Does the individual mandate affect insurance coverage? Evidence from tax returns*

Ithai Z. Lurie & Daniel W. Sacks & Bradley Heim

September, 2019

Abstract

We estimate the effect of the ACAs individual mandate on insurance coverage using regression discontinuity and regression kink designs with tax return data. We have four key results. First, the actual penalty paid per uninsured month is less than half the statutory amount. Second, nonetheless, we find visually clear and statistically significant responses to both extensive margin exposure to the mandate and to marginal increases in the mandate penalty. Third, we find substantial heterogeneity in who responds; men are especially responsive. Fourth, our estimates imply fairly small quantitative responses to the individual mandate, especially in the Health Insurance Exchanges.

JEL codes: G22, H51, I13
Key words: Health insurance, mandate

---

*Lurie: Office of the Treasury, Department of Tax Analysis. Email: ithai.lurie@treasury.gov. Sacks: Kelley School of Business, Indiana University. Email: dansacks@indiana.edu. Heim: O'Neill School of Public and Environmental Affairs, Indiana University. Email: heimb@indiana.edu. We are grateful to Tom DeLeire, Alex Gelber, Josh Gottlieb, Tami Gurley-Calvez, Martin Hackmann, Adam Isen, Robert Kaestner, Sean Lyons, Alex Minicozzi, Shanthi Ramnath, Mark Shepard, Ben Sommers, Sebastian Tello-Trillo, Matthijs Wildenbeest, anonymous referees, and seminar and conference participants at Illinois University (Geis), Indiana University, IUPUI, University of Virginia, ASHEcon, AHEC, MHEC, the Booth Health Economics Initiative Annual Conference, and NTA Spring Symposium for helpful comments and suggestions. The views expressed here are those of the authors and not necessarily those of the U.S. Department of the Treasury.

One of the most contentious features of the Affordable Care Act (ACA) is the individual shared responsibility provision, colloquially known as the individual mandate, which requires that Americans obtain insurance coverage or pay a tax penalty. The individual mandate's supporters argue that it is a critical tool for achieving universal health insurance coverage and reducing adverse selection. Opponents of the individual mandate view it as an unreasonable assault on liberty, an encroachment of government authority, and an onerous burden on low-income Americans.[1]

The individual mandate was effectively repealed by the Tax Cuts and Jobs Act of 2017, which set the penalty for non-coverage to zero.

Despite the controversy, little is known about the effect of the ACA's mandate on insurance coverage. Understanding the coverage response is critical for at least three reasons. First, it tests whether the mandate works as intended—by raising coverage through a greater penalty. Second, the coverage response is critical for forecasting the effects of repeal on coverage, premiums, and government spending on insurance subsidies. Third, in the absence of a federal mandate, three states have enacted their own mandates, and others are considering doing so as well.[2] The ACA experience offers a guide for current policy making. We provide quasi-experimental evidence on the effect of the individual mandate penalty on insurance coverage using administrative data.

Our empirical approach exploits two nonlinearities in the mandate penalty function. The first nonlinearity is that people are exempt from the penalty if their income is below 138 percent of the federal poverty line (FPL) and they live in a state that did not expand Medicaid. This creates a discontinuity in the mandate penalty. Second, the penalty amount is a kinked function of income, with a kink for single taxpayers at $26,550 in 2015 and $38,150 in 2016. We implement regression discontinuity and regression kink designs, looking for discontinuities and kinks in coverage corresponding to the discontinuities and kinks in the penalty amount.

We use income and insurance data derived from the population of tax returns. We

[1]See, e.g., https://thehill.com/blogs/blog-briefing-room/news/59761-kyl-health-bill-a-stunning-assault-on-liberty-

[2]Maryland, Massachusetts, and Washington, D.C. have enacted insurance mandates. California's legislature has passed a law with a mandate. See also https://www.wsj.com/articles/states-look-at-establishing-their-own-health-insurance-mandates.

observe exact income, as well as both self-reported and third-party verified insurance coverage. These comprehensive administrative data offer a clear advantage over previously used data sources. They are free from the measurement error that plagues survey data, and they record all sources of coverage, unlike administrative data from a single employer or Health Insurance Marketplace. Our discontinuity analysis focuses on a sample of about 5 million people filing tax returns in non-expansion states, with income between 133 and 143 percent of FPL in 2015 and 2016. Our kink analysis focuses on a sample of 1.1 million people with income close to the mandate kink point, filing single-person tax returns without signs of offers of employer-sponsored insurance (an imperfect proxy for actual offers).

We begin by documenting a clear first stage. We find large discontinuities in the penalty paid per uninsured month at 138 percent of FPL, and clear kinks at the mandate kink points. However, these discontinuities and kinks are much less than the statutory amounts. For example, for the average person in our sample in 2016, the statutory penalty per month of uninsurance was about $45, but we estimate a discontinuity of only about $10 when people become subject to the mandate penalty. Although the mandate penalty has bite, it is not as sharp as the statutory amount might indicate.

Given the clear nonlinearities in penalty amount, we next turn to estimating coverage effects. We find visually clear, statistically significant discontinuities in coverage at the 2015 and 2016 eligibility thresholds. When people become subject to the individual mandate, insurance coverage increases by about 1 percentage point in 2015, and by about 2.5 percentage points in 2016. These responses imply a semi-elasticity of about 0.06 with respect to the statutory penalty. Given the relatively low penalty actually paid, however, the semi-elasticity with respect to the observed penalty is larger, about 0.2 in 2015 and 0.3 in 2016, meaning each extra dollar of penalty per uninsured month raises coverage by 0.2-0.3 percent. This estimate is on the low end of the literature, which finds semi elasticities from 0.2 to 1.0. The discontinuity evidence therefore shows a clear coverage response to a discontinuous increase of the mandate penalty from zero to a large, positive amount. On a per-dollar basis we find fairly small responses. Some researchers have suggested that the first-dollar response to the individual mandate may be especially large, because some people may exhibit a "taste for compliance" (Saltzman et al., 2015; Saltzman, 2018). Our findings suggest that if present, such a taste for compliance is not large.

We further estimate the response to marginal increases in the mandate penalty amount by examining behavior around the mandate kink points. We find a visually clear kink in coverage at the 2015 kink point, but less clear evidence at the 2016 kink point. The estimates at the 2015 kink point ($26,550) imply a semi-elasticity of about 0.9 with respect to the statutory penalty, and at the 2016 kink point we estimate a statistically insignificant semi-elasticity of 0.35. Thus we find that coverage respond on the margin to a greater mandate penalty, not only to the presence of the penalty.[3]

An important goal of the individual mandate was not only to raise coverage but especially to induce healthy, low-cost people into coverage, to ease adverse selection concerns in the individual market. We therefore investigate heterogenous responses to the individual mandate. We find that men are especially responsive. On the other hand, people with an indicator of poor health—itemized medical expenditures on prior tax returns—seem to respond less to the mandate penalty, although this difference is not statistically significant. We also find that people with lower income over the past three years are more responsive to the mandate penalty. Overall, the evidence suggests that the individual mandate induces low-spending people to obtain insurance coverage.

Our main estimates assume that people can perfectly anticipate their mandate penalty. This might be difficult, however, because insurance is chosen before taxable income, and hence the mandate penalty, are fully realized. Imperfect income forecasts could therefore attenuate our results. To investigate this possibility, we estimate alternative models that use prior year income as the running variable. We continue to find coverage kinks and discontinuities but, surprisingly, no "first stage" kink or discontinuity in the mandate penalty. These results suggest that some people use prior year income to forecast their penalty, but their forecasts are not rational, in the sense that there is a discontinuity in the expected penalty but not in the realized penalty. More generally, this finding suggests

[3]To maximize our limited power in the kink sample, we focus on people without signs of an ESI offer, so these estimate reflect responses among that group only. We do not find a kink in the offer rate, although as we discuss below, this does not rule out the possibility that the mandate affects equilibrium ESI offers. When we look at people with signs of an ESI offer, we find ambiguous and not robust responses at the kinks. In the discontinuity sample, however, we find that the penalty increases ESI coverage.

that imperfect salience of the penalty amount attenuates observed responses, consistent with past evidence on the importance of tax salience (e.g. Chetty et al. (2009)). To account for the responses of people with non-rational forecasts, we develop a simple framework, which implies that our estimate should be scaled up by, at most, about 50 percent.

Even at these larger estimates, however, our estimates imply fairly small quantitative responses to the mandate. For example, among people in non-expansion states with income between 110 and 160 percent of FPL in 2016, our baseline estimates suggests that, had the penalty been zero, overall coverage would have fallen by about a quarter of a million lives, or roughly 2.5 percent. Exchange coverage would have fallen by about 34,000 covered lives or 3.5 percent. Scaling these coverage losses up by even fifty percent would still reveal fairly small coverage losses. We caution, however, that these simple extrapolations may be inaccurate for several reasons. They are based on estimates that are local to our estimation sample, they do not account for insurer responses such as lower premiums, and they are out of sample in the sense that they observe larger penalty changes than we see in the data.

Of course, our estimates can only be interpreted in a causal way under the identification assumption that the counterfactual relationship between income and insurance coverage would be smooth, absent any discontinuity kink in the mandate penalty. This assumption could fail if people manipulate their income to reduce their mandate penalty, or if insurance coverage is inherently kinked or discontinuous. We find no evidence for income manipulation, nor do we find meaningful discontinuities or kinks in predetermined predictors of insurance demand. Using expansion states as a placebo group, we also find no coverage discontinuity at 138 percent of FPL.

We believe this is the first quasi-experimental evidence that the ACA's individual mandate increased insurance coverage. Our work provides new, credible evidence on coverage effects of the mandate. However, by examining the specific sources of coverage gains, we show that the mandate increases coverage along margins not previous considered. Other papers have found coverage effects of the Massachusetts mandate, or looked at the ACA but found no coverage effects. Our estimates complement work by Hackmann et al. (2015) and Jaffe and Shepard (2017), who find the Massachusetts mandate increased coverage in the individual insurance market, as well as structural work by Aizawa and Fang (2018) which estimates and simulates dynamic equilibrium models to predict steady-state effects of the ACA and the individual mandate. Our work is also related to research on the coverage responses to premium subsidies (Tebaldi, 2017; Finkelstein et al., 2019; Saltzman, 2018). Our results are not strictly comparable to these past estimates, however, because we estimate overall coverage effects, rather than effects on individual market or Exchange coverage, which past literature has focused on.

Our results contrast with Frean et al. (2017), who find the individual mandate has a small, negative coverage effect. Several reasons likely explain this difference. First, we use different data and a different identification strategy. Frean et al. (2017) estimate triple difference models using survey data with self-reported income, so their estimates pertain to the population as a whole, whereas our estimates are a local average treatment effect. As it is likely that responses to the mandate vary across the income distribution because of differences in income, tastes, and access to subsidies and ESI, we should not necessarily expect to obtain identical responses as Frean et al. However, differences in data may contribute to our different findings. Measurement error in self-reported income could lead to spurious findings because the mandate penalty is negatively correlated with Medicaid eligibility and PTC coverage, so that the negative response to the mandate could reflect a response to these other provisions of the ACA. Finally, we treat the penalty differently. Frean et al. (2017) assume that the penalty is fully paid among uninsured people; we find, however, that many uninsured people do not pay for the penalty, and failing to account for the penalty underpayment would lead to substantially lower penalty sensitivity.

1 The individual mandate

1.1 The individual mandate

The ACA's individual mandate requires that Americans pay a fee for each month they go without health insurance. The goal of the mandate is to raise insurance coverage, particularly among healthy people, thereby limiting adverse selection in the individual health insurance market. Adverse selection was a concern under the ACA because it included regulations sharply limiting insurers' ability to engage in risk-based pricing. Prior state experiments with similar regulations had resulted in serious adverse selection problems (Buchmueller and DiNardo, 2002; Congdon et al., 2008; Herring and Pauly, 2006; Lo Sasso and Lurie, 2009).

The ACA's individual mandate operates as follows. For each month that a taxpayer (or a dependent) is uninsured, an "individual shared responsibility payment" is owed, unless the individual qualifies for an exemption.[4] We refer to this payment as the mandate penalty. For each month of uninsurance, the mandate penalty is given by

$$Penalty = 1/12 \max \left\{ \min \left\{ [A + .5C]F, 3F \right\}, S(\text{MAGI} - \text{tax filing threshold}) \right\}, \quad (1)$$

where A is the number of uninsured adults, C is the number of uninsured children, F is a fixed dollar amount, S is a share of income, and MAGI is modified adjusted gross income.[5] People with income below the tax filing threshold are exempt. For simplicity, we refer to modified adjusted gross income as "income." The fixed dollar amount (F) was \$95 in 2014, \$325 in 2015, and \$695 in 2016, and is indexed to inflation thereafter. The share of income (S) was 1% in 2014, 2% in 2015, and 2.5% in 2016 and thereafter. For single filers, the filing threshold was \$10,150 in 2014, \$10,300 in 2015 and \$10,350 in 2016. The penalty is waived if a person is uninsured for less than three continuous months, and it is capped at the national average price of the cheapest bronze plan available.[6]

For a single tax payer with no children, the formula for the penalty simplifies considerably. Figure 1 plots the mandate penalty as a function of income for single tax payers in 2015 and 2016. The penalty is a kinked function of income, with a kink occurring where the fixed penalty equals the dollar share of income, \$26,550 (with a kink of 0.02) in 2015, and \$38,150 (with a kink of 0.025) in 2016.[7] We plot vertical lines at 200, 250, 300, and

[4] See Lurie and McCubbin (2016) for more detail.

[5] Modified adjusted gross income is adjusted gross income (AGI) plus tax exempt interest and foreign earned income and housing excluded from AGI.

[6] The penalty is calculated in the instructions to Form 8965, and is paid when the taxpayer files their annual tax return.

[7] As we explain in Appendix A, for multiperson households, the penalty creates more complex incentives.

400 percent of FPL because, as we explain in more detail below, there are other relevant policy nonlinearities at these thresholds. The figure shows that the mandate kink points occur between these other policy threshold. The basic idea in our regression kink models is to look for a corresponding kink in insurance coverage at these income thresholds.

Tax payers may obtain an exemption from the mandate penalty depending on their circumstances. Importantly, taxpayers are exempt from the penalty if their income is below 138 percent of the poverty line and they live in a state that did not expand Medicaid. That is, if a person should be income-eligible for Medicaid based, but her state did not expand Medicaid, she is exempt from the penalty. This exemption creates a discontinuity in the penalty amount at 138 percent of the poverty line in non-expansion states. We illustrate this discontinuity in Figure 1. There is another discontinuity in the penalty: people living in expansion states are exempt from the penalty if their income is below the federal filing threshold. We cannot study this discontinuity because there is a clear discontinuity in the distribution of taxable income here. (This is perhaps unsurprising as taxes need not be filed if income is below the filing threshold.)

1.2 Other income-based insurance inducements

Several other policies also encourage insurance coverage in a way that is tied to income. Some of these policies generate kinks near the mandate thresholds, although the other policy kinks are never collocated with the mandate thresholds.

The Premium Tax Credit and the Advanced Premium Tax Credit: The Premium Tax Credit (PTC) is a subsidy which may be used for purchasing an insurance plan on the Health Insurance Marketplaces. The PTC is equal to the difference between a household's "benchmark premiums"—the second lowest-cost silver-tier health insurance plan available to it in the Health Insurance Marketplaces—and its expected contribution, a percent of income specified by law that ranges from 2 to almost 10 percent. The expected contribution is a kinked and discontinuous function of income, so for the PTC is also a kinked and discontinuous function of income, with potential discontinuities or kinks at 100, 133, 150, 200, 250, 300, and 400 percent of the FPL. There is also a kink in the PTC at the income level at which the expected contribution exactly equals the benchmark premium; this kink point varies across markets, since the benchmark premium varies across markets.[8] To help taxpayers manage liquidity, the PTC is paid in advance, throughout the coverage year, in the form of the Advanced Premium Tax Credit (APTC). APTC payment amounts are based not on realized MAGI, but on project income, which Marketplace enrollees report to the Marketplace at the time of signing up for insurance. If APTC payment are too high (because realized income exceeded projected income), taxpayers must repay the excess, with repayment limits that depend on realized income. These repayment limits are discontinuous functions of income, with discontinuities at 200, 300, and 400 percent of FPL. Heim et al. (2017) provide more detail on the PTC, APTC, and repayment requirements, and document income responses to the premium tax credit at the 400 percent of FPL discontinuity.

Cost-sharing reductions: Whereas the PTC helps subsidizes premiums, cost-sharing reductions (CSRs) subsidize out-of-pocket health care expenses. For every standard silver plan that insurers offer on the Marketplace, they must offer three additional CSR plans, which are identical in all aspects except their cost sharing. A standard silver plan has an actuarial value of 70 percent, meaning it covers 70 percent of expected health care costs. The CSR plans have actuarial value of 73 percent, 87 percent, and 93 percent. Insurers must charge the same premium for these more generous plans as they do for the base silver plan; the government paid for the additional cost-sharing until 2018. Only low-income people are eligible to purchase these more generous plans. People with income between 100 and 150 percent of FPL may purchase the 93 percent actuarial value plans; people with income between 150 and 200 percent of FPL may purchase the 87 percent actuarial value plans; and people with income between 200 and 250 percent of FPL may purchase the 73 percent actuarial value plans.[9]

Other policies: We believe that the PTC and CSRs are the most important potential threats to identification, in the sense that they create meaningful nonlinearities in the

[8]The PTC and the mandate penalty are assessed using slightly different modifications of AGI. The definition of MAGI for the purpose of PTC is similar to the mandate penalty, but also includes the non-taxable Social Security income. We focus on a population aged 27-64 with MAGI between $29,425 and $47,080, so we expect that non-taxable Social Security income is zero for nearly all our sample.

[9]DeLeire et al. (2017) show that these subsidies influence plan choice on the Exchanges.

incentives to obtain insurance near the mandate kink point. Several other programs might also be relevant. Medicaid eligibility of course depends on income, although eligibility is determined in terms of rolling income throughout the year, rather than realized income. The eligibility levels vary across states and people. In some states single, childless adults are not eligible at any income levels, whereas children can be eligible up to 300 percent of FPL or above.[10] Importantly for our analysis, however Medicaid eligibility is assessed on a different income basis than is mandate penalty. Medicaid eligibility depends on a rolling average of income, assess over the previous several months. The mandate penalty depends on realized taxable income. In practice, therefore, a household's income for assessing Medicaid eligibility can be quite different from its income for determining the mandate penalty (or the PTC for that matter). For example, a temporarily low income household can qualify for Medicaid, but have a high enough annual income that it would be subject to the penalty if it were uninsured. This Medicaid qualification would also not disqualify them from receiving a PTC.

Another kink in the incentive to obtain insurance comes from the tax deductibility of employer sponsored insurance, which creates a kink in the incentive to obtain ESI at each kink point in the income tax code. These kinks turn out not to be close to our nonlinearity points, because the income tax applies to taxable income rather than MAGI. For example, for a single taxpayer in 2016, the 15 percent tax bracket ran from $9,275 to $37,650 of taxable income. For a single tax payer with one exemption claiming the standard deduction, this works out to $19,625 to $48,000 in MAGI. Other programs such as SNAP, TANF, and the EITC may affect insurance demand through income effects. These programs do not have discontinuities at 138 percent of FPL, however, and at higher income levels the benefits are generally small.

Summary: The PTC, APTC, repayment requirements, and CSRs all create kinks or discontinuities in the incentive to obtain health insurance coverage.[11] These nonlinearities

[10] See https://www.kff.org/health-reform/state-indicator/medicaid-income-eligibility-limits-for-adults-as-a-percent-of-the-federal-poverty-level/ and https://www.kff.org/medicaid/fact-sheet/where-are-states-today-medicaid-and-chip/.

[11] Tebaldi (2017) studies coverage responses to the PTC in the California Marketplace, and Frean et al. (2017) study coverage responses at a national level using the in the ACS. DeLeire et al. (2017) study coverage responses to the CSRs.

occur at even increments of the FPL: 100, 133, 150, 200, 250, 300, and 400 percent. The mandate discontinuity and kink points occurs between these critical values, as shown in Figure 1. It is therefore possible to separately identify the coverage effect of the individual mandate from the coverage effects of these policies by looking within narrow windows of FPL (133-150 percent, 200-250 percent, and 300-400 percent).

2 Empirical Approach

2.1 Econometric specification

We estimate regression discontinuity and regression kink models The effect of the penalty on coverage can be obtained as the ratio of the kink or discontinuity in coverage to the kink or discontinuity in the penalty (Hahn et al., 2001; Card et al., 2015). We focus on regressions of the form:

$$y_i = \beta_0 + \beta_1 v_i + \beta_2 1\{v_i \geq 0\} + \beta_3 v_i 1\{v_i \geq 0\} + \varepsilon_i, \tag{2}$$

where y_i is an outcome of interest, such as months of insurance coverage, and v_i, our running varaible is income relative to FPL, normalized relative to the kink point or discontinuity. Our estimate of the coverage discontinuity is β_2 and our estimate of the coverage kink is β_3.

For the regression discontinuity estimates, we use a bandwidth of five FPL points. This bandwidth is slightly narrower than the asymptotically optimal bandwidth (Calonico et al., 2014), which ranges from 5.2 to 13 FPL points, depending on the year and outcome. We choose this narrow bandwidth to avoid looking across 133 percent of FPL, where there is a discontinuity in the PTC. However we show robustness to alternative bandwidths, including the asymptotically optimal one.

For the regression kink estimates, we limit the sample to people with income between 200 and 250 percent of FPL when examining the 2015 kink point, and between 300 and 390 percent of FPL in 2016. This is the widest possible range that avoids including any other policy nonlinearities in our estimation window.[12] These windows are wider than the asymptotically optimal bandwidth, which is about 8 FPL points. We chose this wide range after conducting a Monte Carlo study to examine the performance of several different candidate estimators, following Card et al. (2017). We simulated a data generating process that closely resembles our actual data, but imposed a true kink corresponding to a semi-elasticity of 0.5, similar to what the literature estimates. In each of 1000 iterations, we implemented several estimators. We found that the simple piecewise linear specification using the full range of data performed better than other procedures, with 25 percent lower mean squared error, but slightly worse coverage rates. See Appendix B for more details.[13]

Translating our reduced form coverage kinks and discontinuities into penalty sensitivities requires that we scale by the first stage kink or discontinuity in the penalty amount. A natural way to do so would be to scale by the statutory kink or discontinuity, e.g. $325/12 in $20/12 in 2015. We will show, however, that the statutory changes are far larger than the actual changes in penalties paid. Our preferred estimates, therefore, scale by the discontinuities and kinks in the observed penalty. We obtain these observed "first stage" discontinuities and kinks by estimating Equation 2 with the observed penalty amount paid per month of uninsurance as the dependent variable. We note that can only estimate this regression for people with at least one month of uninsurance. For consistency with past research, we report semi-elasticities, the percent increase in insurance coverage caused by a $1 per month increase in the penalty.

2.2 Identifying assumption and tests of validity

We would like to interpret any estimated coverage kink or discontinuity in Equation 2 as caused by the kink or discontinuity in the mandate penalty. Our key identifying assumption is that in the absence of a kink or discontinuity in the mandate penalty, the

[12] We cut the sample off at 390 percent of FPL, rather than 400 percent, because Heim et al. (2017) document bunching in the income distribution at 400 percent of FPL, and we want to exclude the (highly selected) bunchers from our estimation sample.

[13] Our specification choices are generally similar to those used in the empirical RKD literature, such as Gelber et al. (2017); Landais (2015).

relationship between income and insurance coverage would be smooth. This assumption can fail for several reasons. First, other policy nonlinearities near the mandate kink could create a kink in insurance coverage. Second, people may manipulate their income to reduce their mandate penalty. Third, the identification assumption also requires that there is no coincidental kink in the relationship between income and insurance coverage near the mandate kink point.

We attempt to address each of these concerns. We chose our estimation sample to avoid looking across policy kinks, as discussed in Section 1. Below, we test for income manipulation by looking at the income distribution around the mandate kink point (McCrary, 2007), and we test a necessary condition for smoothness of the income-insurance relationship by examining the smoothness of correlates of insurance coverage around the mandate kink point.

2.3 Forecast errors and penalty sensitivity

Our baseline approach for translating the estimated coverage kink or discontinuity into a penalty sensitivity parameter makes an implicit assumption on people's information sets. Specifically we assume that people fully anticipate their mandate penalty at the time they sign up for insurance. To see why this assumption could fail, note that insurance coverage is determined before income or the mandate penalty is fully realized. For example, the open enrollment period in the Health Insurance Marketplaces ran from November 15, 2014 until February 15, 2015, but people may not have learned their penalty until they paid their taxes in early 2016.[14] This issue is likely more acute for the kink design than for the discontinuity design, because the kink point moves from year to year (making it harder to anticipate its exact location) whereas the discontinuity is relatively stable.

Perfect foresight is appealing because it leads to a straightforward specification, but it is a strong assumption. Several considerations suggest that at least some people could anticipate their 2015 mandate penalty in time for their 2015 coverage. First, income is

[14]See https://www.kff.org/health-reform/issue-brief/explaining-the-2015-open-enrollment-period/.

fairly (but not perfectly) stable from year to year. For example, among people whose 2014 income is between 200 and 250 percent of FPL (and who otherwise meet our inclusion criteria), about a third saw their income change by less than 10 percent. Second, at least some tax preparers and tax software systems notified filers in 2014 of their projected 2015 penalty amounts. Finally, although income uncertainty is resolved throughout the year, people also have opportunities throughout the year to respond to the mandate—for example, they can choose to drop Marketplace coverage. Dickstein et al. (2018) document considerable mid-year dropout in California's Marketplace, and argue that this reflects strategic behavior. Nonetheless, it is likely that some people do not perfectly forecast their income or mandate penalty. In that case, our estimates likely understate their true responsiveness, as such forecast errors act as classical measurement error (Dickstein and Morales, 2018). Thus our baseline estimates likely represent a lower bound on the response to the mandate penalty. We return to the important issue of imperfect income expectations in Section 5 below, where we present evidence that some people likely make forecast errors, along with a simple framework to adjust our estimates of the penalty's effect to account for these people's responses.

2.4 Interpretation of the estimand

We emphasize two caveats governing the interpretation of our estimate. First, our estimate is a local average marginal effect of the mandate penalty, specific to people with income near the mandate kink point or discontinuity. The estimated coverage kinks reflect the average marginal effect of a small increase in the mandate penalty. The effect of a small change in the mandate penalty could be different for people at different income levels, and the (per-dollar) effect of fully repealing the mandate could differ from the effect of a small change, for example because of a "taste for compliance" that is independent of the penalty amount (Saltzman et al., 2015). However the coverage discontinuities may include a taste for compliance (if any), as they reflect responses to the first dollar of penalty.

Second, our discontinuity sample is limited to people in non-expansion states, and our kink sample is limited to people without signs of an ESI offer. ESI offers may themselves increase because of the mandate, however, as it raises the compensating differential of ESI (Kolstad and Kowalski, 2016; Aizawa and Fang, 2018). This possibility is not a threat to the internal validity of our estimator if the probability of having a job with an ESI offer is not kinked at the mandate kink point. We test for and fail to find a significant kink in the ESI offer rate.[15] This does not imply, however, that the mandate does not affect ESI offers in equilibrium. It is likely that firms cannot tailor ESI offers to individual employees, and therefore we might not expect to find kinks in the offer rate even if equilibrium ESI provision responds to the penalty. However, our sample selection approach does change the interpretation of our estimates. They reflect the effect of the mandate penalty on insurance coverage, conditional on not receiving an ESI offer. If the penalty increases ESI offers, then our estimates understate the total effect of the individual mandate.

3 Data

For our data, we extract information from the full population of U.S. individual income tax returns in 2015 and 2016, maintained by the Internal Revenue Service (IRS). We supplement this with additional information from prior years and other data sets as described below.[16]

3.1 Variable definitions

Coverage measures: Since 2015, insurers have filed Form 1095, which reports monthly coverage. Marketplace insurers file Form 1095-A, self-insured employers file Form 1095-C, and all other health insurance providers file Form 1095-B. Each form 1095 reports all months of coverage, and also indicates type of coverage. Form 1095-A always represents Marketplace coverage, and Form 1095-C always represents large self-insured employer coverage. Form 1095-B contains insurer identifier codes from which we infer other cov-

[15]See Appendix F, where we also investigate coverage responses in the sample of people with signs of an ESI offer. We find ambiguous evidence consistent with positive, negative, or zero response.

[16]We do not use 2014 because third-party verification did not begin until tax year 2015.

erage types: small employer plans, public plans (Medicaid, Medicare, or VA), and off-Exchange plans. As a third-party information return, Form 1095 likely accurately measures coverage.

We also observe self-reported full year all family health insurance coverage on Form 1040, the main tax return. Filers may check a box to indicate whole year coverage. If they check this box, they might be able to avoid the mandate penalty. This self-reported information is more likely to be misreported. However, Form 1040 can indicate legitimate coverage even without verification from Form 1095, because a person can have coverage that does not meet minimum essential coverage standards (in which case their insurer will not file Form 1095), and because insurers may file Form 1095 late or not at all.

We define months of "any coverage", equal to 12 if the 1040 full-year coverage box is checked, and equal to the number of months with coverage reported on Form 1095 otherwise. We define "verified coverage" as months with 1095-reported coverage. We disaggregate coverage into different types using Form 1095. These coverage types include: Medicaid, Veteran's Affairs (VA), non-VA ESI (including military/Tricare), Exchange, and off-Exchange. Our definition of "Medicaid" coverage pools together all non-Medicare, non-VA governmental insurance plans. Lurie and Pearce (2018) provide more detail on these data, and show their validity. The tax-based coverage measures are quite similar to survey-based measures such as those in the ACS, CPS, or MEPS, although there are some differences. In general the "verified" measure yields a slightly lower coverage rate than survey-based measures, while the "any" measure yields a slightly higher rate. The tax based rate is higher for public and lower for ESI than are survey based measures.

We focus on months of coverage rather than other measures such as an indicator for take-up, because the penalty applies to coverage months, and because months of coverage provides a simple summary of all margins of response, capturing, for example, mid-year take up or drop out. A further advantage of focusing on months of coverage is that all-source coverage is (nearly) additive in coverage type. For example, a person with full-year coverage who switches from Exchange to ESI upon finding a job that offers coverage will have 12 months of any coverage, 6 months of Exchange coverage, and 6 months of ESI coverage. Finally, other recent work on mandates and subsidies has looked at person-months of coverage as the main outcome (e.g. Hackmann et al. (2015); Finkelstein et al. (2019)). However in robustness checks we also look at the extensive margin of take up, showing results for the probability of having at least one month of coverage of a given type.

Mandate penalty data: Individuals who report that they were not covered all year are instructed to fill out a worksheet to determine their mandate penalty. We observe the amount paid, and we define the dummy variable "paid penalty" as an indicator for whether a positive mandate penalty amount was reported on Form 1040.[17] We also calculate the penalty paid per uninsured month. We calculate this at the household level (and report the penalty paid per uninsured person-month) because the penalty is applied at the household level.[18]

Statutory penalty amount: At times we report the statutory penalty per uninsured month. To calculate this, we first use Equation 1 at the household level to obtain the penalty the household would pay if all its members were uninsured for the whole year. We then calculate a statutory penalty per uninsured month by dividing by 12 times the number of people in the household

Income data: Our income measure is modified adjusted gross income (MAGI) as it is calculated for determining the mandate penalty that a person would pay if she were uninsured. To obtain our income measure, we start with adjusted gross income and add tax exempt interest, foreign earned income, and housing income (all reported on Form 1040) to generate reported MAGI.

Covariates: We observe age and gender in the Death Master File from the Social Security Administration, and we merge these fields into our data. In some specifications, we split the sample based on prior income, which we define as average taxable income in the prior three years. We also split the sample based on health status. Although our administrative data lacks direct measures of health, we have one proxy measure: an indicator

[17]Our data are pre-audit and it is possible the IRS would send a notification to people whose mandate payment is suspiciously low (including zero payments).

[18]Because there is no penalty for spells of uninsurance lasting fewer than three months, we also calculated a penalty-per-month measure that does not count short spells in the denominator. However in practice this variable is nearly identical to the penalty-per-month that does count short spells, so for simplicity we count all spells.

for any itemized medical expenses in the past three years. This likely correlated with future health care costs. As a measure of the cost of coverage, we impute the benchmark Exchange premiums using tax return zip codes and premium data from the HIXCompare database.[19]

Signs of ESI offers: We do not observe ESI offers directly, because small employers are not required to report ESI offers to the IRS. Instead, we identify people who have signs of an ESI offer, meaning an employment situation in which an ESI offer is likely. The first sign of an ESI offer is employment in a large employer. Among people working for a large employer in our data, 80 percent had ESI coverage in 2015 and 96 percent in 2016. Second, among people not working for large employers, we observe several signs of ESI offers. Specifically, for each employer, we construct flags indicating whether any of the W2s for its employees contain HSA contributions, employee retirement plan indicators, or health insurance premium payments. We then say that any person with a W2 from a flagged employer has a sign of an ESI offer. [20]

3.2 Sample selection

We construct our samples somewhat differently for the discontinuity and kink analyses, because different populations face the discontinuity and kink. For the discontinuity sample, we take all individuals (defined by taxpayer identification numbers and dependent codes) in 2015 or 2016 with return-level income between 110 and 160 percent of FPL. In nearly all specifications, we limit the sample to our preferred bandwidth of 5 FPL points (i.e. 133 to 143 percent of FPL), and we focus on people living in states that did not expand Medicaid (and so face a discontinuity); in placebo tests we look at expansion states.[21] We exclude people older than 65 as nearly all of them will have coverage through

[19] Available at `hixcompare.org`. We lack premium data for about two percent of tax returns with missing zip code information.

[20] Among people with W2s from flagged employers and not 1095-Cs, about 50 percent have ESI coverage, while people without such W2s or 1095-Cs, the 25 percent have ESI.

[21] The Medicaid expansion states in 2015 are Arkansas, Arizona, California, Colorado, Connecticut, Washington (D.C.), Delaware, Hawaii, Iowa, Illinois, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, North Dakota, New Hampshire, New Jersey, New Mexico, Nevada, New York, Ohio, Oregon, Pennsylvania, Rhode Island, Vermont, Washington, and West Virginia. The 2016 expansion states are all the 2015 expansion states, plus Alaska, Indiana, and Montana.

Medicare. The final discontinuity sample consists of roughly 2.5 million observations in each year (in the 5 FPL-bandwidth range).

For the kink analysis, we focus on tax-returns with single filing status, filed by people with no dependents, meaning no children were claimed on the tax return. We study single-person tax returns because the individual mandate creates complicated, possibly non-salient coverage incentives for larger households. (We explain the issue in Appendix A.) We next limit the sample to individuals age 27-64, dropping people whose age makes them eligible to obtain coverage through their parents or Medicare. We limit the sample to people with income in our estimation windows, i.e. between 200 and 250 percent of FPL in 2015 and 300 to 390 percent of FPL in 2016. We further restrict our kink sample to people without signs of ESI offers. We impose this restriction because we are concerned about power for this sample; the penalty kink is not very large, making it difficult to detect coverage responses. People with ESI offers have access to highly subsidized insurance and had high coverage rates even prior to the ACA (see, e.g, Kaiser Family Foundation and Health Research and Educational Trust (2012)), meaning the mandate likely had a small and difficult-to-detect effect on their coverage decision. Restricting the sample in this way of course means that we potentially miss part of the effect of the penalty if it affects coverage among people with ESI offers or if it affects the probability of receiving an offer. We show in Appendix F, however, that there is no response on either of these margins. After these exclusions, our kink sample consists of 653,891 observations in 2015 and 656,056 observations in 2016.[22]

Table 1 reports summary statistics for the analysis samples. On average people had about 10 months of coverage in the discontinuity sample and 8 months in the kink sample. In the discontinuity sample, ESI and Medicaid are the modal coverage sources; we have a high rate of Medicaid coverage, even in non-expansion states, because our sample includes many low-income children. In the kink sample, exchange coverage and ESI coverage are both fairly common.[23] Medicaid is also a non-trivial coverage source in the kink

[22]The kink sample is smaller than the discontinuity sample for several reasons. The income distribution thins quickly at higher incomes; the discontinuity sample includes non-single filers and children; and the discontinuity sample includes people with signs of ESI offers.

[23]Although we exclude people with signs of ESI offers in our kink sample, we end up with some ESI coverage because, while no one in our sample has ESI in their own name, people can obtain coverage

sample. This is surprising because most states extended Medicaid eligibility for childless adults only up to 138 percent of FPL. The main explanation for the non-trivial Medicaid coverage in our sample is that Medicaid eligibility is assessed on a continuous basis, but we observe income annually. A person with a spell of unemployment in 2014 or 2015, for example, could be temporarily eligible for Medicaid, and remain on its rolls even after her income recovers to above 200 percent of FPL.

Twenty-one percent of our kink sample paid the mandate penalty in 2015 and 13 percent in 2016. Payment rates are lower in the discontinuity sample because many people were exempt. A lower fraction paid the penalty in the discontinuity sample, despite similar coverage rates, because about half our discontinuity sample has income below 138 percent of FPL and is exempt from the penalty. The average penalty paid per uninsured month is low even in the kink sample–about \$15 in 2015 and \$26 in 2016, relative to statutory amounts of \$27 and \$58. From this we conclude that, although many people end up paying the mandate penalty and it therefore likely has real coverage incentives, these incentives may be small relative to the statutory amounts.

4 Results

4.1 Smoothness tests

We begin by examining the validity of our RD and RK designs, showing smoothness of the income distribution and covariates around the mandate discontinuity and kink points. Figure 2 shows the income distribution for each sample; following McCrary (2007) we show under-smoothed densities. One concern with the design would be that the discontinuity and kink create incentives to bunch below or at the thresholds. There is obviously no such bunching (although the kink sample does show round number bunching), as the income distribution is flat or smoothly decreasing just below the threshold, whereas with bunching (or income manipulation more generally) we would expect a steep increase in the density just below thethreshold. More generally there is no

through others, such as domestic partners.

obvious kink or discontinuity in any of the samples. Close inspection, however, reveals perhaps a downward dip in the income distribution in the discontinuity samples. This discontinuity is about -500 (or about 2 percent of the height below the threshold) in each year, and it is statistically significant.

This significant discontinuity in the density might indicate that people are strategically failing to file their taxes (and therefore avoiding paying the penalty). Such strategic behavior would bias our results. However we think this behavior is unlikely because tax payers with income in this range have very strong incentives to file their taxes, even if they end up paying the full penalty. These incentives arise from two redundable tax credits: the earned income tax credit and the child tax credit. A household at 138 percent of FPL with two adults and two children, for example, would be eligible for $2000 in the child tax credit and nearly as much in the earned income credit. To claim these credits, however, the household would have to file a return. Even if the household were fully uninsured, its mandate penalty would only be $2,085, or about half of the tax credits it would receive by filing. We conclude that strategic nonfiling is likely to be unimportant for our results.[24]

Next we turn to showing smoothness of pre-determined covariates. Figure 3 shows the binned scatter plots for several these variables in the discontinuity and kink samples. They are clearly correlated with income, but again there is no obvious kink or discontinuity at the mandate kink point, and the estimated kinks are statistically insignificant. We report smoothness tests for these outcomes in Table 2. We estimate small and, usually, insignificant discontinuities and kinks across these outcomes.[25] We conclude that there is no evidence of income manipulation, and the predictors of insurance demand are smoothly distributed around the mandate kink point.

[24] In Appendix D we provide further evidence that strategic nonfiling is unlikely to be important for our results.

[25] An exception is for prior income in the 2015 discontinuity sample. The explanation for the very large discontinuity is the presence of an extreme outlier, as can be seen in the binned RD plot. The age discontinuity is also significant in the 2015 RD sample, although it is small (roughly one month). As we show in Appendix Tables C.1-C.4 that are results are highly robust to controlling for covariates, this small difference is not enough to account for our discontinuity estimates.

4.2 First stage

Next we show that empirically there are discontinuities and kinks in the strength of the penalty. We plot the observed penalty amount paid per month of uninsurance in Figure 4. The figure reveals several important facts. First, many people pay the penalty: the average penalty amount paid per uninsured month ranges from \$9 to \$27 depending on the sample. Second, there are visually obvious discontinuities in penalty paid in the discontinuity samples, and visually obvious kinks in the kink samples. The second important fact is that the estimated kinks and discontinuities are far below their statutory levels. We would expect statutory discontinuities of \$21 in 2015 and \$45 in 2016, and statutory kinks of \$1.67 and \$2.08. The estimated discontinuities and kinks are less than half as large. These differ from statutory amounts both because people can obtain coverage exemptions, and because some people declined to pay a penalty amount even though they were uninsured and claimed no exemptions. Thus, although exemptions and limited enforcement weaken the bite of the mandate, empirically we do see a sharp increase in penalties around the discontinuities and kinks.

4.3 Coverage effects of the mandate penalty

We now turn to our main results, the estimated discontinuities and kinks in insurance coverage around the mandate discontinuity and kink points. Figure 5 and Figure 6 show months of insurance coverage, by type and year, as a function of income, for the discontinuity and kink samples.[26] We report estimated kinks and discontinuities in Table 3.

The top row of Figure 5 reveals visually clear discontinuities in months of any coverage at the mandate kink points. The discontinuity is about 0.1 months in 2015 and doubles to 0.2 months in 2016. This is a fairly small response—facing the penalty results in only an extra 3 to 6 days of insurance coverage. However, these low responses are

[26] In these figures we plot the four coverage types where we expect the largest responses—any coverage, verified coverage, Exchange, and ESI (for the RD sample) or Medicaid (for the RK sample). The remaining coverage types are in Appendix Figure C.1 (for the RD plots) and Appendix Figure C.2, for the RK plots. We present different outcomes for the RD and RK sample because the RD sample excludes Medicaid expansion states but includes people with signs of ESI offers, whereas the opposite is true for the RK sample.

due in part to the fact that there is a fairly small discontinuity in the actual penalty paid. Scaling these responses by the first stage discontinuities, we obtain semi-elasticities of just over 0.2 in both years. These estimates are similar to but on the low end of those reported in the literature, which generally looks at the semi-elasticity of individual market or Exchange coverage, rather than all-source coverage, which may be less elastic. The estimates of Hackmann et al. (2015) imply a semi-elasticity of about 0.2 in the context of the Massachusetts mandate. Tebaldi (2017) estimates semi-elasticities in California of 0.2 to 0.5 for subsidized households. Jaffe and Shepard (2017) estimate a semi elasticity of about 1, looking at a lower income population in Massachusetts. Finkelstein et al. (2019) also find a semi-elasticity of individual market coverage of about 1, identified off of discontinuities in PTC generosity in Massachusetts. Overall we find statistically significant coverage responses to the individual mandate, but low rates of paying the penalty end up dampening the elasticity of coverage with respect to the statutory amount.

In the next column, we look at verified rather than self-reported coverage. We estimate smaller responses of third party verified coverage, about 80 percent as large in 2015 and half as large in 2016. This suggests that part of the response to the mandate may entail strategic misreporting (although, given the lower baseline rate of verified coverage, we do see a similar semi-elasticity). Consistent with the misreporting story, we see a very sharp discontinuity in any coverage, which can be reported after income is realized, but a less sharp discontinuity in verified coverage. The discontinuities in verified coverage come primarily from ESI coverage, with further responses in Medicaid coverage in 2015 and Exchange coverage in 2016. The Exchange coverage is small on a months-of-coverage basis, but because the base rate of exchange coverage is very low, we again obtain semi-elasticities on the order of 0.2. To illustrate the quantitative implications of our results, we consider the implied effect of setting the penalty to zero for our 2016 discontinuity sample. Here we extrapolate to our full discontinuity sample, people with income between 110 and 160 percent of FPL. The estimates imply that any coverage would fall by about 257,000 person-years (2.6 percent of baseline), verified coverage by 143,000 person years (1.7 percent), and Exchange coverage by 34,000 person-years (3.5 percent).

We turn now to the kink estimates. In 2015, there is a clear upward kink in months of overall insurance coverage at the mandate kink point. The estimated kink, reported in Panel C of Table 3, is a statistically significant 0.051, meaning that for each extra $1000 of income above the mandate kink point, coverage rises by 0.05 months, relative to the trend below. This estimate implies that each $1 of mandate penalty per month raises insurance coverage by about 0.9 percent. About 60 percent of the response comes from verified coverage, which has a marginally significant kink of 0.029 (p=0.09), for a semi-elasticity of 0.6. At the 2016 kink point, there is a smaller and less clear response to the mandate penalty, reported in Panel D of Table 3. We estimate a statistically insignificant kink of 0.020 in overall coverage, corresponding to a semi-elasticity of 0.14. The kink in third-party verified coverage in 2016 is essentially zero, with a large standard error. Although the statutory semi-elasticities are fairly low, the observed semi-elasticities are comparable to past estimates.

We note that these kink estimates are for people without signs of ESI offers; we present results for the complementary sample with signs of ESI offers in Appendix F. The main finding is that the estimates are not always robust. In 2015, the kink is sensitive to bandwidth and, at the Calonico et al. (2014) optimal bandwidth, it is positive and insignificant. In 2016, however, we find estimate a negative and significant kink for the offer sample. Incorporating these responses would imply very small or even negative overall responses in 2016, consistent with our overall findings of a low responsiveness. It is possible that this negative kink in the 2016 ESI offer sample reflects a negative bias due to concavity in the relationship between coverage and income. In that case we might expect that our 2016 kink estimates are also biased downward (and this could explain why we find a smaller kink in 2016 than in 2015).

The overall responses in the kink sample appear to be largely driven by Medicaid, with little response in Exchange coverage or ESI. Indeed, the Exchange kink is negative and insignificant in 2015 and only marginally significant in 2016. This is surprising because the kink sample consists of relatively high income people (200-400 percent of FPL) without offers of ESI; we might have expected this group, if anything, to respond to a larger penalty by buying more insurance coverage. Strikingly, we find in the 2015 kink sample that Medicaid coverage responds sharply to the mandate penalty, accounting for

45 percent of the overall response. There is also a Medicaid response in the discontinuity sample in 2015. But these responses are surprising; the discontinuity sample is limited states that did not expand Medicaid, and the kink sample is limited to income ranges that would disqualify adults from Medicaid, even in expansion states. However the Medicaid responses are easily explained. The key is that in non-expansion states, children are eligible for Medicaid even at high income levels, and in expansion states, eligibility for Medicaid depends not on taxable income (which we use to define our sample) but on recent income, averaged over the last few months. A person with a high annual income but a temporarily low income (because of job loss, for example) could be eligible for Medicaid but still in our sample. We provide more evidence on this point in Appendix G. This may also provide an explanation for the lower responses to the kink in 2016; at the higher kink point, it is less likely that people would be temporarily eligible for Medicaid coverage.

4.4 Heterogeneous responses by observable characteristics

A key goal of the individual mandate is not only to increase coverage, but especially to encourage healthy people to obtain coverage, to ease adverse selection. To look for such heterogeneous responses, we re-estimate our models, stratifying on covariates related to health and spending: sex, age, prior itemized medical expenses, and prior income. o maximize power, we pool 2015 and 2016, and focus on the RD models. The results are in Table 4.

Men are more responsive than women, with a discontinuity that is about a third larger and a semi-elasticity about 20 percent larger. Children (defined here as people younger than 27 and therefore eligible for their parents insurance) are more responsive than adults, but younger adults (below median adult age) and older adults (above median age) are similarly responsive. Age and sex are indirect proxies for health care demand; a more direct measure is prior itemizing of medical expenses. People with prior itemizing have higher coverage and a smaller discontinuity, but the difference is insignificant. We may lack power to detect small differences as we observe relatively few prior itemizers in this low income sample. Overall, although these differences are not dramatic, they point towards the conclusion that groups with lower expected health care expenses respond most to the mandate penalty.

4.5 Further validity and robustness tests

Placebo test based on expansion states: To show the validity of our RD design, we begin with a placebo test looking at expansion states. In these states, there is no discontinuity in the mandate penalty at 138 percent of FPL, so if our model is well-specified, we should also see no coverage discontinuity.[27] We present RD plots for the placebo sample in Figure 7. The top panels show the penalty paid per uninsured month. There is a discontinuity of $0.35 in 2015 and $0.18 in 2016, about a twentieth of the penalty discontinuity in the non-expansion states. Consistent with this small first stage, we find in the remaining panels very small discontinuities in coverage: about 0.01 months for both any coverage and verified coverage, in 2015 and in 2016. This discontinuity is statistically insignificant, often wrong-signed, and again about a twentieth of the estimate in the expansion states. Overall therefore this placebo tests shows no effect in the expansion states.

Permutation test: A key concern with the RKD is the possibility of finding spurious kinks, simply because of curvature in the relationship between income and insurance coverage (Ganong and Jger, 2018). We assess this concern in Appendix E by re-estimating our RKD models, but varying the kink point across a fine grid of placebo locations. If the kink is spurious, then we expect that our estimate is unexceptional in the distribution of placebo estimates. We find that our kink for any coverage in 2015 is larger than 92.2 percent of placebo estimates, and the kink for verified coverage in 2015 is larger than 90.2 percent of the placebo estimates.

Alternative bandwidths: We consider robustness to bandwidth and to alternative specification choices. Figure C.3 shows the estimated discontinuity and its 95% confidence interval, as a function of the bandwidth. We indicate the optimal bandwidth (com-

[27] Of course, in these states, eligibility for Medicaid for childless adults ends at 138 percent of FPL, so we might expect to see a downward discontinuity. However, as we emphasize elsewhere, Medicaid eligibility is not determined by annual taxable income; it is determined by recent income as reported to state Medicaid authorities. We do not expect to see a downward discontinuity in Medicaid coverage at this threshold, and indeed we do not find one.

puted using the procedure of Calonico et al. (2014)) with a vertical line. In general the estimates are not too sensitive to the bandwidth, although with a very small bandwidth we would find insignificant estimates in 2015.

We plot the analogous figure for the 2015 kink estimate in Figure C.4. The point estimate is stable over a wide range of bandwidths. At small enough bandwidths, the point estimate fluctuates and its confidence interval becomes quite large. The mean-squared error optimal bandwidth of Calonico et al. (2014) is about $900. At this bandwidth, the point estimate is 0.17, thrice as large as our main estimate, but the confidence interval is larger still, and so the estimate is marginally significant ($p = 0.08$). Looking across the different coverage types, the point estimates are fairly stable until the bandwidth becomes small, at which point the estimates become less stable and much less precise.[28]

Alternative specifications: We consider robustness to alternative specification choices: allowing for nonlinearities in income (i.e. a quadratic or cubic), controlling for demographics (female dummy and a quadratic in age, plus filing status dummies and dummies for number of exemptions in the RD sample), imposing continuity at the kink point, and excluding people with ESI in the RK sample. The results are in Appendix Tables C.1 to C.4. The RD estimates are generally robust to alternative functional forms. In 2015, the point estimates are similar for the linear and quadratic specifications, but fall a bit with the cubic. In 2016 however the cubic specification yields slightly higher point estimates. The RD results are also unchanged when we include controls.

For the kink sample, the nonlinear income terms generally produce larger estimates, and sometimes substantially larger ones. For example the kink in any coverage in 2015 increases by 60 percent when we include a quadratic term, and it more than doubles when we include a cubic. The standard errors also rise, consistent with the findings from our Monte Carlo analysis that these nonlinear terms substantially increase the MSE of our estimates. We focus on the linear specifications because of these large standard errors, although we find it reassuring that allowing for higher order terms would, if anything, strengthen the conclusion that the mandate penalty increases coverage. In general the

[28]We do not plot the 2016 kink estimates against the bandwidth because that figure only confirms that the 2016 estimates are not statistically significant.

results are not sensitive to other choices of controls. They change little when we impose continuity or add demographic controls. The estimates typically rise when we drop people with ESI coverage.

Extensive margin responses: As a final robustness check, we show that our results are not sensitive to our focus on months of coverage. We re-estimate our models but looking at the extensive margin—the probability of having at least one month of coverage of a given type. Generally we find very similar results. In RD samples, the estimated semi-elasticities are about 0.17 (for any coverage) and 0.13 (for verified coverage), which are similar to but slightly smaller than the estimates using months of coverage. The regression kink estimates show a similar pattern; the semi-elasticities are roughly similar in magnitude whether we examine months of coverage or the extensive margin of any coverage.

5 Imperfect expectations and knowledge

Our results so far show a modest semi-elasticity of coverage with respect to the mandate penalty, on the lower end of the literature's estimates. We interpret this modest semi-elasticity as a low behavioral response to the incentives created by the mandate. In this section, we consider two alternative explanations: forecast errors arising from the difficulty of predicting income, and ignorance of the existence of the mandate penalty. For this section, we focus on the RD sample, where our greater sample size and larger first stage give us much greater power to investigate alternative hypotheses. We find evidence that at least some people fail to perfectly predict their penalty, and accounting for these imperfect expectations could increase our estimated penalty sensitivity by as much as fifty percent. As our baseline sensitivity is fairly low, however, this still implies a fairly small response to the mandate.

5.1 Accounting for imperfect income expectations

An important alternative explanation for our low semi-elasticity is that many people cannot perfectly predict their income, and this attenuates our estimated discontinuity, even if the true response to the penalty is large. This alternative explanation is important because it would suggest a larger coverage response to the mandate's repeal, for example, than what would be implied by our estimates.

Using lagged income as the running variable: If people do not know their taxable income at the time they make their insurance penalty, they may use their prior year income in their forecasts. This suggests that a better running variable might be lagged income, since this is largely known around the time that coverage decisions are made.

We re-estimate our RK and RD models, but using lagged income as the running variables. We present the results in Figure 8. The figure shows two important facts. First, there is no "first stage:" the actual penalty paid per uninsured month in 2016, for example, is continuous in 2015 income. Second, however, there are visually clear but fairly small coverage responses to the statutory thresholds. We estimate a discontinuity of 0.02 months of any coverage in 2015, and 0.06 months of any coverage in 2016. These discontinuities are small relative to our main estimates—only 20 to 30 percent as large. This small response might indicate that few people use lagged income to predict the penalty. However, the proper comparison is not the reduced form discontinuities but the discontinuity per dollar of penalty. Indeed, the discontinuities here are very large given that we do not observe any discontinuity in penalty paid.

How are we to interpret the clear but small coverage discontinuity along with the zero first stage discontinuity? Together they likely indicates that some people misperceive the mandate penalty or miss forecast their income. Specifically, if some people use their past income to predict their mandate penalty, then we would expect to see a coverage discontinuity but not a penalty discontinuity only if people's expectations (or forecast process) are not rational (in the sense of agreeing with the realized data on average). Clearly, however, not everyone miss forecasts their income in this way, or else we would not observe a discontinuity with respect to current year income.

Accounting for forecast errors The fact that some people miss-forecast their income implies that our semi-elasticities may be too small. To get some quantitative indication of how large the bias could be, we introduce a simple model. Consistent with the evidence so far, we assume that a fraction α of the population correctly forecasts their penalty and responds with a sensitivity β_1. This sensitivity is the coverage response per dollar of penalty. For this group the first stage penalty discontinuity is Δ, which we can estimate. The remaining $(1-\alpha)$ uses their prior year income to forecast the penalty. Their penalty sensitivity is β_2, and we assume they perceive a penalty discontinuity of $\tilde{\Delta}$ with respect to past income. We cannot estimate $\tilde{\Delta}$ because we do not observe expected penalties. Only the first group contributes to our estimated discontinuity when we use current income as the running variable because there is no first stage for the second group. Our estimated penalty sensitivity is therefore

$$\beta^{est} = \alpha\beta_1$$

but the population penalty sensitivity is

$$\beta = \alpha\beta_1 + (1-\alpha)\beta_2.$$

Our estimates of the population penalty are attenuated by $(1-\alpha)\beta_2$. We can gauge the size of this bias with additional assumptions. To see how, note the discontinuity in coverage with respect to lagged income is

$$(1-\alpha)\beta_2\tilde{\Delta}.$$

We observe this discontinuity—it is about 0.06 months of any coverage. If we knew $\tilde{\Delta}$, we could back out $(1-\alpha)\beta_2$, the bias in our main estimate. We cannot estimate $\tilde{\Delta}$ because we lack data on the perceived penalty. However we can obtain a reasonable range on $(1-\alpha)\beta_2$ by considering a reasonable range of values for $\tilde{\Delta}$. Specifically, the largest plausible value of $\tilde{\Delta}$ is the maximal statutory penalty, \$695/12 in 2016 for a single adult. This is an upper bound because it is much larger than the average actual penalty discontinuity paid (and some people below the threshold pay a positive amount), but it is possible that people overestimate the penalty they face. A reasonable lower bound is the actual empirical discontinuity in 2015, or about \$5. This would be the discontinuity in experienced penalty paid. This is a lower bound because it requires people to not realize that the penalty amount rises from 2015 to 2016. These alternative values for $\tilde{\Delta}$ give a range for the bias of about 0.001 to 0.012. The observed penalty sensitivity is about 0.022, so this implies a bias of 5 to 55 percent. Although this is a fairly wide range in percent terms, our baseline penalty sensitivity is fairly low, so even at the upper end of the range, there is not a hugely different coverage effect of repeal. For example, scaling up the all-source coverage losses by 55 percent would imply a coverage decline of about 400,000 covered lives, rather than 257,000—certainly a larger change, but still not an enormous one.[29]

Rational but imperfect expectations: Our model of accounting for imperfect expectations allows for only two possibilities: perfect expectations for some, and non-rational expectations for others. It is possible that some people have imperfect but rational expectations, but these people would not contribute to the discontinuity when we use lagged income as the running variable, because they would not expect a first stage discontinuity. However, we do not find evidence that some people have imperfect but rational expectations. We test for such people by stratifying our sample according to measures of income uncertainty: the number of W2s and the presence of unemployment insurance income. The idea here is that income is much easier to predict for people with a single job (i.e. a single W2) and no job loss (i.e. no unemployment insurance income). We begin by defining "predictable" earners, people in households with at most one W2 per earner, and no unemployment insurance income. About three quarters of the sample meets this definition. If people make imperfect but rational forecasts of the penalty, then we would expect to find much larger penalty sensitivities among predictable earners, because for them realized income, which we observe, is likely close to expected income, the unobserved but theoretically desirable running variable

We provide two pieces of evidence to suggest that our measure of predictability is valid. First, the change in income is more tightly concentrated around zero for predictable earners than for non-predictable earners, as can be seen in Appendix Figure C.5,

[29] Analogous exercises for verified and exchange coverage reveal that our estimates should be scaled by about 100 percent for verified coverage, but scaled down by 100 percent for Exchange, since we estimate a negative discontinuity in lagged income for Exchange coverage.

which shows the distribution of income growth rates for predictable and non-predictable earners. Second, income is indeed much more predictable for this group than for non-predictable earners. For example, when we regress income in year t on a degree 7 polynomial in lagged income, interacted with year and filing status dummies, we obtain an R^2 of 60 percent for predictable earners but only 37 percent for non-predictable earners.

Of course, both these facts indicate that income and family structure alone do not perfectly predict next year's income, even among predictable earners. Nonetheless we think that looking at predictable earners is useful for two reasons. First, if expectations error in the running variable are a serious source of bias, we should still see larger discontinuities for predictable earners, for whom there is less measurement error. Second, predictable earners as we define them may experience idiosyncratic increases in income, for example coming from raises or promotions. If these idiosyncratic income changes are predictable to the people in our data (given their private information), then realized income is still the appropriate running variable, even if we cannot predict these income changes given our limited information.

To test the importance of income expectations for our low semi-elasticity, we re-estimate our main RD models, separately for predictable and non-predictable earners. The results are in Table 5, and we show the RD plots in Figure 9. We see large and statistically significant differences between predictable and non-predictable earners but, surprisingly, it is the predictable earners who are less responsive to the penalty. Their semi elasticity is 50 to 100 percent lower than non-predictable earner's, depending on the specification. This low responsiveness does not arise from high baseline coverage rates; predictable earners' coverage rates are only 3 percent higher. Nor is it an artifact of self-reported coverage or access to ESI; we find similar differences across coverage categories.

These results suggest that income unpredictabiltiy does not necessarily explain our low semi-elasticity. This conclusion, however, is potentially sensitive to our (admittedly arbitrary) classification of "predictable" and "non-predictable." As an alternative approach, we divide our data up into cells defined by unemployment insurance receipt (yes/no), number of W2s (zero, one, or two or more), filing status (single, married filing jointly, head of household, other), and year. Note that our "predictable" measure is a coarsening of these cells. For each cell, we estimate the predictability of income (wit the R^2 from a regression of income on a degree 7 polynomial in its lag), and we estimate our main RD model for coverage (continuing to use a 5-FPL point bandwidth). In Appendix Figure C.6, we plot the estimated semi-elasticities against income predictability for each cell. The figure shows no apparent relationship between the semi-elasticity and income predictability, although a weak relationship would be hard to see because some of the estimates are noisy. To more precisely measure the relationship between penalty responsiveness and income predictability, we regress cell-level semi-elasticity against cell-level R^2. We obtain a constant of 0.19 and a slope of -0.59; that is, semi-elasticities and predictability are negatively correlated.[30]

These results show that people with more predictable income do not exhibit a substantially greater coverage discontinuity. Of course, an important caveat is that there may be unobserved factors correlated with income predictability that lead to low responsiveness; such factors would muddy the interpretation of these results. Nonetheless, this finding is consistent with the view that imperfect but rational forecast errors are not a major factor in explaining our low semi-elasticity estimates.

5.2 Imperfect knowledge of the penalty

Another possible explanation for our relatively low semi-elasticity is that people are simply unaware of the mandate's existence. If this ignorance is permanent, then our estimates would still reflect the long-run effect of the mandate. If, however, people have learned about the mandate in recent years, then our estimates might understate the long-run effect of penalty. To provide some evidence on the role of knowledge of the penalty, we consider whether people who paid the penalty in the past—and were therefore likely aware of the penalty—respond more to penalty in a given year. Of course, people who paid the penalty in the past could not have been insured for the whole year. So to make the comparison clean, we look only at people who were not insured for the whole year, comparing responsiveness to the mandate among those who did and did not pay the

[30] We weight by the inverse of the standard error, to down weight the cells where we have less statistical precision.

penalty in the past. Specifically, we merge in to the 2015 and 2016 data information on penalty paying in 2014, along with an indicator for self-reported whole-year coverage, i.e. for checking the coverage box on Form 1040.

Table 6 shows the RD estimates for 2015 and 2016 coverage for three groups: people who "checked the box" reporting full year coverage, people who did not check the box but did not pay a penalty, and people who did not check the box and paid the penalty. We present the RD plots in Figure 10. The results show that, among the uninsured in 2014, people who did and did not pay the penalty had fairly similar coverage rates in 2015 and 2016, but people who paid the penalty were if anything less responsive than people who did not pay the penalty; they have lower discontinuities and, especially, lower semi-elasticities.

Under the assumption that people who previously paid the penalty are fully informed of the penalty, these results do not support the view that widespread ignorance of the penalty explains our low semi-elasticities. That is, if people who paid the penalty are fully aware of it, then people who didn't pay the penalty should also be fully aware (because they are more responsive to the penalty) so ignorance of the penalty cannot be widespread. Of course, an alternative interpretation of these results is that people who paid the penalty are not particularly aware of it. We cannot rule out this possibility—it could be that people simply answer questions in their tax software, and pay the final amount. Nonetheless, the evidence here is consistent with the view that people are reasonably aware of the penalty.

6 Conclusions and implications

The individual mandate has generated considerable attention and controversy, with protests, legal challenges, and numerous repeal efforts ultimately culminating in the Tax Cuts and Jobs Act of 2017, which set the mandate penalty to zero. Despite all this attention, relatively little is known about whether the individual mandate actually causes people to obtain coverage. We provide new evidence on this question using regression discontinuity and regression kink designs and data derived from the population of tax returns.

We find visually clear and statistically significant discontinuities and kinks in coverage corresponding to the discontinuities and kinks in the mandate penalty. A variety of specification checks support a causal interpretation of our estimates. These results indicate that the penalty affects coverage decisions. Quantitatively, however, our estimates imply fairly small responses to the penalty. For example, setting the penalty to zero would reduce coverage only by 2.5-5 percent in the 2016 discontinuity sample, holding fixed premiums and noting that this estimate is local to people with income near 138 percent of FPL. Part of the reason for this small response is that the actual penalty paid is fairly low—only about $20 per uninsured month in 2016 (among people subject to the penalty), relative to a statutory amount of about $45. We caution, however, that this simple extrapolation could be inaccurate because it is out-of-sample and does not account for insurer responses.

Our findings have several further implications beyond the direct coverage effects. First, some commentators have suggested that the penalty's greatest effect is by inducing a "taste for compliance." Our evidence rules out one version of this theory. If the taste for compliance arises because the first dollar of the penalty signals that a person should obtain coverage, then we should see a larger per-dollar response to smaller penalties and to smaller penalty changes. This is not what we see: the 2015 and 2016 discontinuities generate similar per-dollar responses, but the 2016 discontinuity is twice as large. We also see similar or larger responses to the kink, which is a marginal increase in the penalty amount. Our evidence does not, however, let us rule out stronger versions of the "taste for compliance" in which the mere existence of a penalty that some people must pay affects coverage for everyone, including those with no statutory penalty.

Second, as we find heterogeneous responses to the penalty, our evidence implies that the mandate affects the composition of the insured pool. Men, young adults, and low prior-income people are all especially responsive to the penalty. People with prior itemized expenses are also less responsive, although this difference is not significant. To the extent that these groups have low health care spending, the compositional shifts induced by the mandate would result in lower costs of the insured and lower premiums. Reducing adverse selection in this way in the non-group insurance market, especially the Exchanges, was a key policy motivation for the mandate. We find, however, that the mandate induces fairly small coverage responses in the Exchanges. For example, our estimates imply that repeal would reduce Exchange coverage in the 2016 discontinuity sample by about 3.5 percent.

Finally, our results have several implications that may be of interest to policy makers, including states considering implementing their own mandate. First, our results suggest that the verifiability of coverage is an important consideration: verified coverage responses are only about half as large as unverified responses. As some unverified coverage does not represent true insurance gains, states may wish tighten verification requirements. Of course, the costs of verification must be weighted against and additional coverage it generates. Second, as much of the response to the mandate is in Medicaid (for expansion states), state mandates may result in higher than expected Medicaid spending. Last, our results highlight the importance of the effective penalty as well as the statutory penalty. The effective penalty paid is often less than half of the statutory amount, which ultimately limited the coverage effects of the federal amount. States must balance the desirability of exemptions and limited enforcement against the cost of lower coverage.

References

Aizawa, Naoki and Hanming Fang, "Equilibrium Labor Market Search and Health Insurance Reform," 2018. unpublished working paper.

Buchmueller, Thomas and John DiNardo, "Did Community Rating Induce and Adverse Selection Death Spiral? Evidence from New York, Pennsylvania and Connecticut," *American Economic Review*, 2002, *92* (1), 280–294.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econonetrica*, 2014, *82* (6), 2295–2326.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber, "Inference on Causal Effects in a Generalized Regerssion Kink Design," *Econometrica*, 2015, *83* (6), 2453–2483.

_ , _ , _ , and _ , "Regression Kink Design: Theory and Practice," in Matias D. Cattaneo and Juan Carlos Esanciano, eds., *Advancse in Econometrics*, Vol. 38, Oxford University Press, 2017.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99*, 1145–1177.

Congdon, William J., Amanda Kowalski, and Mark H. Showalter, "State Health Insurance and Regulations and the Price of High-Deductible Health Policies," *Forum for Health Economics & Policy*, 2008, *11.*

DeLeire, Thomas, Andre Chappel, Ken Finegold, and Emily Gee, "Do individuals respond to cost-sharing subsidies in their selections of marketplace health insurance plans?," *Journal of Health Economics*, 2017, *56*, 71–86.

Dickstein, Michael and Eduardo Morales, "What do exporters know?," *Quarterly Journal of Economics*, 2018, *133*, 1753–1801.

_ , Rebecca Diamond, Tim McQuade, and Petra Persson, "Take-Up, Drop-Out, and Spending in ACA Marketplaces," 2018. NBER Working Paper No. 24668.

Finkelstein, Amy, Nathaniel Hendren, and Mark Shepard, "Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts," *American Economic Review*, 2019, *109*, 1530–1567.

Frean, Molly, Jonathan Gruber, and Benjamin D. Sommers, "Premium subsidies, the mandate, and Medicaid expansion: Coverage Effects of the Affordable Care Act," *Journal of Health Economics*, 2017, *53*, 72–86.

Ganong, Peter and Simon Jger, "A Permutation Test for the Regression Kink Design," *Journal of the American Statistical Association*, 2018, *113* (522), 494–504.

Gelber, Alexander M., Timothy Moore, and Alexander Strand, "The Impat of Disabiltiy Insurance on Beneficiaries' Earnings," *American Economic Journal: Economic Policy*, 2017, *9* (3), 229–261.

Hackmann, Martin B., Jonathan T. Kolstad, and Amanda E. Kowalski, "Adverse Selection and an Individual Mandate: When Theory Meets Practice," *America Economic Review*, 2015, *105* (3), 1030–1066.

Hahn, Jinyong, Petra Todd, and Wilbert Van Der Klaauw, "Mandate-Identification and estimation of treatment effects with a regression discontinuity design," *Econometrica*, 2001, *69* (1), 201–209.

Heim, Bradley T., Gillian Hunter, Adam Isen, Ithai Z. Lurie, and Shanthi P. Ramnath, "Income Responses to the Affordable Care Act: Evidence from the Premium Tax Credit," 2017. Unpublished working paper.

Herring, Bradley and Mark V. Pauly, "The Effect of Community Rating Regulations on Premiums and Coverage on the Individual Health Insurance Market," 2006. NBER Working Paper No. 12504.

Imbens, Guido and Karthik Kalyanaraman, "Optimal bandwidth choice for the regression discontinuity estimator," *Review of Economic Studies*, 2012, *79* (3), 933–959.

Jaffe, Sonia and Mark Shepard, "Price-Linked Subsidies and Health Insurance Markups," January 2017. Unpublished working paper.

Kaiser Family Foundation and Health Research and Educational Trust, "Employer Health Benefits, 2012 Annual Survey," 2012.

Kolstad, Jonathan T. and Amanda E. Kowalski, "Mandate-based health reform and the labor market: Evidence from the Massachusetts Reform," *Journal of Health Economics*, 2016, *47*, 81–106.

Landais, Camille, "Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design," *American Economic Journal: Economic Policy*, 2015, *7* (4), 243–278.

Lo Sasso, Anthony T. and Ithai Z. Lurie, "Community rating and the market for private non-group health insurance," *Journal of Public Economics*, 2009, *93*, 264–279.

Lurie, Ithai Z. and James Pearce, "Health Insurance Coverage from Administrative Tax Data," June 2018.

_ and Janet McCubbin, "What Can Tax Data Tell Us About the Uninsured? Evidence from 2014," 2016. Office of Tax Analysis Working Paper 106.

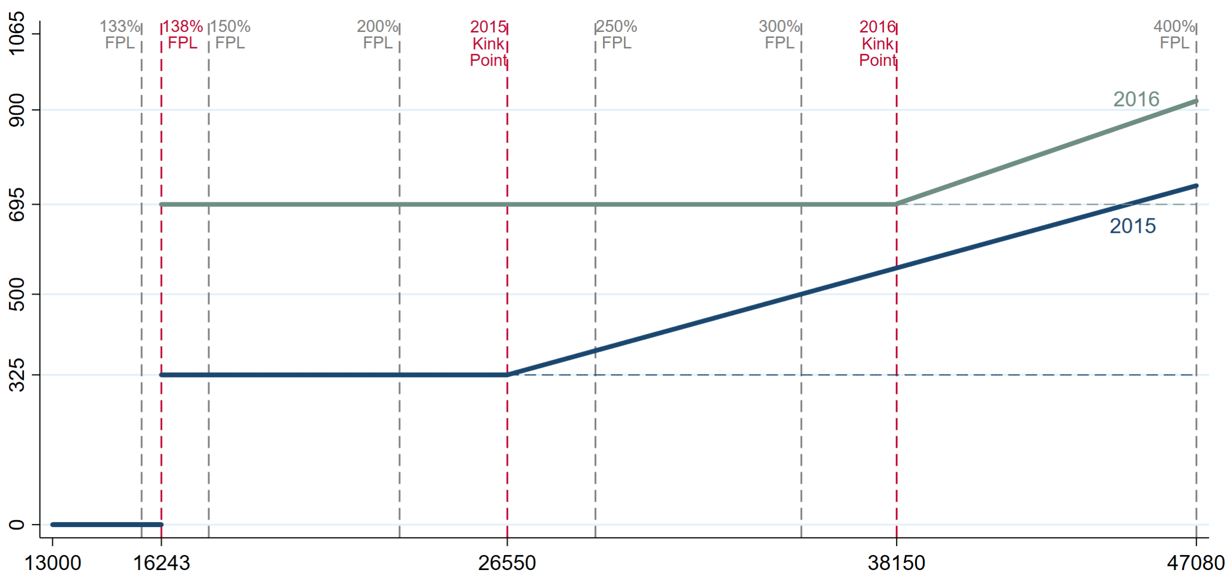
McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2007, *142*, 698–714.

Saltzman, Evan, "Demand for Health Insurance: Evidence from California and Washington ACA Exchanges," 2018. Unpublished working paper available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3189548`.

Saltzman, Evan A., Christine Eibner, and Alain C. Enthoven, "Improving the Affordable Care Act: An Assessment of Policy Options For Providing Subsidies," *Health Affairs*, 2015, *34* (12), 2095–2103.

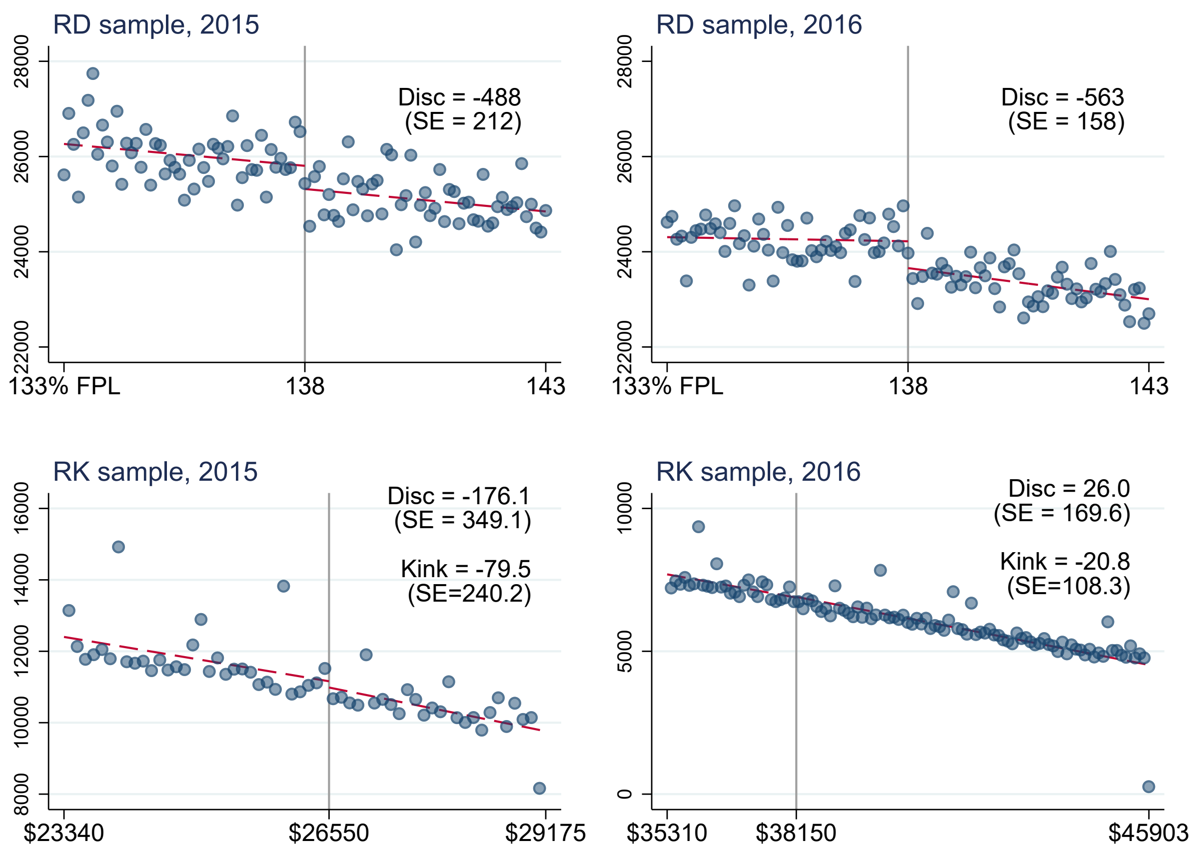
Tebaldi, Pietro, "Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA," 2017. Unpublished working paper.

Figure 1: Mandate penalty and as a function of income



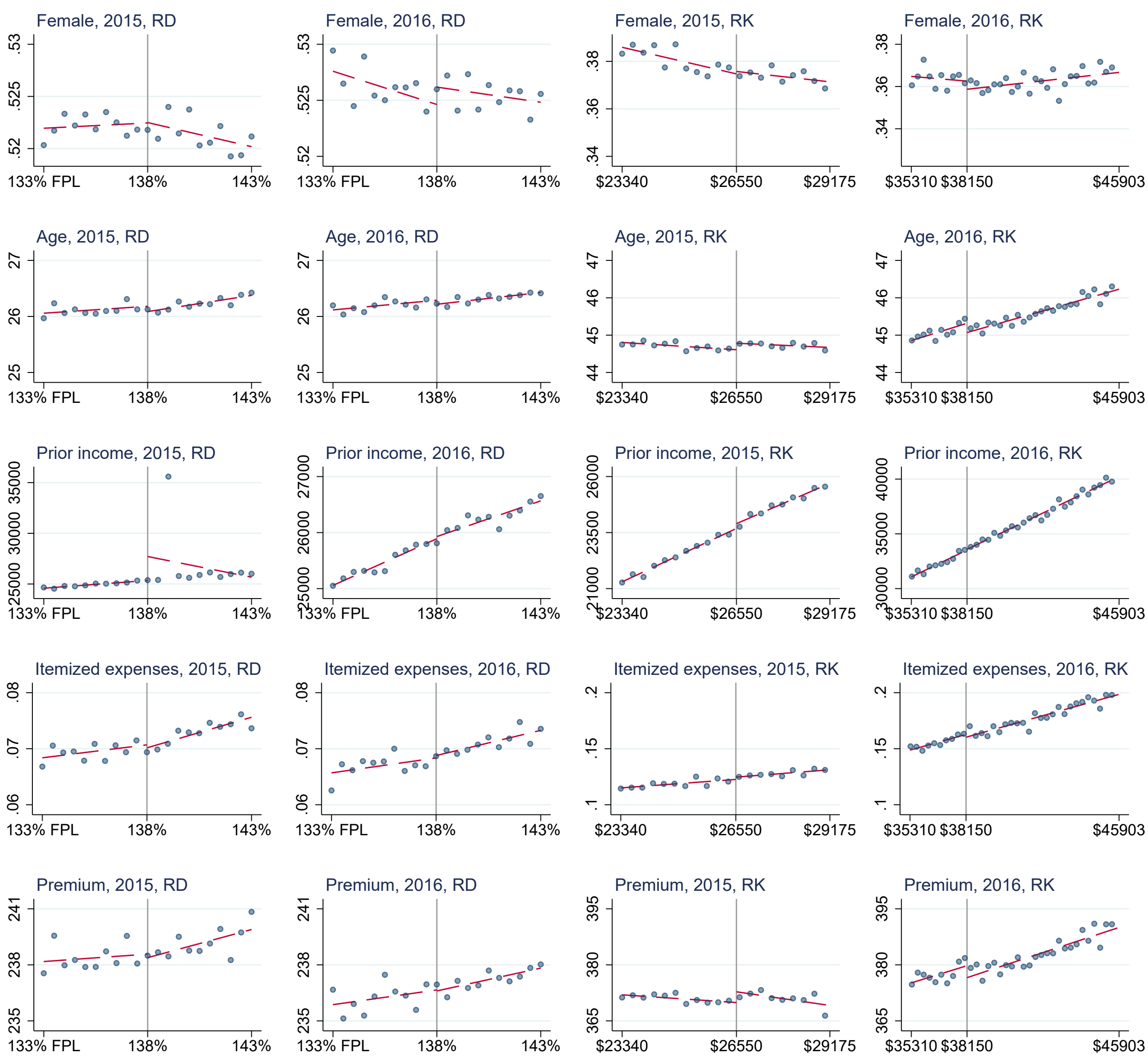
Notes: Figure shows the 2015 and 2016 mandate penalty for a single person per 12 months of uninsurance, as a function of income. The light dashed lines indicates thresholds for cost-sharing reductions and premium tax credits (given the 2016 poverty line of $11,770).

Figure 2: Income distribution



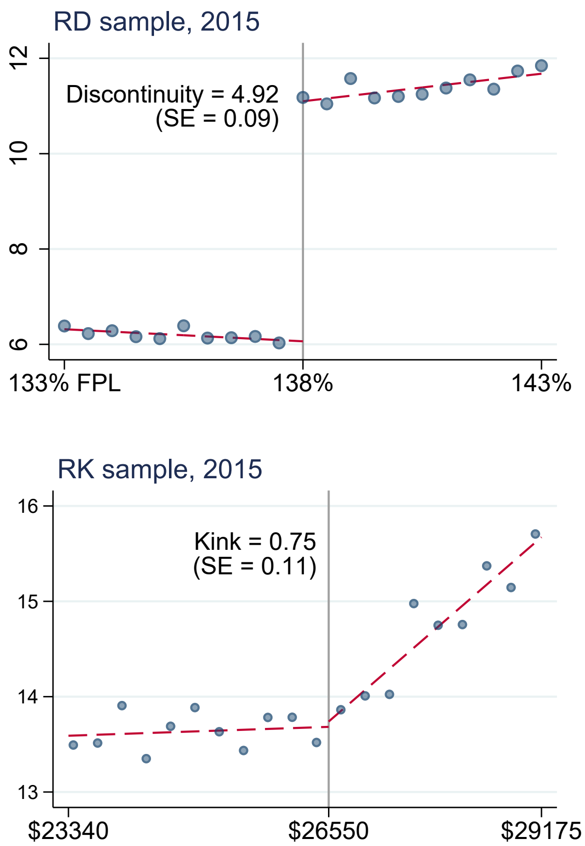
Notes: Each figure shows the number of observations (in 1000s) in each bin of the running variable. We plot the fit on either side of the threshold, and we report the difference in the log height of the bins (McCrary, 2007). The discontinuity sample consist of people aged 0-64 in the indicated year, living in states that did not expand Medicaid. The kink sample consists of people aged 27-64 in the indicated year, who filed single tax returns with one exemption, without signs of ESI offers. The vertical line shows the mandate discontinuity or kink point.

Figure 3: Smoothness of covariates



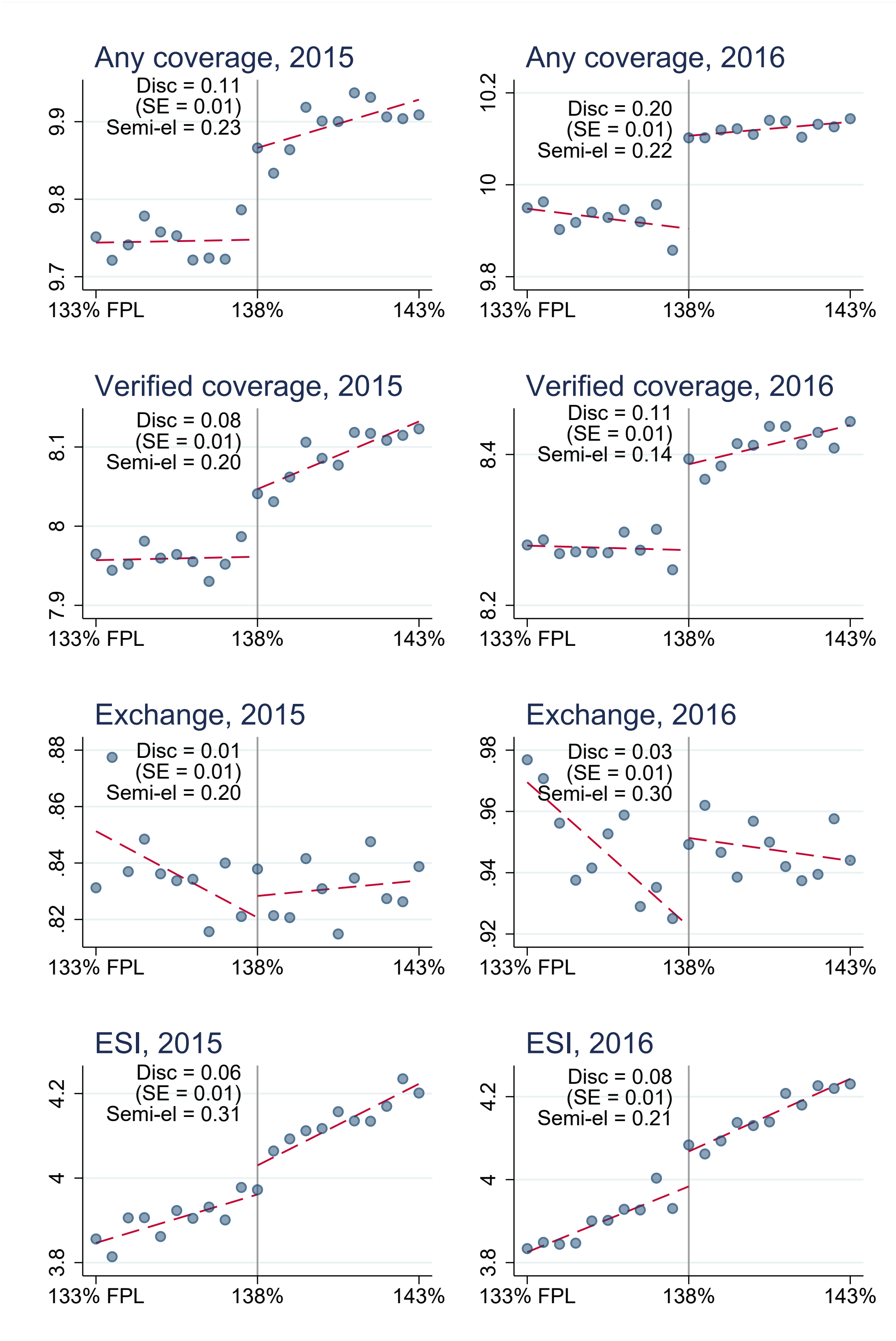
Notes: Figure shows the average of the indicated variable in each bin of FPL or income. “RD” and “RK” refer to the discontinuity and kink samples. The discontinuity sample consist of people aged 0-64 in the indicated year, living in states that did not expand Medicaid. The kink sample consists of people aged 27-64 in the indicated year, who filed single tax returns with one exemption, without signs of ESI offers. The vertical line shows the mandate discontinuity or kink point.

Figure 4: Penalty paid per uninsured month



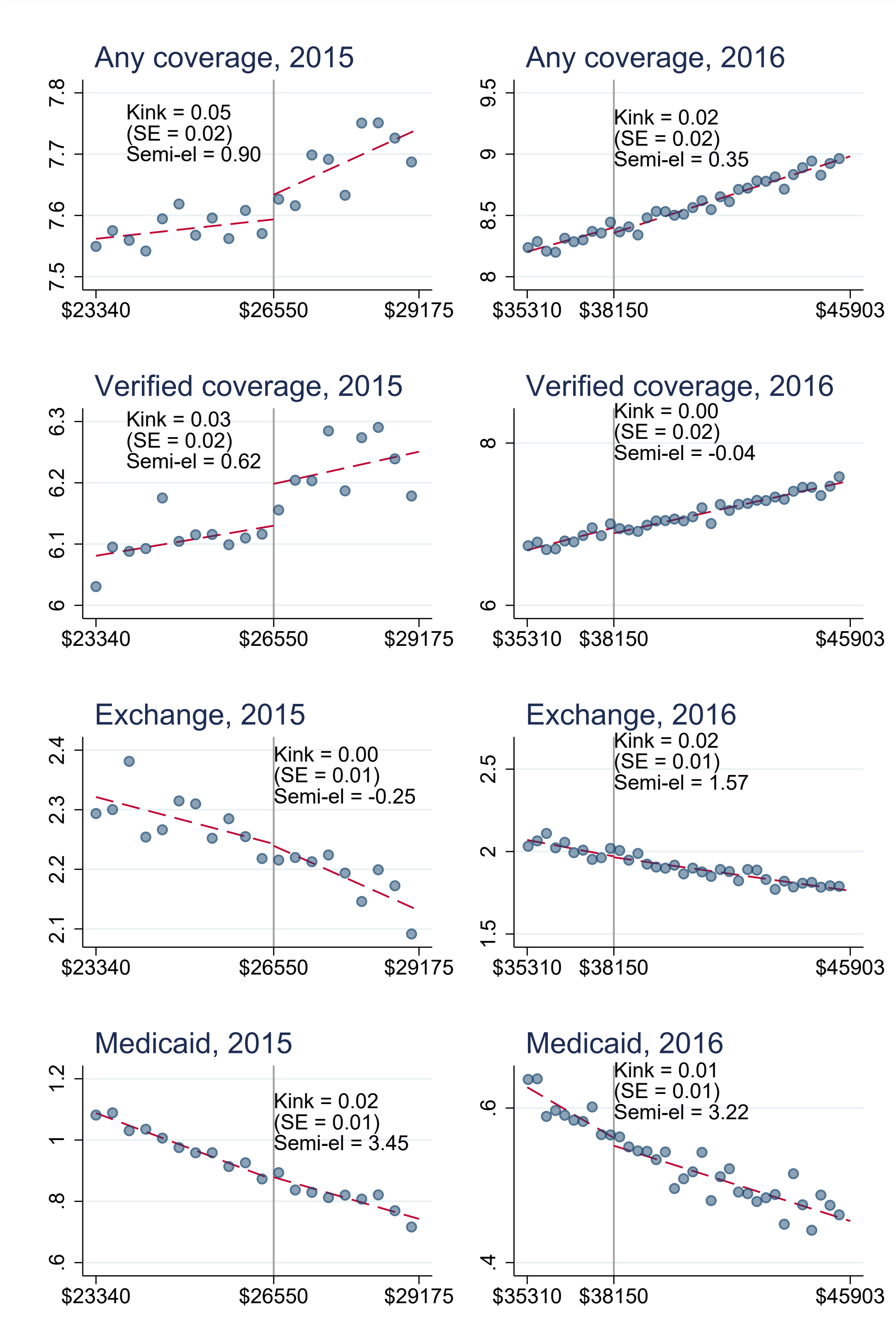
Notes: Figures plots the observed penalty paid per month of uninsurance in each bin of income, in 2015 and 2016, averaging over people with at least one month of insurance. The discontinuity sample consist of people aged 0-64 in the indicated year, living in states that did not expand Medicaid. The kink sample consists of people aged 27-64 in the indicated year, who filed single tax returns with one exemption, without signs of ESI offers. The vertical line shows the mandate discontinuity or kink point.

Figure 5: Discontinuities in months insured, by coverage type



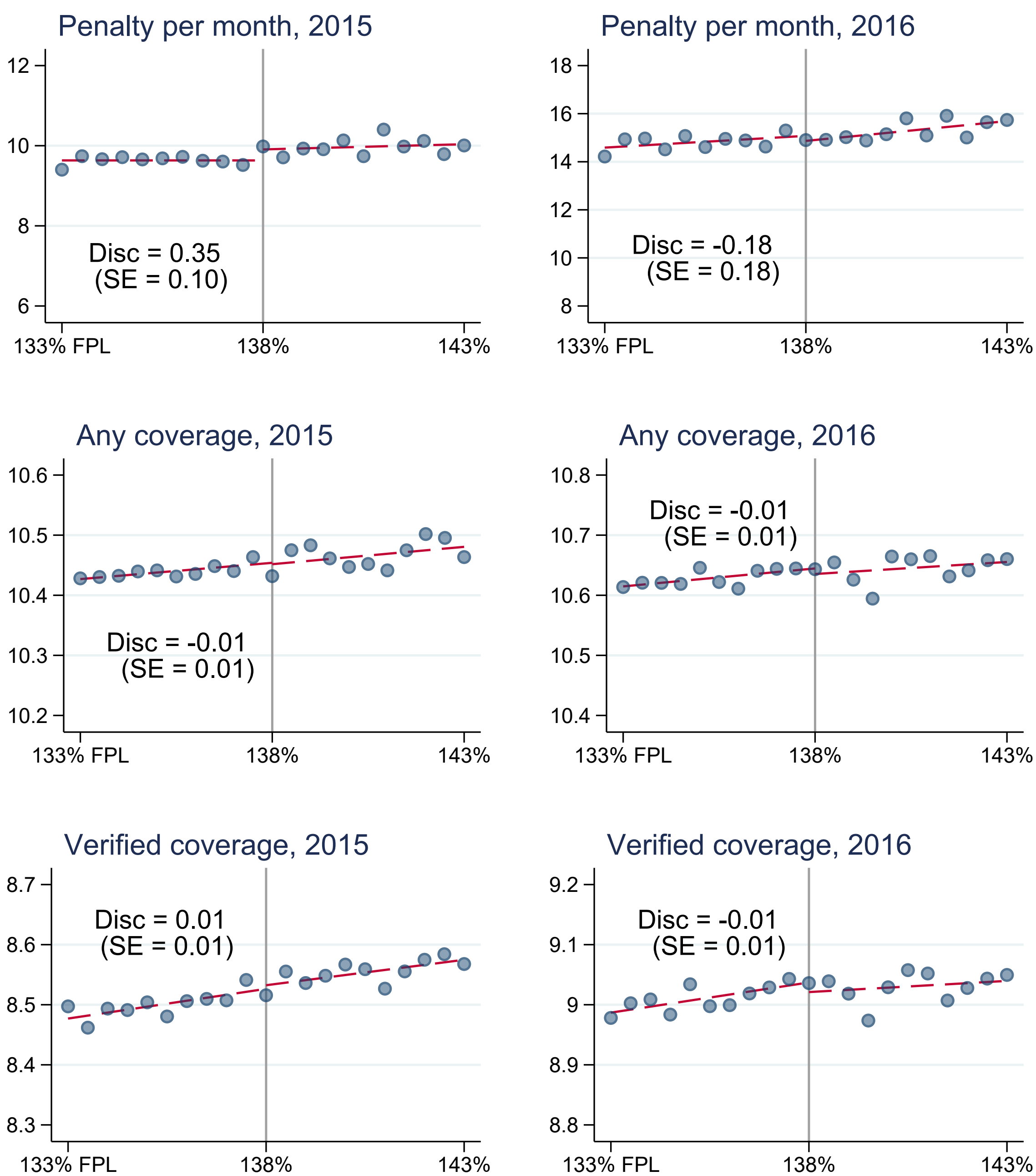
Notes: Figure shows the average number of months of insurance of the indicated type, in each 0.5 FPL point bin. Sample consists of people aged 0-64 in the indicated year, living in non-expansion states. Figure reports the estimated discontinuity, its standard error, and the implied semi-elasticity with respect to the penalty paid.

Figure 6: Kinks in months insured, by coverage type



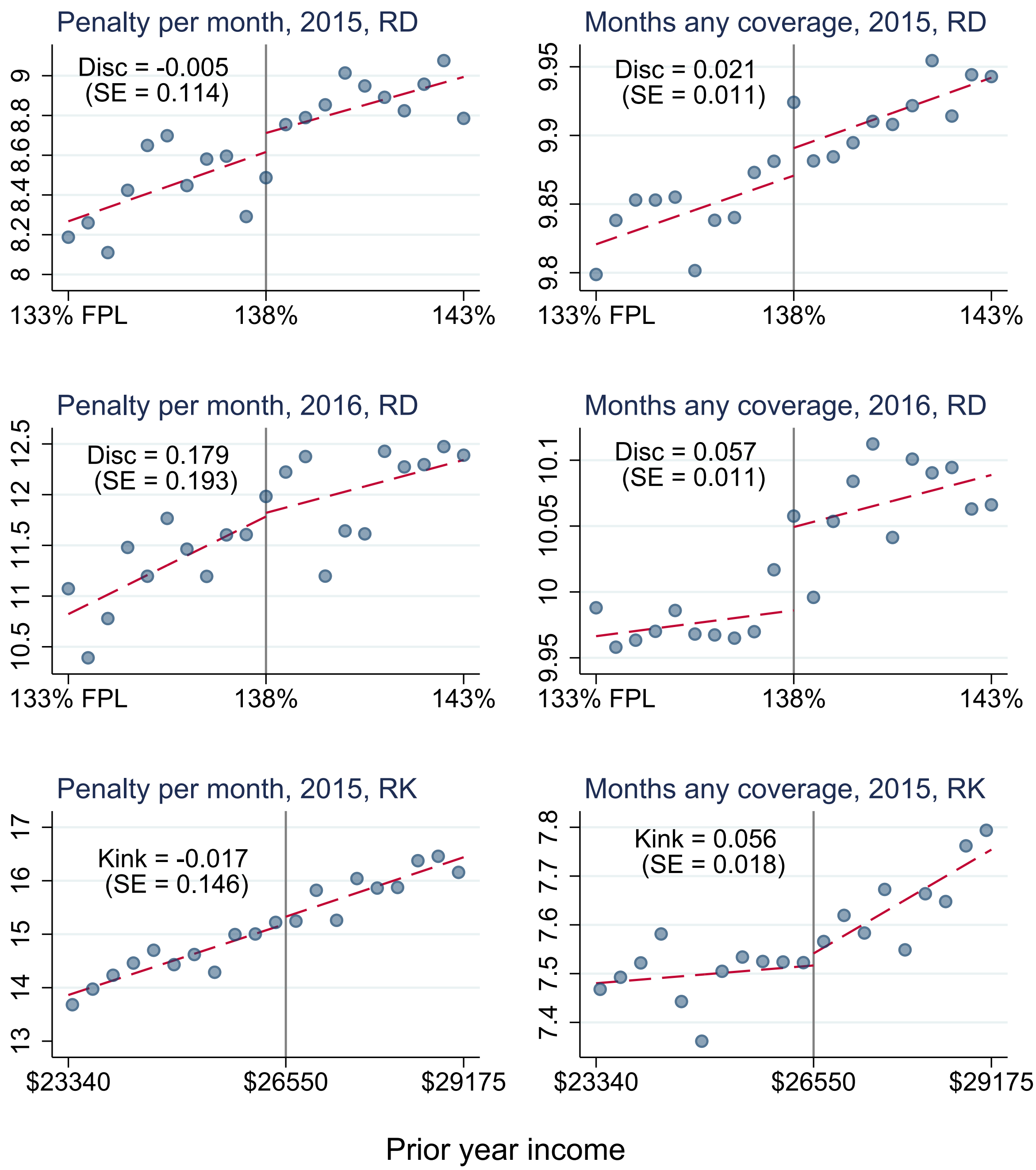
Notes: Figure shows the average number of months of insurance of the indicated type, in each $300 bin. Sample consists of people aged 27-64 in the indicated year, without signs of ESI offer, with single person tax returns and no dependents. Figure reports the estimated kink, its standard error, and the implied semi-elasticity with respect to the penalty paid.

Figure 7: Placebo test in expansion states



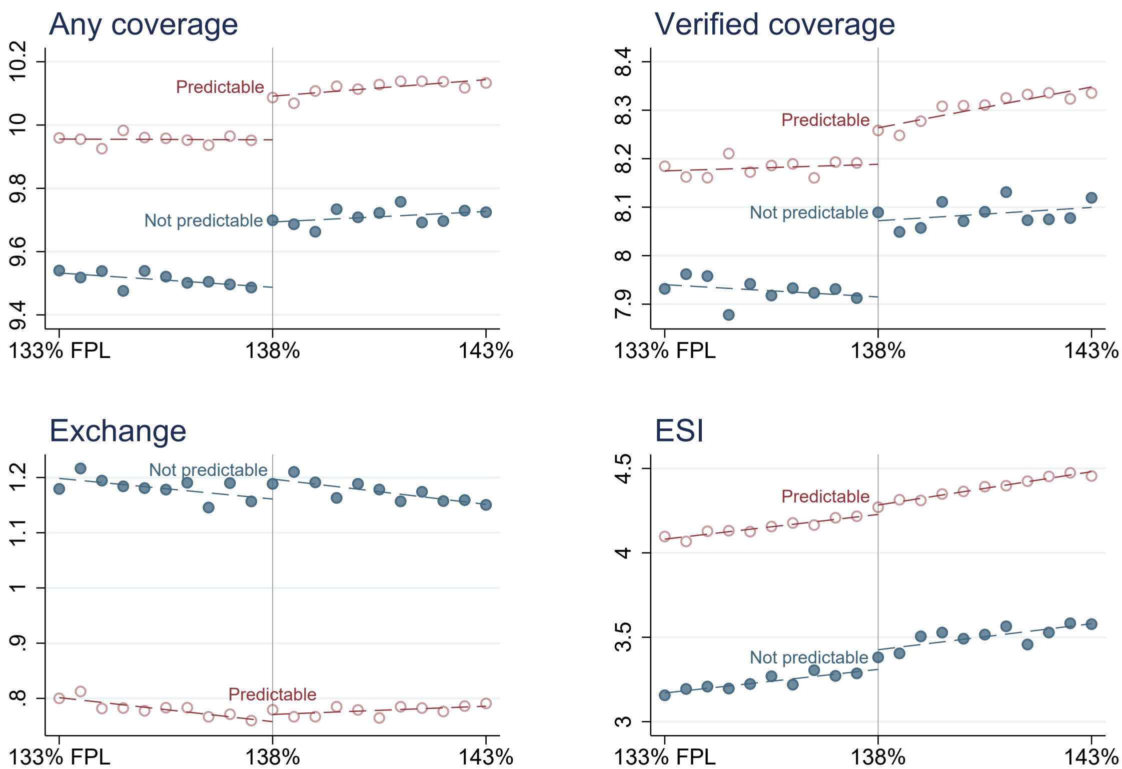
Notes: Figure shows the average of the indicated outcome in each 0.5 FPL point bin. This is a placebo test because, in expansion states, people on both sides of 138 percent of FPL are subject to the mandate. Sample consists of people aged 0-64 in the indicated year, living in Medicaid expansion states. Figure reports the estimated discontinuity and its standard error.

Figure 8: Using lagged income as the running variable



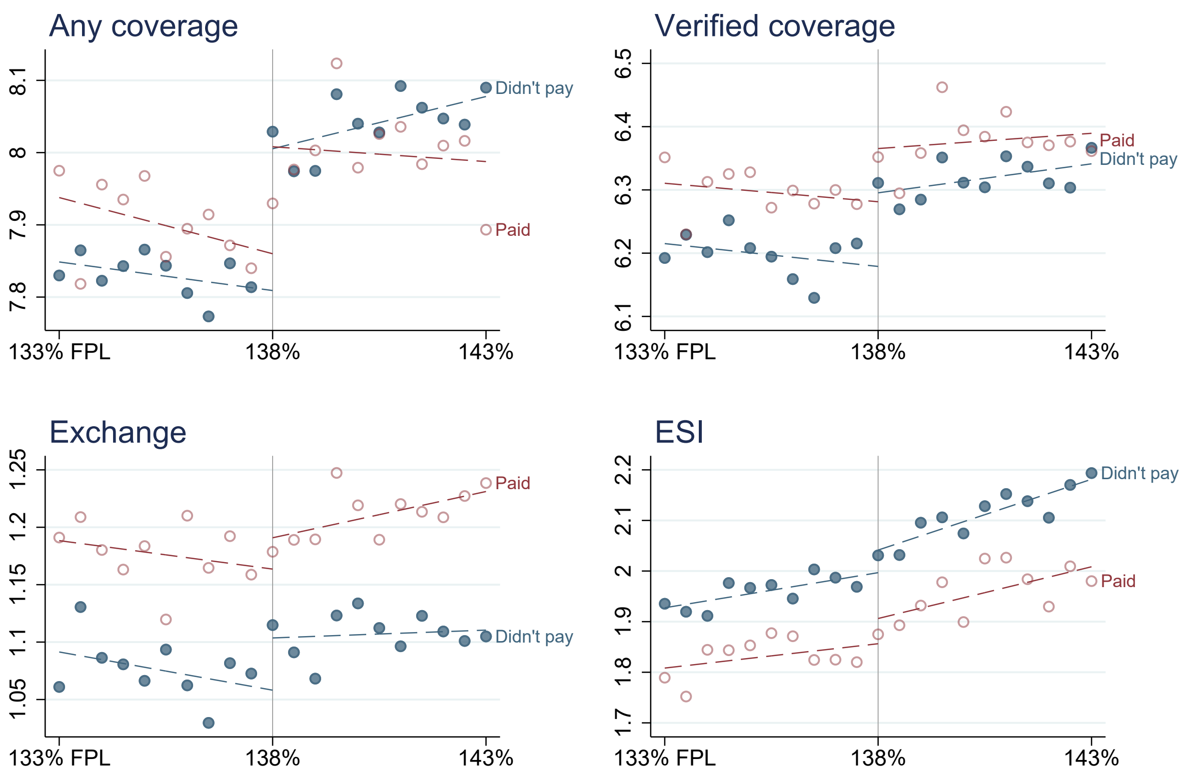
Notes: Figure shows months of coverage in the indicated year against income in the prior year. The kink sample consists of adults aged 27-64 filing single-person tax returns without signs of ESI offers. The discontinuity samples consist of people aged 0-64 living in non-expansion states.

Figure 9: Months of coverage by type, stratifying on income predictability



Notes: Figure shows the average number of months of insurance of the indicated type, in each 0.5 percent of FPL bin, separately for predictable earners and not predictable earners. Predictable earners are defined as people living in households with no unemployment insurance income and at most one W2 per earner. Sample consists of people aged 0-64, pooling 2015-2016.

Figure 10: Penalty paid and months of coverage, stratifying on past experience with the penalty



Notes: The figures plot the indicated outcome against income, separately for people who did and did not pay the mandate penalty in 2014. Sample is limited to people who indicated in 2014 that they did not have full-year coverage. The sample pools 2015 and 2016, and consists of people aged 0-64.

Table 1: Summary statistics

Sample	Discontinuity		Kink	
Year	2015	2016	2015	2016
	(1)	(2)	(3)	(4)
A. Coverage information				
Months of...				
Any coverage	9.82	10.02	7.62	8.53
Verified coverage	8.02	8.34	6.15	7.06
Exchange	0.83	0.95	2.25	1.92
Off-Exchange	0.20	0.17	0.89	1.29
Employer-sponsored insurance	4.01	4.02	1.80	3.11
Medicaid	3.19	3.47	0.92	0.54
Medicare	0.21	0.21	0.36	0.28
VA	0.16	0.15	0.33	0.28
Paid penalty	0.13	0.09	0.21	0.13
Penalty paid per uninsured month	8.39	12.10	14.59	25.92
B. Statistics of income				
10th Percentile	16,200	16,200	23,900	36,200
25th Percentile	21,600	21,600	24,700	37,500
50th Percentile	27,500	27,500	26,100	40,000
75th Percentile	33,700	33,700	27,600	42,700
90th Percentile	39,800	39,800	28,500	44,600
Mean	28,153	28,453	26,150	40,185
C. Covariates				
Female	0.52	0.53	0.38	0.36
Age	26.16	26.25	44.73	45.42
Prior income	25,826	25,812	23,304	35,135
Prior itemized medical expenses?	0.07	0.07	0.12	0.17
Benchmark premium	238.66	236.67	371.20	380.90
# Exemptions	3.15	3.13	1	1
# Observations	2,556,318	2,380,342	653,891	656,064

Notes: Panels A and C report averages of the indicated variables. Panel B reports the indicated statistics of modified adjusted gross income. The discontinuity sample people consists of people age 0-64 with income between 133 and 143 percent of FPL, living in non-expansion states. The kink sample consists of people who filed single tax returns with one exemption, aged 27-64, with no with no signs of ESI offers and income between 200-250 percent of FPL in 2015 or 300-390 percent of FPL in 2016. Income percentiles are percentiles of income in $100 bins, to avoid disclosing individual taxpayers' income.

Table 2: Smoothness tests

Dep. var.	Female	Age	Prior income	Any prior itemized expenses	Benchmark Premium	Number of Exemptions
	(1)	(2)	(3)	(4)	(5)	(6)
			A. 2015 discontinuity sample			
Discontinuity	0.000	-0.089	2862	-0.001	-0.049	-0.002
	(0.001)	(0.043)	(1389)	(0.001)	(0.334)	(0.004)
			B. 2016 discontinuity sample			
Discontinuity	0.002	-0.077	63.4	0.000	-0.028	0.004
	(0.001)	(0.045)	(56.8)	(0.001)	(0.341)	(0.004)
			C. 2015 kink sample			
Kink	0.003*	0.058	28.091	0.001	0.154	–
	(0.001)	(0.037)	(74.079)	(0.001)	(0.460)	–
			D. 2016 kink sample			
Kink	0.001	0.021	60.522	0.000	0.398	–
	(0.001)	(0.034)	(71.339)	(0.001)	(0.474)	–

The kink sample consists of single tax returns, with one exemption claimed, aged 27-64, no signs of ESI offers, and with income between 200 and 250 percent of FPL (in 2015) or 300 and 390 percent of FPL (in 2016). The discontinuity sample consists of people aged 0-64 with income between 133 and 143 percent of FPL, living in non-expansion states. Robust standard errors in parentheses. We do not report the kink in number of exemptions in the kink sample because everyone in that sample has one exemption. * As the kink is rounded up and the standard error is rounded down, the the p-value here is 0.07.

Table 3: Estimated discontinuities and kinks in months of insurance coverage

Coverage type	Any (1)	Verified (2)	Exchange (3)	Off-Exchange (4)	ESI (5)	Medicaid (6)	VA (7)	Medicare (8)
	A. 2015 discontinuity sample							
Discontinuity	0.110	0.080	0.008	-0.005	0.060	0.024	0.005	-0.001
	(0.011)	(0.013)	(0.007)	(0.004)	(0.013)	(0.013)	(0.004)	(0.003)
Semi-elasticity	0.23	0.20	0.19	-0.46	0.31	0.15	0.51	-0.16
	B. 2016 discontinuity sample							
Discontinuity	0.198	0.110	0.026	-0.001	0.078	0.006	0.007	0.003
	(0.011)	(0.013)	(0.008)	(0.004)	(0.014)	(0.013)	(0.004)	(0.003)
Semi-elasticity	0.22	0.14	0.30	-0.06	0.21	0.02	0.38	0.19
	C. 2015 kink sample							
Kink	0.051	0.029	-0.004	-0.003	0.016	0.023	0.010	0.002
	(0.017)	(0.017)	(0.013)	(0.010)	(0.013)	(0.009)	(0.006)	(0.006)
Semi-elasticity	0.918	0.630	-0.251	-0.426	1.193	3.511	4.200	0.857
	D. 2016 kink sample							
Kink	0.020	-0.002	0.021	0.007	-0.046	0.012	0.001	0.000
	(0.015)	(0.016)	(0.012)	(0.010)	(0.014)	(0.007)	(0.005)	(0.005)
Semi-elasticity	0.284	-0.031	1.268	0.651	-1.833	2.608	0.378	0.137

Table reports the RD or RK estimate, obtained from a regression of the indicated coverage type on income. The discontinuity samples consist of people with income between 133 and 143 percent of FPL in the indicated year, living in non-expansion states, aged 0-64. The kink sample consists of people with income between 200 and 250 percent of FPL (in 2015) or 300 and 390 percent of FPL (in 2016), filing single tax returns with no dependents, aged 27-64, and no signs of ESI offers. Robust standard errors in parentheses.

Table 4: Heterogeneous coverage discontinuities

	# in group (1)	Mean months (2)	Discontinuity (3)	Standard error (4)	Semi-elasticity (5)
A. Split on sex					
Male	2,351,049	9.68	0.175	(0.012)	0.248
Female	2,585,635	10.13	0.132	(0.010)	0.197
p-value of equality			0.006		
B. Split on age					
Child	2,660,091	10.44	0.112	(0.009)	0.153
Young adult	1,130,014	8.98	0.218	(0.019)	0.320
Older adult	1,146,579	9.62	0.184	(0.017)	0.311
p-value (Child vs. YA)			0.000		
p-value (Child vs. OA)			0.000		
p-value (YA vs. OA)			0.181		
C. Split on prior itemized medical expenses					
Has none	4,548,541	9.86	0.153	(0.008)	0.220
Has some	342,080	10.92	0.129	(0.022)	0.219
p-value of equality			0.307		
D. Split on prior income					
Below median	2,468,604	9.74	0.181	(0.011)	0.258
Above median	2,468,080	10.09	0.121	(0.011)	0.181
p-value of equality			0.000		

Notes: Table reports the estimated discontinuity in months of any coverage, for the indicated groups, as well as the group size average coverage rate (for months of any coverage), and the p-value for the test that discontinuities in each group are equal. Sample consists of people aged 0-64 in 2015 or 2016, with income between 133 and 143 percent of FPL, living in non-expansion states. Here a child is a defined as younger than 27 (and therefore eligible for parent's insurance), a young adult is below median adult age, and an older adult is above median adult age.

Table 5: Coverage discontinuities among predictable and non-predictable earners

Group	# in group (1)	Mean months (2)	Discontinuity (3)	Standard error (4)	Semi-elasticity (5)
			A. Any coverage		
Not predictable	1,352,328	8.00	0.201	(0.015)	0.364
Predictable	3,584,356	8.24	0.132	(0.009)	0.230
p-value of equality				0.000	
			B. Verified coverage		
Not predictable	1,352,328	8.00	0.157	(0.018)	0.284
Predictable	3,584,356	8.24	0.069	(0.011)	0.121
p-value of equality				0.000	
			C. Exchange		
Not predictable	1,352,328	8.00	0.037	(0.012)	0.067
Predictable	3,584,356	8.24	0.011	(0.006)	0.019
p-value of equality				0.044	
			D. ESI		
Not predictable	1,352,328	8.00	0.112	(0.017)	0.203
Predictable	3,584,356	8.24	0.047	(0.012)	0.081
p-value of equality				0.002	

Notes: Table reports the estimated discontinuity in coverage, as well as the semi-elasticity, estimated separately for predictable earners and not predictable earners, and the p-value for the test that discontinuities in each group are equal. Sample consists of people aged 0-64 living in non-expansion states with income between 133 and 143 percent of FPL. To maximize power, we pool 2015 and 2016 and include a year dummy variable.

Table 6: Coverage discontinuities, stratifying on past experience with the mandate

Group	# in group (1)	Mean months (2)	Discontinuity (3)	Standard error (4)	Semi-elasticity (5)
			A. Any coverage		
Checked box	2,874,296	9.60	0.092	(0.007)	0.155
No box, no penalty	1,058,434	6.25	0.194	(0.021)	0.456
No box, penalty	605,647	6.34	0.131	(0.027)	0.248
p-value of equality			0.067		
			B. Verified coverage		
Checked box	2,874,296	9.60	0.054	(0.010)	0.091
No box, no penalty	1,058,434	6.25	0.117	(0.022)	0.276
No box, penalty	605,647	6.34	0.081	(0.028)	0.154
p-value of equality			0.304		
			C. Exchange		
Checked box	2,874,296	9.60	0.010	(0.007)	0.016
No box, no penalty	1,058,434	6.25	0.042	(0.013)	0.098
No box, penalty	605,647	6.34	0.026	(0.017)	0.050
p-value of equality			0.473		
			D. ESI		
Checked box	2,874,296	9.60	0.045	(0.013)	0.076
No box, no penalty	1,058,434	6.25	0.046	(0.016)	0.107
No box, penalty	605,647	6.34	0.048	(0.021)	0.090
p-value (penalty vs. not)				0.939	

Notes: Table reports the estimated discontinuity in coverage in 2015 and 2016, as well as the semi-elasticity, estimated separately for three groups: people who checked the box in 2014 to indicate full-year coverage; people who did not check the box and did not pay a mandate penalty; and people who did not check the box and paid a penalty. Table also reports the p-value for the test that discontinuities in each group are equal Sample consists of people aged 0-64 living in non-expansion states with income between 133 and 143 percent of FPL. To maximize power, we pool 2015 and 2016 and include a year dummy variable.

For online publication only

Appendices

A Further details on the mandate penalty

We restrict our sample to single, childless adults because the mandate penalty is more complicated for larger families. Here we describe these complications. Recall from Section 1 that the monthly mandate penalty is

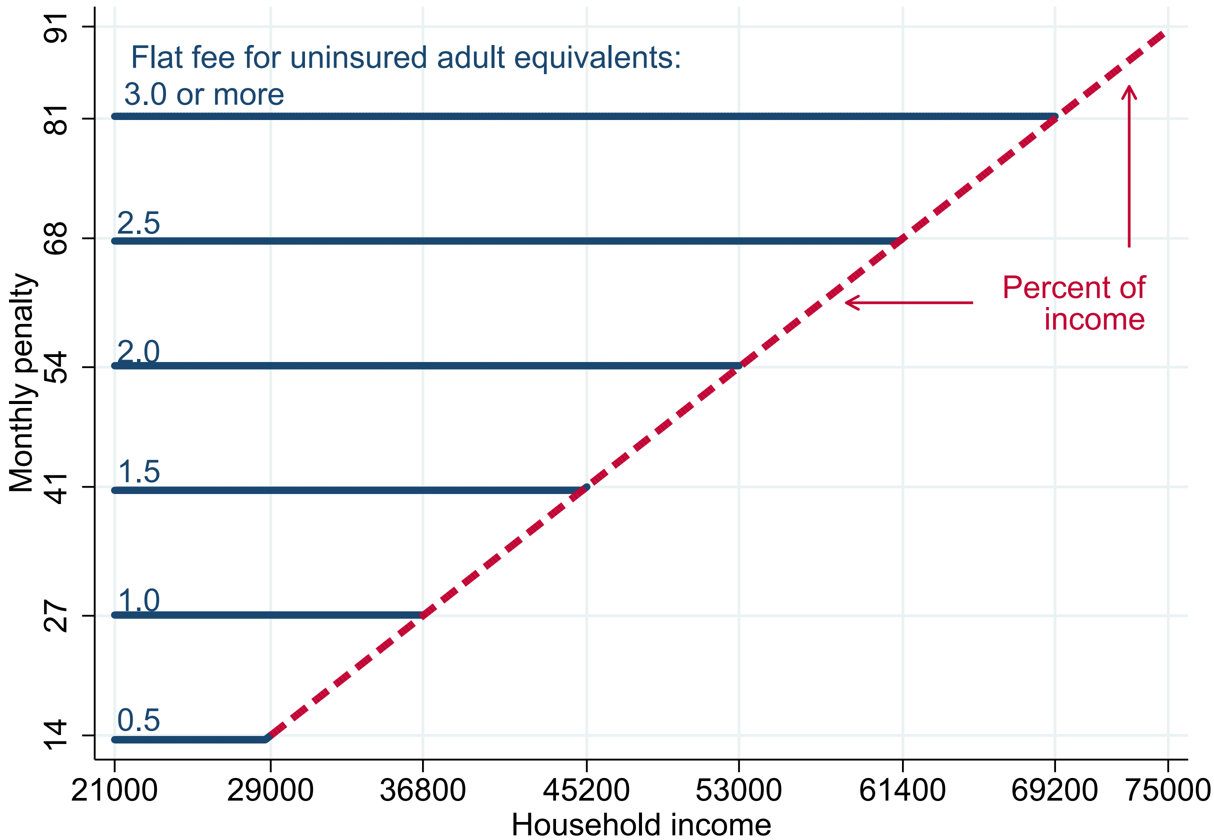
$$Penalty = 1/12 \max\left\{\min\left\{[A + .5C]F, 3F\right\}, S(\text{MAGI} - \text{tax filing threshold})\right\},$$

where F is the flat fee, A the number of uninsured adults that month, C the number of uninsured children, and S the percent of income. In 2016, F was \$325 and S was 0.02. For a given filing threshold, the penalty is therefore a function of income and the number of uninsured adults and children. Each child counts as half an adult for the purposes of determining the mandate penalty so we refer to the number of uninsured "adult equivalents", equal to $A + .5C$. Note that the number of uninsured adult equivalents affects the flat fee (the first term in brackets) but not the percent of income (the second term).

Appendix Figure A.1 plots the monthly mandate penalty in 2015 as a function of income and the number of uninsured adult equivalents, for a married, filing jointly tax return (which has a filing threshold of \$20,600). There are six kink points, equal to each intersection of $F(A + .5C)$ with the percent of income, for $A + .5C \in \{0.5, 1, 1.5, 2, 2.5\}$. At each of the kink points, the mandate penalty increases for some margins of coverage, but not for all margins. For example, consider a household with two adults and two children, and income of \$68,000. For this household, the percent of income payment is \$79. If there were one uninsured household member, the penalty would be \$79, because the percent of income exceeds the flat fee. A second uninsured household member would not increase the penalty because the percent of income exceeds the flat fee for two uninsured members. Only if the entire household were uninsured would the flat fee, \$81.25, exceed the percent of income. However, if the household's income were \$70,000, the percent of income would always exceed the flat fee, regardless of the number of uninsured adult equivalents. Thus there is a kink in the incentive to have the entire family covered, but no kink in the marginal incentive to cover the first, second, or third family member.

This example shows that, for multiperson households, the mandate penalty creates complex and fairly subtle incentives to increase coverage. In principle it is possible to examine coverage responses at each of the six coverage kinks, focusing on the relevant margin of coverage (e.g. a kink in the probability of having three or more uninsured adult equivalents at \$69,200). In practice we are concerned that households may not understand the specific incentives for monthly coverage generated by the individual mandate. We therefore focus on single households. For these households the penalty is relatively simple–it is linear in their number of uninsured months.

Figure A.1: Monthly mandate penalty for multi-person households, 2015



Notes: Source: Figure shows the monthly mandate penalty in 2015 as a function of income and the number of uninsured adult equivalents, for a household with married filing jointly tax return. The number of uninsured adult equivalents is the number of uninsured adults plus half the number of uninsured children. For some income levels and numbers of uninsured adult equivalents, insuring an additional adult equivalent does not change the mandate penalty, because the percent of income penalty is the same for any number of uninsured adult equivalents.

B Monte Carlo study of RKD estimators

We conducted a Monte Carlo simulation study to assess the performance of alternative RKD estimators. The canonical RKD specification is

$$y_i = \sum_{d=0}^{D} \alpha_d (v_i - c)^d + \sum_{d=0}^{D} \beta_d (v_i - c)^d \times D_i + \varepsilon_i,$$

where v_i is the running variable, D_i is indicator for the running variable exceeding the cutoff, and $\hat{\beta}_1$ is the kink estimate (Card et al., 2015). Our estimating equation is identical to this, with $D = 1$. In estimating a regression kink design, researchers must make several specification choices: the choice of degree D, the bandwidth h, the kernel, and whether to allow for a discontinuity.

The theoretical econometric literature recommends using a triangular kernel for boundary estimation problems such as this one. For estimating a kink, the theoretical literature also recommends $D \geq 2$, and it has developed plug-in estimators for bandwidth choice based on minimizing asymptotic mean squared error of the kink estimate (Imbens and Kalyanaraman, 2012; Calonico et al., 2014). However applied researchers have favored the uniform kernel—as the regression can then be estimated with OLS—and have found that high degree terms and asymptotically optimal bandwidths do not necessarily perform well in finite samples. Applied researchers also sometimes impose continuity (i.e. dropping the $(v_i - c)^0 D_i$ term).

To determine our baseline specification choices, we conducted a Monte Carlo following the suggestions in Card et al. (2017). The overall idea is to simulate many data sets using a data generating process that closely resembles our data, and then compare the performance of alternative RKD estimators across the data sets. Our data generating process is based on a high-order polynomial approximation to the data, with a true kink imposed. To do so, we first "dekink" the data by estimating the following regression, separately for 2015 and 2016:

$$y_i - \hat{\tau}_t D_i v_i = \sum_{d=0}^{5} \beta_d v_i^d + \sum_{d \neq 1}^{5} \theta_d v_i^d D_i + \epsilon_i. \tag{3}$$

where $\hat{\tau}_t$ is the estimated kink in year t, 0.05 in 2015 and 0.02 in 2016. Let $\hat{y}(v)$ be the predicted value from this regression when the running variable is v.

We simulate data with a known kink τ. We consider two cases: τ corresponding to a semi-elasticity of 0.5, which we consider to be the middle of past estimates, and τ corresponding to a semi elasticity of 0.2, which is low but consistent with the Massachusetts evidence. For each year and each of 1000 simulation data sets, we sampling with replacement from the empirical distribution of v and ϵ in that year. Each simulation dataset is the same size as our estimation dataset. Given the draw of v and ϵ, we form the outcome y as $\hat{y}(v) + \epsilon + D\tau v$, where τ is the assumed kink. We then estimate several different RKD specifications on the simulated data. For each simulated data set, we considered the power set of the following specification choices: bandwidth equal to the full range of income, the Fan-Gijbels bandwidth selector (as proposed by Card et al. (2015)), or Calonico et al. (2014) bandwidth selector (without scale regularization); polynomial degree $D = 1$ or $D = 2$; and imposing continuity or not. Throughout we use a uniform kernel, for consistency with the applied literature. We do not consider the bias-corrected estimator of Calonico et al. (2014) because it is computationally costly and initial simulations suggested that it lead to dramatically higher variance without large reductions in bias or improvements in coverage rates (a result also reported by Card et al. (2017)).

Appendix Table B.1 summarizes the performance of the various estimators in the 2015 sample. The linear estimator performs well: it achieves its nominal coverage rate, and rejects a false null 96-97 percent of the time. The Fan-Gijbels and CCT bandwidth selectors choose fairly small bandwidths, \$1,248 to \$1,806, relative to a maximal bandwidth of about \$2,900. Relative to using the full range of the data, they have a higher RMSE and a lower rejection rate; the coverage rate is slightly higher for the discontinuous estimator and slightly lower for the continuous estimator. The linear estimator using the full range of the data has the lowest RMSE. The estimators that use only relatively local information give up some power without reducing bias. Allowing for a discontinuity results in slightly higher bias and variance. The quadratic estimators perform substantially worse than the linear estimators: they have higher (absolute) bias, much higher variance, and worse coverage.[31] We conclude that the linear estimator using the full range of data is likely to perform better than the alternatives, although none of the estimators achieves the nominal coverage rate, and this estimator has the worst coverage.

Appendix Table B.2 summarizes the performance of the estimators in the 2016 sample. Here too we find that the linear specification using the full range of the data has the lowest mean squared error, again with somewhat higher confidence intervals. In this case the coverage rate of the linear estimator is below the nominal rate when we impose continuity.

Because our 2016 estimates were statistically insignificant, we also investigated the power of our estimator to detect small kinks. Specifically, we re-ran our Monte Carlo simulations, but assuming a semi-elasticity of 0.2 instead of 0.5, and assuming a semi-elasticity of 0.14. The 0.2 semi-elasticity corresponds to the estimate that Hackmann et al. (2015) find using the Massachusetts mandate. They look at a sample of relatively high income adults, with income above 300 percent of FPL, so we believe this is a useful benchmark. We report the results of this simulation in Appendix Tables B.3 and B.4. The semi-elasticity of 0.14 corresponds to our main estimate. Consistent with our other simulation results, we find that the linear estimator using the full range of data outperforms estimators with higher order terms or tighter bandwidths. However, even for this estimator, we find somewhat limited power. When we do not impose continuity, we reject a false null in only 73 percent of iterations. Imposing continuity improves power. At the smallest semi-elasticity we considered, 0.14, we find limited power even when imposing continuity; we reject the false null in 74 percent of iterations. Without continuity we reject in less than half of all iterations.

[31] The FG bandwidth usually ends up exceeding the range of data in the quadratic case, so its performance is the same as the estimator using the full range of data.

Table B.1: Summary of performance of RKD estimators in Monte Carlo Simulation, 2015

Estimator	Median Bandwidth	$\frac{RMSE}{\tau}$	Coverage Rate	$\frac{Bias}{\tau}$	$\frac{Variance}{\tau^2}$	Rejection Rate
	(1)	(2)	(3)	(4)	(5)	(6)
A. Linear estimators						
BW = full, continuous	–	0.261	0.958	0.001	0.261	0.971
BW = FG, continuous	1806	0.339	0.953	0.005	0.339	0.857
BW = CCT, continuous	1248	0.428	0.939	0.012	0.428	0.697
BW = full, discontinuous	–	0.267	0.951	0.005	0.267	0.963
BW = FG, discontinuous	1806	0.347	0.956	0.009	0.347	0.844
BW = CCT, discontinuous	1248	0.436	0.936	0.019	0.436	0.686
B. Quadratic estimators						
BW = full, continuous	–	1.043	0.900	0.780	1.043	0.393
BW = FG, continuous	4787	1.043	0.899	0.779	1.043	0.393
BW = CCT, continuous	1462	1.519	0.922	0.802	1.519	0.233
BW = full, discontinuous	–	1.071	0.891	0.839	1.071	0.408
BW = FG, discontinuous	4787	1.071	0.891	0.839	1.071	0.408
BW = CCT, discontinuous	1462	1.560	0.916	0.870	1.560	0.247

Table summarizes the performance of 12 RKD estimators, which differ in the degree of the underlying polynomial (linear or quadratic), bandwidth selector (full range of data, Fan-Gjbels, or Calonico et al. (2014)), and whether a discontinuity is imposed. The data are generated using a true kink of $\tau = 6.33 \times 10^{-5}$, corresponding to a semi-elasticity of 0.5 at the mandate kink point. The coverage rate is the fraction of confidence intervals containing this kink.

Table B.2: Summary of performance of RKD estimators in Monte Carlo Simulation, 2016

Estimator	Median Bandwidth	$\frac{RMSE}{\tau}$	Coverage Rate	$\frac{Bias}{\tau}$	$\frac{Variance}{\tau^2}$	Rejection Rate
	(1)	(2)	(3)	(4)	(5)	(6)
A. Linear estimators						
BW = full, continuous	–	0.133	0.921	0.068	0.133	1.000
BW = FG, continuous	2308	0.184	0.949	0.066	0.184	1.000
BW = CCT, continuous	1304	0.255	0.945	0.066	0.255	0.963
BW = full, discontinuous	–	0.164	0.950	0.035	0.164	0.999
BW = FG, discontinuous	2308	0.234	0.947	0.044	0.234	0.987
BW = CCT, discontinuous	1304	0.329	0.947	0.043	0.329	0.881
B. Quadratic estimators						
BW = full, continuous	–	0.544	0.778	-0.625	0.544	0.109
BW = FG, continuous	7598	0.557	0.797	-0.624	0.557	0.109
BW = CCT, continuous	1649	0.964	0.890	-0.632	0.964	0.066
BW = full, discontinuous	–	0.666	0.721	-0.929	0.666	0.039
BW = FG, discontinuous	7598	0.681	0.754	-0.929	0.681	0.037
BW = CCT, discontinuous	1649	1.186	0.866	-0.892	1.186	0.030

Table summarizes the performance of 12 RKD estimators, which differ in the degree of the underlying polynomial (linear or quadratic), bandwidth selector (full range of data, largest symmetric band, Fan-Gjbels, or Calonico et al. (2014)), and whether a discontinuity is imposed. The data are generated using a true kink of $\tau = 8.75 \times 10^{-5}$, corresponding to a semi-elasticity of 0.5 at the mandate kink point. The coverage rate is the fraction of confidence intervals containing this kink.

Table B.3: Summary of performance of RKD estimators in Monte Carlo Simulation, 2016, assuming low semi-elasticity

Estimator	Median Bandwidth (1)	$\frac{RMSE}{\tau}$ (2)	Coverage Rate (3)	$\frac{Bias}{\tau}$ (4)	$\frac{Variance}{\tau^2}$ (5)	Rejection Rate (6)
A. Linear estimators						
BW = full, continuous	–	0.332	0.921	0.171	0.332	0.946
BW = FG, continuous	2308	0.461	0.949	0.166	0.461	0.698
BW = CCT, continuous	1304	0.639	0.945	0.166	0.639	0.445
BW = full, discontinuous	–	0.410	0.950	0.088	0.410	0.731
BW = FG, discontinuous	2308	0.584	0.947	0.109	0.584	0.467
BW = CCT, discontinuous	1304	0.822	0.947	0.107	0.822	0.288
B. Quadratic estimators						
BW = full, continuous	–	1.361	0.778	-1.562	1.361	0.008
BW = FG, continuous	7598	1.392	0.797	-1.559	1.392	0.008
BW = CCT, continuous	1649	2.410	0.890	-1.580	2.410	0.019
BW = full, discontinuous	–	1.665	0.721	-2.324	1.665	0.002
BW = FG, discontinuous	7598	1.703	0.754	-2.322	1.703	0.001
BW = CCT, discontinuous	1649	2.965	0.866	-2.229	2.965	0.011

Table summarizes the performance of 12 RKD estimators, which differ in the degree of the underlying polynomial (linear or quadratic), bandwidth selector (full range of data, largest symmetric band, Fan-Gjbels, or Calonico et al. (2014)), and whether a discontinuity is imposed. The data are generated using a true kink of $\tau = 3.5 \times 10^{-5}$, corresponding to a semi-elasticity of 0.2 at the mandate kink point. The coverage rate is the fraction of confidence intervals containing this kink, and the rejection rate is the fraction of confidence intervals that exclude zero.

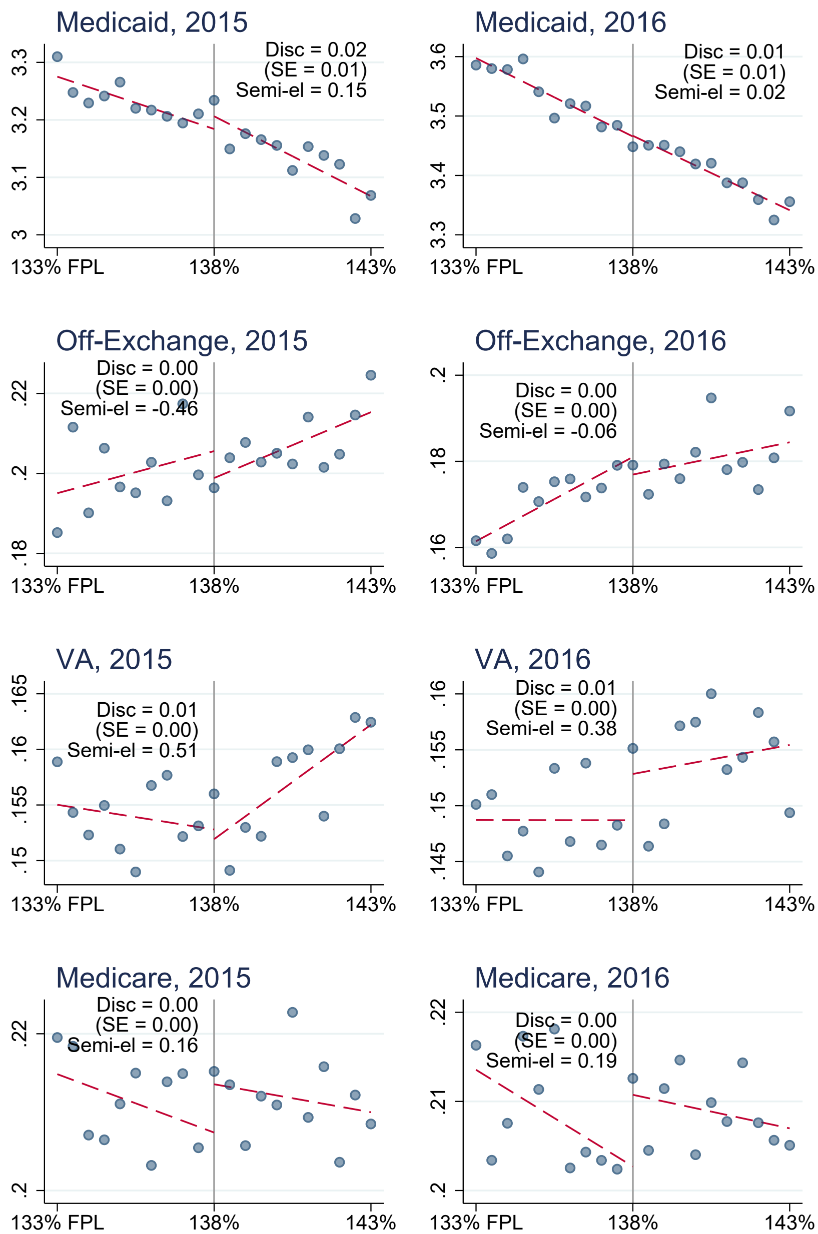
Table B.4: Summary of performance of RKD estimators in Monte Carlo Simulation, 2016, assuming very low semi-elasticity

Estimator	Median Bandwidth (1)	$\frac{RMSE}{\tau}$ (2)	Coverage Rate (3)	$\frac{Bias}{\tau}$ (4)	$\frac{Variance}{\tau^2}$ (5)	Rejection Rate (6)
A. Linear estimators						
BW = full, continuous	–	0.474	0.921	0.244	0.474	0.731
BW = FG, continuous	2308	0.658	0.949	0.237	0.658	0.448
BW = CCT, continuous	1304	0.912	0.945	0.237	0.912	0.283
BW = full, discontinuous	–	0.586	0.950	0.126	0.586	0.475
BW = FG, discontinuous	2308	0.834	0.947	0.156	0.834	0.281
BW = CCT, discontinuous	1304	1.174	0.947	0.153	1.174	0.183
B. Quadratic estimators						
BW = full, continuous	–	1.944	0.778	-2.231	1.944	0.006
BW = FG, continuous	7598	1.988	0.797	-2.227	1.988	0.007
BW = CCT, continuous	1649	3.443	0.890	-2.257	3.443	0.014
BW = full, discontinuous	–	2.379	0.721	-3.320	2.379	0.000
BW = FG, discontinuous	7598	2.432	0.754	-3.317	2.432	0.000
BW = CCT, discontinuous	1649	4.235	0.866	-3.184	4.235	0.006

Table summarizes the performance of 12 RKD estimators, which differ in the degree of the underlying polynomial (linear or quadratic), bandwidth selector (full range of data, largest symmetric band, Fan-Gjbels, or Calonico et al. (2014)), and whether a discontinuity is imposed. The data are generated using a true kink of $\tau = 3.5 \times 10^{-5}$, corresponding to a semi-elasticity of 0.2 at the mandate kink point. The coverage rate is the fraction of confidence intervals containing this kink, and the rejection rate is the fraction of confidence intervals that exclude zero.

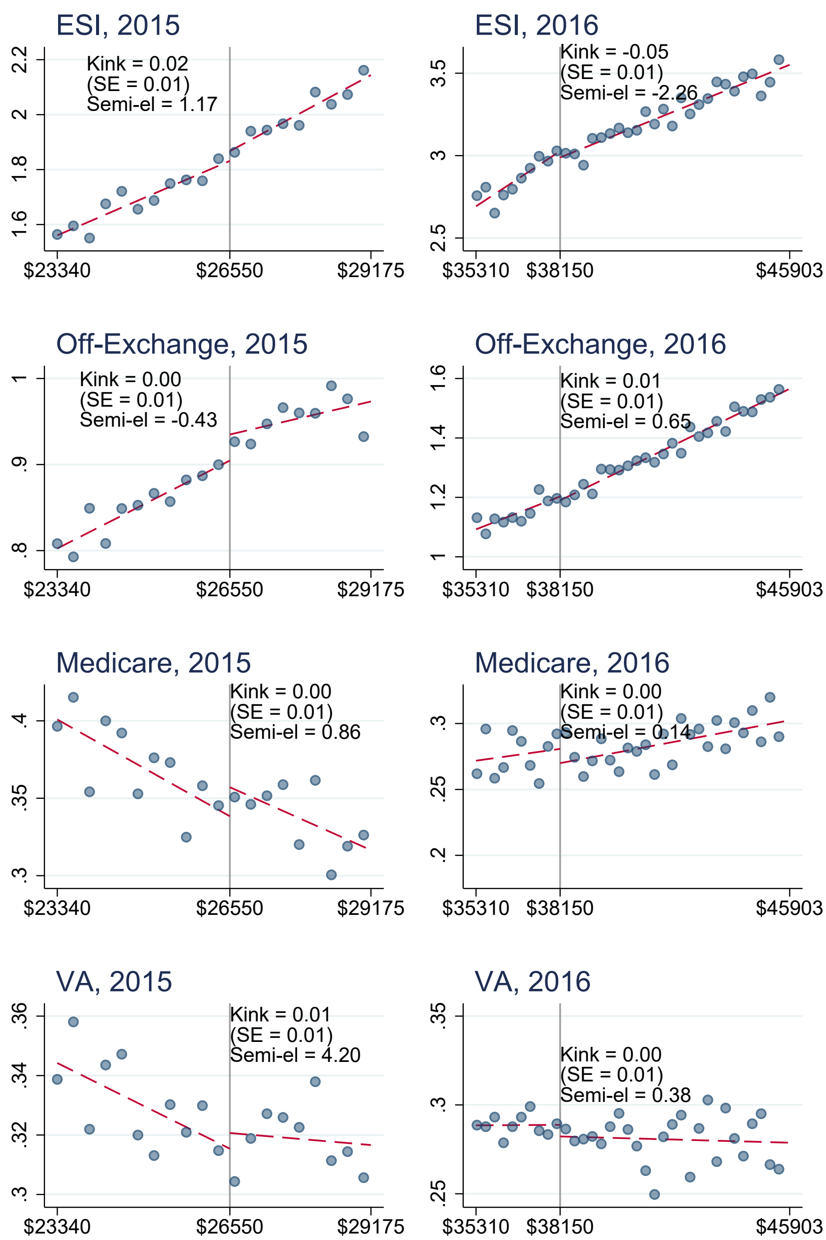
C Additional tables and figures

Figure C.1: Discontinuities in months insured, additional coverage types



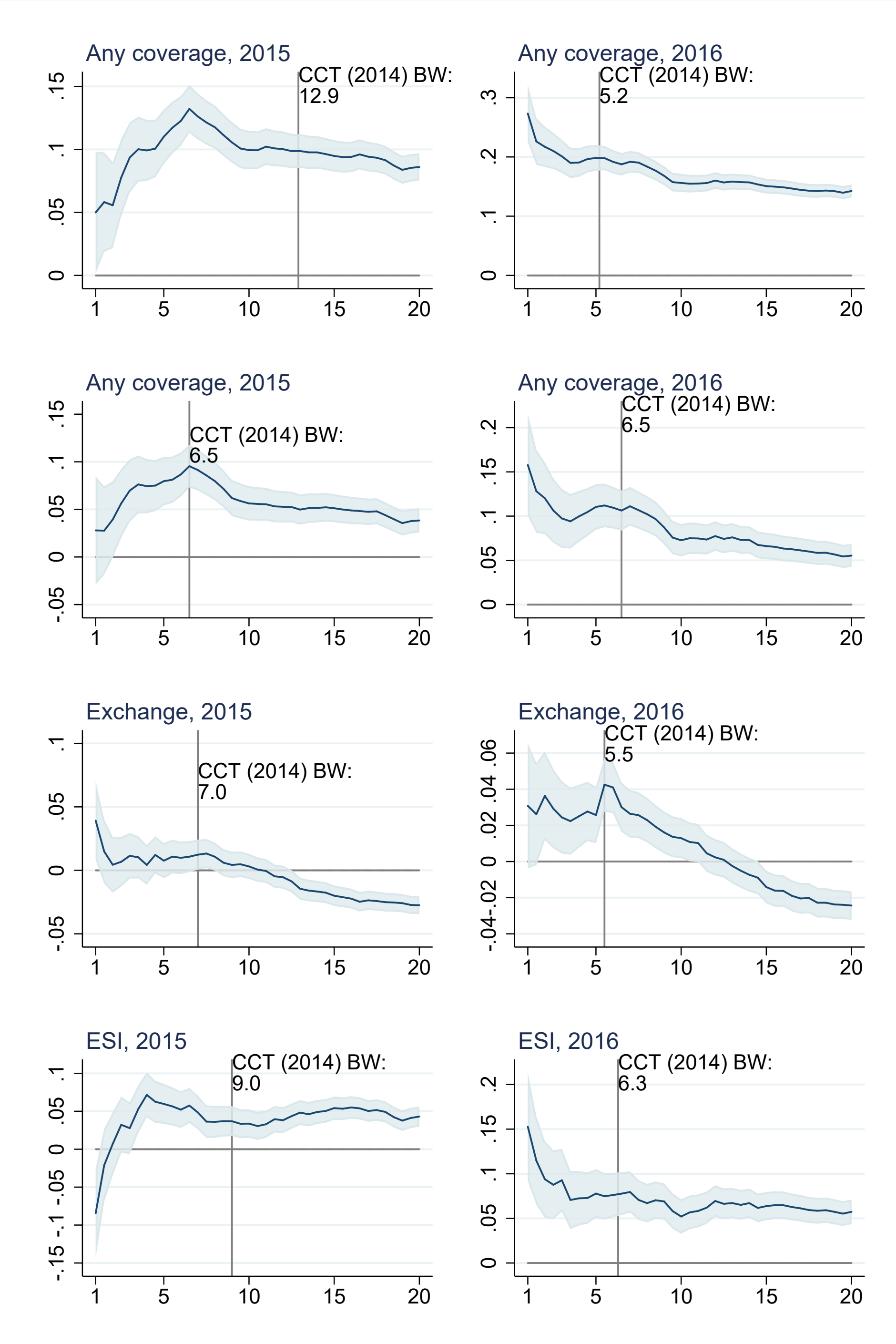
Notes: Figure shows the average number of months of insurance of the indicated type, in each 0.5 FPL point bin. Sample consists of people aged 0-64 in the indicated year, living in non-expansion states. Figure reports the estimated discontinuity, its standard error, and the implied semi-elasticity with respect to the penalty paid.

Figure C.2: Kinks in months insured, additional coverage types



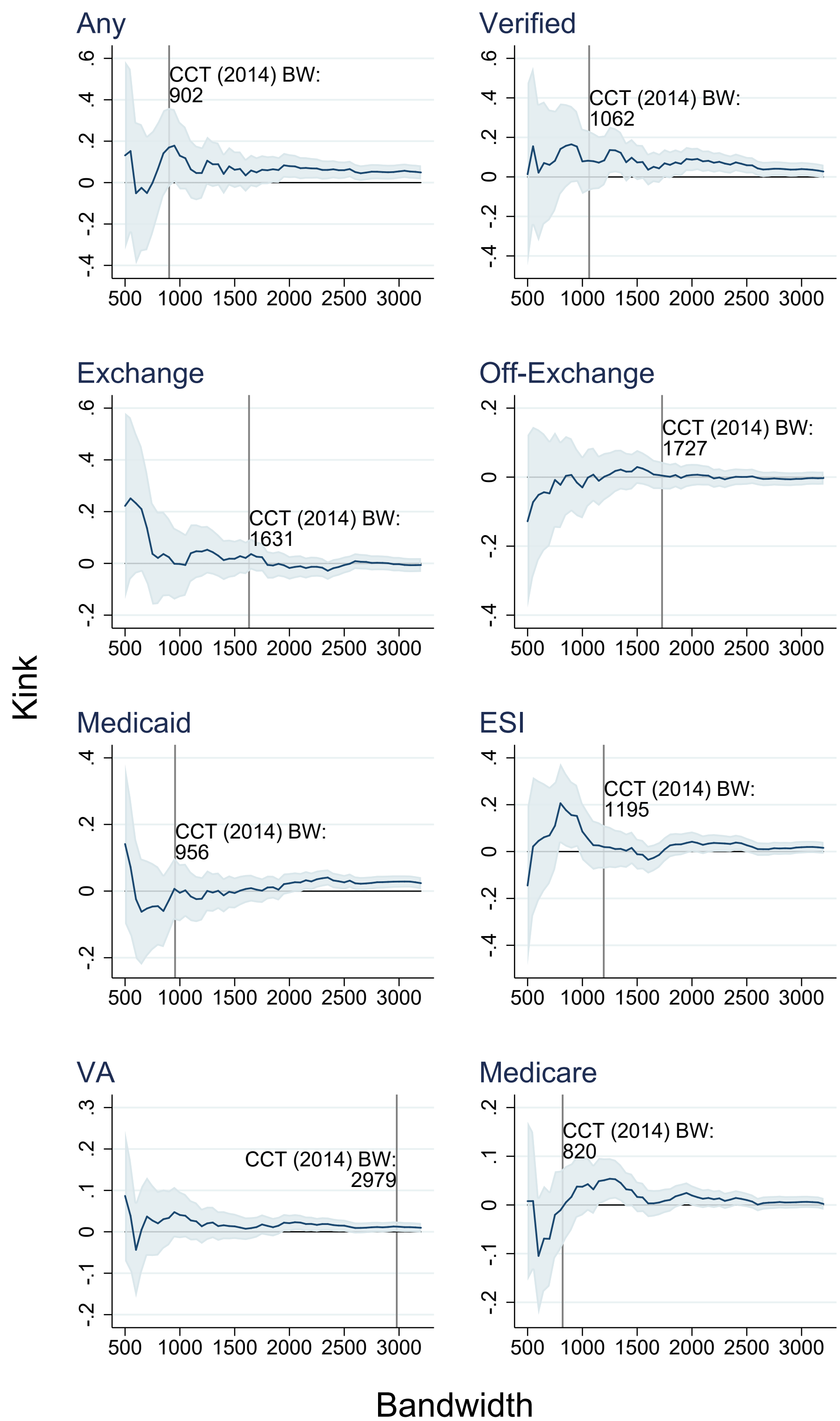
Notes: Figure shows the average number of months of insurance of the indicated type, in each $300 bin. Sample consists of people aged 27-64 in the indicated year, without signs of ESI offer, with single person tax returns and no dependents. Figure reports the estimated kink, its standard error, and the implied semi-elasticity with respect to the penalty paid.

Figure C.3: Estimated discontinuity as a function of bandwidth



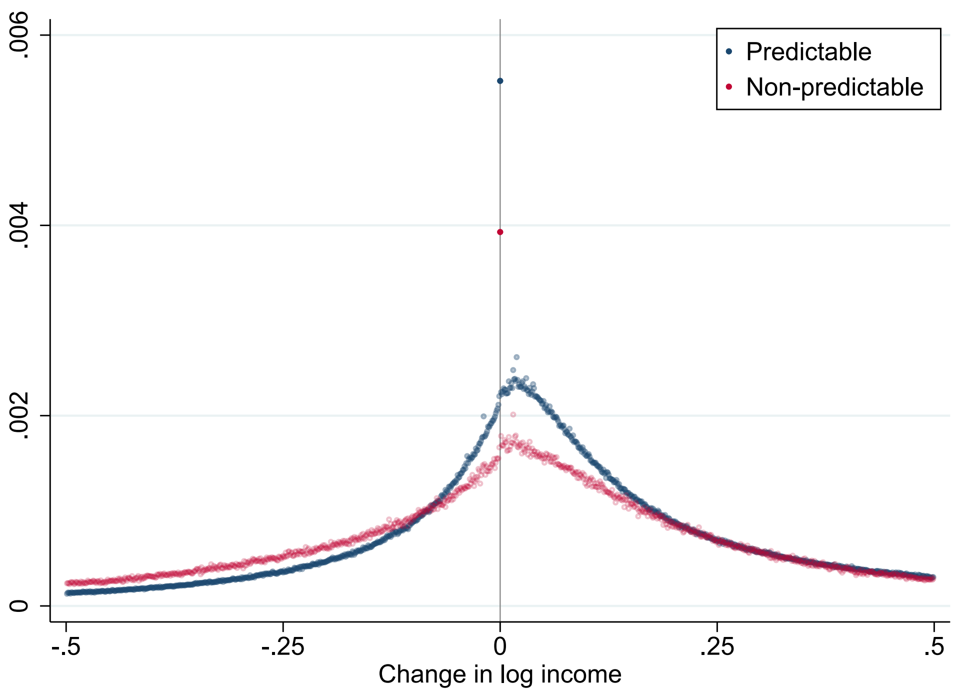
Notes: Figure shows estimated discontinuity (and 95% confidence interval) in months of coverage of the indicated type, as a function of bandwidth, for the indicated types of coverage. The vertical line is the MSE-optimal bandwidth of Calonico et al. (2014). The estimates in the paper use a bandwidth of 5, the largest symmetric bandwidth that avoids looking across the 133 percent of FPL discontinuity in the premium tax credit.

Figure C.4: Estimated kink as a function of bandwidth, 2015



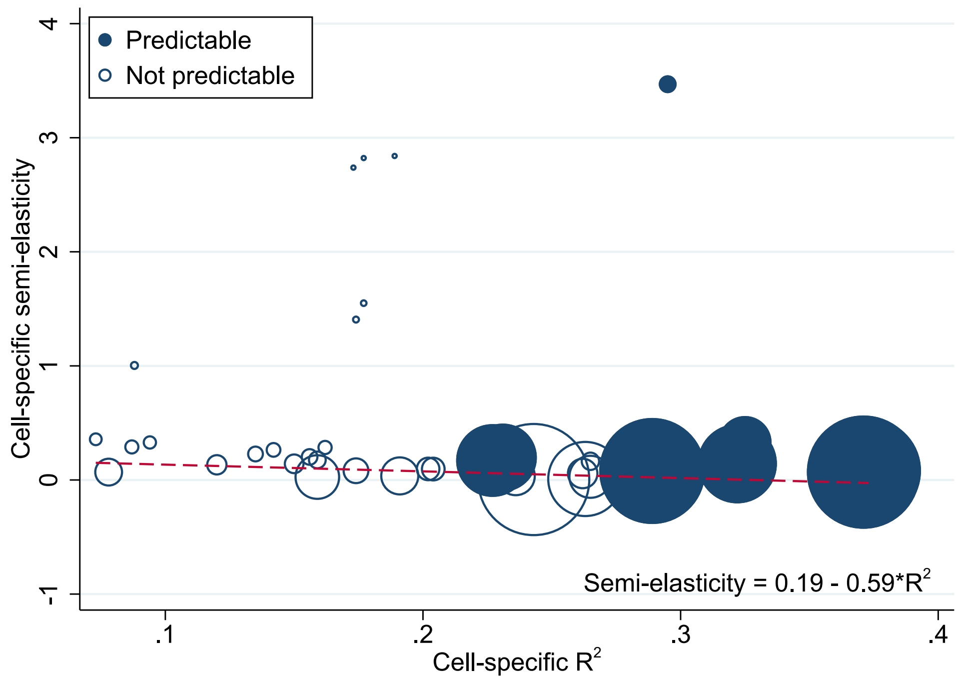
Notes: Figure shows estimated kink (and 95% confidence interval) in 2015 months covered insured as a function of bandwidth, for the indicated types of coverage. The CCT (2014) bandwidth is the MSE-optimal bandwidth of Calonico et al. (2014).

Figure C.5: Change in log income, predictable and non-predictable earners



Notes: Figure shows the distribution of the change in log income for predictable and non-predictable earners. Each bin is 0.001 points wide. Not shown are changes greater than 50 percent in absolute value. About 18 percent of the non-predictable are in this category, and 17 percent of predictable earners. The sample consists of people aged 0-64 in 2015 or 2016, with income between 110 and 160 percent of FPL, living in non-expansion states. Predictable earners are defined as having at most one W2 per earner and no unemployment insurance income.

Figure C.6: Income predictability and responsiveness to the mandate penalty



Notes: Each point is a cell defined by uninsurance receipt (yes/no), number of w2s (zero, one, or two or more), filing status (single, married filing jointly, head of household, other), and year. For each cell, we estimate semi-elasticities using our RD model and a 5-FPL point bandwidth, and we plot the estimated semi-elasticity on the y-axis. On the x-axis, we plot income predictability, measured as the R^2 from a regression of income on a degree seven polynomial in its lag. The size of each point is proportional to the inverse of the standard error of the semi-elasticity estimate. We also report the estimate from a weighted least squares regression, where the weights are the inverse of the standard error.

Table C.1: Robustness of 2015 RD estimates to specification choices

	(1)	(2)	(3)	(4)
A. Y = Months any coverage				
Discontinuity	0.110	0.072	0.058	0.105
	(0.011)	(0.017)	(0.022)	(0.011)
Semi-elasticity	0.230	0.147	0.118	0.200
B. Y = Months verified coverage				
Discontinuity	0.080	0.055	0.028	0.074
	(0.013)	(0.019)	(0.026)	(0.013)
Semi-elasticity	0.203	0.138	0.069	0.160
C. Y = Months Exchange				
Discontinuity	0.008	0.012	0.009	0.012
	(0.007)	(0.011)	(0.014)	(0.007)
Semi-elasticity	0.188	0.287	0.218	0.552
D. Y = Months ESI				
Discontinuity	0.060	0.034	-0.040	0.065
	(0.013)	(0.020)	(0.027)	(0.013)
Semi-elasticity	0.307	0.172	-0.199	0.683
E. Y = Months Medicaid				
Discontinuity	0.024	-0.002	0.041	0.006
	(0.013)	(0.019)	(0.025)	(0.011)
Semi-elasticity	0.154	-0.012	0.258	0.018
F. Y = Months off-Exchange				
Discontinuity	-0.005	-0.007	-0.012	-0.004
	(0.004)	(0.006)	(0.007)	(0.004)
Semi-elasticity	-0.456	-0.650	-1.174	-0.271
Degree	Linear	Quadratic	Cubic	Linear
Controls	No	No	No	Yes

Notes: Table shows robustness of the regression discontinuity coverage estimates to alternative specifications (polynomial degree) or controls. Column (1) is the base estimates. The degree specification controls for polynomials of the indicated degree, allowed to vary on either side of the discontinuity. The controls in column (4) are a female dummy, a quadratic in age, and dummies for filing status and number of exemptions. The sample consists of people with income between 133 and 143 percent of FPL, aged 0-64, living in non-expansion states.

Table C.2: Robustness of 2016 RD estimates to specification choices

	(1)	(2)	(3)	(4)
A. Y = Months any coverage				
Discontinuity	0.198	0.203	0.243	0.194
	(0.011)	(0.016)	(0.022)	(0.011)
Semi-elasticity	0.217	0.235	0.287	0.194
B. Y = Months verified coverage				
Discontinuity	0.110	0.098	0.128	0.106
	(0.013)	(0.020)	(0.026)	(0.013)
Semi-elasticity	0.145	0.137	0.180	0.120
C. Y = Months Exchange				
Discontinuity	0.026	0.029	0.029	0.031
	(0.008)	(0.012)	(0.016)	(0.008)
Semi-elasticity	0.301	0.360	0.359	0.642
D. Y = Months ESI				
Discontinuity	0.078	0.088	0.118	0.085
	(0.014)	(0.021)	(0.028)	(0.014)
Semi-elasticity	0.212	0.254	0.350	0.549
E. Y = Months Medicaid				
Discontinuity	0.006	-0.020	-0.034	-0.015
	(0.013)	(0.020)	(0.027)	(0.011)
Semi-elasticity	0.018	-0.067	-0.115	-0.021
F. Y = Months off-Exchange				
Discontinuity	-0.001	-0.002	0.002	-0.000
	(0.004)	(0.005)	(0.007)	(0.004)
Semi-elasticity	-0.061	-0.157	0.147	-0.002
Degree	Linear	Quadratic	Cubic	Linear
Controls	No	No	No	Yes

Notes: Table shows robustness of the regression discontinuity coverage estimates to alternative specifications (polynomial degree) or controls. Column (1) is the base estimates. The degree specification controls for polynomials of the indicated degree, allowed to vary on either side of the discontinuity. The controls in column (4) are a female dummy, a quadratic in age, and dummies for filing status and number of exemptions. The sample consists of people with income between 133 and 143 percent of FPL, aged 0-64, living in non-expansion states.

Table C.3: Robustness of 2015 RK estimates to specification choices

	(1)	(2)	(3)	(4)	(5)	(6)
A. Y = Months any coverage						
Kink	0.052	0.112	0.073	0.050	0.047	0.056
	(0.017)	(0.069)	(0.171)	(0.017)	(0.017)	(0.019)
Semi-elasticity	0.930	1.030	1.503	0.887	0.602	1.095
B. Y = Months verified coverage						
Kink	0.029	0.210	0.061	0.030	0.023	0.027
	(0.017)	(0.070)	(0.174)	(0.017)	(0.017)	(0.019)
Semi-elasticity	0.647	2.394	1.571	0.655	0.304	0.677
C. Y = Months Exchange						
Kink	-0.004	0.021	0.067	-0.007	-0.008	-0.001
	(0.013)	(0.054)	(0.135)	(0.013)	(0.013)	(0.016)
Semi-elasticity	-0.247	0.641	4.692	-0.428	-0.517	-0.033
D. Y = Months ESI						
Kink	0.017	0.054	0.103	0.020	0.016	0.007
	(0.013)	(0.051)	(0.129)	(0.013)	(0.013)	(0.005)
Semi-elasticity	1.232	2.061	8.852	1.449	0.441	5.174
E. Y = Months Medicaid						
Kink	0.023	0.027	-0.169	0.021	0.023	0.020
	(0.009)	(0.036)	(0.089)	(0.009)	(0.009)	(0.010)
Semi-elasticity	3.504	2.122	-29.318	3.283	3.846	2.670
F. Y = Months off-Exchange						
Kink	-0.003	0.026	0.004	-0.001	-0.003	0.003
	(0.010)	(0.038)	(0.095)	(0.009)	(0.010)	(0.011)
Semi-elasticity	-0.409	1.986	0.625	-0.160	-0.446	0.384
Degree	Linear	Quadratic	Cubic	Linear	Linear	Linear
Controls	No	No	No	No	Yes	No
Discontinuity	Yes	Yes	Yes	No	Yes	Yes
Include ESI?	Yes	Yes	Yes	Yes	Yes	No

Notes: Table shows robustness of the regression kink coverage estimates to alternative specifications (polynomial degree), controls, and samples. Column (1) is the base estimates. The degree specification controls for polynomials of the indicated degree, allowed to vary on either side of the kink. In column (4) we impose continuity. The controls in column (5) are a female dummy and a quadratic in age. The sample consists of people with income between 200 and 250 percent of FPL, aged 27-64, without signs of ESI offers, single filing status, and one exemption.

Table C.4: Robustness of 2016 RK estimates to specification choices

	(1)	(2)	(3)	(4)	(5)	(6)
A. Y = Months any coverage						
Kink	0.020	-0.079	-0.161	0.026	0.016	0.050
	(0.015)	(0.060)	(0.150)	(0.012)	(0.014)	(0.018)
Semi-elasticity	0.282	-0.646	-4.227	0.376	0.158	0.844
B. Y = Months verified coverage						
Kink	-0.002	-0.164	-0.029	0.008	-0.007	0.030
	(0.016)	(0.063)	(0.158)	(0.013)	(0.015)	(0.018)
Semi-elasticity	-0.036	-1.605	-0.935	0.146	-0.077	0.667
C. Y = Months Exchange						
Kink	0.021	0.056	-0.126	0.014	0.019	0.013
	(0.012)	(0.048)	(0.120)	(0.009)	(0.012)	(0.015)
Semi-elasticity	1.292	1.997	-14.230	0.891	1.409	0.610
D. Y = Months ESI						
Kink	-0.046	-0.193	0.005	-0.036	-0.049	-0.004
	(0.014)	(0.056)	(0.140)	(0.011)	(0.014)	(0.005)
Semi-elasticity	-1.850	-4.281	0.332	-1.486	-0.850	-1.940
E. Y = Months Medicaid						
Kink	0.012	-0.024	0.044	0.014	0.012	0.015
	(0.007)	(0.027)	(0.067)	(0.005)	(0.007)	(0.009)
Semi-elasticity	2.604	-2.840	17.318	3.185	2.120	2.610
F. Y = Months off-Exchange						
Kink	0.006	-0.063	-0.062	0.012	0.006	0.008
	(0.010)	(0.039)	(0.098)	(0.008)	(0.010)	(0.013)
Semi-elasticity	0.628	-3.495	-11.203	1.251	0.743	0.614
Degree	Linear	Quadratic	Cubic	Linear	Linear	Linear
Controls	No	No	No	No	Yes	No
Discontinuity	Yes	Yes	Yes	No	Yes	Yes
Include ESI?	Yes	Yes	Yes	Yes	Yes	No

Notes: Table shows robustness of the regression kink coverage estimates to alternative specifications (polynomial degree), controls, and samples. Column (1) is the base estimates. The degree specification controls for polynomials of the indicated degree, allowed to vary on either side of the kink. In column (4) we impose continuity. The controls in column (5) are a female dummy and a quadratic in age. The sample consists of people with income between 300 and 390 percent of FPL, aged 27-64, without signs of ESI offers, single filing status, and one exemption.

Table C.5: Estimated kinks and discontinuities in extensive margin of coverage

Coverage type	Any (1)	Verified (2)	Exchange (3)	Off-Exchange (4)	ESI (5)	Medicaid (6)	Medicare (7)	VA (8)
	A. 2015 Discontinuity Sample							
Discontinuity	0.007	0.005	0.000	-0.001	0.005	0.002	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Semi-elasticity	0.17	0.13	0.01	-0.58	0.27	0.12	0.60	-0.17
	B. 2016 Discontinuity Sample							
Discontinuity	0.014	0.010	0.002	-0.000	0.007	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Semi-elasticity	0.17	0.13	0.25	-0.07	0.20	0.03	0.45	0.19
	C. 2015 Kink Sample							
Kink	0.0048	0.0033	-0.0008	-0.0000	0.0002	0.0022	0.0009	0.0012
	(0.0014)	(0.0015)	(0.0013)	(0.0009)	(0.0005)	(0.0009)	(0.0005)	(0.0012)
Semi-elasticity	0.960	0.770	-0.458	-0.071	0.870	2.966	4.248	0.934
	D. 2016 Kink Sample							
Kink	0.0014	0.0003	0.0016	0.0010	-0.0000	0.0014	0.0001	-0.0039
	(0.0012)	(0.0013)	(0.0011)	(0.0009)	(0.0004)	(0.0007)	(0.0004)	(0.0012)
Semi-elasticity	0.237	0.058	1.004	1.129	-0.038	2.595	0.693	-1.694

Notes: The dependent variable is an indicator for at least one month of coverage of the indicated type. The table reports RD and RK estimates. The RD sample consists of people aged 0-64 with income between 133 and 143 percent of FPL, living in non-expansion states. The RK sample consists of people with single tax returns, one exemption, aged 27-64, without sings of an ESI offer, and income between 200 and 250 percent of FPL (in 2015) or 300 and 390 percent of FPL (in 2016).

D Further evidence on missing density

In our main discontinuity samples, we find statistically significant discontinuities in the density of the running variable. The densities do not indicate manipulation of the running variable, because there is no evidence for excess mass above the threshold. However they do indicate missing observations above the threshold. If this missing mass is driven by strategic non-filing of tax returns by uninsured people trying to avoid paying the penalty, then our RD estimates will be biased. Such strategic non-filing is unlikely to drive our results, however, because most households have a strong incentive to file a tax return, even if it means paying the penalty. Specifically, households can only claim their refundable tax credits if they file a tax return. The two most important of these credits are the earned income tax credit (EITC) and the child tax credit. The size of these credits depends on income and family structure. For households with income near 138 percent of the poverty line, we report in Appendix Table D.1 the sum of these credits, along with the maximum mandate penalty the household could face in 2016.[32] The table shows that, for many household structures, the refundable tax credits exceed the mandate penalty even in the worst case, sometimes by a large amount. Therefore there is no incentive to avoid filing, at least for some households.

This logic suggests that we should see lower filing discontinuities in households with a stronger incentive to file. In column (2) of the table, we report the discontinuity in the distribution of the running variable (in percents, so that differences are comparable across groups) in 2016. (We focus on 2016 because we have the largest first stage here.) Households with children have small distribution discontinuities, especially single-parent households. Often these discontinuities are statistically insignificant and in some cases they are positive. If differential selection explained our results, we would expect to see no coverage discontinuity where there is no density discontinuity. But as the table shows, nearly every group exhibits a coverage discontinuity, and in fact the group with the largest coverage discontinuity shows an upward discontinuity in the density of the running variable.

[32]These calculations assume that all income in MAGI is earned income and that all children in the household are 17 or younger. The credits vary with household composition for two reasons. First, the child tax credit pays $1000 per child. Second, the EITC depends on absolute income, not income relative to FPL. As the family size increases but we fix income at 138 percent of FPL, absolute income increases, because the poverty line increases. This reduces the EITC.

Table D.1: Tax credits, density discontinuities, and coverage discontinuities, 2016, by family type

Family composition		Credits & Penalty		Density discontinuity		Coverage discontinuity		
# Adults (1)	# kids (2)	Credits (3)	Penalty (4)	Estimate (5)	(SE) (6)	Estimate (7)	(SE) (8)	Semi-elasticity (9)
1	0	0	695	-0.028	(0.008)	0.308	(0.029)	0.371
1	1	3789	1043	-0.001	(0.006)	0.090	(0.027)	0.068
1	2	5622	1390	-0.005	(0.008)	0.010	(0.025)	0.007
1	3	6139	1738	-0.020	(0.013)	0.143	(0.049)	0.182
1	4	5262	2085	0.046	(0.022)	0.283	(0.113)	0.895
2	0	1894	1390	-0.065	(0.008)	0.265	(0.038)	0.316
2	1	3442	1738	-0.032	(0.008)	0.249	(0.036)	0.184
2	2	3959	2085	-0.016	(0.010)	0.186	(0.032)	0.161
2	3	3779	2085	-0.039	(0.011)	0.324	(0.039)	0.434
2	4	4000	2085	-0.044	(0.014)	0.132	(0.048)	0.237

Notes: Table shows the maximum tax credits (EITC plus CTC) that a family of the indicated type would be eligible for in 2016, if they had income at 138 percent of FPL, along with the estimated discontinuity in the density of the running variable (in percent) and coverage. Robust standard errors in parentheses. The sample consists of people with income between 133 and 143 percent of FPL, aged 0-64, living in non-expansion states.

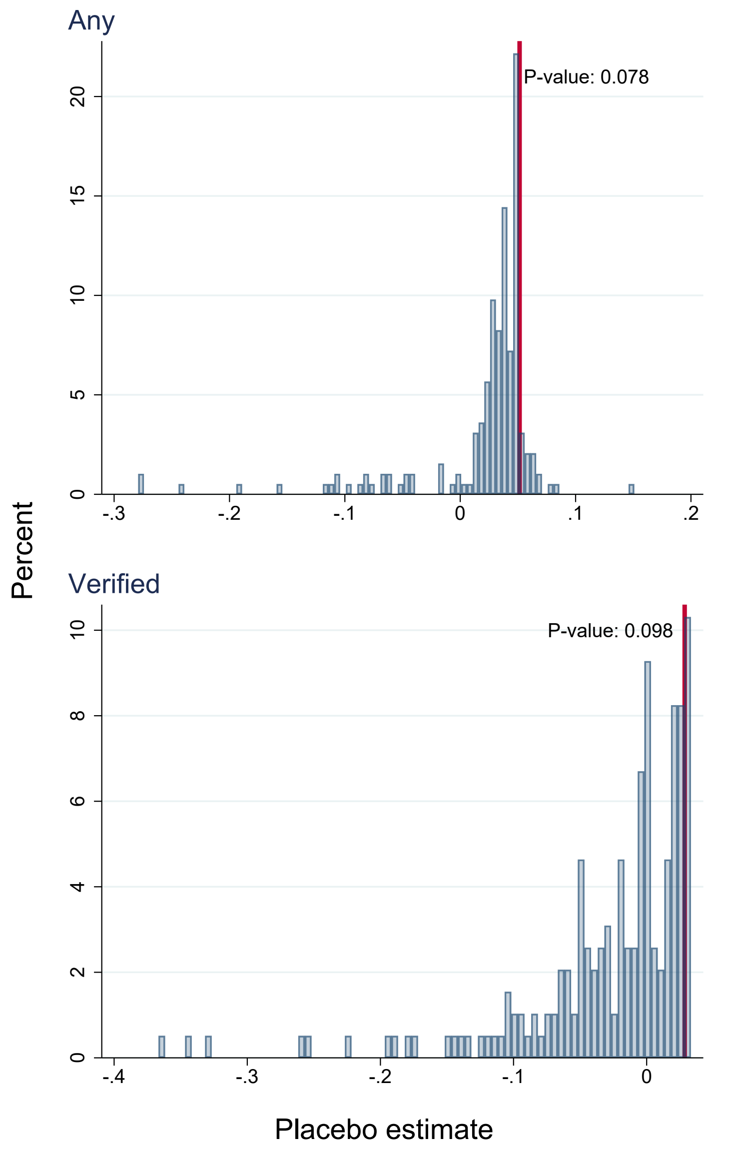
E Permutation tests

We have found clear kinks in months of any insurance coverage and months of verified coverage in 2015. One concern with the regression kink approach, however, is that it may detect spurious kinks, simply due to curvature in the relationship between the outcome and running variable (Ganong and Jger, 2018). We assess this concern by re-estimating our RKD models, but varying the kink point across a fine grid of placebo locations. If the kink is spurious, then we expect that our estimate is unexceptional in the distribution of placebo estimates.

Figure E.1 shows the distribution of placebo kink estimates, for any coverage and for verified coverage. We consider permutation kinks every \$25, starting from \$500 above 200 percent of FPL, and ending at \$500 below 250 percent of FPL. We look in this range because we do not believe looking elsewhere in the income distribution would be informative about the possibility of a false positive at our income level. There are likely to be other policy-induced kinks elsewhere in the income distribution (for example, because of the PTC). We exclude kink points near the boundaries because estimating a kink near the boundary produces very large, very noisy estimates, because there is very little data with which to estimate a slope on one side of the kink.

The histograms show, first, a long left tail of placebo kink points. This is generated by the fact that placebo kink locations near the boundaries tend to produce large, negative placebo estimates. Second, the estimated kink, shown with the vertical line, is larger than all but a handful of the placebo kinks. The implied p-value—the fraction of placebo point estimates that exceed the true point estimate—is 0.078 for any coverage and 0.098 for verified coverage. The reader may worry that these p-values are small in part because of the inclusion of the many very negative placebo kink points estimated near the boundary. If we instead estimate p-values, but excluding placebo kink points within \$1000 of the boundary, we obtain p-values of 0.058 for any coverage and 0.118 for verified coverage.

Figure E.1: Estimated kink in months insured at placebo kink points, 2015



Notes: Figure shows estimated kink in 2015 months insured at placebo kink points. The p-value is the fraction of placebo kinks that exceed the true estimate.

F Exploring the ESI offer sample

Our main kink sample excludes people with signs of an ESI offer. Hence, our main estimates do not reflect two channels through which a greater mandate penalty could affect insurance coverage: by changing ESI offer rates, or by changing take up among people offered ESI. Here we present evidence that neither of these channels is quantitatively important, at least in the context of the kink. To do so, we expand our kink samples to include people with signs of ESI offers.

We begin by showing that there is no kink in the probably of having a sign of ESI offer at the mandate kink point. Appendix Figure F.1 plots the fraction of people who have a sign of an ESI offer, as a function of income. In both 2015 and 2016 the offer rate is increasing in income but essentially smooth through the mandate kink point. In 2015 the estimated kink is -0.06, with a standard error of 0.18, meaning that each thousand dollars of income above the kink point reduces the offer rate by 0.06 percentage points—an economically small, statistically insignificant and wrong-signed amount. In 2016, the estimated kink is -0.21 (standard error=0.13)—larger in absolute value, but still small and statistically insignificant. This evidence shows that a kink in the ESI offer rate does not bias our kink sample results (through endogenously changing our sample in the neighborhood of the kink). We emphasize that this evidence is not particularly informative about the effect of the individual mandate on ESI offers, as it is unlikely that employers could tailor ESI offer to individual employees, and so we should not expect to see a sharp kink in the ESI offer rate.

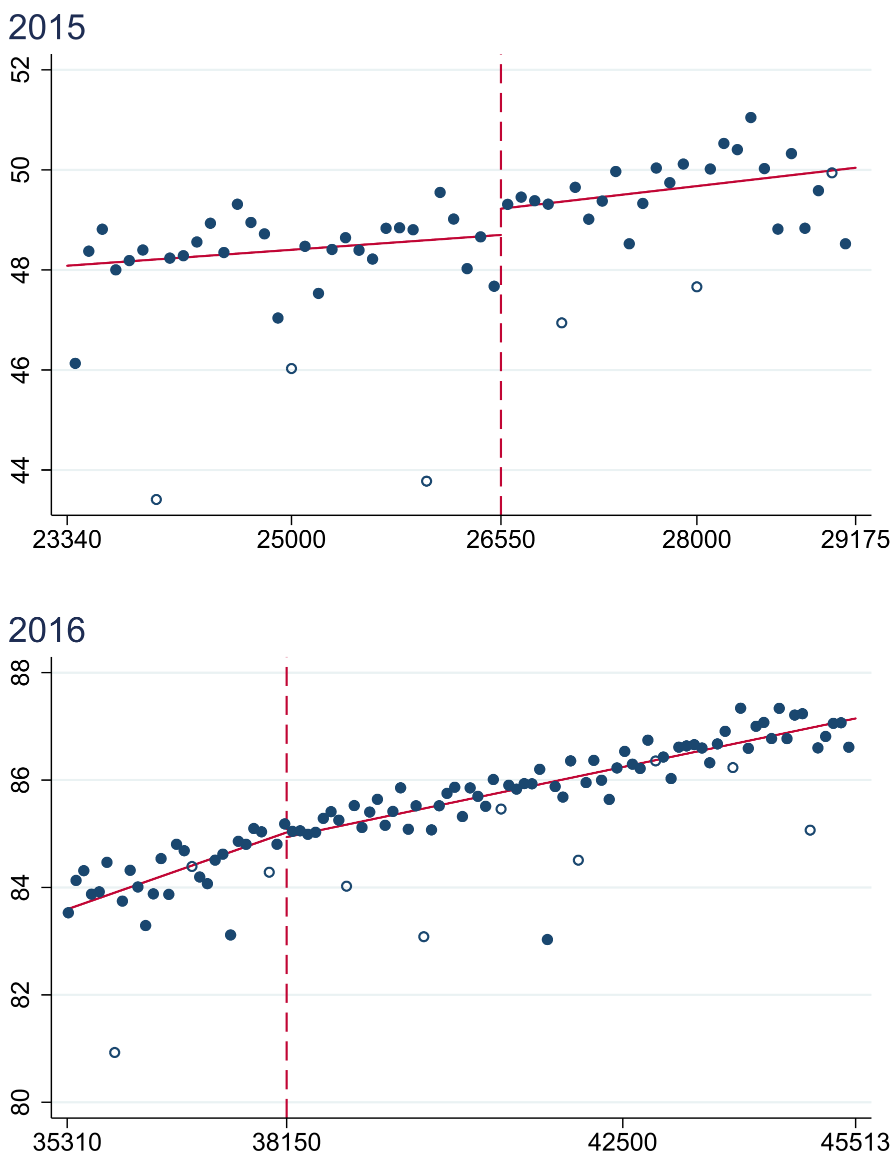
Next we show the effect of a greater mandate penalty among people with signs of an ESI offer. Appendix Figure F.2 shows months of any coverage among people with signs of an ESI offer. The figure reveals several important patterns. First, and unsurprisingly, coverage is much higher in the offer sample than in the no-offer sample. Second, there is some evidence curvature. Looking above the kink point only, for example, the slope seems to decline as income rises. Third, looking locally to the kink point, the slope in 2015 appears flatter below the kink than above it, although in 2016 the pattern is ambiguous. Taken together, these patterns indicate that perhaps there is a slight kink in 2015, but any kink will be difficult to detect and may be sensitive to the assumed polynomial and bandwidth.

In Appendix Table F.1, we report estimated RK models for the offer sample. We begin in column (1) with the linear specification using the full range of data, which is our main specification for the no-offer sample. The estimates are negative and statistically significant in both years. However, the linear/full range of data specification may not be appropriate for the ESI offer sample, with the evident curvature and larger sample size. When we use the asymptotically optimal bandwidth, the 2015 kink becomes positive and statistically insignificant (but not small, with a semi-elasticity of 0.33); the 2016 kink remains similar (because the optimal bandwidth uses nearly all the data).

One potential concern with the specifications in columns (1) and (2) is that they do not adequately control for the concavity evident in Figure F.2. This concavity could bias us toward finding a negative kink. In the remaining columns we therefore control for quadratic and cubic functions of the running variable. The quadratic specification produces kinks and positive point estimates in 2015 and negative estimates in 2016, both insignificant. However the 2016 first stage disappears when we use control for a quadratic in income. Likewise we see in columns (5) and (6) that the first stage is gone in both 2015 and in 2016 when we control for a cubic in income. It is therefore impossible to interpret the reduced form kink in these specifications.

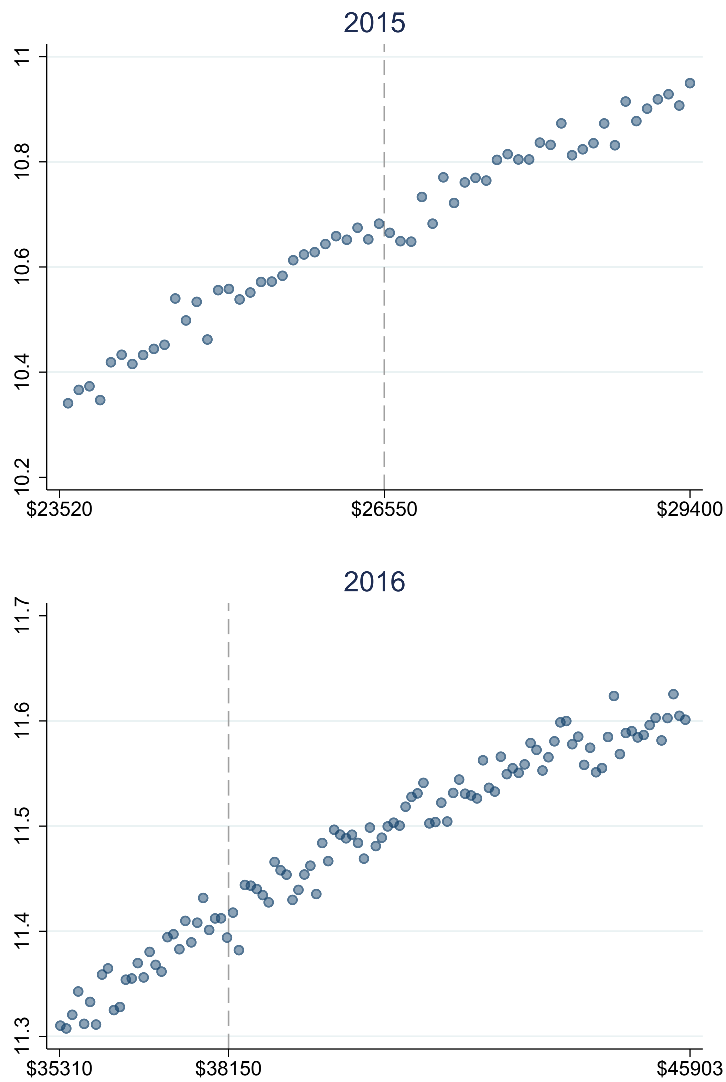
Overall, therefore, the ESI results are sensitive to specification choices. Based on the concavity in Figure F.2, we would prefer specifications that control for higher order terms. These specifications yield insignificant first stages and coverage kinks that are very noisy. Given this ambiguous evidence, we remain agnostic about the effect of the penalty among people with ESI offers at higher incomes. At lower incomes, we pool the offer and non-offer sample, and we find significant and positive effects on ESI coverage, suggesting that the mandate raises coverage even in ESI, at least at lower income levels.

Figure F.1: Probability of sign of ESI offer as a function of income



Notes: Figure shows the fraction of people with a sign of ESI offer, in each \$100 bin of modified adjusted gross income. Sample consists of people aged 27-64 in the indicated year, who filed single tax returns with no dependents. The hollow circles indicate round number incomes (\$1000 multiples); such incomes are much more common among the self-employed, who lack signs of an ESI offer. We include dummies for "round number incomes" when estimate RKD models for signs of an ESI offer.

Figure F.2: Months insured as a function of income, signs-of-ESI-offer sample



Notes: Figure shows the average number of months of any insurance, in each $100 bin of modified adjusted gross income. Sample consists of people aged 27-64 in the indicated year, who filed single tax returns with no dependents, but with signs of ESI offers.

Table F.1: RK estimates in the ESI offer sample

Degree	Linear		Quadratic		Cubic	
Bandwidth	Full (1)	CCT (2)	Full (3)	CCT (4)	Full (5)	CCT (6)
A. 2015, Y= penalty per month						
Kink	0.801	0.801	1.017	1.017	0.056	0.056
	(0.070)	(0.070)	(0.270)	(0.270)	(0.675)	(0.675)
BW	–	5603	–	5123	–	4696
B. 2015, Y= Months any coverage						
Kink	-0.020	0.017	0.022	0.117	0.111	0.145
	(0.005)	(0.015)	(0.019)	(0.042)	(0.047)	(0.049)
Semi-elasticity	-0.242	0.331	0.208	-45.402	19.163	7.801
BW	–	1287	–	1664	–	2914
C. 2015, Y= Months verified coverage						
Kink	-0.031	0.000	-0.005	0.047	0.118	0.133
	(0.006)	(0.018)	(0.023)	(0.048)	(0.057)	(0.061)
Semi-elasticity	-0.432	0.008	-0.060	2.245	23.473	6.994
BW	–	1342	–	1750	–	2839
D. 2016, Y= penalty per month						
Kink	0.780	0.297	0.020	0.020	-0.963	-0.963
	(0.161)	(0.210)	(0.629)	(0.629)	(1.442)	(1.442)
BW	–	3026	–	8557	–	14040
E. 2016, Y= Months any coverage						
Kink	-0.022	-0.019	-0.023	-0.022	-0.004	-0.004
	(0.004)	(0.004)	(0.014)	(0.015)	(0.035)	(0.035)
Semi-elasticity	-0.258	-0.256	-10.702	1.044	0.041	0.040
BW	–	5399	–	4891	–	7374
F. 2016, Y= Months verified coverage						
Kink	-0.036	-0.035	-0.051	-0.056	-0.057	-0.062
	(0.005)	(0.010)	(0.019)	(0.022)	(0.047)	(0.048)
Semi-elasticity	-0.452	-1.105	-25.428	1.507	0.580	0.554
BW	–	2072	–	3860	–	6168

Table reports the estimated kink for the indicated year and outcome. For each outcome, we report linear, quadratic, and cubic specifications (meaning we control for separate linear, quadratic, or cubic specifications), and we report estimates using the full range of data and the bandwidth of Calonico et al. (2014) ("CCT"). When calculate coverage semi-elasticities, we use the same bandwidth for calculating the first stage as we use for estimating the kink. The sample consists of single tax returns in 2015 or 2016 with one exemption claimed and a sign of an ESI offer, aged 27-64, with income between 200 and 250 percent of FPL (in 2015), or 300 and 390 percent of FPL (in 2016). Robust standard errors in parentheses. The bandwidth we report in the full section is half the full range, but it is not symmetric.

G Digging into the Medicaid response

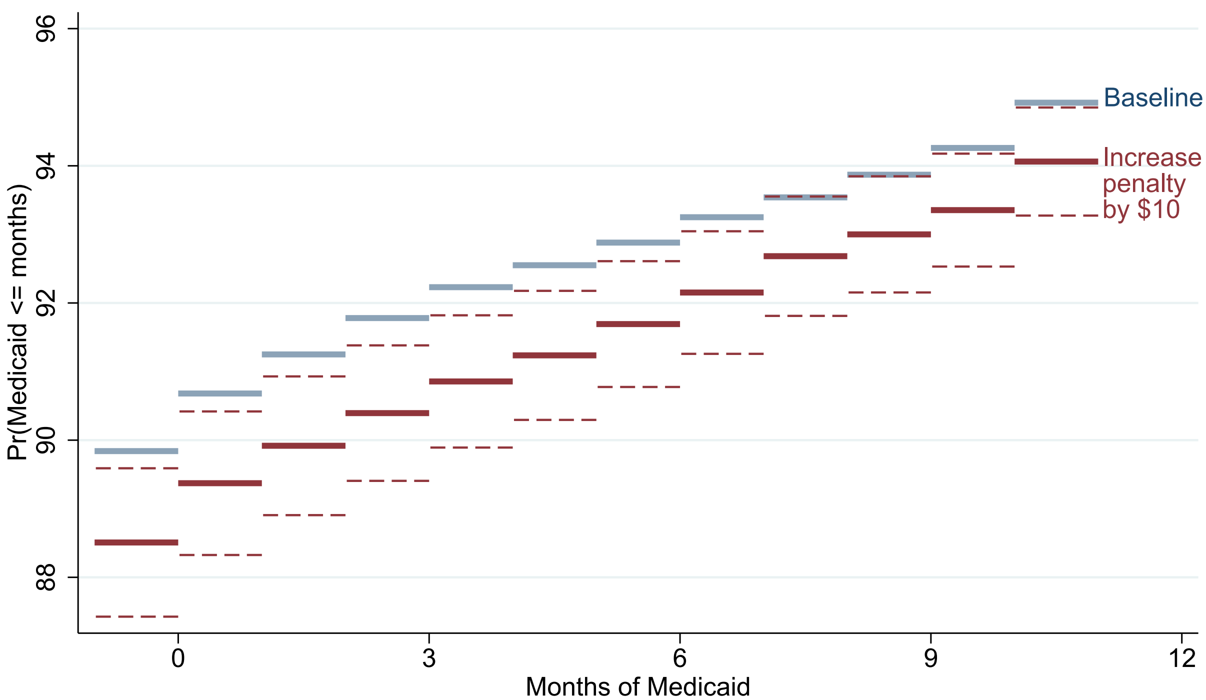
We dig further into the large Medicaid response and small individual market response by re-estimating our RK models, stratifying on Medicaid expansion status. We expect larger Medicaid responses in states that expanded Medicaid, and larger individual market responses in non-expansion states. We report the estimates in Appendix Table G.1. The Medicaid response occurs almost entirely in Medicaid expansion states, in both 2015 and in 2016. Likewise, in 2016, we find the Exchange and off-Exchange responses are concentrated in non-expansion states. In 2015, however, we find no individual market response even in non-Expansion states. The point estimates are all statistically insignificant, fairly small, and some are wrong signed. These estimates suggest that people respond to a greater mandate penalty by obtaining Medicaid coverage if at all possible. Only at fairly high income levels and in non-expansion states do we see an individual market response.

The substantial Medicaid response raise an important question: how is it that people with income above 200 percent of FPL obtain Medicaid coverage? Medicaid eligibility is assessed based on rolling income, with infrequent recertification, rather than on realized annual income. It is likely that people in our sample obtain Medicaid coverage because they found or lost a job during the year, and were temporarily eligible for Medicaid. We expect to see the biggest increases in partial year Medicaid coverage—people with a few more months of coverage, rather than an increase from 0 to 12 months of coverage. To test this hypothesis, we estimate regressions of the form

$$Pr(\text{Medicaid Months}_i <= m) = \beta_0^m + \beta_1^m v_i + \beta_2^m 1\{v_i \geq 0\} + \beta_3^m v_i 1\{v_i \geq 0\} + \varepsilon_i^m. \quad (4)$$

This is an RKD where the dependent variable is an indicator for having at most m months of Medicaid coverage. We expect to find larger effects on the probability having an intermediate number of months of coverage (1-11). This implies that we should find less negative kinks as m grows larger. We present the estimates graphically in Appendix Figure G.1, and we report the estimated kinks in Appendix Table G.2. The effect is largest for 0-5 months of coverage. Specifically we show the baseline CDF at the 2015 kink point, and the new CDF induced by a \$10 per month increase in the mandate penalty, along with the new CDF's 95% confidence interval. The baseline CDF is given by the estimates of β_m^0 from Equation 4. We obtain the new CDF by adding the implied effect of a \$10 penalty increase to the baseline CDF. The new CDF is lower everywhere than the old CDF, implying that the penalty shifts people towards more months of Medicaid. However the distance between the CDFs is greatest for relatively low months of coverage. The mandate penalty increases months of Medicaid coverage primarily at the bottom end of the coverage spectrum, pulling people up from zero months of coverage to 1-6 months coverage, with a relatively smaller effect higher up.

Figure G.1: CDF of months of Medicaid coverage, 2015



Notes: Source: Figure shows the CDF of months of Medicaid coverage at the 2015 mandate kink point ("baseline") and the counterfactual CDF induced by a $10 increase in the monthly mandate penalty, along with the 95% confidence interval.

Table G.1: Kinks in months of insurance coverage, all categories, by Medicaid expansion status

Coverage type	Any		Verified		Medicaid		Exchange		Off-Exchange	
Expanded Medicaid?	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	A. Coverage year 2015									
Kink	0.047	0.044	0.012	0.029	0.002	0.033	-0.028	0.008	0.018	-0.018
	(0.028)	(0.022)	(0.028)	(0.022)	(0.006)	(0.014)	(0.022)	(0.017)	(0.015)	(0.012)
P-value (difference)	0.934		0.643		0.048		0.193		0.074	
	B. Coverage year 2016									
Kink	0.020	0.025	-0.006	0.005	0.006	0.017	0.047	0.006	0.018	-0.001
	(0.026)	(0.018)	(0.027)	(0.020)	(0.006)	(0.010)	(0.020)	(0.015)	(0.016)	(0.012)
P-value (difference)	0.889		0.746		0.339		0.107		0.349	

Table reports the estimated kink obtained from a regression of the indicated coverage type on income, allowing for a kink and discontinuity at the mandate kink point, estimated separately by whether the state of residence has expanded Medicaid by the indicated coverage year. The sample consists of single tax returns in 2015 (Panel A) or 2016 (Panel B) with one exemption claimed and no signs of ESI offers, aged 27-64, with income between 200 and 250 percent of FPL (in 2015), or 300 and 390 percent of FPL (in 2016). Robust standard errors in parentheses. The reported p-value is the p-value of the hypothesis that the kinks are the same for the expansion and non-expansion states.

Table G.2: RKD estimates for CDF of Medicaid months, 2015

	Kink (1)	Standard error (2)	Penalty effect (3)
Pr(Medicaid Months $\leq$0)	-0.222	(-0.092)	-0.133
Pr(Medicaid Months $\leq$1)	-0.218	(-0.089)	-0.131
Pr(Medicaid Months $\leq$2)	-0.222	(-0.086)	-0.133
Pr(Medicaid Months $\leq$3)	-0.231	(-0.084)	-0.139
Pr(Medicaid Months $\leq$4)	-0.229	(-0.082)	-0.137
Pr(Medicaid Months $\leq$5)	-0.219	(-0.080)	-0.131
Pr(Medicaid Months $\leq$6)	-0.198	(-0.078)	-0.119
Pr(Medicaid Months $\leq$7)	-0.183	(-0.076)	-0.110
Pr(Medicaid Months $\leq$8)	-0.143	(-0.074)	-0.086
Pr(Medicaid Months $\leq$9)	-0.145	(-0.072)	-0.087
Pr(Medicaid Months $\leq$10)	-0.151	(-0.070)	-0.091
Pr(Medicaid Months $\leq$11)	-0.143	(-0.067)	-0.086

The sample consists of single tax returns with 2015 income between 200 and 250 percent of FPL, with one exemption claimed, no signs of ESI offers, aged 27-64. Each row is a separate regression; the outcome is an indicator for having at most the indicated number of months of Medicaid coverage on income (multiplied by 100). The independent variable is 2015 income (in thousands), allowing for a kink and discontinuity at the 2015 kink point. Column (1) shows the estimated kink, column (2) shows the standard error, and column (3) shows the implied effect of an extra dollar of penalty per month, which is $kink/20 * 12$.