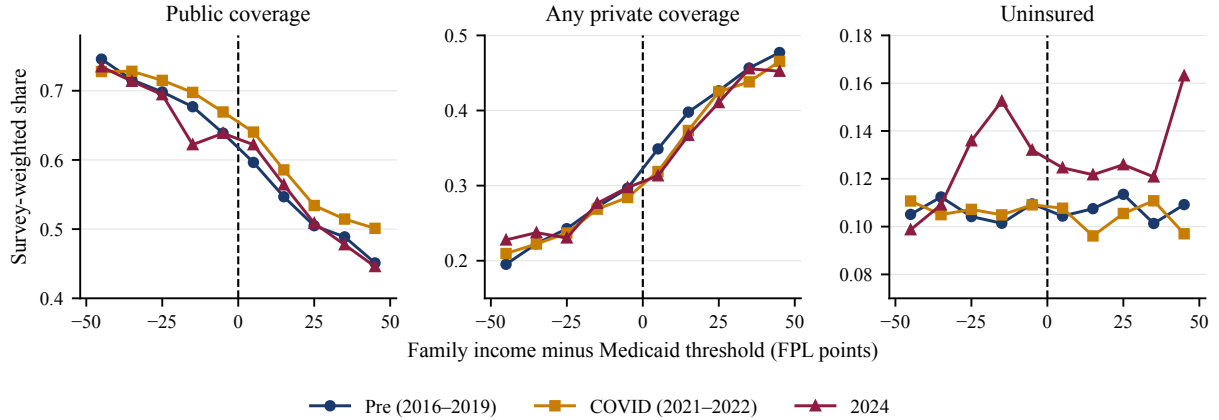


Figure S5: Coverage Profiles around the Constructed Medicaid Threshold in the American Community Survey

Notes. Points are survey-weighted coverage shares in 10-FPL bins; lines connect the bin means within each period. The dashed vertical line marks the constructed Medicaid threshold. The panels use different vertical-axis ranges. Public coverage is the IPUMS `HINSCAID` measure, which includes Medicaid and CHIP. Any private coverage is IPUMS `HCOVPRIV` and can include dual public–private coverage. Uninsured status is derived from IPUMS `HCOVANY`. The sample includes all ACS children ages 6–17 in strict separate-CHIP state-years among historically non-CE states. The profiles are descriptive and are not local RD estimates. Values reproduce `output/acs/acs_profile.csv`. ACS = American Community Survey; IPUMS = Integrated Public Use Microdata Series; RD = regression discontinuity; CHIP = Children’s Health Insurance Program; CE = continuous eligibility; FPL = federal poverty level.

Across the full 100-FPL window, public coverage declines as income rises, private coverage increases, and uninsurance is comparatively flat (Figure S5). Those broad slopes are expected to capture the income gradient in coverage and should not be read as discontinuities. Near zero, the binned profiles do not show a common abrupt change across the three coverage measures or periods. The local estimates provide the formal threshold comparisons.

Table S36: Coverage Discontinuities at the Constructed Medicaid Threshold in the American Community Survey

Outcome	Pre (2016–2019)	COVID (2021–2022)	2024
Public coverage	–0.0012 [–0.0283, 0.0260]	0.0269 [–0.0151, 0.0689]	0.0342 [–0.0526, 0.1209]
Any private coverage	0.0122 [–0.0160, 0.0404]	–0.0089 [–0.0423, 0.0244]	–0.0639 [–0.1242, –0.0036]
Uninsured	–0.0058 [–0.0277, 0.0161]	–0.0072 [–0.0409, 0.0264]	0.0057 [–0.0511, 0.0625]
N_h	145,433	66,541	29,910

Notes. Entries are robust bias-corrected local-linear RD estimates, with 95 percent confidence intervals in brackets. Estimates use ACS person weights, a triangular kernel, equal effect and bias-correction bandwidths of $h = b = 50$ FPL points, and standard errors clustered on running-variable support points. N_h is the unweighted positive-kernel-weight count inside the effect bandwidth and is common across the three outcomes. The strict separate-CHIP, historically non-CE sample includes all ACS children ages 6–17. Public coverage is `HINSCAID`; any private coverage is `HCOVPRIV`, which can include dual public–private coverage; and uninsured status is derived from `HCOVANY`. The 2024 any-private estimate is the only one of the nine exploratory estimates that rejects zero at the unadjusted 5 percent level. None rejects at the 5 percent level after Benjamini–Hochberg adjustment across the nine estimates. Values reproduce `output/acs/acs_rd.csv`. ACS = American Community Survey; RD = regression discontinuity; CHIP = Children’s Health Insurance Program; CE = continuous eligibility; FPL = federal poverty level.

Eight of the nine ACS estimates do not reject zero at the unadjusted 5 percent level (Table S36). The exception is a 6.4-percentage-point decrease in any private coverage in 2024, with a 95 percent confidence interval from –12.4 to –0.4 points. That isolated result does not reject after adjustment for the nine exploratory tests. Public coverage changes by approximately zero before the pandemic, 2.7 percentage points during COVID, and 3.4 points in 2024; none of these estimates rejects zero at the 5 percent level. The uninsurance estimates are also small and do not reject zero in any period.

The ACS results narrow one alternative explanation without identifying the paper’s proposed mechanism. They provide no evidence of an abrupt loss of total public coverage or rise in uninsurance during the pre-pandemic and COVID periods, when the NSCH medication estimates are positive. That finding weighs against a large total-coverage cliff as the explanation for the medication pattern. A coverage-mediated treatment change remains possible, however, because children can remain publicly insured while moving between Medicaid and separate CHIP or while facing different plans, provider networks, formularies, or authorization rules. The ACS cannot observe those within-public-coverage changes.

Direct evidence on the mechanism would require linked enrollment and claims data that identify Medicaid and separate CHIP separately, record the date and direction of program or plan changes, and measure

associated changes in prescribers, therapists, networks, formularies, and treatment use. Until such data are available, the ACS is best read as an exploratory total-coverage check rather than evidence of program switching or treatment substitution.

AI-Assistance Disclosure

From May through August 2026, the author used Anthropic Claude and OpenAI Codex as writing, coding, and verification assistants. Because both services were continuously updated during this period, exact model build identifiers were not preserved. The tools helped draft and edit exposition and caveat language, check internal consistency across the manuscript and supplement, write reproducibility and validation scripts, and compare regenerated numerical outputs with reported results. The author reviewed and verified all AI-assisted prose, code, citations, and numerical claims, made all substantive research and interpretation decisions, and accepts full responsibility for the submitted work. No AI system is listed as an author.

References

- Armstrong, T.B., Kolesár, M., 2020. Simple and honest confidence intervals in nonparametric regression. *Quantitative Economics* 11, 1–39. doi:10.3982/QE1199.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.
- Bensken, W.P., Dankovchik, J., Fein, H.L., Duhon, G., Sills, M.R., 2025. Unwinding of continuous Medicaid coverage among pediatric community health center patients. *JAMA Network Open* 8, e2458155. doi:10.1001/jamanetworkopen.2024.58155.
- Bisgaier, J., Rhodes, K.V., 2011. Auditing access to specialty care for children with public insurance. *New England Journal of Medicine* 364, 2324–2333. doi:10.1056/NEJMs1013285.
- Blake, S.C., Song, M., Gaydos, L., Cummings, J.R., 2019. Prior authorization policies and preferred drug lists in Medicaid plans: Stakeholder perspectives on the implications for youth with ADHD. *Administration and Policy in Mental Health and Mental Health Services Research* 46, 580–595. doi:10.1007/s10488-019-00937-y.
- Cattaneo, M.D., Jansson, M., Ma, X., 2020. Simple local polynomial density estimators. *Journal of the American Statistical Association* 115, 1449–1455. doi:10.1080/01621459.2019.1635480.
- Centers for Medicare & Medicaid Services, 2026a. CHIP cost sharing. Medicaid.gov. URL: <https://www.medicaid.gov/chip/chip-cost-sharing>. accessed July 22, 2026.
- Centers for Medicare & Medicaid Services, 2026b. Cost sharing. Medicaid.gov. URL: <https://www.medicaid.gov/medicaid/cost-sharing>. accessed July 22, 2026.
- Chua, K.P., Constantin, J., Kenney, G.M., Conti, R.M., Simon, K., 2025. Changes in chronic medication dispensing to children and young adults during Medicaid unwinding. *Pediatrics* 155, e2024070380. doi:10.1542/peds.2024-070380.
- Higa-McMillan, C.K., Francis, S.E., Rith-Najarian, L., Chorpita, B.F., 2016. Evidence base update: 50 years of research on treatment for child and adolescent anxiety. *Journal of Clinical Child & Adolescent Psychology* 45, 91–113. doi:10.1080/15374416.2015.1046177.
- Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6, 65–70. doi:10.2307/4615733.
- Hu, J.C., Cummings, J.R., Ji, X., Wilk, A.S., 2023. Evaluating Medicaid managed care network adequacy standards and associations with specialty care access for children. *Health Affairs* 42, 759–769. doi:10.1377/hlthaff.2022.01439.

- Hyman, S.L., Levy, S.E., Myers, S.M., Council on Children with Disabilities, Section on Developmental and Behavioral Pediatrics, 2020. Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics* 145, e20193447. doi:10.1542/peds.2019-3447.
- Kaiser Family Foundation, 2016–2024. Medicaid and CHIP eligibility, enrollment, and cost-sharing policies: Annual 50-state surveys, January 2016–May 2024 [dataset]. Kaiser Family Foundation reports (annual). URL: <https://www.kff.org/series/medicaid-eligibility-and-enrollment-survey/>. Series of annual 50-state surveys for January 2016 through May 2024, including the January 2021 survey; the 2024 survey’s threshold table is dated May 2024. Threshold values used in this paper are drawn from each survey’s Table 1 (Income Eligibility Limits for Children’s Health Coverage as a Percent of the Federal Poverty Level). The January 2021 Table 1 extract used for threshold coding does not contain the comparable continuous-eligibility table used for the year-by-year policy cross-walk. Accessed April 2026.
- Kolesár, M., Rothe, C., 2018. Inference in regression discontinuity designs with a discrete running variable. *American Economic Review* 108, 2277–2304. doi:10.1257/aer.20160945.
- Murphy, T.K., Lewin, A.B., Storch, E.A., Stock, S., AACAP Committee on Quality Issues, 2013. Practice parameter for the assessment and treatment of children and adolescents with tic disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 52, 1341–1359. doi:10.1016/j.jaac.2013.09.015.
- Pringsheim, T., Okun, M.S., Müller-Vahl, K., Martino, D., Jankovic, J., Cavanna, A.E., Woods, D.W., Robinson, M., Jarvie, E., Roessner, V., Oskoui, M., Holler-Managan, Y., Piacentini, J., 2019. Practice guideline recommendations summary: Treatment of tics in people with Tourette syndrome and chronic tic disorders. *Neurology* 92, 896–906. doi:10.1212/WNL.00000000000007466.
- Riley, M., Ahmed, S., Locke, A., 2016. Common questions about oppositional defiant disorder. *American Family Physician* 93, 586–591. URL: <https://www.aafp.org/pubs/afp/issues/2016/0401/p586.html>.
- Ruggles, S., Flood, S., Sobek, M., Backman, D., Cooper, G., Rivera Drew, J.A., Richards, S., Rodgers, R., Schroeder, J., Williams, K.C.W., 2025. IPUMS USA: Version 16.0 [dataset]. IPUMS, Minneapolis, MN. doi:10.18128/D010.V16.0.
- Schmid, I., Burcu, M., Zito, J.M., 2015. Medicaid prior authorization policies for pediatric use of antipsychotic medications. *JAMA* 313, 966–968. doi:10.1001/jama.2015.0763.
- Steiner, H., Remsing, L., 2007. Practice parameter for the assessment and treatment of children and adolescents with oppositional defiant disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 46, 126–141. doi:10.1097/01.chi.0000246060.62706.af.

- U.S. Code of Federal Regulations, 2026. Cumulative cost-sharing maximum: 42 CFR Section 457.560. Code of Federal Regulations, Title 42, Section 457.560. URL: <https://www.ecfr.gov/current/title-42/part-457/section-457.560>. accessed July 22, 2026.
- Walter, H.J., Abright, A.R., Bukstein, O.G., Diamond, J., Keable, H., Ripperger-Suhler, J., Rockhill, C., 2023. Clinical practice guideline for the assessment and treatment of children and adolescents with major and persistent depressive disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 62, 479–502. doi:10.1016/j.jaac.2022.10.001.
- Walter, H.J., Bukstein, O.G., Abright, A.R., Keable, H., Ramtekkar, U., Ripperger-Suhler, J., Rockhill, C., 2020. Clinical practice guideline for the assessment and treatment of children and adolescents with anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 59, 1107–1124. doi:10.1016/j.jaac.2020.05.005.
- Wolraich, M.L., Hagan, J.F., Allan, C., Chan, E., Davison, D., Earls, M., Evans, S.W., Flinn, S.K., Froehlich, T., Frost, J., Holbrook, J.R., Lehmann, C.U., Lessin, H.R., Okechukwu, K., Pierce, K.L., Winner, J.D., Zurhellen, W., 2019. Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics* 144, e20192528. doi:10.1542/peds.2019-2528.
- Zhu, J.M., Charlesworth, C.J., Polsky, D., McConnell, K.J., 2022. Phantom networks: Discrepancies between reported and realized mental health care access in Oregon Medicaid. *Health Affairs* 41, 1013–1022. doi:10.1377/hlthaff.2022.00052.