

What Did Medicaid Do? Re-examining the Oregon Health Insurance Experiment

David Crabtree

October 2026

Summary

In 2008, Oregon could afford to expand Medicaid to only some of the low-income adults who wanted it, so it held a lottery. Because winners were chosen at random, the Oregon Health Insurance Experiment is one of the few randomized tests of what health insurance does. Its best-known clinical results, published by Baicker et al. in the *New England Journal of Medicine* in 2013, found that Medicaid reduced depression and financial hardship but produced “no significant improvements in measured physical health outcomes.”

I rebuilt the analysis from the public-use data, reproducing every headline estimate in the paper, and then examined how well the evidence supports the way those results are described. **Some of the paper’s summary claims go further than its evidence. The physical-health null comes from a study too small to detect moderate effects, while the positive findings for depression and financial strain are described by their point estimates rather than their much wider confidence intervals, a more lenient standard than the paper applies to its null results.**

1. **The blood pressure “null” cannot distinguish no effect from a meaningful one.** The study had a 67% chance of detecting a reduction the authors themselves call clinically significant, and 80% power only for reductions of at least 8.3 points, about half the baseline rate. Its 95% confidence interval runs from a 44% reduction to a 27% increase.
2. **The positive findings are stated more precisely than the data allow.** “Nearly eliminated” catastrophic medical spending describes a point estimate whose 95% confidence interval is consistent with a reduction as small as 13%. The “30% reduction” in depression has an interval running from 5% to 56%.
3. **The paper’s definition of elevated blood pressure is misstated.** Table 2’s footnote requires systolic pressure of at least 140 *and* diastolic of at least 90 mm Hg; the published results use *or*. Under the stated definition, 7.0% of the control group would have elevated blood pressure, not the reported 16.3%.

None of these overturns the paper’s conclusions. Together they suggest the study is better read as **“Medicaid improved mental health and financial security, and the study was too small to say much about physical health”** than as evidence that Medicaid does not improve physical health.

The experiment

Lottery winners won the chance to apply for Medicaid; many did not apply or did not qualify. Winning raised Medicaid enrollment from 18% to 42%, an increase of 24.1 percentage points. Comparing winners with non-winners therefore understates what coverage itself does.

To estimate the effect of *having* Medicaid, I use the lottery as an instrument for enrollment (two-stage least squares). Intuitively, since winning raised enrollment by about a quarter, the effect of coverage is roughly four times the effect of winning. The estimate applies to people who enrolled because they won, and it rests on two assumptions: winning affected health only through Medicaid, and nobody was made less likely to enroll by winning.

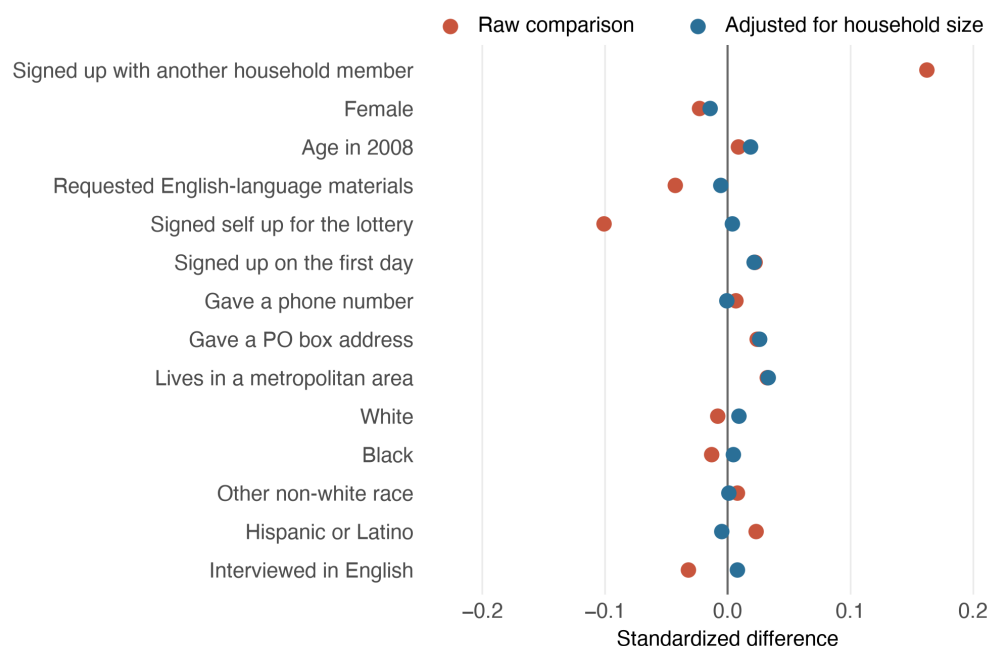
The analysis uses the 12,229 people who completed an in-person health exam from about two years after the lottery. Following the original study, all estimates use survey weights, cluster standard errors by household, and control for the number of household members on the lottery list.

Was the lottery fair?

The lottery drew households, not people, so people who signed up with a household member had more chances to win. After accounting for that, winners and non-winners are nearly identical on every characteristic recorded before the drawing and at the exam (all standardized differences below 0.04 standard deviations). Together the characteristics cannot predict who won (joint test $p = 0.80$), and winners and non-winners completed the exam at the same rate (73.4% vs. 73.0%, $p = 0.86$).

After accounting for household size, lottery winners and non-winners look alike

Standardized difference, winners minus non-winners. Households with more members on the lottery list had more chances to win.

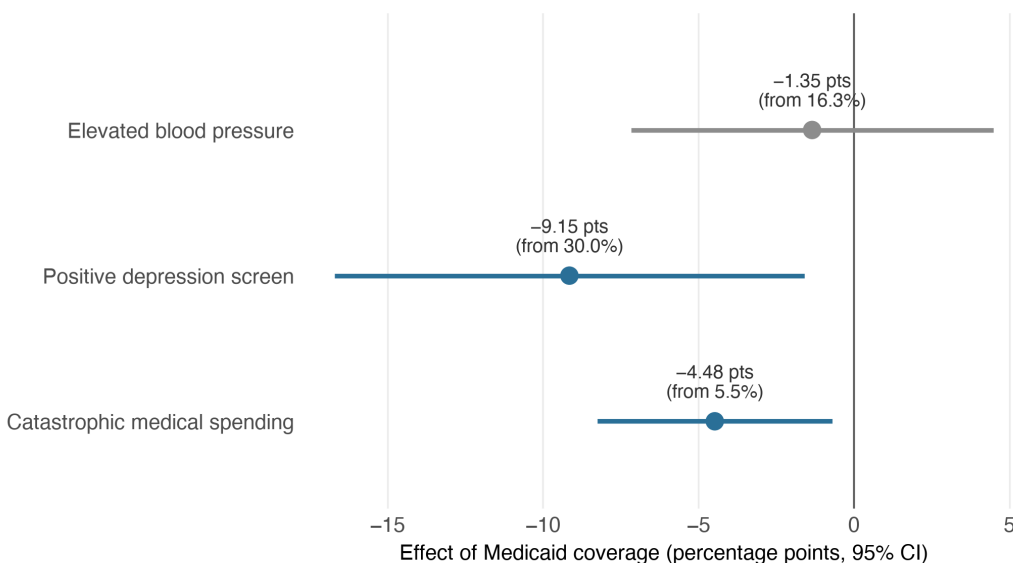


What Medicaid did

Outcome	Rate among non-winners	Effect of Medicaid, points (95% CI)
Elevated blood pressure	16.3%	−1.35 (−7.17 to 4.48)
Positive depression screen	30.0%	−9.15 (−16.70 to −1.60)
Catastrophic medical spending	5.5%	−4.48 (−8.26 to −0.69)

Having Medicaid reduced positive depression screens by 9.15 points (95% CI 1.60 to 16.70) and catastrophic medical spending (out-of-pocket costs above 30% of income) by 4.48 points (95% CI 0.69 to 8.26). The estimated effect on elevated blood pressure was small and not statistically significant: −1.35 points (95% CI −7.17 to 4.48). These reproduce the published estimates (see [Reproducibility](#)).

Medicaid reduced depression and financial catastrophe, with no detectable effect on high blood pressure



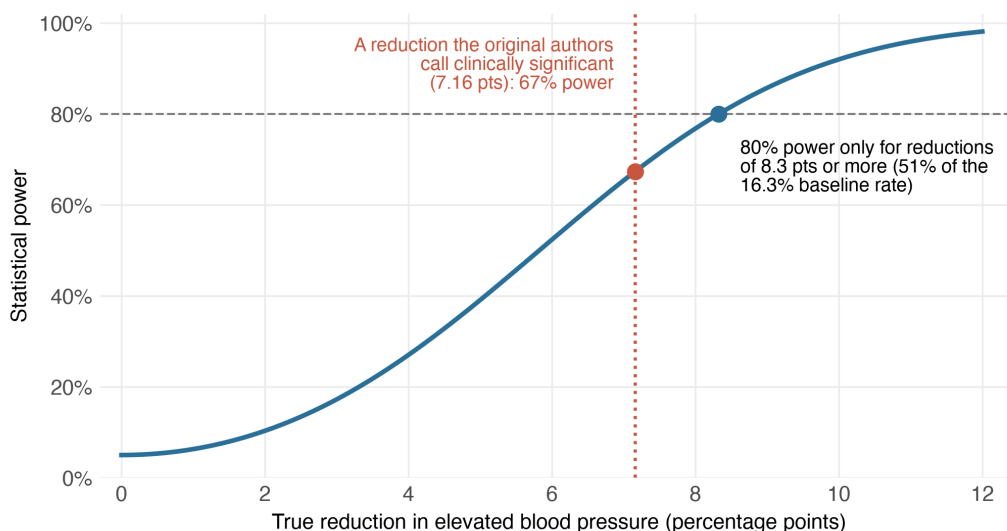
Re-examining the evidence

A null result from a study too small to detect it

The original authors acknowledge that their power to detect changes in physical health was limited, and note that their confidence interval includes a clinically significant 7.16-point reduction in hypertension. Below I analyze precisely what magnitude of effects the authors were capable of capturing.

The study could only detect large effects on blood pressure

Chance of a statistically significant result if Medicaid truly reduced elevated blood pressure by a given amount, at the study's realized precision.



Because only a quarter of winners gained coverage, the study's precision for the effect of Medicaid is far lower than its 12,000-person sample suggests. Detecting a 20% reduction in elevated blood pressure (3.3 points) with 80% power would have required about 80,000 people, 6.5 times the actual sample. The study can rule out very large effects; it cannot rule out moderate ones.

Headline magnitudes versus their uncertainty

The paper's abstract describes catastrophic expenditures as "nearly eliminated," and the results section reports a relative reduction of "more than 80%." That is the point estimate. Expressed relative to the 5.5% rate among non-winners, the 95% confidence interval runs from a 13% reduction to complete elimination. The depression estimate, reported as a 30% reduction, has an interval from 5% to 56%.¹

These are real effects: both intervals exclude zero. But the summary language describes the most likely value, while the null findings for physical health are, appropriately, discussed in terms of their uncertainty. Applying the same standard to both gives a more cautious picture of the financial and mental-health benefits.

A misstated definition

Definition	Control-group rate	Effect of Medicaid (points)
Systolic at least 140 OR diastolic at least 90 (replication code)	16.3%	−1.35

¹Relative intervals divide the confidence limits for the effect by the control-group rate, treating that rate as fixed. They are approximate but adequate for comparing with the paper's own relative statements.

Definition	Control-group rate	Effect of Medicaid (points)
Systolic at least 140 AND diastolic at least 90 (Table 2 footnote)	7.0%	−1.83

Table 2 of the paper defines elevated blood pressure as systolic pressure of at least 140 *and* diastolic pressure of at least 90 mm Hg. The reported control-group rate of 16.3% is reproduced only with *or*, which is the standard clinical threshold and the definition in the authors’ replication code. This is a documentation error rather than an analytical one, but a reader applying the stated definition would get different results.

Could better analysis recover the lost precision?

In A/B testing, adjusting for characteristics measured before treatment is a standard way to tighten estimates. I re-estimated each effect adding sex, age, race, ethnicity, language, and lottery sign-up characteristics as controls.

Outcome	Variance explained by covariates	Reduction in standard error
Elevated blood pressure	8%	−0.1%
Positive depression screen	5%	1.2%
Catastrophic medical spending	1%	1.2%

It barely helps. Standard errors shrink by at most about 1%, because the available covariates explain only 1%–8% of the variation in outcomes. Adjustment pays off when a covariate strongly predicts the outcome, typically a measurement of the outcome itself before treatment. The study never measured participants’ blood pressure before the lottery, and no demographic characteristic sufficiently substitutes for it.

Limitations

- **Who the estimates describe.** The effects apply to low-income Portland-area adults who signed up for the lottery and enrolled because they won, and to roughly 17 months of coverage on average.
- **Scope.** I examine three of the paper’s outcomes. The original study reports many more, including diabetes diagnosis and medication, where it did find effects.
- **Power calculations.** These use the study’s realized standard errors, so they describe the design as executed rather than as planned.

Reproducibility

The data are the Oregon Health Insurance Experiment public-use files, available from NBER under a data-use agreement. The full analysis is implemented in both R and Python, and the two agree on every estimate to at least four decimal places. Both also reproduce the published results:

Estimate	This analysis	Baicker et al. (2013)
Lottery's effect on Medicaid enrollment	24.1	24.1
Lottery's effect on exam response	0.28 (p = 0.86)	0.28 (p = 0.86)
Elevated blood pressure	−1.35	−1.33
Positive depression screen	−9.15	−9.15
Catastrophic medical spending	−4.48	−4.48

The blood pressure estimate differs from the published value by 0.02 points; the source of the gap is unidentified.

Reference. Baicker, K., Taubman, S. L., Allen, H. L., et al. (2013). The Oregon Experiment — Effects of Medicaid on Clinical Outcomes. *New England Journal of Medicine*, 368(18), 1713–1722.

Data. Oregon Health Insurance Experiment, public-use data. National Bureau of Economic Research. <https://www.nber.org/research/data/oregon-health-insurance-experiment-data>