CARRA Working Paper Series

Working Paper 2017-07

Labor Market Effects of the Affordable Care Act: Evidence from a Tax Notch

Kavan Kucko
Cornerstone Research

Kevin Rinz
U.S. Census Bureau

Benjamin Solow
ECARES, Université Libre de Bruxelles

Center for Administrative Records Research and Applications
U.S. Census Bureau
Washington, D.C. 20233

Paper Issued: July, 2017

Disclaimer: This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Labor Market Effects of the Affordable Care Act: Evidence from a Tax Notch*

Kavan Kucko,† Kevin Rinz,‡ and Benjamin Solow§

July 12, 2017

Abstract

States that declined to raise their Medicaid income eligibility cutoffs to 138 percent of the federal poverty level (FPL) under the Affordable Care Act (ACA) created a "coverage gap" between their existing, often much lower Medicaid eligibility cutoffs and the FPL, the lowest level of income at which the ACA provides refundable, advanceable "premium tax credits" to subsidize the purchase of private insurance. Lacking access to any form of subsidized health insurance, residents of those states with income in that range face a strong incentive, in the form of a large, discrete increase in post-tax income (i.e. an upward notch) at the FPL, to increase their earnings and obtain the premium tax credit. We investigate the extent to which they respond to that incentive. Using the universe of tax returns, we document excess mass, or bunching, in the income distribution surrounding this notch. Consistent with Saez (2010), we find that bunching occurs only among filers with self-employment income. Specifically, filers without children and married filers with three or fewer children exhibit significant bunching. Analysis of tax data linked to labor supply measures from the American Community Survey, however, suggests that this bunching likely reflects a change in reported income rather than a change in true labor supply. We find no evidence that wage and salary workers adjust their labor supply in response to increased availability of directly purchased health insurance.

Keywords: Health Reform, Affordable Care Act, Bunching, Elasticity of Taxable Income, Labor Supply
JEL Classification Codes: H42, I18, J22

*This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. We thank Matthew Fiedler, Emily Gee, Daniel Hungerman, Johannes Schmieder, Johannes Spinnewijn, Victoria Udalova, John Voorheis, and Abigail Wozniak for thoughtful comments. Any remaining errors are our own.

†Cornerstone Research; kavan.kucko@gmail.com.
‡U.S. Census Bureau; kevin.rinz@census.gov.
§ECARES, Université Libre de Bruxelles; ben.solow@gmail.com.

1 Introduction

Behavioral responses to taxation are of central interest in modern public finance. Indeed, the optimal size of the public sector and progressivity of the tax system depend crucially on the compensated elasticity of labor supply. Feldstein (1995, 1999) argues that, under certain conditions, the elasticity of taxable income (ETI) is a sufficient statistic for efficiency of the tax system. Of course, labor supply responses to public policy are also of great interest to labor economists, who have also devoted decades of attention to understanding the causes and consequences of these decisions.[1] The Affordable Care Act (ACA), the largest new tax and transfer program in the United States since the Great Society, created a variety of incentives, many of which operate through the tax code, that could lead to changes in how much people work and earn. Estimating the degree to which income and work respond to those incentives is necessary to understanding how the programs that give rise to them affect the wellbeing of their participants. In this paper, we quantify the behavioral response to a specific unintended feature of the ACA that provides an incentive for certain workers to increase their earnings.

The ACA provides tax credits to subsidize the purchase of private health insurance for households with income between 100 percent and 400 percent of the federal poverty level (FPL). This form of subsidy was designed to overlap with newly increased Medicaid eligibility income limits, which the law originally raised to 138 percent FPL nationwide. Following the Supreme Court's ruling in *National Federation of Independent Business v. Sebelius*, which made the expansion of Medicaid eligibility optional for states, many states declined to expand eligibility for Medicaid as the law originally required. As a result, some groups of taxpayers in those states remain subject to Medicaid income limits far below the FPL. This created a region of the income distribution in which those taxpayers have access to neither Medicaid nor the tax credit subsidy for private health insurance, leaving them with a very strong incentive to increase their earnings to obtain subsidized health insurance. The unintended introduction of this gap in subsidized coverage provides a unique opportunity to evaluate how people respond to health insurance subsidies that operate through the tax code.

We investigate the effect of the premium tax credit on taxable income using the universe of Internal Revenue Service (IRS) Form 1040 tax returns. In states that did not expand Medicaid eligibility, the premium tax credit provides a large, discrete increase in post-tax income at the FPL, creating a large upward notch in taxpayers' post-tax budget constraints. If taxpayers relocate to that notch in order to obtain the premium tax credit, we should observe excess mass, or bunching, in the income distribution at the FPL, which we can use to estimate the ETI via a bunching estimator.

[1] See handbook chapters by Killingsworth and Heckman (1986), Pencavel (1986), and Blundell and MaCurdy (1999) for extensive, focused reviews of research on labor supply.

Visual inspection of our data reveals that some groups of taxpayers do appear to exhibit bunching behavior that leads to excess mass in the distribution of income around the premium tax credit eligibility threshold. This behavior is limited to filers with self-employment income, consistent with Saez (2010). These filers likely have greater flexibility to adjust their earnings in response to new tax incentives. In particular, bunching is most pronounced among single filers without children, the group with the most limited set of subsidized health insurance options before the ACA, and is also evident among married filers with three or fewer children. However, visual inspection also reveals other tax-induced bunching at nearby points in the income distribution. This bunching poses challenges for standard bunching approaches that identify excess mass by comparing the empirical income distribution with a cross-sectionally estimated counterfactual distribution. To overcome these challenges, we develop a new, longitudinal approach to estimating counterfactual distributions that prevents longstanding features of the tax code from distorting our estimates. Although the data requirements are intense, and the resulting elasticities do not have the same clearly-defined theoretical interpretation as those produced by other estimators, we believe that incorporating prior years' distributional information from the same locations into the generation of counterfactual distributions and the development of an approach that remains practical in the presence of other distortions in the income distribution represent useful methodological advances.

Ultimately, the ETIs we estimate range from about 0.6 to 1.0 in our preferred specification. We find no evidence that the existence of the notch we consider changes the share of filers who report positive self-employment income, suggesting that our estimates likely represent an intensive margin response. Our estimates are similar in magnitude to those reported by Saez based on analysis of the first kink in the Earned Income Tax Credit (EITC) schedule, despite substantial differences between policy and methodological details of our two settings.

We then link tax data for self-employed filers to their American Community Survey (ACS) responses regarding employment, hours worked, and weeks worked to further investigate whether the bunching behavior we observe reflects a true change in labor supply, or simply a change in reported income.[2] We believe this represents an important contribution of our paper, as we are aware of no other bunching study that moves beyond estimating the ETI to consider direct measures of labor supply. Grouping people according to their income in 2013, the final year before the introduction of the premium tax credits, we look for increases in measures of labor supply among those who would have fallen into the coverage gap in 2014 and 2015 had their income not changed. We find no evidence of such a response, suggesting that the bunching we observe is unlikely to reflect a true change in labor supply. Similarly, we find no visual evidence of changes in labor market outcomes for wage and salary workers (i.e. those who have no self-employment income and do not

[2]For more information on the ACS, see https://www.census.gov/programs-surveys/acs/.

exhibit bunching), despite changes in health insurance coverage.

A great deal of research, summarized by Currie and Madrian (1999) and Gruber (2000), considered the question of how public health insurance affects labor market outcomes, but conclusions differ across outcomes and studies. More recently, Dague et al. (2014) and Garthwaite et al. (2014) consider pre-ACA changes in Medicaid coverage for non-disabled childless adults and find substantial effects on employment outcomes. However, in their broader study of the Oregon Health Insurance Experiment, Finkelstein et al. (2012) find little effect on employment after one year. Studies of the ACA itself have tended not to find significant effects on labor market outcomes (e.g. Gooptu et al., 2016; Kaestner et al., 2016; Levy et al., 2016; Moriya et al., 2016), with the possible exception of involuntary part-time work (Dillender et al., 2016; Even and Macpherson, 2015). We focus on a different labor market incentive created by the ACA than previous studies, which have tended to consider the consequences of broader eligibility for Medicaid or the law's employer mandate. Major changes associated with those components of the law could plausibly lead to reductions in labor supply or labor demand.[3] Despite this difference, we again see little evidence of changes in labor market outcomes.

A broader set of studies have examined the behavioral responses to nonlinearities in tax liability more generally. The most closely related paper is the study of the Danish tax scheme for foreigners by Kleven et al. (2013). They study a time-limited preferential tax scheme for high-income foreign workers and find large migration elasticities, but relatively small intensive earnings elasticities. Similarly, we study an upward notch in the post-tax income schedule, but find large elasticities of taxable income on the intensive margin. Kleven et al. (2013) conclude that the small intensive response indicates little adjustment in work effort or tax avoidance. On the contrary, we find little change in labor supply relative to the changes in income in response to the tax notch.[4]

The distinction between elasticities of labor supply and taxable income is an important one for analysis of this policy. Lockwood (2016) argues that the sufficient statistics approach of Feldstein (1999) requires correcting for changes in tax revenue in a setting with tax notches.[5] Our results indicate that the welfare computations using only taxable income may overestimate the distortionary effects of the policy; even self-employed workers (those who bunch) seem not to be reducing leisure while increasing their taxable income.

[3]In the case of Medicaid, the income effect associated with receiving health insurance through the program could lead to a reduction in labor supply. The employer mandate increases the cost of full-time employees for firms of a certain size by requiring the provision of health insurance, possibly reducing labor demand.

[4]We focus on graphical evidence in this paper. We report regression discontinuity results using the preceding year's taxable income to proxy for the estimated income which determines eligibility for the premium tax credits. These estimates likely suffer from attenuation bias despite the fact that previous year's income is used as a baseline for estimating income by the health care exchanges when applying.

[5]This result stems from the fact that notches have first-order effects on revenue, while kinks do not. The Congressional Budget Office (2016) estimates subsidies from the federal government to people who procure health insurance on the non-group market (which includes the policy we study) amounts to around $900 billion, or approximately 10 percent of the total net federal health insurance subsidy over the years 2017-2026.

A few previous studies have considered the possibility that taxpayers may adjust their income in response to incentives created by health insurance subsidies provided to people with low to moderate income through the tax code. Finkelstein et al. (2017) and Gallagher et al. (2017) look for evidence of bunching in the income distribution as a preliminary matter before conducting regression discontinuity analyses of insurance take-up and home payment delinquency, respectively, and find no evidence that taxpayers are able to precisely manipulate their income. However, changes in the dollar value of the subsidy associated with crossing the thresholds Finkelstein et al. consider are much smaller than the the dollar value of becoming eligible for the premium tax credit. Also, even the larger ACS sample Gallagher et al. use may be too small to detect bunching that is concentrated in smaller groups of taxpayers. In a paper more directly related to ours, Heim et al. (2017) estimate the ETI in response to the notch induced by the fact that the value of premium tax credit falls to zero at 400 percent FPL. However, unlike the notch we consider, the 400 percent FPL notch creates an incentive for taxpayers to reduce their income in order to access the subsidy. We are aware of no prior work that finds evidence of upward income adjustment in response to tax incentives.

The rest of this paper proceeds as follows. Section 2 discusses the details of the Affordable Care Act that are relevant to our analysis. Section 3 describes our data and provides some graphical analysis. Section 4 details our methodology. Section 5 discusses our results, and Section 6 concludes.

2 Background

As originally enacted, the ACA created a comprehensive system for providing subsidized health insurance to people with income at or below 400 percent of the FPL. The law required all states to expand eligibility for their Medicaid programs to include all adults with incomes up to 138 percent FPL, with the federal government covering the vast majority of the cost of the expansion.[6]

The law also created marketplaces (sometimes called "exchanges") on which private insurers could sell policies to individuals who wished to purchase their own health insurance. Individuals with income between 100 percent and 400 percent FPL who did not have access to another affordable form of health insurance and chose to purchase health insurance on these exchanges would be eligible for refundable, advanceable "premium tax credits" designed to limit the cost of obtaining a baseline level of coverage.[7] That baseline level of coverage is referred to as a silver plan, which covers ten essential health benefits and has an actuarial

[6]The increases the Medicaid eligibility threshold to 133 percent FPL, but it includes a five percent earnings disregard, bringing eligibility to 138 percent FPL. The federal government paid 100 percent of the cost of this expansion for the first three years. That rate is scheduled to gradually decline to 90 percent in 2020.

[7]Individuals who have access to affordable employer-sponsored insurance or who are eligible for coverage through another government program such as Medicare, Medicaid, TRICARE, or CHIP are not eligible for the premium tax credit. In 2014, an employer-sponsored plan was considered affordable if the employee's share of the premium for minimally acceptable self-only coverage does not exceed 9.5 percent of household income (9.56 percent in 2015).

value of 70 percent.[8] The value of the tax credit adjusts based on household income and area of residence to limit the cost of the premium for the second-least expensive silver plan available to a given customer. Those with income at or just above the FPL pay a premium equal to 2.0 percent of household income. That percentage increases with income, reaching about 9.5 percent for those at 400 percent FPL. The value of the tax credit cannot exceed the cost of the plan it is used to purchase. Between Medicaid and the premium tax credits, the ACA made subsidized health insurance available at all levels of income up to 400 percent FPL.

In 2012, however, the Supreme Court's ruling in *National Federation of Independent Business v. Sebelius* altered this system before it went into effect. The court held that the provisions of the ACA that required states to expand eligibility for their Medicaid programs (under threat of the loss of all federal Medicaid funding if they did not) were unduly coercive and therefore unconstitutional. As a result, states were instead left with the option of expanding Medicaid eligibility, with the federal government bearing the cost of the expansion as the law originally required. Figure 1 shows which states exercised that option. As of the beginning of 2017, 31 states and the District of Columbia have expanded Medicaid eligibility. In 24 of those states and DC, the expansion took effect on January 1, 2014. The other seven expanded eligibility between April 1, 2014 (Michigan) and July 1, 2016 (Louisiana). The remaining 19 states had not yet expanded Medicaid eligibility as of July 2017.

Since most states had Medicaid income eligibility cutoffs below 100 percent FPL for non-pregnant, non-disabled adults prior to the ACA, those that decided not to expand Medicaid eligibility created regions of the income distribution between their eligibility cutoffs and the FPL in which such adults could not receive subsidized health insurance. People with income below the FPL are not eligible for premium tax credits because Medicaid coverage at that level of income was designed to be universal. This has become known as the "Medicaid gap" or "coverage gap."[9] Figure 2 illustrates this gap in stylized terms. As open enrollment in plans offered through the ACA's marketplaces began in October 2013, about 5 million uninsured adults fell into the coverage gap (Kaiser Commission on Medicaid and the Uninsured, 2013a,b).[10]

This unforseen combination of the Supreme Court striking down the requirement that Medicaid expansion be universal and political gridlock preventing Congress from adjusting the law in response created an unusually strong incentive for people with pre-tax income just below 100 percent FPL to reach that threshold, as illustrated in Figure 3. A single, non-disabled adult with no children living in a state that had not expanded eligibility for its Medicaid program and earning $11,400 in 2014 had income just under 99 percent

[8] For individuals with income between 100 percent and 250 percent FPL who purchase a silver plan, an additional Cost Sharing Reduction (CSR) subsidy is provided to limit out-of-pocket expenditures, effectively improving the actuarial value of the plan.

[9] Wisconsin has not expanded Medicaid eligibility under the ACA, but its income eligibility limit is equal to 100 percent FPL. Therefore, there is no coverage gap in Wisconsin, and we exclude it from all of our analysis.

[10] As of October 2016, that number had fallen to 2.6 million (Garfield and Damico, 2016), thanks in part to Medicaid expansion taking effect in additional states.

FPL and was ineligible for both Medicaid and the premium tax credit. If that person earned an additional \$270 and reached 100 percent FPL, she would have become eligible for a tax credit worth nearly \$3,000 on average, increasing her after-tax income by more than 25 percent through only a small change in her pre-tax income.

Because the assistance the premium tax credit provides is needed throughout the year, eligibility for it must initially be determined using a prospective measure of income. Thus, eligibility is determined based on projected income at the time a plan is selected on the marketplace website. Though marketplace customers could conceivably submit projections that differ from their actual income in order to maximize the value of the credit, multiple mechanisms limit the appeal of that approach. First, submitted projections are compared to available information about past income, and additional documentation is requested to explain large differences. Second, when tax credit recipients file their taxes for the year in question, they may be required to repay some or all of the credit they received if their realized income is sufficiently greater than the income they projected when they initially signed up for insurance. These two factors encourage alignment between projected and realized income.

Figure 4 provides a simple illustration of how the discrete increase in post-tax income that the premium tax credit provides could lead to excess mass in the distribution of income at the eligibility threshold. Suppose consumers' utility is increasing in consumption and decreasing in work effort (through which they earn pre-tax income). They choose consumption and effort to maximize utility subject to their post-tax budget constraint. In the absence of the premium tax credit, a consumer with indifference curve $U_{b,0}$ would choose effort level $z_{b,0}$. However, after the introduction of the premium tax credit, that consumer can move up to indifference curve $U_{b,1}$, which is now feasible thanks to the upward notch that the tax credit creates in the budget constraint. To reach $U_{b,1}$, the consumer chooses effort level $z_{b,1}$, the level at which pre-tax income coincides with the premium tax credit eligibility threshold. Another consumer with indifference curve $U_{a,0}$ may also choose effort level $z_{b,1}$ in the presence of the premium tax credit. However, choosing that level of effort does not give this consumer increased utility relative to the situation that would obtain in the absence of the premium tax credit. The indifference curve $U_{a,0}$ is tangent to the budget constraint below the notch and intersects the budget constraint at the notch, leaving the consumer indifferent between effort levels $z_{a,0}$ and $z_{b,1}$. This indifference curve represents a marginal buncher.

A growing literature uses excess mass near changes in key tax parameters to assess the impact of those changes on outcomes of interest (e.g. Saez, 2010; Chetty et al., 2011; Kleven and Waseem, 2013; Hungerman and Ottoni-Wilhelm, 2016; Gelber et al., 2017). Given the potential for the premium tax credit to induce bunching at the eligibility threshold, we will investigate the extent to which excess mass is in fact observed in the income distribution near that threshold and use that mass to estimate the elasticity of reported income with respect to the subsidy the credit provides. The data requirements of this type of analysis are intense; even in relatively large surveys, it is often difficult to discern bunching behavior from sampling variation. Fortunately, we are able to analyze the universe of IRS Form 1040 tax returns for the years surrounding the introduction of the premium tax credit. We are also able to link individual tax return data to responses to the ACS to obtain measures of labor supply, allowing us to analyze the labor market effects of the tax credit directly. We next describe our data and provide some graphical analysis that informs the methodological choices laid out in the following section.

3 Data

Our bunching analysis utilizes the universe of Form 1040 tax returns from tax years 2000 through 2015 as provided by the IRS to the Census Bureau's Center for Administrative Records Research and Applications. These data are ideal for this analysis for two major reasons. First, the data include all returns filed in those years, eliminating concerns about sampling error and making it easier to identify bunching behavior. Second, the income concept that is relevant to eligibility for the premium tax credit, modified adjusted gross income (MAGI) relative to the FPL, is based on income that is reported on Form 1040. MAGI is the sum of adjusted gross income, non-taxable social security benefits, tax-exempt interest income, and excluded foreign income. Our data include all of these components except excluded foreign income. Fewer than 500,000 taxpayers reported excluded foreign income in 2014, and our analysis includes only returns filed from US addresses, so our measure should match MAGI for virtually all taxpayers in our sample. We can also determine the number of people in each tax unit using information on Form 1040, allowing us to assign an FPL value to each return and construct MAGI relative to the FPL. We infer the number of adults in each tax unit from the filing status listed on the return (one for single or head of household, two for married) and the number of children from the number of exemptions claimed for children at or away from home.[11]

We also obtain information on state of residence and the presence of self-employment income from the 1040s. State of residence is necessary in order to identify which taxpayers were exposed to the coverage gap. Self-employment income is relevant because previous work on bunching around the kinks in the EITC schedule (Saez, 2010) indicates that only the self-employed exhibit bunching behavior. Our data does not contain the amount of self-employment income reported on each return, but it does contain a flag that indicates whether the form used to report self-employment income (Schedule SE of Form 1040) was filed.

[11] There are some limitations to this approach. For instance, the FPL is officially determined according to household size, not tax unit size. Moreover, we do not have a way to identify dependent adults in our data. As a consequence, we may assign incorrect FPL values to households that contain multiple tax units or dependent adults. We believe, however, that this is the best available approach, and visual inspection of the resulting distribution of income relative to poverty suggests that mis-assigned FPL values are likely not a major problem.

This is sufficient for our purposes, as the presence of self-employment income is enough to identify filers who are more likely to respond, and the dollar amount of self-employment income is not strictly required to analyze the policies in question. Just under 13 percent of returns reported self-employment income in 2014 and 2015.

When our analysis requires survey-based measures of labor market activity, we link tax records to restricted ACS data at the individual level using Protected Identification Keys (PIKs), unique individual identifiers assigned by the Center for Administrative Records Research and Applications (CARRA)'s Person Identification Validation System (PVS). This system assigns PIKs to both survey and administrative records based on personally identifiable information like social security numbers, date of birth, place of birth, name, and address. The ACS provides measures of labor supply that are unavailable in the tax data. For example, respondents provide the number of hours they usually work per week and the number of weeks they worked in the last year, as well as their employment status at the time of the survey. We also obtain information on age, race, sex, and education from the ACS for use as controls in our regression analyses.

3.1 Graphical Analysis of the Tax Data

Before we discuss the approach we take to our empirical analysis, we begin by inspecting the tax data visually to assess the extent to which taxpayers exhibit bunching behavior in response to the introduction of the coverage gap in non-expansion states, as the degree of bunching influences our methodological choices. We plot the distribution of returns over income relative to the FPL, which determines eligibility for the premium tax credit. Because the FPL depends on family size, we group returns according to filing status and number of children, and we analyze each family structure separately. This ensures that each figure is based on filers located in the same region of the absolute income distribution (i.e. the distribution reported in dollars, not relative to the poverty level) and income bins correspond to fixed dollar values, although those values and the relevant region of the absolute income distribution vary across figures.

We first consider filers who report self-employment income, as the literature on bunching around kinks in the EITC schedule suggests that their reported income is most likely to be sensitive to changes in tax incentives like the one we consider. For each year from 2013 to 2015, Figure 5 shows the share of returns in each income bin between 50 percent and 150 percent FPL for single filers without children who report self-employment income, by Medicaid expansion status of their states of residence. All years of data for both expansion and non-expansion states show notable bunching around 85-90 percent FPL and a smaller amount around 115-120 percent FPL. For these taxpayers, bunching in these regions corresponds to the points in the raw income distribution at which a taxpayer claiming the standard deduction and the personal exemptions to which they are entitled (i.e. typically one exemption, or two if over the age of 64) reach positive federal income tax liability. The distribution of returns is fairly smooth around the FPL, which lies in between these two regions, in 2012 (and in previous years, which are not shown in the figure for the sake of simplicity). In 2013, the distribution largely resembles the 2012 distribution, with a small amount of additional mass observed in the 100 and 101 percent FPL bins, just above the soon-to-take-effect premium tax credit eligibility threshold. Although the premium tax credit was not available in 2013, individuals filed their 2013 taxes in early 2014, and publicity surrounding the ACA, health insurance marketplaces, and the premium tax credit could have influenced how some people prepared their returns. This phenomenon could also have spilled over onto the 97 percent FPL bin, which contains $11,000 in 2013 and also sees additional mass that year relative to 2012.[12]

In 2014, the first year in which this population was exposed to the coverage gap, a clear pattern of bunching near the premium tax credit eligibility threshold emerges in non-expansion states. The bunching remains evident in 2015. No similar pattern of bunching emerges in states that expanded Medicaid immediately, that is, those where the Medicaid expansion took effect by January 1, 2014, where no coverage gap exists. This comparison between expansion and non-expansion states makes clear that at least some taxpayers began bunching near the premium tax credit eligibility threshold at the time when and in the places where it was relevant to policy.

Critically, we are aware of no other contemporaneous policy changes that could give rise to this pattern of bunching. The generosity of other major transfer programs generally does not change sharply near the premium tax credit eligibility threshold.The gross income eligibility cutoff for the Supplemental Nutrition Assistance Program (SNAP) is 130 percent FPL. Similarly, the free lunch component of the National School Lunch Program phases out at 130 percent FPL, while the reduced-price component goes up to 185 percent FPL, as does eligibility for the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). Changes to the slope of the EITC and Child Tax Credit (CTC) schedules do not coincide with the FPL. Moreover, these programs are all administered federally, so if they did induce bunching in this region of the income distribution, one would expect to find it in both Medicaid expansion and non-expansion states. Eligibility for Temporary Assistance for Needy Families (TANF) is determined at the state level, but income limits are typically well below the FPL. The Supplemental Security Income (SSI) income limit is also below the FPL. Section 8 housing income thresholds are based on local median incomes, and in some cases, some thresholds coincide with the FPL, but the number of people affected by these is small relative to the other programs mentioned here, and these areas can be found in both expansion and non-expansion states, so we

[12] Bunching at round numbers is a phenomenon commonly observed in both survey and administrative data (e.g. Dominitz and Manski, 1996; Kleven and Waseem, 2013; Hungerman and Ottoni-Wilhelm, 2016). In the years shown in Figure 5, $11,000 falls in the 99 percent (2012), 97 percent (2013), 95 percent (2014) and 94 percent (2015) FPL bins.

would not expect them to induce bunching only where Medicaid eligibility has not been expanded.

Exactly which taxpayers exhibit bunching behavior remains in question. Figure 6 indicates that the bunching phenomenon is not limited to single filers; married filers without children who report self-employment income also appear to bunch near the FPL in non-expansion states, and it again emerges when the premium tax credit becomes available. Figures 7 and 8, which shows distributions of returns by number of children for self-employed single and married filers in 2015, also suggests that bunching likely occurs among some filers with children, particularly married filers. Figures 9 and 10, however, show that the bunching phenomenon is not universal. Single and married filers with no children who live in non-expansion states and do not report self-employment income (i.e. those who are subject to the coverage gap but who likely have less ability to easily adjust their earnings) show no signs of movement toward the premium tax credit eligibility threshold in 2014 and 2015.[13] Indeed, visual inspection indicates there is essentially no bunching among taxpayers who do not report self-employment income, regardless of family structure. Additional figures are available upon request.

4 Methodology

Overall, visual inspection reveals that some groups of taxpayers with self-employment income do appear to respond to the notch in after-tax income created by states that did not expand Medicaid under the ACA by bunching near the premium tax credit eligibility threshold. Taxpayers who are not self-employed, however, exhibit no such response. As a result, we will focus our bunching analysis on workers with self-employment income.

A growing literature infers the magnitude of behavioral responses to changes in the tax code from the degree of bunching observed in regions of the distribution affected by changing policy parameters. This type of analysis requires credible specification of the counterfactual distribution of income. While several papers have laid out approaches that are well suited to the questions they study, we believe our setting differs from those examined in the previous literature in ways that make implementation of those methods difficult. Before proceeding to our analysis, we provide a brief description of a typical approach to estimating behavioral responses to policy changes using bunching, lay out the details of our setting that we feel make this approach ill-suited for the analysis we wish to perform, and describe the modified approach we employ to overcome these challenges.

[13]The single points of excess mass visible in Figure 9 for single filers just above the FPL correspond to the the bin containing $12,000 (i.e. $1,000 per month) in each year. A small increase in mass is visible for married filers in the 100 percent FPL bin in 2015; that bin contains $16,000 that year.

4.1 A Typical Approach to Bunching Analysis

We begin by describing a common approach to specifying counterfactual distributions in bunching analyses employed in several recent papers. We focus on the strategy employed by Kleven and Waseem (2013) in analyzing small tax notches in Pakistan, but others (e.g. Chetty et al., 2011; Hungerman and Ottoni-Wilhelm, 2016) have used components of the same strategy to analyze tax kinks.[14] These analyses typically seek to estimate elasticities of the outcomes in question (e.g. taxable income) with respect to price, as varied by some policy change, using cross-sectional distributions of those outcomes that exhibit bunching in regions affected by changes in policy parameters. Of course, each region of the overall distribution is only observed under the policy regime that actually applies to it. Determining the degree to which the distribution of outcomes has been changed by some policy requires comparing the empirical distribution of outcomes to the (unobserved) distribution that would have been realized in the absence of the policy in question.

In practice, an estimated counterfactual distribution stands in for the unobserved, true counterfactual distribution in that comparison. The estimated counterfactual distribution is typically generated using a regression analogous to

$$M_b = \sum_{i=0}^{p} \beta_i^0 \cdot (Y_b)^i + \sum_{j=-R}^{R} \gamma_j^0 \cdot 1[Y_b = j] + \varepsilon_b^0 \tag{1}$$

where M_b gives the mass in bin b of outcome Y (i.e. the count or share of observations in that bin), the summed $\beta \cdot Y$ terms are a polynomial of order p, and summed γ terms are coefficients on dummy variables that indicate whether bin b is in the region $[-R, R]$ in which excess mass due to bunching is expected.[15] This excluded region is likely determined at least in part based on visual inspection, as bunching is often sharp on at least one end of the region. The value of the counterfactual distribution in each bin, then, is the fitted value from this regression, excluding the influence of the γ coefficients, which represent the amount of excess mass in the bin.

Additional adjustments may be made to this procedure as circumstances dictate. The counterfactual distribution is then compared to the empirical distribution to identify the location and magnitude of excess mass due to bunching, and that excess mass value is typically then converted into an elasticity. In the case of Kleven and Waseem (2013), the elasticity is calculated by comparing the earnings response implied by the excess mass in the bunching region to the change in implicit net-of-tax rate associated with the notch.

[14] We focus on the tax notch approach because our setting is dominated by a discrete increase in after-tax income rather than a change in marginal tax rates. Chetty et al. (2011) and Hungerman and Ottoni-Wilhelm (2016) also estimate counterfactual distributions similarly to Kleven and Waseem (2013). Ultimately, all three estimate at least some elasticities similarly to Saez (2010), an approach that was developed for analyzing tax kinks.

[15] If the outcome Y is such that bunching is common to a lesser extent at various non-policy relevant values throughout the distribution (e.g. Y is income and some bunching is observed at round numbers like \$5,000, \$10,000, etc.), additional dummy variables may be included in the regression to prevent bins containing irrelevant excess mass from influencing the estimation of the counterfactual distribution.

They also use the mass that remains in the region that is theoretically dominated by the bunching region to produce elasticity estimates that are not attenuated by optimization frictions that prevent some taxpayers from adjusting their income.

4.2 Challenges in this Setting

There are several challenges to utilizing the approach just described in our setting. As Figure 5 illustrates, groups of taxpayers who respond to the coverage gap by bunching near the FPL also respond to other policy parameters by bunching at points in the income distribution near the FPL. This other bunching is not close enough to our region of interest to render bunching estimation impossible, but it is close enough to make it difficult to fit even a high-order polynomial through the cross-sectional distribution. Cloyne et al. (2017) face a similar challenge in their analysis of the effect of home prices on household borrowing. Although we could adapt the typical approach by including dummy variables for bins in these other bunching regions, this would leave us with a counterfactual distribution based on only a few of the 20 closest income bins in either direction from the premium tax credit eligibility cutoff, which is conceptually unappealing. Moreover, if the income distribution is examined only in the cross section, as is typical of past bunching analyses, it is not clear how one would identify which bins should have their influence excluded from estimation of the polynomial due to their membership in one of the diffuse bunching regions that exist in this part of the distribution. Finally, the presence of these other bunch points could play a role in taxpayers' decisions about whether to adjust their income in order to gain eligibility for the premium tax credit, so using a counterfactual distribution that ignores them may be improper even if it is feasible.

Setting aside the difficulty of estimating a counterfactual distribution in the presence of other nearby bunching for a moment, two other factors complicate our setting. First, premium tax credits are refundable and advanceable, because they must be available for people to use to pay for health insurance throughout the year. As such, eligibility for the tax credits is based on projected income rather than actual income. For people with stable employment situations, accurate projections are relatively easy to make. Unexpected deviations from projected income, however, could lead some people who received premium tax credits to end up reporting income somewhat below the eligibility threshold at the end of the year.[16] This could be one cause of the somewhat diffuse bunching we observe around the FPL. The fact that bunching is not sharp makes it difficult to visually identify precise boundaries for the bunching region. The ambiguity of these boundaries also complicates both the Saez (2010) approach to estimating elasticities using the average of the density above and below the bunching region and the estimation of the counterfactual distribution via

[16] People facing this circumstance are not required to repay money they received through the tax credit over the course of the year. People who earn more than expected may be required to repay a portion of their subsidy, but the repayment amount is capped at $300 for single filers and $600 for married filers as long as actual income remains below 200 percent FPL.

high-order polynomial.

Finally, we consider a large, "upward" notch; the incentive it creates is for people in a certain region of the distribution to increase their income.[17] For our analytical purposes, the fact that ours is an upward notch is important because its existence does not create a dominated region of the distribution. Very strong disutility from work, for example, could rationalize remaining just below the eligibility threshold, even as doing so means foregoing an increase of thousands of dollars of after-tax income. This complicates analyses based on the amount of "missing mass" in regions bunchers likely moved from, including the production of frictionless elasticity estimates in the style of Kleven and Waseem (2013). The magnitude of the subsidy we consider also calls into question the production of elasticity estimates based on strategies developed for analysis of tax kinks (or small notches) around which the income effects are negligible, especially when they explicitly assume no income effect. Here, the income effect of the notch is the dominant consideration.

4.3 A Longitudinal Approach to Estimating the Counterfactual Distribution

We address the challenges to estimating counterfactual distributions described in the previous section by switching from cross-sectional to longitudinal analysis. Instead of estimating counterfactual distributions by fitting high-order polynomials across the distribution of income within single years of data, we analyze each income bin individually over time.[18] We aggregate the universe of IRS Form 1040 returns from 2000 through 2015 into cells defined by year, state Medicaid expansion status, income relative to poverty in 1 percent increments, filing status (single/married), number of children, and presence of self-employment income, retaining the number of returns in each cell. We perform our analysis separately for groups defined by expansion status, filing status, number of children, and presence of self-employment income. As mentioned above, this allows us to avoid comparing different regions of the income distribution and to fix the size of income bins within groups.

Because visual inspection (e.g. Figure 5) indicates that the income distribution in Medicaid expansion states largely resembles the distribution in non-expansion states in the relevant region prior to 2014 and expansion states do not exhibit bunching after 2014, we use a difference-in-differences regression to estimate the share of returns that would have appeared in each income bin in 2014 and 2015 in the absence of the

[17]This differs from the "downward" notch Kleven and Waseem consider, which discourages people from earning more by discretely increasing their tax liability.

[18]Cloyne et al. (2017) also take advanatge of longitudinal data to deal with the close proximity of the interest rate notches they consider to each other, using panel mortgage data to compare borrowers' actual loan-to-value choices to those that would have been realized if their loans had been passively rolled over.

coverage gap in non-expansion states.[19] For each income bin within each group, the regression has the form

$$S_{s,t} = \beta_0 + \beta_1 \left(NonExp_s \cdot Y2014_t\right) + \beta_2 \left(NonExp_s \cdot Y2015_t\right) + \beta_3 NonExp_s + \beta_4 Y2014_t + \beta_5 Y2015_t + \beta_6 NonExp_s \cdot Year_t + \beta_7 Exp_s \cdot Year_t + \varepsilon_{s,t} \tag{2}$$

where $S_{s,t}$ gives the share of tax returns within expansion status group s that appear in the bin at time t, $NonExp_s = 1$ for the group of non-expansion states, $Exp_s = 1$ for the group of immediate expansion states, $Y2014_t$ and $Y2015_t$ are equal to one in 2014 and 2015, respectively, $Year_t$ is a linear time trend, and $\varepsilon_{s,t}$ is an error term. The coefficients β_1 and β_2 give the amount of excess mass in non-expansion states within the bin in 2014 and 2015, respectively. We generate our counterfactual distribution by producing fitted values from this regression that exclude the influence of β_1 and β_2.[20] Figure 11 shows examples of counterfactual distributions generated using this approach for one group in which we observe bunching around the FPL.

Conveniently, this approach to generating counterfactual distributions also provides a data-driven way to identify the bunching region, despite its diffuse nature. We define the bunching region as the income bins between 90 and 110 percent FPL in which the difference-in-difference estimate of excess mass (i.e. β_1 for 2014 or β_2 for 2015) is statistically significant at the ten percent level.[21] We allow the bunching region to differ across groups and across years. In order to make the bunching region continuous, we also include in it gaps as large as three consecutive bins without statistically significant excess mass located between bins with statistically significant excess mass.

We believe that our approach has some key advantages. The use of longitudinal comparisons to generate counterfactual distributions brings more information to bear on the question of what the distribution of income would have looked like in the absence of the notch than a standard cross-sectional approach would. This additional information is worth utilizing even when other nearby distortions do not complicate cross-sectional analysis, as it could help identify regions of the distribution vacated by bunchers and help generate more plausible counterfactual distributions. Our approach also eliminates the need to make potentially arbitrary, subjective decisions about what constitutes the bunching region based on visual evidence. This is especially valuable in settings like ours in which bunching can be somewhat diffuse, but a similar approach could also be applied to verify visual judgments in settings with more sharply defined bunching regions. Of

[19] As a robustness check, we also generate counterfactual distributions using regressions that do not rely on comparing expansion states to non-expansion states, instead identifying excess mass by comparing 2014 and 2015 to within-bin trends using polynomials of various degrees. These approaches produce similar results.

[20] In order to ensure that the counterfactual distribution has the same total mass as the empirical distribution over the relevant region, we calculate the difference between the total mass of each distribution between 50 and 150 percent FPL. We then add that difference to the counterfactual distribution's mass, distributed evenly across all income bins in that region. Finally, we recalculate excess mass in the bunching region using this adjusted counterfactual distribution and use that figure to calculate elasticities. We do not use the β_1 and β_2 coefficients from the regression to calculate elasticities.

[21] We limit the bunching region to this area for two reasons. First, this is where the bunching appears to occur based on visual inspection. Second, bunching outside of this region is more likely attributable to other policy parameters.

course, the data demands of this type of analysis are intense, and the elasticities it produces do not have the same clearly-defined theoretical interpretation as those produced by other estimators. Nonetheless, it does provide an option for analyzing bunching in response to newly introduced policies, especially in regions of the income distribution that are crowded with other distortions.

With the bunching region identified, we estimate elasticities of taxable income with respect to the premium tax credit (e) for each group using the total excess mass in the region (B), the counterfactual distribution's mass in the region (C), the value of the premium tax credit at the FPL (S), and the amount of post-tax income a household has at the FPL (FPL), adjusting for federal and state taxes and credits using the National Bureau of Economic Research's TAXSIM program (Feenberg and Coutts, 1993). Elasticities are estimated according to $e = (B/C)/(S/FPL)$. The values of the premium tax credit are assigned using household size to determine the relevant FPL and the number of adults to determine how many people are covered through the insurance purchased, since children are generally covered by CHIP or Medicaid at this level of income and therefore not eligible for the credit.[22] As Saez (2010) notes, these elasticities will be a mix of compensated and uncompensated elasticities because they are estimated using a large notch. We generate standard errors by taking the standard deviation of 500 elasticity estimates derived from a bootstrap process in which we randomly reallocate estimates of excess mass across bins before calculating the elasticities using the original bunching region, which we hold fixed throughout.

5 Results

Informed by the graphical analysis above, we begin by presenting elasticities of taxable income based on our bunching analysis for taxpayers with self-employment income. We then provide some suggestive evidence to address the question of whether or not those elasticities reflect true changes in labor supply. Finally, we discuss what can be said about labor market changes among wage and salary workers.

5.1 Bunching Estimates of the Elasticity of Taxable Income

Table 1 reports estimates of the elasticity of taxable income for taxpayers with self-employment income that are derived from the approach described in Section 4.3. We report elasticities for all family structure groups, estimated separately for 2014 and 2015. Table 2 reports the same set of estimates for taxpayers who do not have any self-employment income (who we will also refer to as wage and salary workers). Where the

[22] The dollar values of the premium tax credit are obtained from the Kaiser Family Foundation's online subsidy calculator for 2014 and 2015. We assume that all adults are 40 years old (the middle age for which premiums are published for all plans in the federal marketplace, and roughly the median age in the United States) and no adults or children use tobacco, and we use the national average of the subsidy value.

procedure described in Section 4.3 does not identify any statistically significant bunching, we use all income bins between 95 percent and 105 percent FPL to estimate elasticities.

As expected based on our graphical analysis, we consistently identify bunching and estimate significant elasticities only for groups of taxpayers who report self-employment income. No group of wage and salary workers exhibits significant bunching in either 2014 or 2015. Among self-employed workers, single filers without children and married filers with three or fewer children exhibit significant bunching in both 2014 and 2015, and they are similarly responsive to the premium tax credit; elasticities range from roughly 0.7 to 0.9 for filers without children and from 0.6 to 1.0 for married filers with one to three children.

Though the differences in elasticities over time are generally not statistically significant, the point estimate for 2015 is larger than the point estimate for 2014 in four of these five groups. Married filers with four or more children also newly exhibit significant bunching in 2015 after not doing so in 2014. This provides an early suggestion that taxpayers may become more aware of the existence of the coverage gap or the availability of the premium tax credits over time. This is consistent with the Saez (2010) result that elasticities implied by bunching around the first EITC kink grew over time.

In addition to quantifying the elasticities implied by the bunching we observe, it is also worth understanding the magnitude of the bunching response in terms of the number of returns involved. Across all groups that exhibited significant bunching in 2014, the excess mass we identify in their bunching regions represents about 15,000 additional returns. For groups that exhibit bunching in 2015, that number rises to about 23,000 returns. In both years, single filers without children account for more than two-thirds of the returns we observe in excess of what is predicted by our counterfactual distributions. Given these numbers, it seems plausible that the bunching we observe could have arisen organically, despite the fact that this seems like a fairly sophisticated filing situation (self-employed filers precisely locating at a low level of income that yields tax benefits), and its existence and benefits are not necessarily obvious at first glance. Awareness could have been raised and the bunching we observe enabled through outreach efforts by groups dedicated to promoting enrollment in health insurance plans offered on the marketplaces, or groups that prodvide tax preparation assistance to low-income filers, or even by trial and error by users of the marketplace website.

It is important to note that, even among people with self-employment income, the number of points in the income distribution at which significant bunching is observed in the U.S. is very limited. Prior to the ACA, research had identified the first kink point in the EITC schedule and the point at which federal income tax liability becomes positive as the primary points at which identifiable bunching associated with tax policy occurs among taxpayers with self-employment income (Saez, 2010). The fact that taxpayers regard the premium tax credits as important enough to justify adjusting their income in order to obtain them is interesting in its own right, and this paper joins Heim et al. (2017) in providing evidence that certain groups of taxpayers at both the low and high ends of the income range in which the premium tax credit is available value it highly enough to adjust their incomes in order to obtain it.

One might argue that the bunching we observe is driven not by the income effect associated with the subsidy but by the ACA's individual shared responsibility provision, also known as the individual mandate. This provision of the law requires all taxpayers to carry minimum essential health coverage or pay a penalty when they file their taxes. However, people who fell into the coverage gap created by states that did not expand Medicaid eligibility were explicitly exempt from the individual mandate penalty. Many would have qualified for an exemption based on their low levels of income in any case. Moreover, even if these taxpayers believed they were subject to the penalty, the cost associated with not having health insurance was capped at 1 percent of income in 2014 and 2 percent of income in 2015, amounts that are small in comparison to the value of the subsidy.

It is, however, worth considering the interpretation of these elasticities carefully, because our setting differs from those considered in previous bunching studies in important ways. As mentioned above, we consider a large upward tax notch, rather than a kink or a small downward notch. As such, the elasticities we estimate are not the clearly defined intensive margin compensated elasticities of Saez (2010) and others. The income effect of the subsidy we consider and the non-convexity it creates in taxpayers budget constraints dominate any substitution effect created by the small change in the marginal tax rate created by the gradual reduction in the value of the premium tax credit. The non-convexity is likely to be particularly important, as one would expect a pure income effect to lead to a reduction in labor supply if leisure is a normal good. Moreover, the incentives created by this subsidy scheme may be large enough to create a response on the extensive margin. Theoretically, our estimates might best be characterized as total elasticities, incorporating income and substitution effects, the non-convexity of the budget constraint, and responses on both the intensive and extensive margins. Empirically, however, we find no evidence that the existence of this notch changes the rate at which filers report positive self-employment income, as shown across a variety of specifications in Table 3. This suggests that our elasticity estimates primarily represent an intensive margin response.

Another important difference between our setting and others is that the subsidy provided by the premium tax credit is only available for the purchase of health insurance. This fact has two important consequences. First, not all taxpayers in the income range covered by the premium tax credit use it to purchase health insurance, whether because they remain uninsured or obtain health insurance from another source. Take-up of the premium tax credit is therefore less than complete. This differs from analyses of bunching around point at which taxable income exceeds zero (a situation faced at some level of income for all taxpayers), but is similar to analysis of the EITC, which also has less than complete take-up. As such, our estimates are best interpreted as elasticities with respect to subsidies offered rather than subsidies received. From the perspective of understanding the consequences of this particular subsidy, this "intent to treat" style estimate is of direct interest to policymakers, who have limited control over program take-up. Second, the fact that the premium tax credit is used to purchase health insurance means that receiving it represents more than simply a change in a taxpayer's net-of-tax rate; the recipient also has access to (newly inexpensive) health insurance. To the extent that taxpayers' valuations of access to health insurance are not fully captured by the dollar value of the subsidies that enable that access, our estimates may differ from those that would be obtained from analysis of a refundable tax credit in the same amount the use of which was not restricted to the purchase of health insurance. To the extent that recipients value subsidized insurance at less than its cost, as some research suggests (e.g. Finkelstein et al., 2015, 2017), our estimates will be smaller in magnitude than estimates that use average valuation of the health insurance received as the value of the premium tax credit instead of the credit's average dollar amount.

Differences in interpretation notwithstanding, it remains instructive to compare the magnitude of our elasticity estimates to Saez's EITC estimates. Over the 10-year period he analyzes (1995-2004), Saez reports an elasticity of reported income for self-employed taxpayers with one child of 1.101 based on bunching around the first EITC kink. For taxpayers with two or more children, he finds an elasticity of 0.755. This corresponds closely with the range of estimates we find for married filers with children.[23]

There are reasons to think, though, that taxpayers may ultimately prove more responsive to the premium tax credit than to the first kink in the EITC schedule. As Saez notes, responsiveness to the EITC kink grew over time, and elasticities estimated using the second half of his data are larger than those estimated using the first half. Moreover, Saez's data begin in 1995, 20 years after the EITC was originally enacted. We estimate elasticities for only the first two years of the premium tax credit's existence. If awareness of the premium tax credit and the benefits of locating at the notch grows over time, take-up of the credit and bunching at the notch could increase.

The question of take-up raises another point related to interpreting these elasticities. While both our estimates and the EITC estimates are with respect to benefits offered, differences in take-up rates between the two programs could lead to somewhat different interpretations. Across all relevant income levels, about 80 percent of eligible taxpayers receive the EITC.[24] A comparison between enrollment figures from the Centers for Medicare and Medicaid Services and a Kaiser Family Foundation analysis of the eligible population suggests that about 45 percent of eligible taxpayers in states with a coverage gap took up the premium tax credit by the end of 2015. If these programs' take-up rates for self-employed workers close to the relevant

[23]When we pool taxpayers in groups analogous to those used by Saez (i.e. based only on number of children), we again find point estimates in line with those he reports, though they are not statistically significant. The value of the EITC for workers without children is small and phases out at relatively low levels of income. Saez does not include it in his analysis, so there is no estimate comparable to our elasticities for filers without children from his paper.

[24]This figure comes from IRS and Census analysis for tax year 2013 (Internal Revenue Service, 2017).

earnings thresholds are similar to their overall take-up rates, a back-of-the-envelop calculation suggests that the taxable income of taxpayers who respond to the premium tax credit could be about 75 percent more responsive than that of taxpayers who who locate at the EITC kink. This likely represents something of an upper bound on the extent to which the magnitude of the income response to the premium tax credit could exceed responsiveness to the first EITC kink, as the value of the premium tax credit is greatest at the eligibility threshold and take-up is probably higher than average there. If take-up rates are similar across programs at the relevant income levels, this difference in interpretation would be eliminated.[25]

It is also worth considering how our estimates compare with those of Heim et al. (2017), who analyze the 400 percent FPL notch. The bunching they find implies an elasticity (0.11) that is below our range of estimates. Given differences in our analyses, this is not surprising. Their estimate is based on the full population of taxpayers, including wage and salary workers, for whom they do find bunching via reduced earnings in response to the potential loss of the premium tax credit at that income level. Our elasticities are based on filers with self-employment income, who both we and they find to be more responsive to the premium tax credit. Moreover, conditional on geography, the value of the premium tax credit is much greater at 100 percent FPL than at 400 percent FPL, and Heim et al. find more substantial bunching responses among taxpayers exposed to larger subsidies.

Although we believe the difference-in-differences approach to generating counterfactual distributions used in our baseline analysis is the best option because it takes fullest advantage of the available variation, we also produce elasticities using alternative approaches that are not based on comparisons between non-expansion and expansion states. Table 4 compares our preferred baseline estimates to analogous estimates from four alternative approaches. The columns labeled linear, quadratic, and cubic estimate within-income bin time trends of the indicated degree to identify bunching and predict counterfactual distributions. The 2012 (2013) column uses the empirical income distribution from 2012 (2013) for these purposes, with bunching identified via difference-in-differences regressions. Our results for single and married filers without children, as well as those for married filers with one child, are generally robust to these alternative approaches. For these groups, all approaches show significant bunching and elasticities, and their magnitudes are roughly similar across methods. Estimates for married filers with two or three children are less robust. The methods that are based only on comparison to trends within group and do not take advantage of comparisons to immediate expansion states have a harder time identifying bunching for these groups. Overall, the elasticities that are most robust to alternative empirical approaches to identifying bunching and estimating counterfactual distributions are those that are most easily identified visually.

[25]It is also possible that premium tax credit take-up could exceed EITC take-up at their respective, relevant income levels. In this case, the magnitude of the income response among bunchers would be greater among those locating at the first EITC kink.

Across all approaches to identifying bunching, our assumptions about the appropriate measures of baseline income and subsidy value are held constant. It is worth noting, however, that these assumptions have important implications for the elasticity estimates that emerge. Table 5 presents elasticity estimates using our baseline approach to identifying bunching that vary those assumptions. In particular, we vary assumptions about whether the baseline income measure is adjusted to account for taxes and tax credits and about how subsidy values are assigned.

The first column presents our baseline estimates, in which the baseline income measure is adjusted for payroll taxes, federal and state income taxes, and federal and state tax credits, and subsidies are assigned that cover the purchase of insurance for the adults in each type of household. This approach implicitly assumes full take-up of the EITC and CTC, and that no children receive subsidized health insurance through the marketplace. Although these assumptions are not completely accurate, we believe this is a reasonable baseline approach, because take-up rates for these credits are high, and the vast majority of children at this level of income are covered by CHIP or Medicaid and therefore ineligible for the premium tax credit.[26] The other columns of Table 5 explore how the elasticity estimates change if we use assumptions opposite to our baseline assumptions. The second column instead uses a "pre-tax" baseline income measure that is not adjusted for taxes or credits, implicitly assuming no one takes up EITC or CTC (or pays taxes). The third and fourth columns reproduce the first two columns, but assign subsidy values that assume all members of the household, including children, are covered by marketplace plans subsidized by the premium tax credit.

Variation in our estimates across columns of Table 5 highlights the importance of these assumptions.[27] Ignoring taxes and transfer programs that operate through the tax code can lead to important differences in estimated elasticities of taxable income. These differences are smaller for filers without children, who do not benefit from the EITC or CTC, but ignoring taxes still leads to estimated elasticities that are six to eight percent higher. For filers with children the differences are larger, since these tax credits are refundable, allowing filers to more than eliminate their tax liability, and their value is large relative to pre-tax income in this region of the distribution. Estimated elasticities for filers with children that do not account for the additional resources these programs provide are 12 to 19 percent lower than those that do. The subsidy assignment assumptions also have important implications for the magnitude of the elasticities for households with children. When we assume, counterfactually, that all children receive subsidized insurance through the

[26] EITC take-up is noted above. Estimates of CTC take-up among eligible filers do not appear to be available, but our calculations based on IRS Statistics of Income tables show that 47 percent of all returns that claimed exemptions for children at or away from home in tax year 2014 also claimed the CTC. The participation rate among eligible households is likely much higher. For households with income near the FPL, only dependent children who are over the age of 19 or subject to waiting periods of up to three months in a limited number of states could be covered under insurance purchased through a marketplace and subsidized by the premium tax credit.

[27] Note that within each row of Table 5 the four elasticities presented differ only in terms of the scalars used to convert the group's bunching estimate into an elasticity. Bunching is estimated only once within each row.

marketplaces, and the dollar value of the premium tax credit credit is correspondingly higher, elasticities for households with children are 21 to 45 percent smaller.

5.2 Do Changes in Reported Income Reflect Changes in Labor Supply?

The previous section provides estimates of the elasticity of reported income. Though this is the parameter typically estimated in bunching analyses using individual tax returns and is a parameter of primary interest from a public finance perspective, it does not definitively answer the question of how taxpayers respond to the coverage gap.

We observe bunching at the premium tax credit eligibility threshold among self-employed taxpayers. This could, of course, be the result of these taxpayers working harder in order to reach the premium tax credit threshold and gain access to subsidized health insurance. But since these workers do not have forms characterizing the full amount of their earnings filed to the IRS for them by their employers, they are precisely the ones who are most able to adjust the amount of income they report without changing the amount of time they spend working. In order to understand the broader consequences of the bunching response we observe, it is important to know whether it reflects a change in actual labor supply or merely a change in reported income.

This question often goes unanswered because it is difficult to answer with only tax data, in which labor supply is not observed beyond the binary measure of having wage and salary income or not. However, by matching tax return data to measures of labor supply from the ACS, we can provide suggestive evidence to address it. Working with restricted data at the U.S. Census Bureau, we link ACS responses to 1040s at the individual level using PIKs as described in Section 3.[28]

The ACS provides measures of labor supply that are unavailable in the tax data. For example, respondents provide the number of hours they usually work per week and the number of weeks they worked in the last year, as well as their employment status at the time of the survey. If the bunching response that we see in the tax data reflects a true change in labor supply, we would also expect to see changes in measures of labor supply taken from the ACS for people who would have been most likely to fall into the coverage gap. In particular, people who had income in 2013 that was close to but below the level that would become the premium tax credit eligibility threshold in 2014 see most likely to adjust their labor supply to obtain the tax credit.

In order to determine whether this response in fact occurs, we merge our measure of income relative to poverty from 2013 tax forms onto ACS data from 2012 through 2015. We then estimate changes in labor

[28]ACS data are based on a sample and subject to both sampling and non-sampling error. For more information, see https://www.census.gov/programs-surveys/acs/.

supply measures due to the introduction of the coverage gap using a difference-in-differences approach, and we allow those effects to vary by level of income relative to poverty in 2013. Estimates that show an increase in labor supply concentrated among people with 2013 income close to but below the premium tax credit eligibility threshold would be suggestive evidence that the bunching response observed in the tax data reflects a true change in labor supply.[29]

We estimate the equation

$$\begin{aligned} y_{ist} = \sum_{b=50}^{150} [\alpha_b(After_t \cdot NonExp_s \cdot Bin_b) + \beta_b(After_t \cdot Bin_b) + \gamma_b(NonExp_s \cdot Bin_b) + \delta_b Bin_b] + \\ \kappa(After_t \cdot NonExp_s) + \lambda After_t + \mu NonExp_s + \zeta X_i + \eta_s + \theta_t + \varepsilon_{ist} \end{aligned} \tag{3}$$

where y_{ist} is labor supply of individual i in state s at time t, $After_t = 1$ in 2014 and 2015, and $NonExp_s = 1$ in states that have not expanded Medicaid. Bin_b indicates that an individual's 2013 income was in income bin b, and interaction with the $After_t$ and $NonExp_s$ terms allows effects of exposure to the coverage gap to vary across 2013 income. To accommodate smaller sample sizes in the ACS matched data, we use wider income bins (5 percent FPL in width) than in our analysis of the tax data. X_i contains individual demographic characteristics such as age, race, sex, educational attainment, and family structure, η_s contains state fixed effects, θ_t contains year fixed effects, and ε_{ist} is an error term. We exclude individuals whose 2013 was less than 50 percent FPL or greater than 150 percent FPL. Regressions are estimated using ACS sample weights that have been adjusted to account for systematic differences across groups in the probability of being assigned a PIK.[30]

Coefficients from these regressions are summarized in Figure 12. The shaded area represents the 95 percent confidence interval of each coefficient estimate. Standard errors are clustered at the state level. For the sake of simplicity, the figure shows coefficients from regressions that pool individuals from all groups that exhibit bunching (reported in Table 1).[31]

Suppose there is a true change in labor supply in response to the coverage gap. Graphically, we would then expect to see larger positive effects on some or all of the hours, weeks, and employment outcomes (or larger negative effects on unemployment) in the region below premium tax credit eligibility threshold, while seeing smaller or no effects elsewhere in the income distribution.[32] This is not, however, what we find. The

[29]Although there is certainly income volatility in this region of the income distribution, there is enough stability that it is reasonable to expect the region below the FPL to be treated more intensively by exposure to the coverage gap than the region above the FPL. In the years preceding the introduction of the premium tax credit, about two-thirds of people who with income between 50 percent and 100 percent FPL one year also had income in that range the next year. A similar share of those in the 100 percent to 150 percent range one year also had income in that range the next year.

[30]See Appendix A for more details.

[31]We ran these regressions separately for each group, and the results were consistent with the summary figures shown here.

[32]Though bunching estimation is not designed to capture changes in labor supply on extensive margin, we consider them here because the change in income at the notch is large and could induce them.

pattern of point estimates is fairly consistent across the range of income considered, exhibiting little trend and giving no indication of larger magnitude effects in the region most likely to respond to the coverage gap. Virtually none of the coefficients we estimate are statistically significant. Confidence intervals are fairly wide for some outcomes, so we cannot completely rule out the possibility of a pattern of economically meaningful effects that could be consistent with a true labor supply response, but our estimates do not suggest that such a response is likely. We interpret these results as suggestive evidence that the bunching behavior we observe in the tax data more likely represents a simple change in reported income rather than a true change in labor supply.

5.3 Does Labor Supply Change among Wage and Salary Workers?

The fact that we find no bunching among workers without self-employment income, who we call wage and salary workers, suggests that the set of labor market responses to the premium tax credit that could be taking place within this group is likely somewhat circumscribed. Specifically, any labor market changes could not lead to meaningful changes in the distribution of income near the premium tax credit eligibility threshold. However, given the potential increase in labor market flexibility provided by access to health insurance not tied to an employer (i.e. a reduction in job lock a la Madrian, 1994), this set is not empty. Freed from the need to maintain a connection to an employer that provides health insurance, wage and salary workers could concentrate their labor in fewer weeks of the year by working more hours each week, spending the remaining time outside the labor force engaged in leisure activities or non-market work. Alternatively, these workers could spend more time unemployed while searching for higher-quality jobs. Critically, responses along these lines could make individual workers better off without leading to changes in the distribution of income.

Given that eligibility for the premium tax credit is determined based on income and changes sharply at the FPL, a regression discontinuity (RD) or difference-in-discontinuities approach seems at first glance well-suited to identify effects of the credit on labor market outcomes. However, the income measure that determines eligibility is not MAGI as reported on Form 1040, but rather projected MAGI for the year for which health insurance is being purchased, as reported to the Marketplace website.[33] As such, we do not observe the income measure that determines premium tax credit eligibility. We do have access to realized MAGI from the previous year, which is perhaps the most reasonable available proxy for projected income and the measure Healthcare.gov recommends starting with when coming up with an estimate of projected income. Indeed, Gallagher et al. (2017) use the previous year's MAGI as the running variable in their RD analysis of the effects of the premium tax credit on home payment delinquency.

[33]Reported MAGI is subject to verification, and reports that differ from verified past income by sufficiently large margins require additional documentary support, though marketplace customers are granted presumptive eligibility based on their reported income for a limited time while these discrepancies are resolved.

Before we proceed with analyzing health insurance and labor market outcomes according to previous year's MAGI, we offer a word of caution. Without access to the actual projected MAGI measure for at least some sample of marketplace customers, we are unable to assess how closely the previous year's income approximates projected income. To the extent that differences between the two measures are consistent across the income distribution, the likely consequence is attenuation in estimates of changes at the eligibility threshold. This limitation affects both graphical and regression-based analyses. For the sake of transparency, we focus on graphical analysis of changes in labor market outcomes near the premium tax credit eligibility threshold.[34]

We begin by considering changes in health insurance coverage, because any changes in labor market outcomes likely arise from changes in access to insurance. Figure 13 plots rates of coverage under Medicaid and directly purchased health insurance for 2013-2015 by income in the previous year.[35] The figure indicates that the ACA did increase coverage by both Medicaid and directly purchased health insurance. The Medicaid coverage increase is concentrated almost entirely in expansion states and extends to income levels ranges both below and above that directly covered by the expansion. This likely reflects a combination of a publicity or "woodwork" effect (as suggested by Sommers et al., 2012 in the context of state Medicaid expansions and consistent with Frean et al., 2017 in the ACA context), through which efforts to make the newly eligible aware of the expansion also alerted previously eligible non-participants to their eligibility, and the fact the incomes change from year to year and health insurance coverage is measured at a point in time in the ACS, so some respondents legitimately covered by Medicaid will have income outside the eligibility range under the measure used in this figure.

The increase in directly purchased health insurance coverage is concentrated in non-expansion states, and though it is largest in the income region in which the premium tax credit covers the largest share of the cost of insurance, coverage under this type of insurance also increases below the income range in which the premium tax credit is available. While this could again be partially due to measurement challenges, the lack of other coverage options appears to be an important contributor. In immediate expansion states, the bottom-left panel of Figure 13 shows little increase in directly purchased health insurance coverage below roughly 138 percent FPL, the same income range that sees a large increase in Medicaid coverage under the expansion, shown in the top left panel. In non-expansion states, the increase in health insurance coverage in this income range is driven by directly purchased insurance. The increase in the prevalence of this type of coverage is larger than in immediate expansion states and occurs both above and below the premium

[34]Because changes near the premium tax credit eligibility threshold can be difficult to discern visually from noise in the data, we also present difference-in-discontinuities estimates for these same outcomes in Appendix B. We offer these regressions simply as descriptive analyses of how how labor market outcomes empirically differ across the threshold.

[35]We have also produced versions of Figures 13 and 14 that are based on realized MAGI from the same year as the survey measures. These figures are similar to those shown here.

tax credit eligibility threshold, though the increase is larger just above the threshold than just below it, especially by 2015.

Figure 14 depicts changes in various labor market outcomes across the income distribution. We consider the same set of outcomes shown in Figure 12. Adjacent panels plot the same outcome for immediate expansion and non-expansion states. Despite the clear changes in health insurance coverage we observe in the previous figure, comparisons across time and expansion status provide little visual evidence that any of the labor market outcomes we consider changed meaningfully in non-expansion states near the premium tax credit eligibility threshold after the credit became available.

These figures partially alleviate some concerns about attenuation associated with using the previous year's income as a proxy for reported income. The fact that we see clear visual evidence of differential changes in coverage across insurance types and expansion status indicates that the previous year's income does contain some signal of eligibility for the premium tax credit. The absence of any corresponding change in labor market outcomes suggests little connection between them and the changes in health insurance coverage. However, we also see why some attenuation concerns remain. Figure 13 shows increased coverage by directly purchased health insurane below the subsidized income region in non-expansion states, likely due to some combination of measurement error and increased rates of unsubsidized health insurance purchase. To the extent that measurement error is responsible, this figure also illustrates the potential for attenuation in RD estimates. Without access to data on premium tax credit take-up, we cannot assess the importance of this possibility. This is the reason for our hesitancy to focus on such estimates, though both the presence of a change in health insurance coverage and the absence of a change in labor market outcomes are visually clear in the figures we present.

In their home payment delinquency analysis, Gallagher et al. focus on a set of people that is particularly likely to be affected by the credit: those who do not have a form of health insurance that makes them ineligible to receive it. Our main analysis of wage and salary workers does not impose any sample restrictions based on insurance coverage because our primary aim is to consider labor market effects of the credit on the scale of the economy as a whole, but also because coverage under some other form of insurance is potentially endogenous to the subsidy regime being studied. However, this group of workers that may be more intensely exposed to the incentive created by the premium tax credit notch is also worth considering.

We reproduce Figures 13 and 14 for a sample constructed to more closely match that used by Gallagher et al. in Appendix C. This involves excluding people covered by employer-sponsored insurance, Medicare, Medicaid, or TRICARE/VA insurance and including people aged 19-25. We again see no evidence of meaningful changes in labor market outcomes in non-expansion states upon introduction of the premium tax credit.[36]

6 Conclusion

States that declined to expand Medicaid eligibility under the Affordable Care Act created a large upward notch in their residents' after-tax budget constraints at the FPL. Thousands of dollars in refundable, advanceable premium tax credit subsidies for the purchase of private health insurance were made available to those with income at or above the FPL, while those with income just below the FPL are not eligible for any form of subsidized health insurance. Consequently, small changes in pre-tax income can lead to much larger changes in after-tax income for workers near the FPL. It is natural to ask, then, to what extent workers adjust their incomes in order to gain eligibility for the premium tax credit.

A more general form of this question – how does reported income respond to changes in the net-of-tax rate? – has long been of interest in public finance. The elasticity of taxable income is a parameter of primary interest in academic debates over the optimal design of income tax systems. Researchers beginning with Feldstein (1999) have argued that it can in fact be a sufficient statistic for welfare analysis in this context, while others have since considered the exact conditions under which this claim holds.[37]

We use the notch at the FPL induced by the ACA's coverage gap and the universe of IRS Form 1040 tax returns to estimate the elasticity of taxable income with respect to the premium tax credit subsidy via analysis of excess mass in the income distribution near the notch. Our approach has the advantage of identifying excess mass by comparison to counterfactual distributions that are estimated longitudinally, so the nearby presence of other bunching induced by longstanding features of the tax code does not distort our estimates. Consistent with Saez (2010), we find that only taxpayers with self-employment income bunch near the notch. Among the self-employed, single and married filers without children and married filers with one or two children are particularly responsive to the availability of the premium tax credit. Estimated elasticities of taxable income for these groups range from about 0.6 to 1.0 in our preferred specification. These estimates are comparable in magnitude to those Saez finds using the first kink in the EITC tax schedule, despite the fact that there are important policy and estimation differences between our setting and his.

[36]While our study and Gallagher et al. (2017) focus on different outcomes, both consider changes in directly purchased, or "private," health insurance coverage at the premium tax credit eligibility threshold. Gallagher et al. find an 11 percentage point increase in such coverage at that threshold. We find a much smaller increase. A few factors may account for this difference. Sample construction could be an important contributor. Our figures focus on wage and salary workers, while their analysis also includes the self-employed, whose earnings we show are more responsive to the credit. Also, their sample consists entirely of households that filed their income tax returns electronically through a program that provides free access to tax preparation software. The set of relatively low-income people who seek out and use such a program may be more aware of the availability of the premium tax credit and/or capable of successfully navigating the marketplace website to shop for and purchase insurance. Of course, sampling variation could also be a contributing factor, as their sample includes survey responses from about 5,000 taxpayers, while our ACS sample is much larger.

[37]Lockwood (2016), for instance, finds that the sufficient statistic approach to welfare analysis does not apply to tax systems where notches have first-order revenue effects.

To our knowledge, our paper constitutes the first evidence that some groups of taxpayers adjust their incomes upward in response to tax incentives. However, changes in reported income may or may not reflect changes in actual work, especially among self-employed workers, whose earnings are not reported to the IRS for them by employers. Whether the taxable income response we find is due to changes in reporting or changes in work has public finance implications and is also an important labor question in its own right. We are aware of no study of tax-induced income bunching that considers effects on direct measures of labor supply, which are generally not available on tax forms. We are able to match tax return data to survey-based measures of labor supply and fill this hole in the literature. On top of that, potential effects of the ACA on the labor market have been of interest since before the law was even passed in 2010. A great deal of attention has been paid to how other aspects of the law, such as the Medicaid expansion or the employer mandate, have influenced labor market outcomes, but we know of no prior study that considers the effects of the notch created at the FPL by the premium tax credit in non-expansion states.

Our analysis of matched tax-ACS data indicates that the elasticities of taxable income we estimate in our bunching analysis are unlikely to reflect true changes in labor supply. If they did, one might have expected to see changes in employment outcomes concentrated among those whose 2013 income would have left them without access to subsidized health insurance in 2014, but we find no differential effects for this group. We also consider the possibility that increased access to directly purchased health insurance could cause changes in the labor market outcomes of wage and salary workers even though the distribution of income is little changed for this group. Visual comparisons across time and expansion status, however, reveal no evidence of meaningful changes in labor market outcomes near the premium tax credit eligibility threshold in non-expansion states.

Our results have important implications for future efforts to provide subsidized health insurance. Despite the magnitude of the subsidy provided by the premium tax credit, we find no evidence that workers change the amount of labor they supply in order to obtain it. Moreover, the number of people represented by the excess mass we find at the premium tax credit eligibility threshold among filers with self-employment income is small relative to the number of people in the coverage gap. Incentives, even strong ones, may be insufficient to induce people to move up from the unsubsidized region of the distribution to the subsidized region. If policymakers want to subsidize health insurance coverage for those currently ineligible for both Medicaid and the premium tax credit, it may be necessary to change the rules of one program or the other to close the coverage gap rather than relying on workers to adjust their earnings in order to qualify for subsidies under the current rules. Conversely, if policymakers do close the coverage gap, it is unlikely that eliminating the notch it created will lead to major changes in labor supply.

References

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. *Handbook of Labor Economics*, 3:1559–1695.

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records. *The Quarterly Journal of Economics*, 126(2):749–804.

Cloyne, J., Huber, K., Ilzetzki, E., and Kleven, H. (2017). The effect of house prices on household borrowing: A new approach. Working paper.

Congressional Budget Office (2016). Federal subsidies for health insurance coverage for people under age 65: 2016 to 2026.

Currie, J. and Madrian, B. C. (1999). Health, health insurance, and the labor market. *Handbook of Labor Economics*, 3:3309–3416.

Dague, L., DeLeire, T., and Leininger, L. (2014). The effect of public insurance coverage for childless adults on labor supply. NBER Working Paper No. 20111.

Dillender, M., Heinrich, C., and Houseman, S. (2016). Health insurance reform and part-time work: Evidence from Massachusetts. *Labour Economics*, 43:151–158.

Dominitz, J. and Manski, C. F. (1996). Perceptions of economic insecurity: Evidence from the survey of economic expectations. Institute for Research on Poverty Discussion Paper no. 1105-96.

Even, W. E. and Macpherson, D. A. (2015). The Affordable Care Act and the growth of involuntary part-time employment. IZA Discussion Paper No. 9324.

Feenberg, D. and Coutts, E. (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management*, 12(1):189–194.

Feldstein, M. (1995). The effect of marginal tax rates on taxable income: A panel study of the 1986 Tax Reform Act. *Journal of Political Economy*, 103(3):551–572.

Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics*, 81(4):674–680.

Finkelstein, A., Hendren, N., and Luttmer, E. F. P. (2015). The value of Medicaid: Interpreting results from the Oregon Health Insurance Experiment. NBER Working Paper No. 21308.

Finkelstein, A., Hendren, N., and Shepard, M. (2017). Subsidizing health insurance for low-income adults: Evidence from Massachusetts. Working paper.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., and Baicker, K. (2012). The Oregon Health Insurance Experiment: Evidence from the first year. *The Quarterly Journal of Economics*, 127(3):1057–1106.

Frean, M., Gruber, J., and Sommers, B. D. (2017). Premium subsidies, the mandate, and Medicaid expansion: Coverage effects of the Affordable Care Act. *Journal of Health Economics*, 53:72–86.

Gallagher, E. A., Gopalan, R., and Grinstein-Weiss, M. (2017). The effect of health insurance on home payment delinquency: Evidence from ACA marketplace subsidies. Working paper.

Garfield, R. and Damico, A. (2016). The coverage gap: Uninsured poor adults in states that do not expand Medicaid.

Garthwaite, C., Gross, T., and Notowidigdo, M. J. (2014). Public health insurance, labor supply, and employment lock. *The Quarterly Journal of Economics*, 129(2):653–696.

Gelber, A. M., Jones, D., Sacks, D. W., and Song, J. (2017). Using kinked budget sets to estimate extensive margin responses: Evidence from the social security earnings test. NBER Working Paper No. 23362.

Gooptu, A., Moriya, A. S., Simon, K. I., and Sommers, B. D. (2016). Medicaid expansion did not result in significant employment changes or job reductions in 2014. *Health Affairs*, 35(1):111–118,1–12.

Grembi, V., Nannicini, T., and Troiano, U. (2012). Policy responses to fiscal restraints: A difference-in-discontinuities design. IZA Discussion Paper No. 6952.

Gruber, J. (2000). Health insurance and the labor market. *Handbook of Health Economics*, 1:645–706.

Heim, B., Hunter, G., Isen, A., Lurie, I. Z., and Ramnath, S. (2017). Income responses to the Affordable Care Act: Evidence from the Premium Tax Credit notch. Working paper.

Hungerman, D. and Ottoni-Wilhelm, M. (2016). What is the price elasticity of charitable giving? toward a reconciliation of disparate estimates. Working paper.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Internal Revenue Service (2017). EITC participation rate by states. `https://www.eitc.irs.gov/EITC-Central/Participation-Rate`.

Kaestner, R., Garrett, B., Gangopadhyaya, A., and Fleming, C. (2016). Effects of ACA Medicaid expansions on health insurance coverage and labor supply. NBER Working Paper No. 21836.

Kaiser Commission on Medicaid and the Uninsured (2013a). The coverage gap: Uninsured poor adults in states that do not expand Medicaid.

Kaiser Commission on Medicaid and the Uninsured (2013b). The coverage gap: Uninsured poor adults in states that do not expand Medicaid.

Killingsworth, M. R. and Heckman, J. J. (1986). Female labor supply: A survey. *Handbook of Labor Economics*, 1:103–204.

Kleven, H. J., Landais, C., Saez, E., and Schultz, E. (2013). Migration and wage effects of taxing top earners: Evidence from the foreigners tax scheme in Denmark. *The Quarterly Journal of Economics*, 129(1):333–378.

Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Levy, H., Buchmueller, T. C., and Nikpay, S. (2016). Health reform and retirement. *J Gerontol B Psychol Sci Soc Sci*.

Lockwood, B. (2016). Malas notches. Working paper.

Madrian, B. C. (1994). Employment-based health insurance and job mobility: Is there evidence of job-lock? *Quarterly Journal of Economics*, 109(1):27–54.

Moriya, A. S., Selden, T. M., and Simon, K. I. (2016). Little change seen in part-time employment as a result of the Affordable Care Act. *Health Affairs*, 35(1):119–123,1–14.

Pencavel, J. (1986). Labor supply of men: A survey. *Handbook of Labor Economics*, 1:3–102.

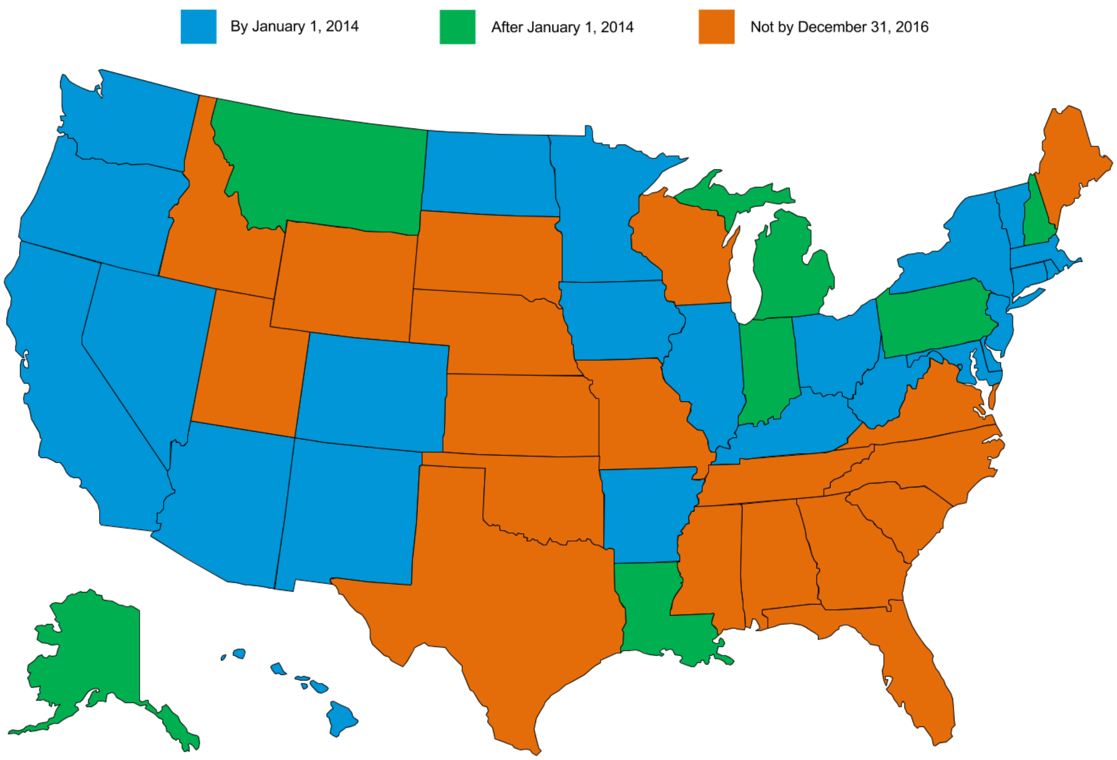
Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212.

Sommers, B. D., Baicker, K., and Epstein, A. M. (2012). Mortality and access to care among adults after state Medicaid expansions. *New England Journal of Medicine*, 367(11):1025–1034.

Van der Klaauw, W. (2008). Regression–discontinuity analysis: A survey of recent developments in economics. *Labour*, 22(2):219–245.

Figures

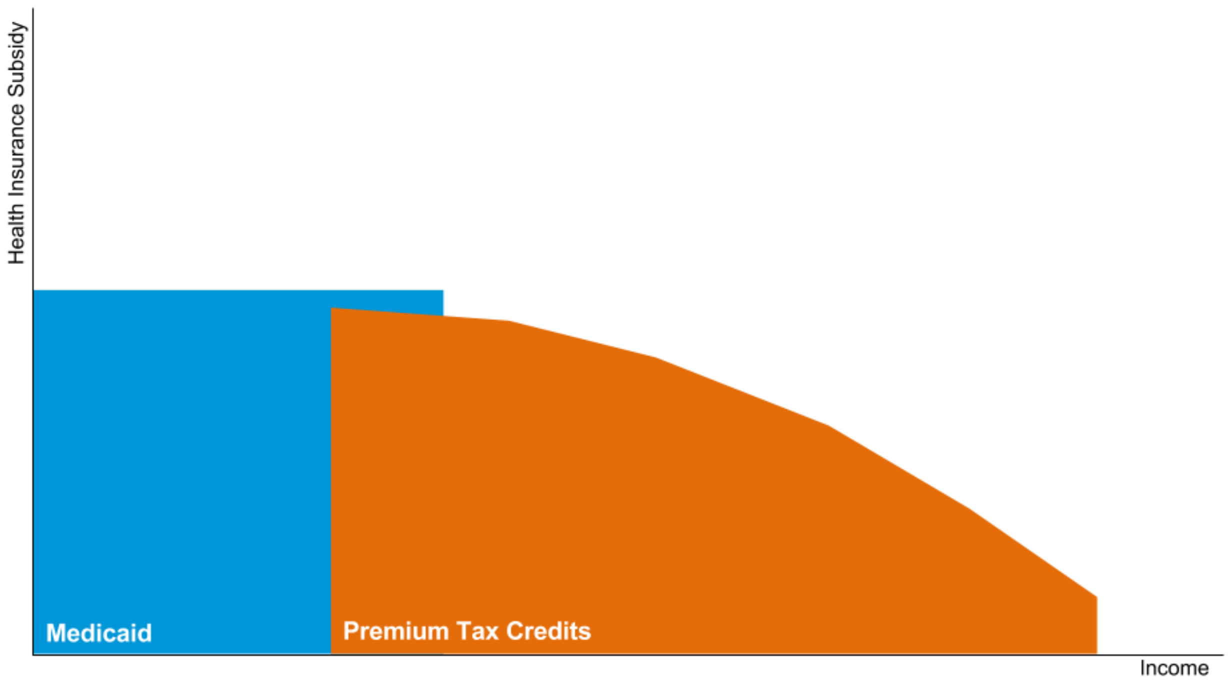
Figure 1: Effective Date of ACA Medicaid Expansion by State



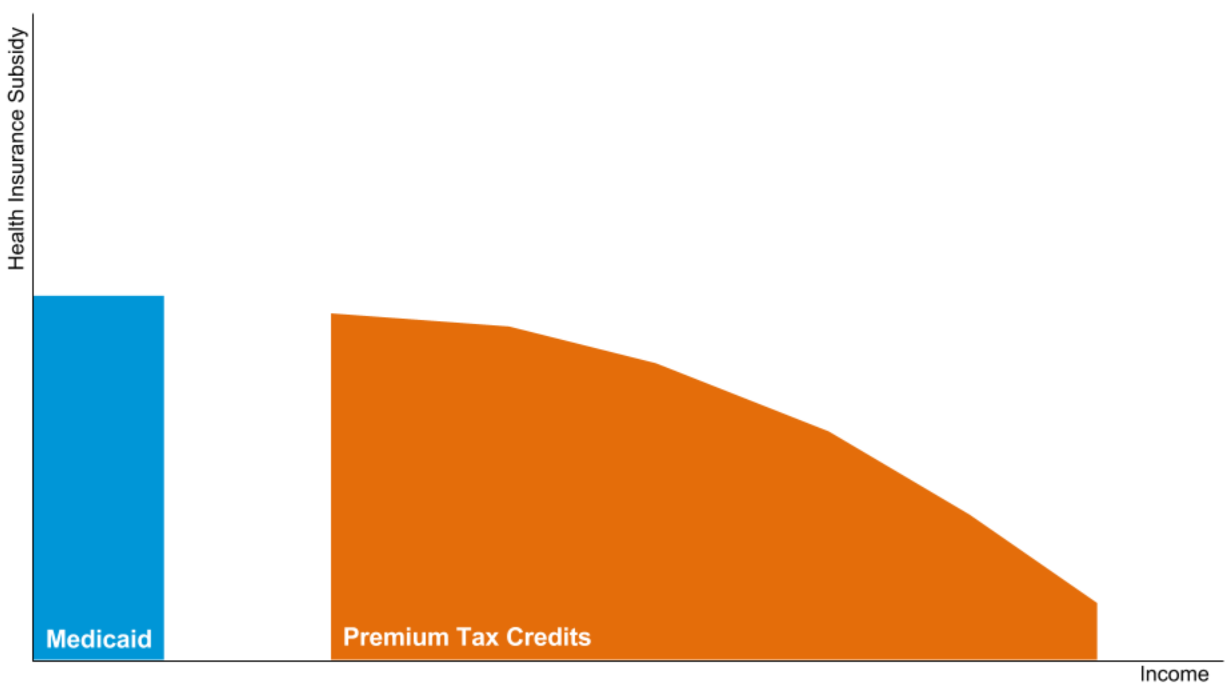
Source: Kaiser Family Foundation
Note: Arizona, California, Connecticut, Delaware, the District of Columbia, Hawaii, Massachusetts, Minnesota, New York, and Vermont had taken significant steps to provide coverage to low-income childless adults prior to 2014 and subsequently adopted the Medicaid expansion under the ACA or took advantage of a provision allowing for early expansion under the ACA. Two other states shifted residents from state programs to Medicaid under the ACA prior to January 1, 2014, but did not increase their Medicaid income eligibility limits until that date. Wisconsin has not expanded Medicaid eligibility under the ACA, but its income eligibility limit is equal to 100 percent of the federal poverty level. Therefore, there is no coverage gap in Wisconsin, and we exclude it from all of our analysis. See Levy et al. (2016) for more details.

Figure 2: Stylized Health Insurance Subsidies and the Medicaid Coverage Gap

(a) With Medicaid Expansion

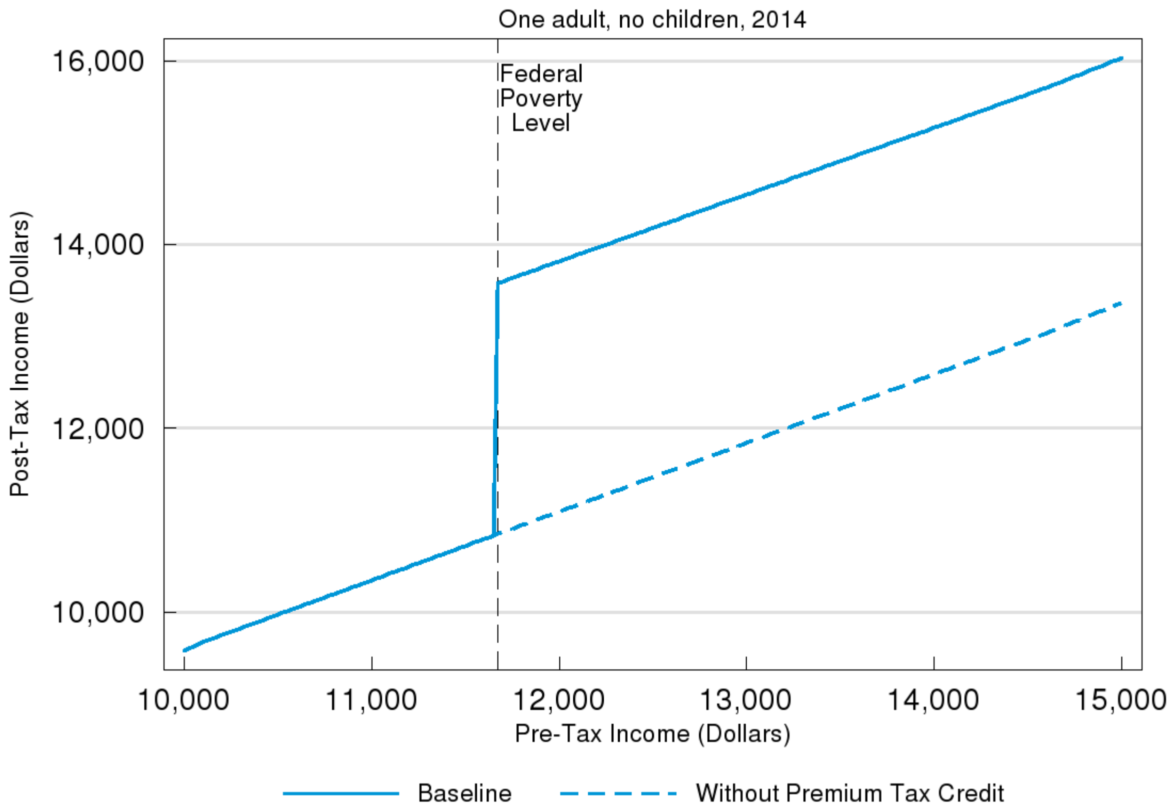


(b) Without Medicaid Expansion



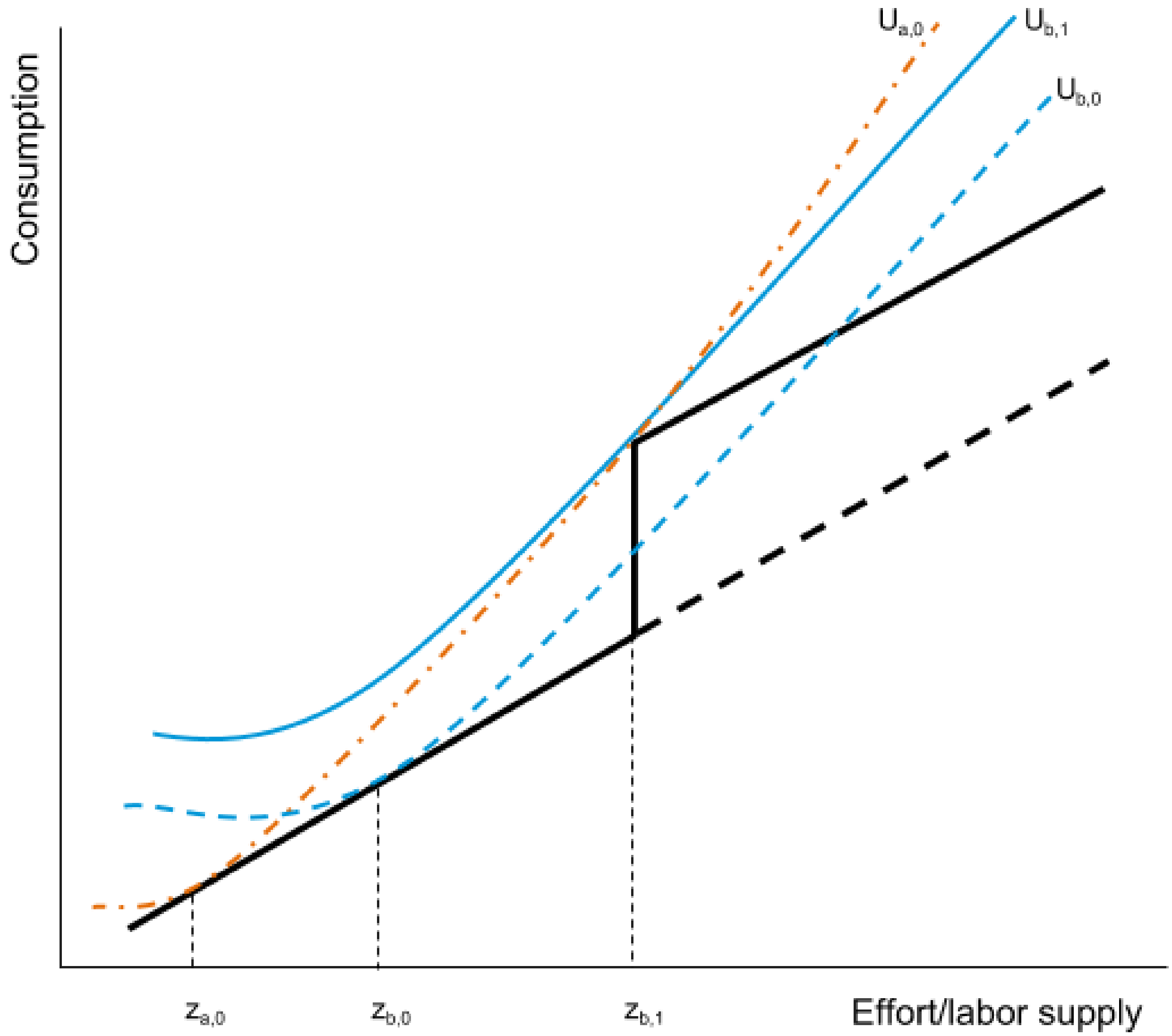
Note: The overlap between Medicaid and the premium tax credit in panel (a) and the space between them in panel (b) illustrate how the Medicaid coverage gap arises. Other than that, the dimensions of the polygons depicting those two forms of subsidy are not meant to precisely reflect any particular policy paramter.

Figure 3: Example Post-Tax Income Schedule in Non-Expansion States



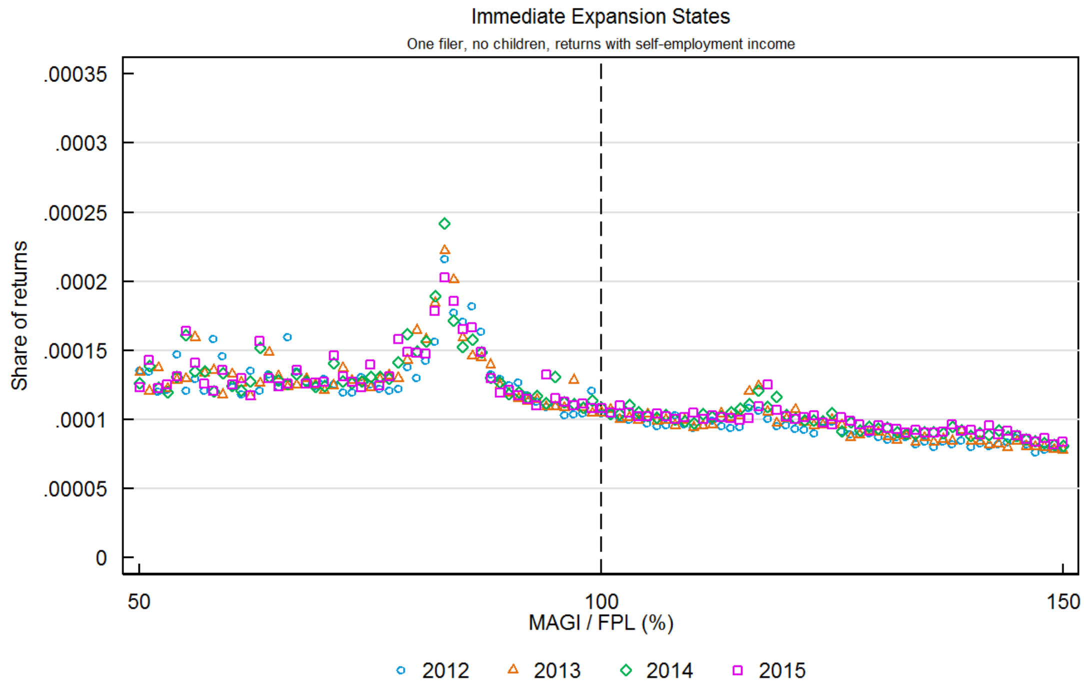
Note: Post-tax income plotted here reflects adjustment for federal income taxes, employee FICA taxes, the EITC, and the premium tax credit. It does not reflect state income tax liability or state supplements to the EITC. The value of the premium tax credit is obtained from the Kaiser Family Foundation's subsidy calculator for 2014 and is the national average for a 38-year old adult with income at the federal poverty level who does not use tobacco.

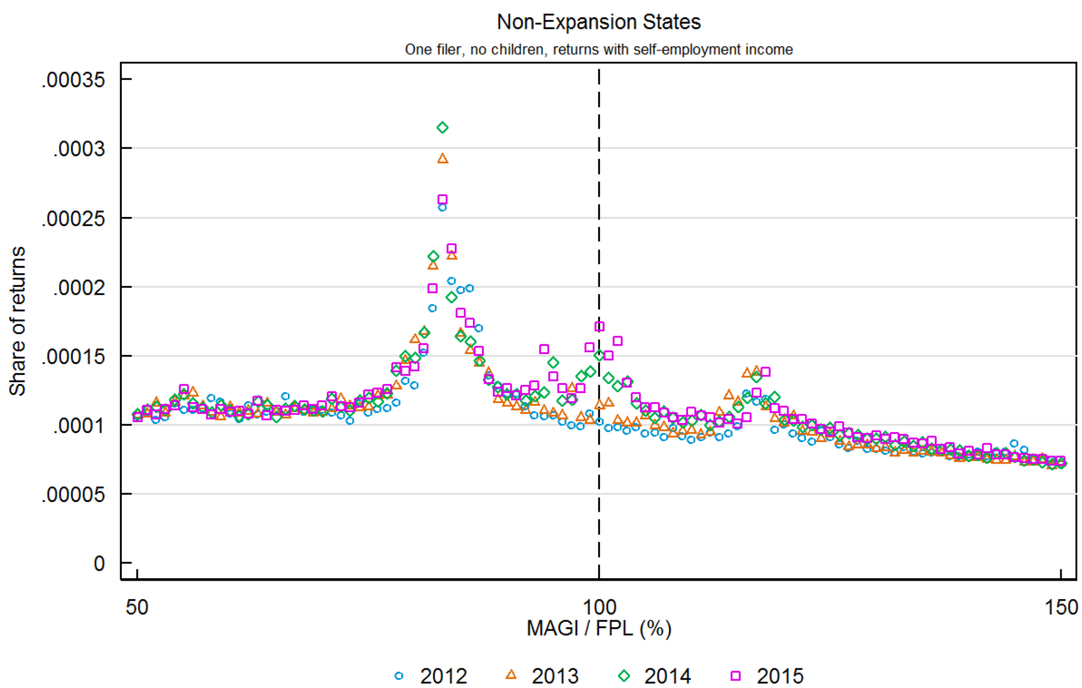
Figure 4: Indifference Curves that Induce Bunching at the Notch from Below



Note: The solid black line plots the budget constraint in the presence of the premium tax credit. The dashed black line plots the budget constraint in the absence of the premium tax credit. The dashed orange line depicts the indifference curve of the marginal buncher, who achieves the same level of utility with or without the premium tax credit. The blue lines plot indifference curves for someone who reaches I higher level of utility by bunching at the premium tax credit eligibility threshold after the introduction of the credit. The dashed line depicts the choice of effort in the absence of the premium tax credit (z_0), and the solid line depicts the choice of effort in the presence of the premium tax credit (z_1). Utility is increasing in the northwest direction.

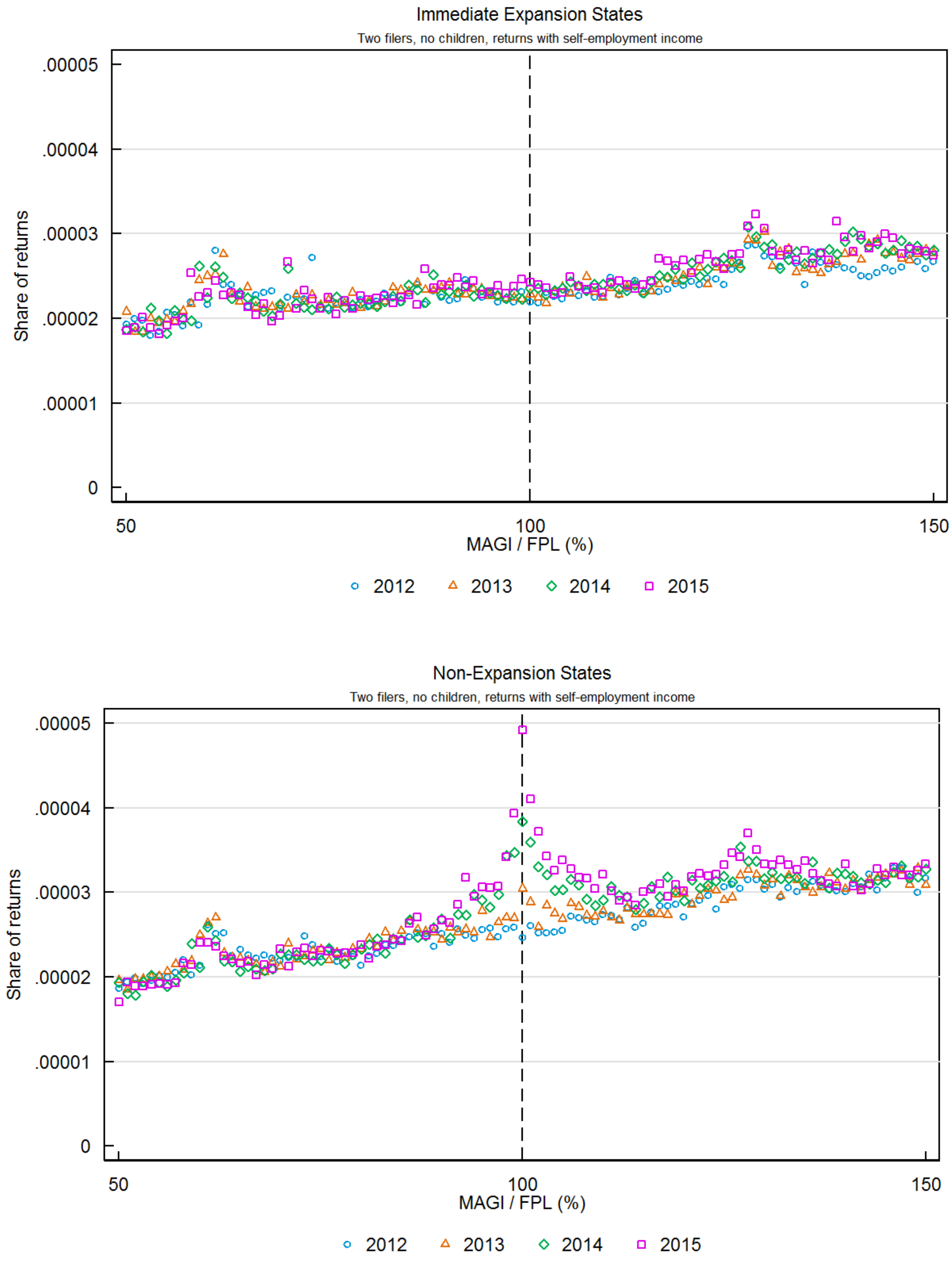
Figure 5: Distribution of Returns with Self-Employment Income by Year and Medicaid Expansion Status, Single Filers, No Children





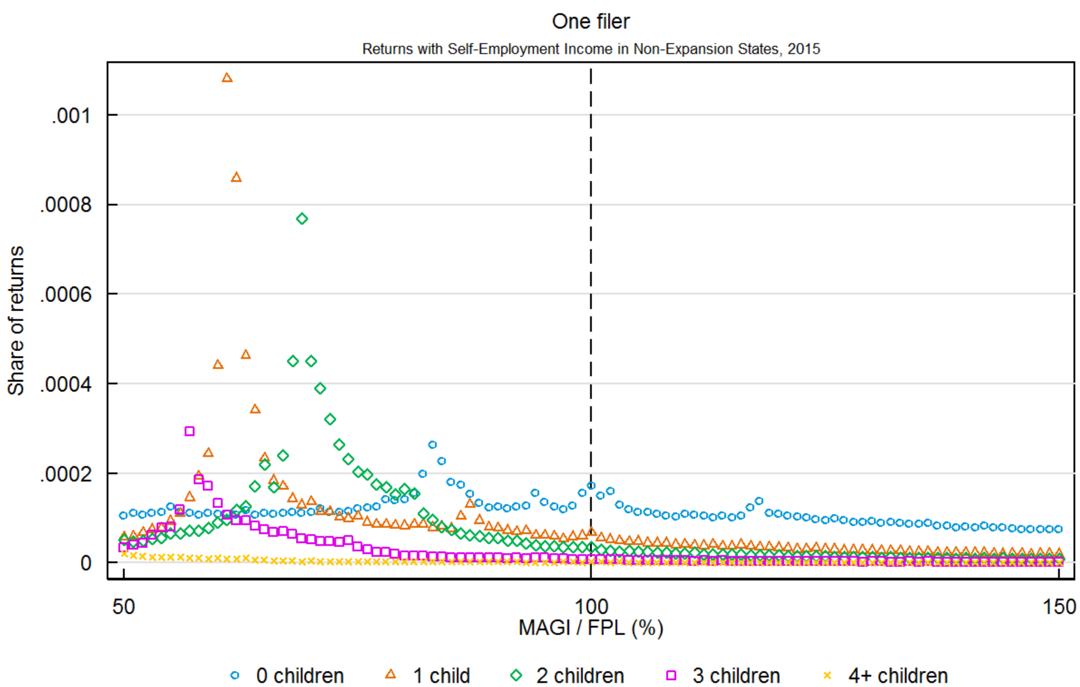
Source: Authors' calculations, IRS Form 1040 data, 2012 through 2015. Estimates in the income range and states shown are based on 5,275,428 total records.
Note: Each plotted point represents the share of all tax returns within groups of states defined by Medicaid expansion status that falls in a given income bin. Bins are defined using modified adjusted gross income (MAGI) relative to the federal poverty level (FPL). Bins are 1 percent FPL in width.

Figure 6: Distribution of Returns with Self-Employment Income by Year and Medicaid Expansion, Married Filers, No Children



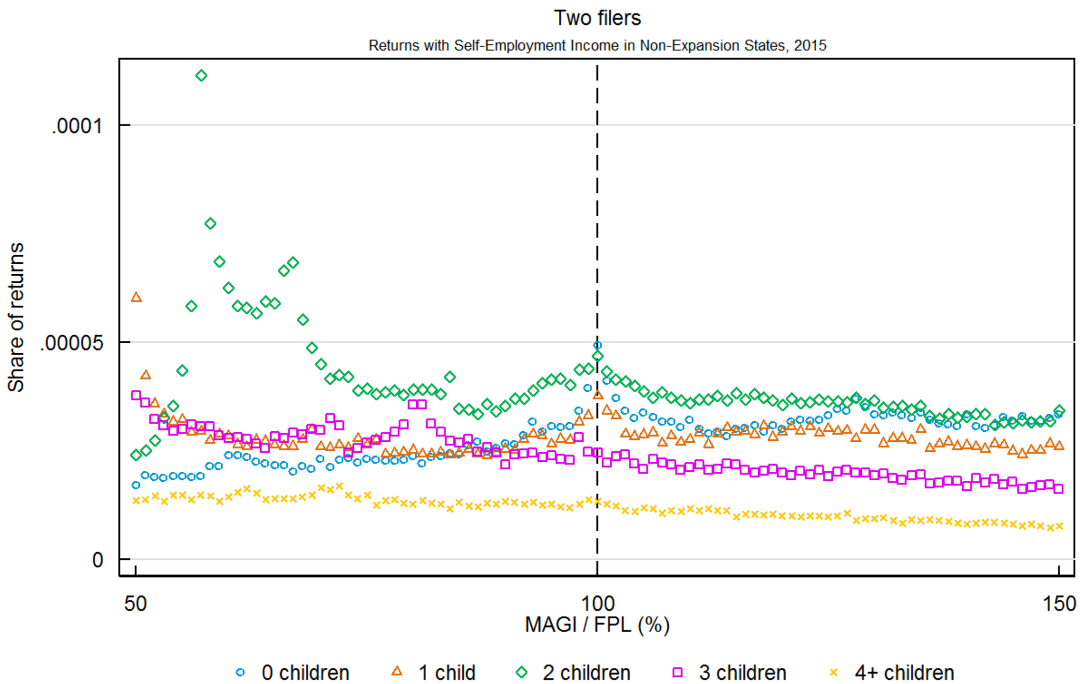
Source: Authors' calculations, IRS Form 1040 data, 2012 through 2015. Estimates in the income range and states shown are based on 1,184,054 total records.
Note: Each plotted point represents the share of all tax returns within groups of states defined by Medicaid expansion status that falls in a given income bin. Bins are defined using modified adjusted gross income (MAGI) relative to the federal poverty level (FPL). Bins are 1 percent FPL in width.

Figure 7: Distribution of Returns with Self-Employment Income by Number of Children, Single Filers, 2015



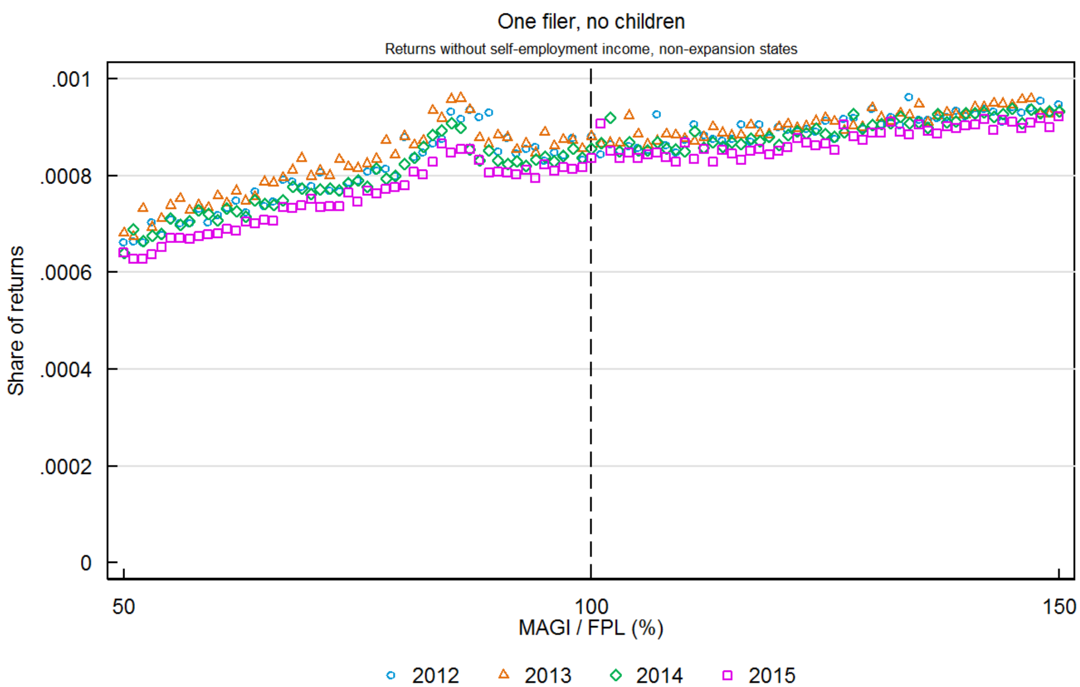
Source: Authors' calculations, IRS Form 1040 data, 2015. Estimates in the income range and states shown are based on 1,553,697 total records.
Note: Each plotted point represents the share of all tax returns within Medicaid non-expansion states that falls in a given income bin. Bins are defined using modified adjusted gross income (MAGI) relative to the federal poverty level (FPL). Bins are 1 percent FPL in width.

Figure 8: Distribution of Returns with Self-Employment Income by Number of Children, Married Filers, 2015



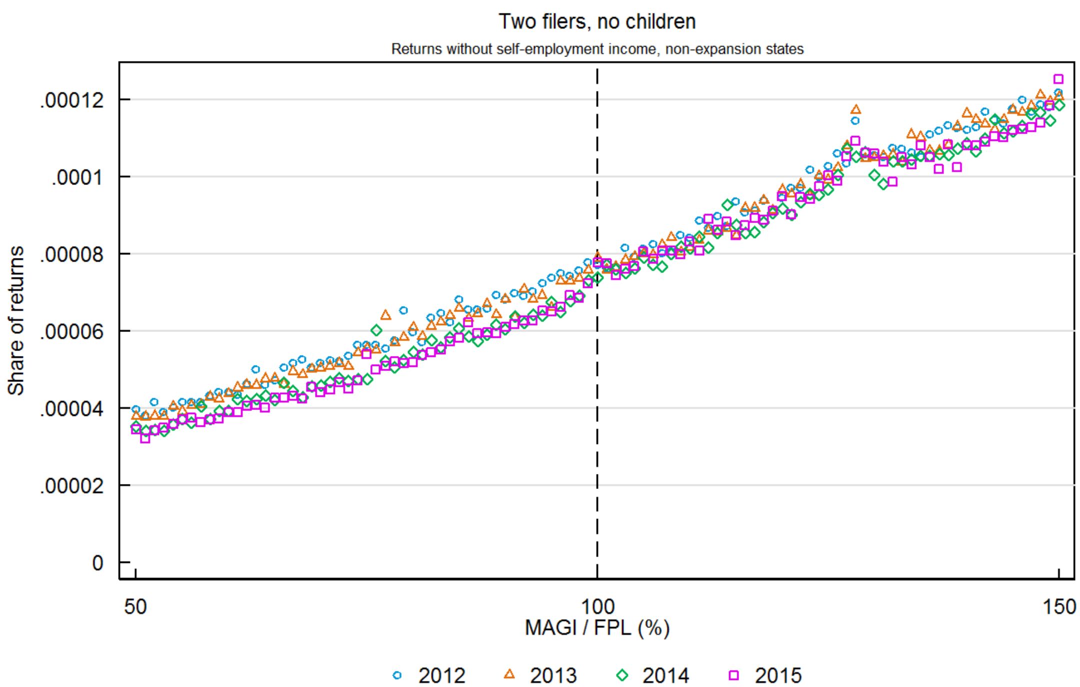
Source: Authors' calculations, IRS Form 1040 data, 2015. Estimates in the income range and states shown are based on 654,938 total records.
Note: Each plotted point represents the share of all tax returns within Medicaid non-expansion states that falls in a given income bin. Bins are defined using modified adjusted gross income (MAGI) relative to the federal poverty level (FPL). Bins are 1 percent FPL in width.

Figure 9: Distribution of Returns without Self-Employment Income, Single Filers without Children



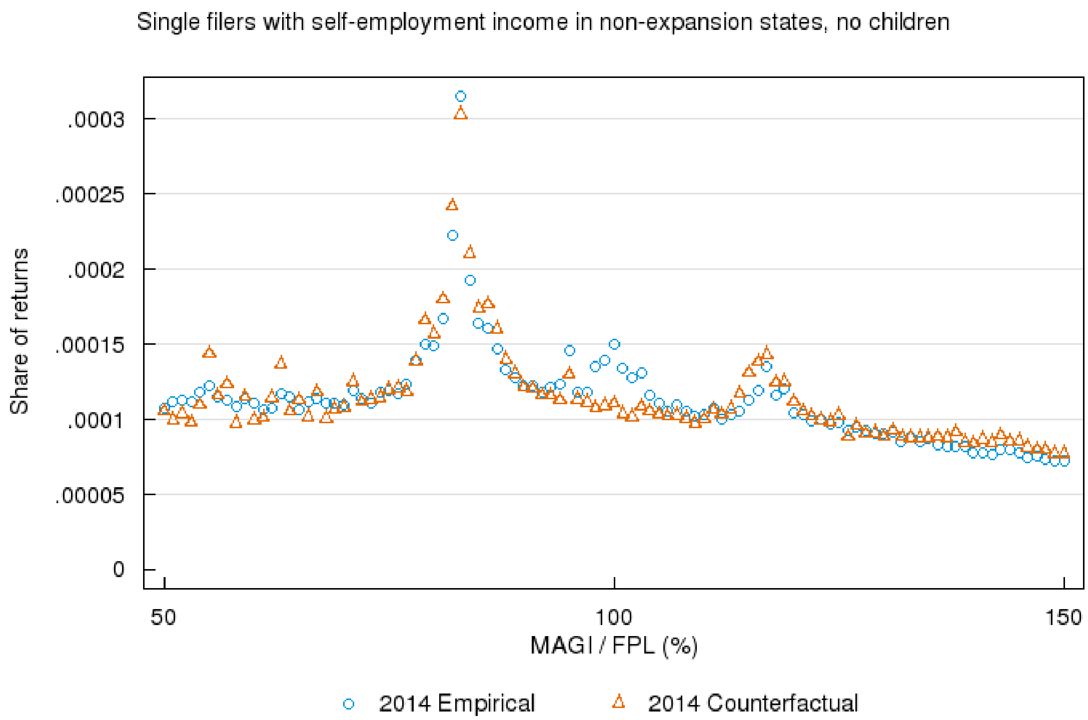
Source: Authors' calculations, IRS Form 1040 data, 2012 through 2015. Estimates in the income range and states shown are based on 16,202,875 records.
Note: Each plotted point represents the share of all tax returns within Medicaid non-expansion states that falls in a given income bin. Bins are defined using modified adjusted gross income (MAGI) relative to the federal poverty level (FPL). Bins are 1 percent FPL in width.

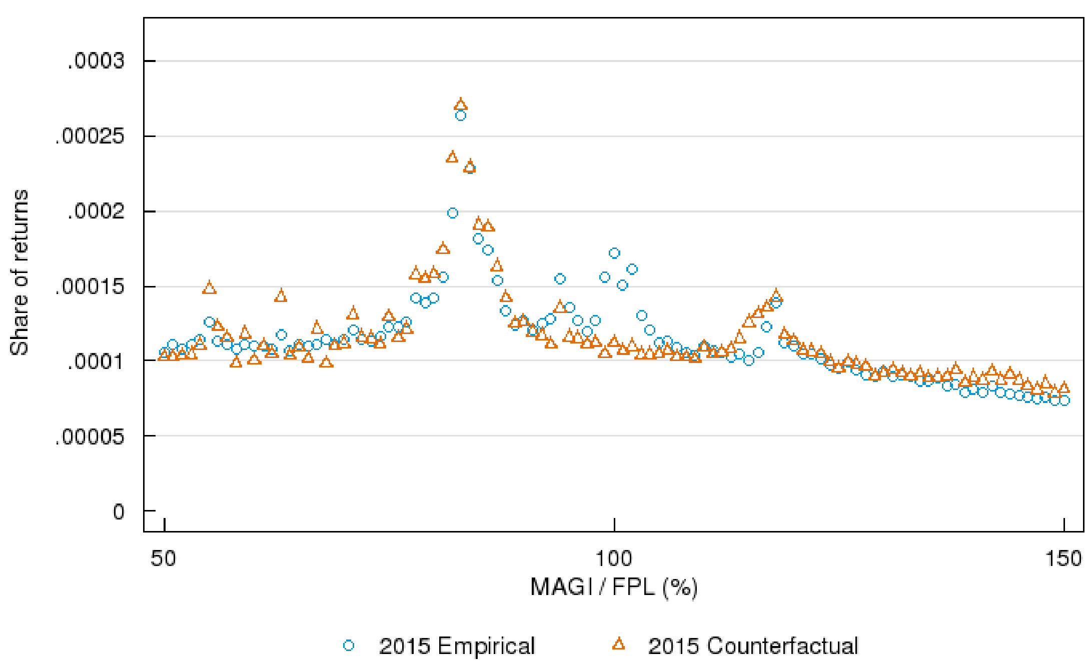
Figure 10: Distribution of Returns without Self-Employment Income, Married Filers without Children



Source: Authors' calculations, IRS Form 1040 data, 2012 through 2015. Estimates in the income range and states shown are based on 1,455,157 records.
Note: Each plotted point represents the share of all tax returns within Medicaid non-expansion states that falls in a given income bin. Bins are defined using modified adjusted gross income (MAGI) relative to the federal poverty level (FPL). Bins are 1 percent FPL in width.

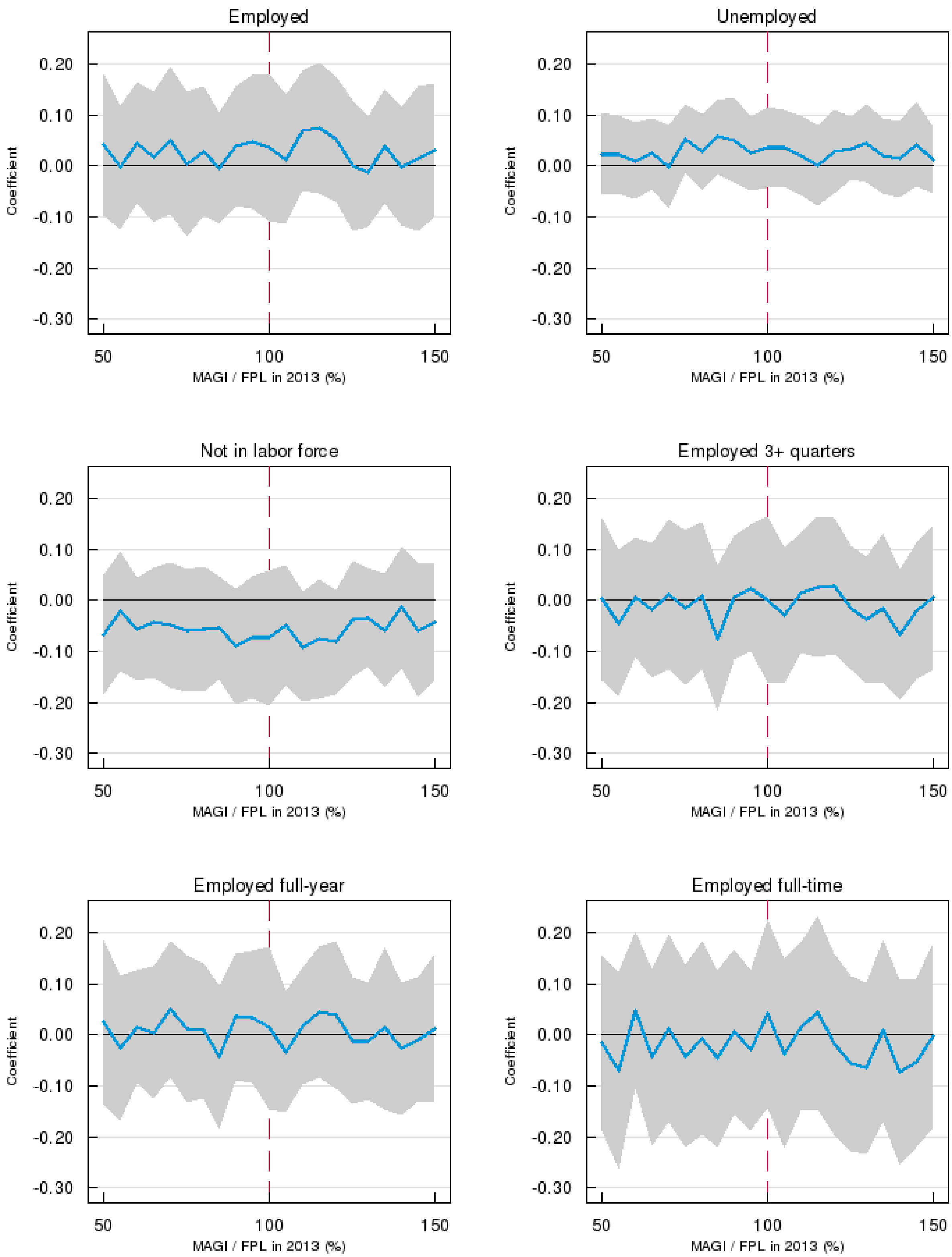
Figure 11: Counterfactual Distributions from Longitudinal Difference-in-Differences Approach



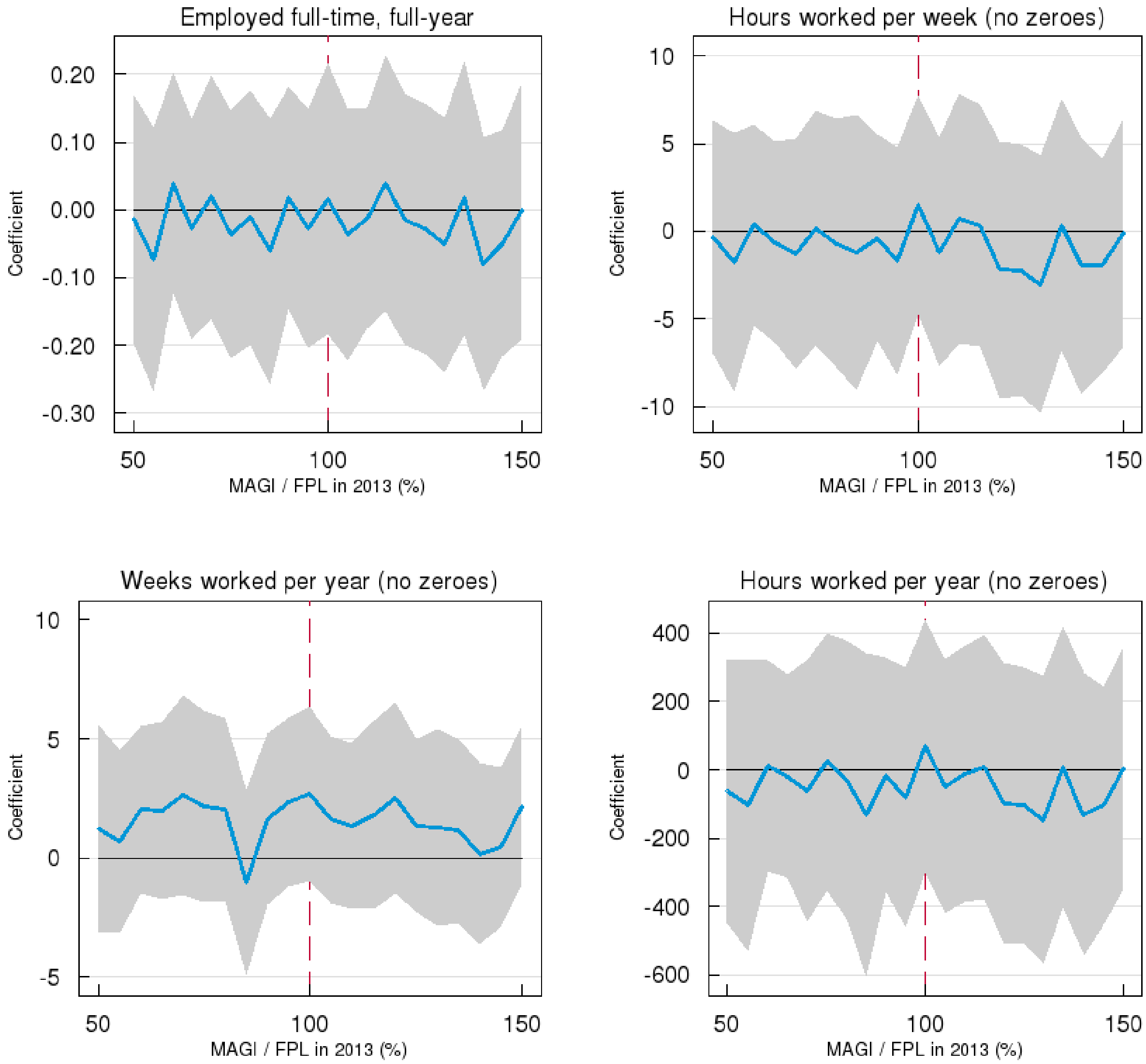


Source: Authors' calculations, IRS Form 1040 data, 2014 through 2015. Estimates of the empirical distribution in the income range and states shown are based on 1,116,378 total records.
Note: Each plotted point represents the share of all tax returns within Medicaid non-expansion states that falls in a given income bin. Bins are defined using modified adjusted gross income (MAGI) relative to the federal poverty level (FPL). Bins are 1 percent FPL in width. Counterfactual distributions are estimated as described in Section 4.3.

Figure 12: Changes in Employment Outcomes Near FPL in Non-expansion States Post-ACA

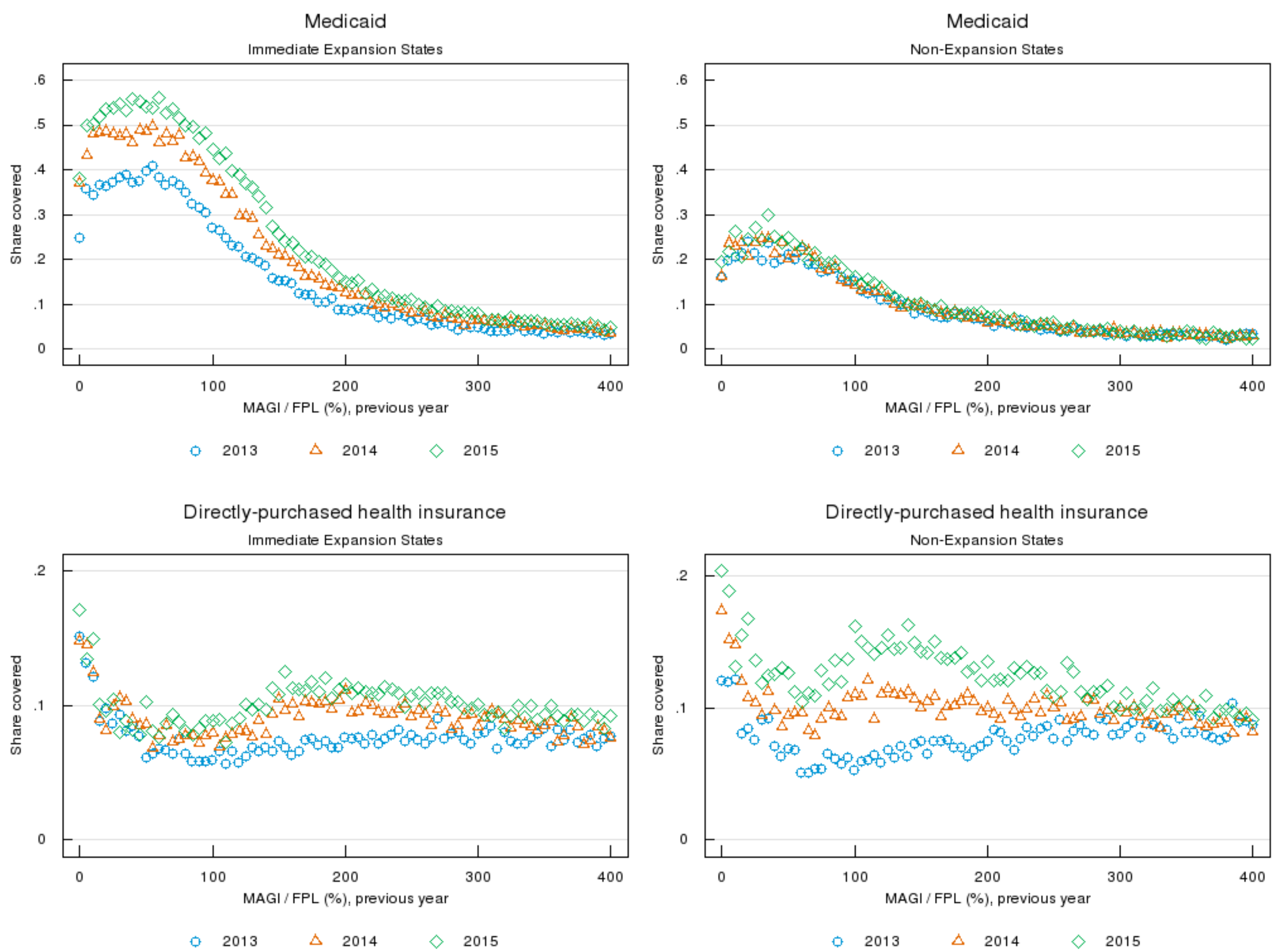


Continued – Changes in Employment Outcomes Near FPL in Non-expansion States Post-ACA



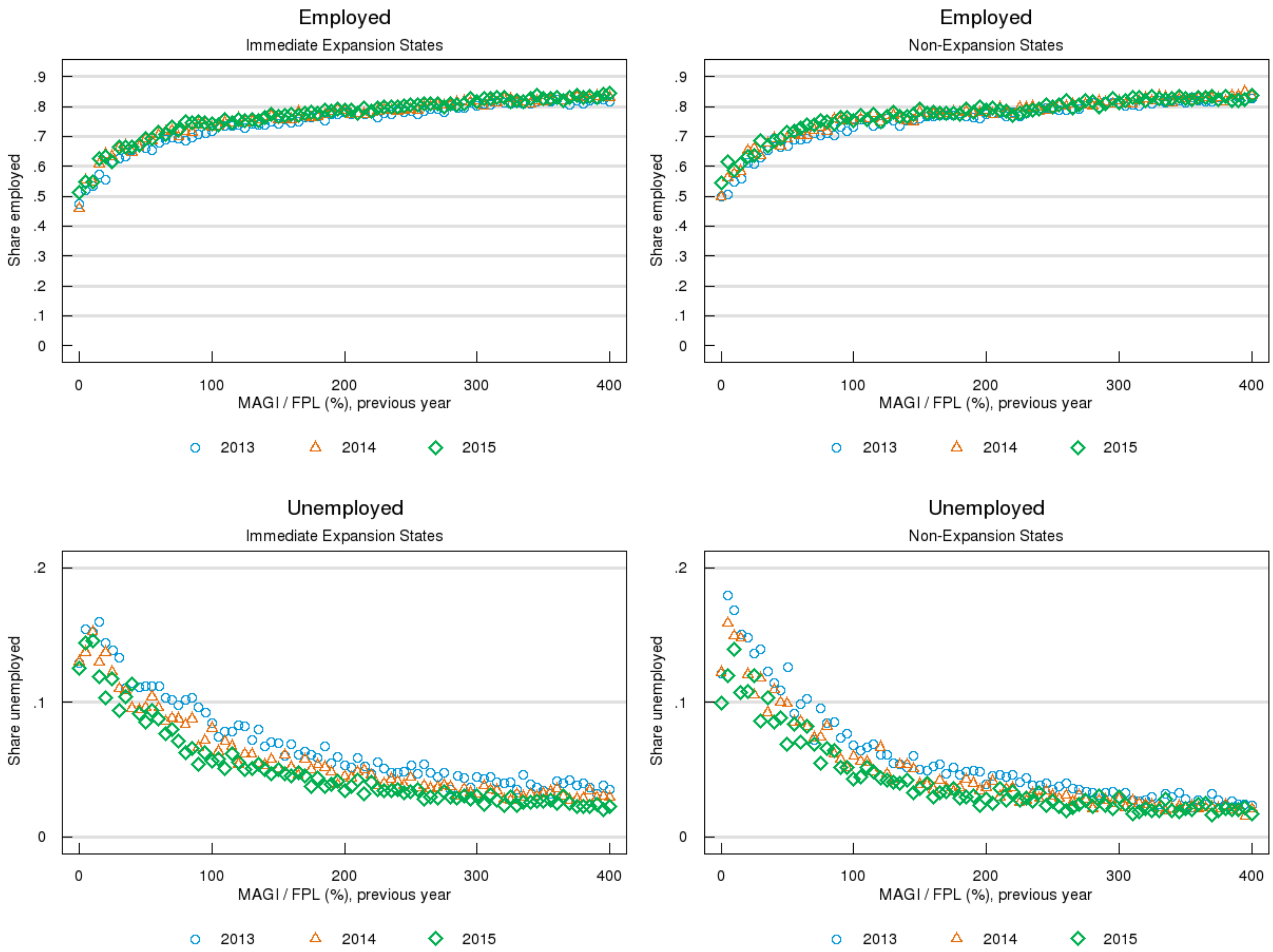
Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2012 through 2015. Estimates in the income range shown are based on 199,837 observations.
Note: The blue line plots the α_b coefficients estimated in equation 3. The shaded gray region represents the 95 percent confidence interval. Standard errors are clustered at the state level. Sample includes all ACS observations that were matched to 1040s with 2013 income between 50 and 150 percent of the federal poverty level (FPL) for groups that exhibit bunching behavior when the counterfactual distributions are generated using within-income group difference-in-differences regressions that include state group-specific linear trends (those listed in Table 1). Regressions are estimated using sample weights that have been adjusted for differences in match probabilities across groups. Full time is defined as at least 35 hours per week. Full year is defined as at least 50 weeks per year. Weeks per year is reported in bins in ACS. For this analysis, we use the midpoint of each bin. Analysis of hours and weeks outcomes excludes individuals who report not working. Hours per year is the product of hours per week and weeks per year. Coefficients are estimated by income bins that are 5 percent FPL in width and plotted at the bottom of the income range (i.e. the coefficient plotted at 95 percent FPL is based on people with income between 95 and 99 percent FPL, inclusive).

Figure 13: Health Insurance Coverage among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status

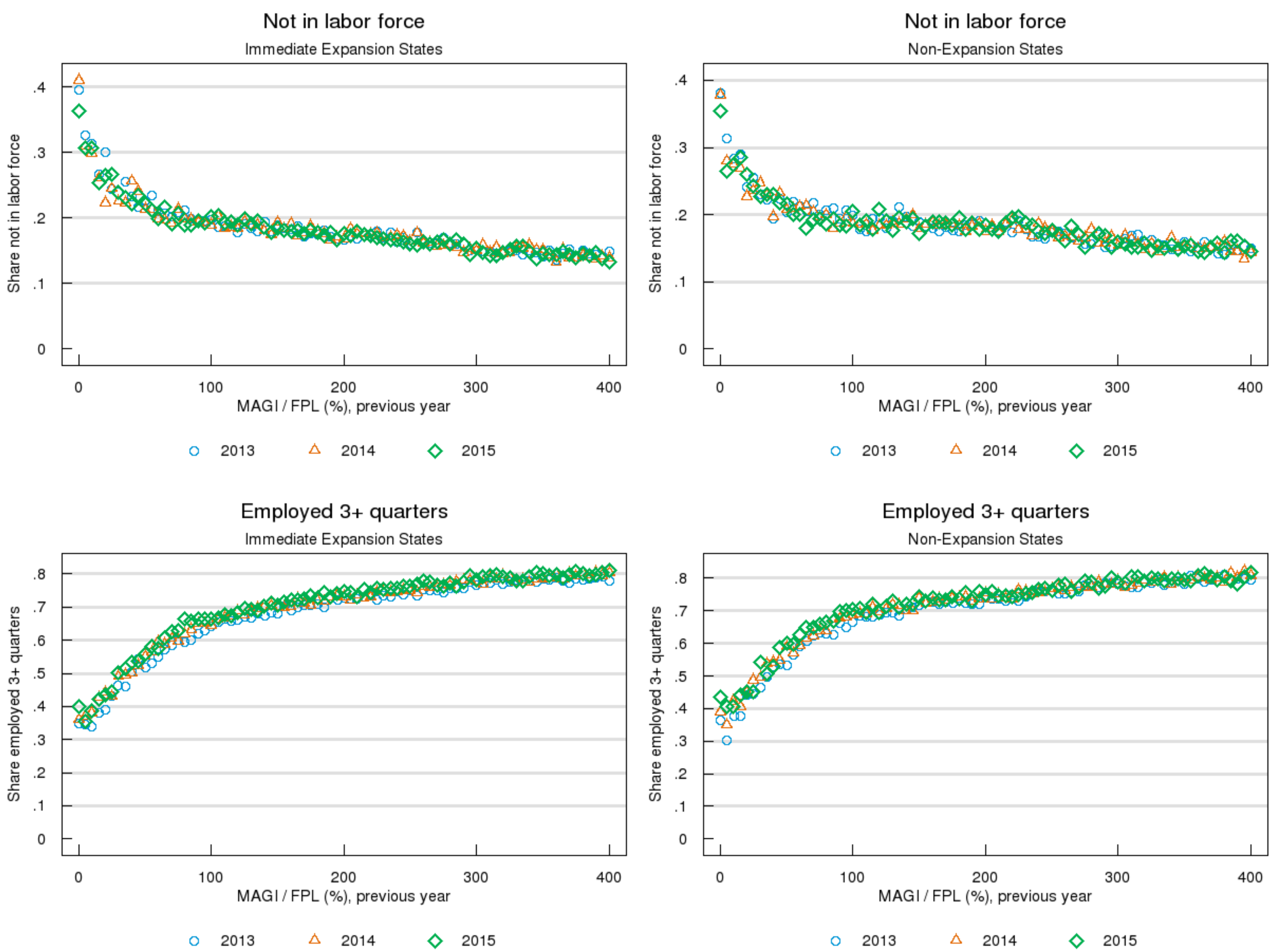


Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2012 through 2015. Estimates in the range shown are based on 1,533,091 observations.
Note: Sample includes individuals aged 26 to 64 without self-employment income. Income is obtained from the tax return on which each individual appeared for the year prior to responding to the ACS. Each point represent the mean outcome within an income bin that is 5 percent of the federal poverty level in width, plotted at the minimum value contained in the bin.

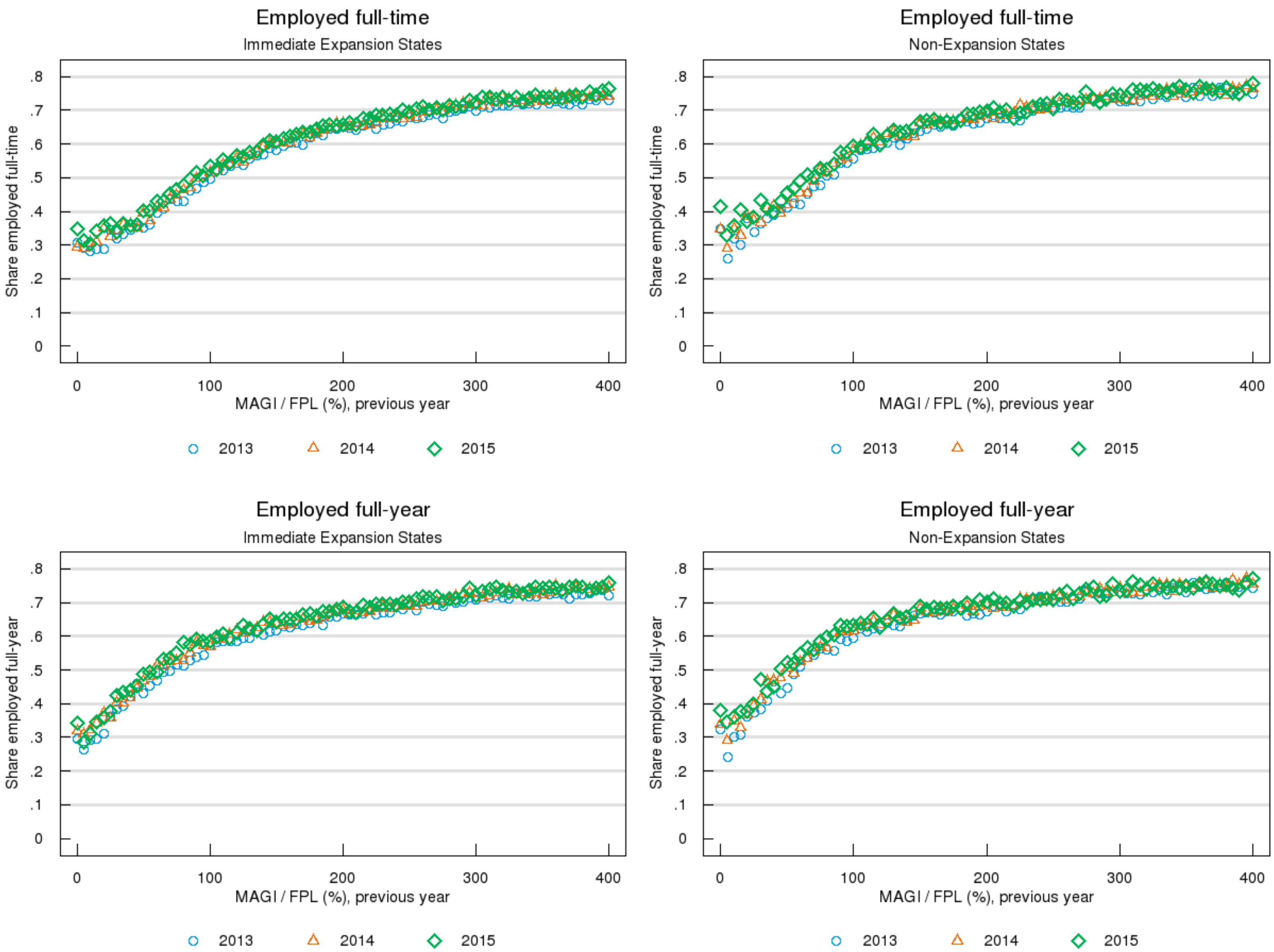
Figure 14: Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



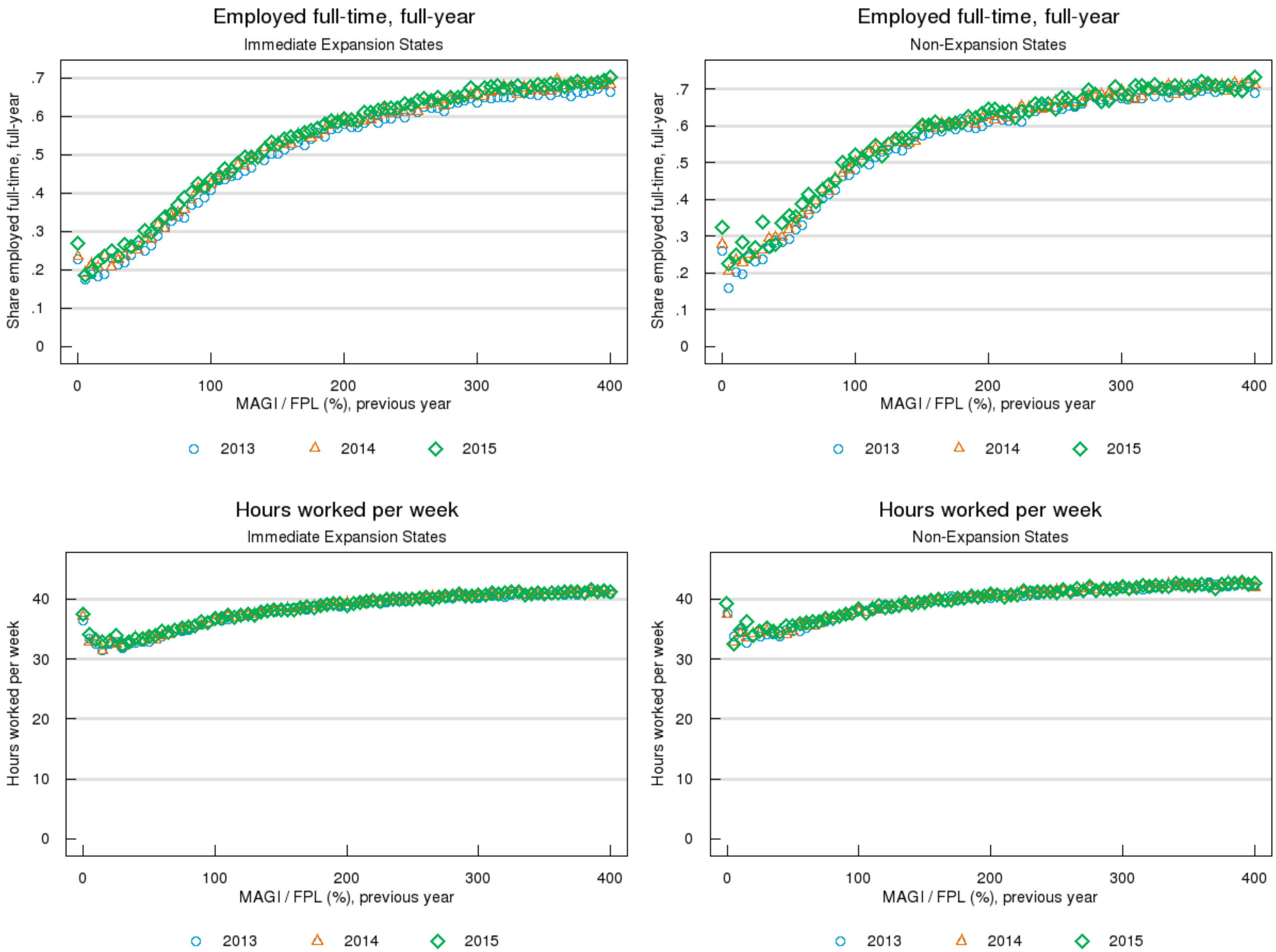
Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



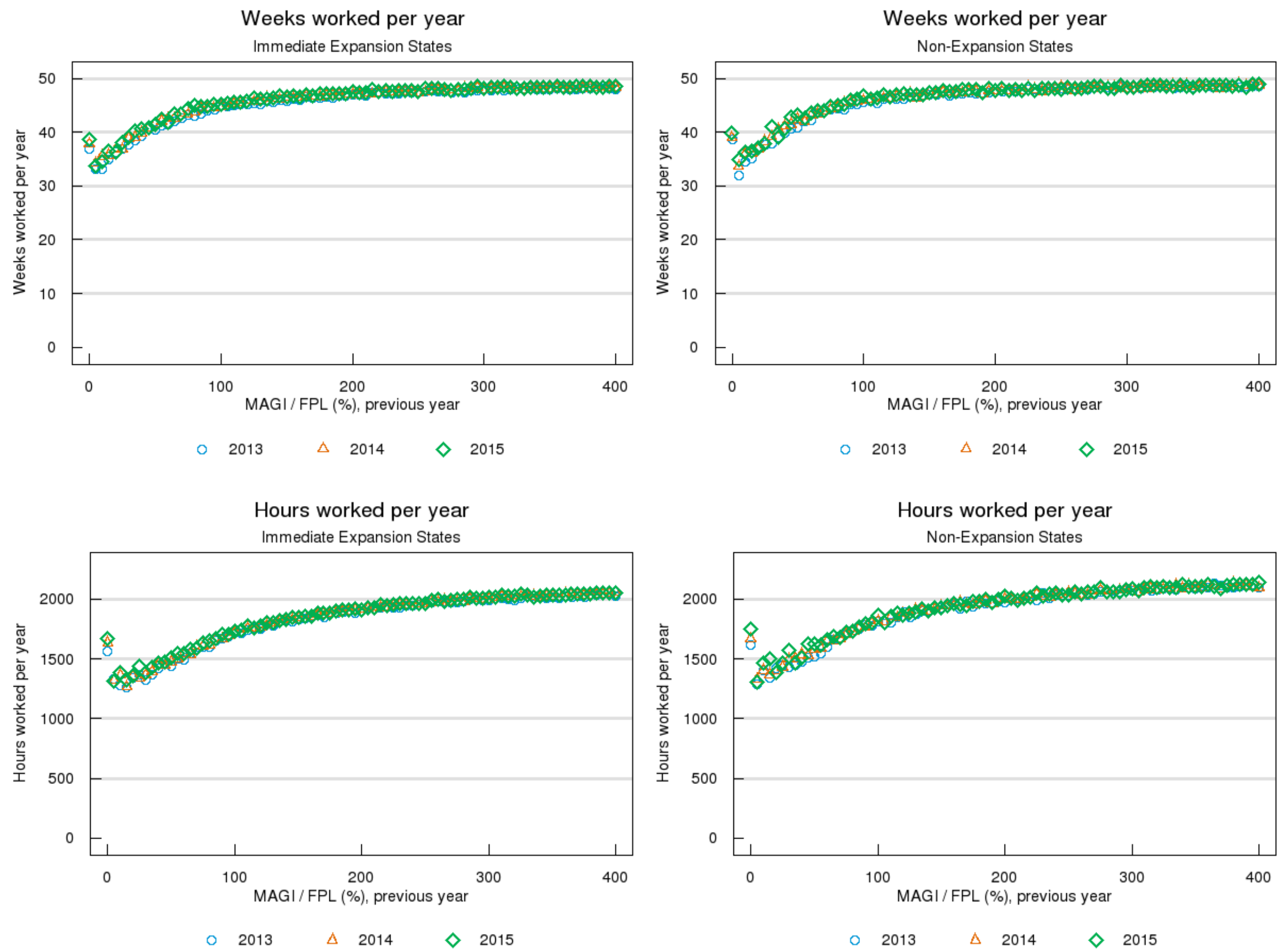
Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2012 through 2015. Estimates in the income range and states shown are based on 1,533,091 observations.
Note: Sample includes individuals aged 26 to 64 without self-employment income. Income is obtained from the tax return on which each individual appeared for the year prior to responding to the ACS. Full time is defined as at least 35 hours per week. Full year is defined as at least 50 weeks per year. Weeks per year is reported in bins in ACS. For this analysis, we use the midpoint of each bin. Analysis of hours and weeks outcomes excludes individuals who report not working. Hours per year is the product of hours per week and weeks per year. Each point represent the mean outcome within an income bin that is 5 percent of the federal poverty level in width, plotted at the minimum value contained in the bin.

Tables

Table 1: Elasticities of Taxable Income, Bunching Estimates, Filers with Self-employment Income

	(1)	(2)	(3)	(4)
	One Filer		Two Filers	
Children	2014	2015	2014	2015
0	0.730†***	0.897†***	0.704†***	0.888†***
	(0.107)	(0.118)	(0.097)	(0.122)
1	0.300	0.359	0.641†***	0.683†***
	(0.976)	(0.991)	(0.143)	(0.144)
2	0.700	0.981	0.997†***	1.061†*
	(1.498)	(1.947)	(0.333)	(0.546)
3	-1.334	-1.528	0.879†***	0.797†*
	(2.427)	(2.720)	(0.264)	(0.453)
4+	2.788†	-0.661	0.340***	0.980†***
	(5.022)	(1.116)	(0.115)	(0.232)

Source: Authors' calculations, IRS Form 1040 data, 2000 through 2015. Bunching estimates are based on 2,173,586 total returns filed in non-expansion states with income between 50 and 150 percent of the federal poverty level (FPL) that report self-employment income in 2014 and 2,208,635 such returns in 2015.
Note: Estimates are produced using counterfactual distributions for each group generated via within-income group difference-in-differences regressions that include state group-specific linear trends. Elasticities are estimated separately for groups defined by year, presence of self-employment income, number of filers, and number of children (topcoded at 4). Standard errors reported here are the standard deviation of 500 bootstrap estimates of the elasticities in question. Elasticities are calculated using subsidy values that assume all children receive health insurance coverage through CHIP. The baseline post-tax income measure incorporates federal and state taxes and credits, calculated using the National Bureau of Economic Research's TAXSIM program. Significant bunching at the 10 percent confidence level or greater is indicated by †. Asterisks indicate statistically significant elasticities at the 10 percent (*), 5 percent (**), and 1 percent (***) levels. In groups that do not exhibit statistically significant bunching, elasticities are calculated using the 95-105 percent FPL income range.

Table 2: Elasticities of Taxable Income, Bunching Estimates, Filers without Self-employment Income

	(1)	(2)	(3)	(4)
	One Filer		Two Filers	
Children	2014	2015	2014	2015
0	0.050	0.076*	-0.028	0.043
	(0.041)	(0.045)	(0.026)	(0.034)
1	-0.055	-0.042	0.001	-0.003
	(0.068)	(0.066)	(0.034)	(0.033)
2	-0.101	-0.079	-0.010	-0.011
	(0.133)	(0.161)	(0.039)	(0.052)
3	-0.067	-0.159**	-0.096**	-0.069
	(0.085)	(0.079)	(0.050)	(0.047)
4+	-0.427***	-0.285*	0.068	0.035
	(0.150)	(0.170)	(0.068)	(0.066)

Source: Authors' calculations, IRS Form 1040 data, 2000 through 2015. Bunching estimates are based on 8,620,062 total returns filed in non-expansion states with income between 50 and 150 percent of the federal poverty level (FPL) that do not report self-employment income in 2014 and 8,529,606 such returns in 2015. Note: Estimates are produced using counterfactual distributions for each group generated via within-income group difference-in-differences regressions that include state group-specific linear trends. Elasticities are estimated separately for groups defined by year, presence of self-employment income, number of filers, and number of children (topcoded at 4). Standard errors reported here are the standard deviation of 500 bootstrap estimates of the elasticities in question. Elasticities are calculated using subsidy values that assume all children receive health insurance coverage through CHIP. The baseline post-tax income measure incorporates federal and state taxes and credits, calculated using the National Bureau of Economic Research's TAXSIM program. Significant bunching at the 10 percent confidence level or greater is indicated by †. Asterisks indicate statistically significant elasticities at the 10 percent (*), 5 percent (**), and 1 percent (***) levels. In groups that do not exhibit statistically significant bunching, elasticities are calculated using the 95-105 percent FPL income range.

Table 3: Effect of Premium Tax Credit Notch on Self-employment Income Reporting

	(1) 1040 Share	(2) 1040 Share	(3) Population Share	(4) Population Share	(5) Count	(6) Count	(7) Log Count	(8) Log Count
Non-expansion, Post-ACA	0.00143	0.00114	-0.000355	-0.000128	19,321	-1,245	0.0184	-0.00999
	(0.00230)	(0.00106)	(0.000918)	(0.000396)	(25,809)	(3,382)	(0.0229)	(0.00739)
Non-expansion	0.00462***	-4.178***	0.00268***	-1.917***	-9,206***	-6.228e+07***	-0.198***	-37.82***
	(0.000288)	(0.0874)	(0.000115)	(0.0327)	(3,226)	(279,108)	(0.00287)	(0.610)
Post-ACA	0.00950***	-0.000377	0.00613***	0.000240	85,789***	2,039	0.240***	0.00146
	(0.00239)	(0.000510)	(0.000993)	(0.000273)	(21,346)	(1,987)	(0.0237)	(0.00475)
Observations	816	816	816	816	816	816	816	816
R-squared	0.913	0.987	0.936	0.990	0.986	1.000	0.998	1.000
State Trends	None	Linear	None	Linear	None	Linear	None	Linear

Source: Authors' calculations, IRS Form 1040 data.

Note: Regressions are estimated on data for all filers, aggregated to the state level for 2000 through 2015. All regressions include state and year fixed effects. The dependent variables are the share of 1040 returns that report self-employment income (columns 1-2), the number of returns reporting self-employment income divided by the state population (3-4), the number of returns reporting self-employment income (5-6), and the log of the number of returns reporting self-employment income (7-8). Results are similar when the data are restricted to filers with income between 1 percent and 500 percent of the federal poverty level. Asterisks indicate statistically significant elasticities at the 10 percent (*), 5 percent (**), and 1 percent (***) levels.

Table 4: Elasticities of Taxable Income, Taxpayers with Self-Employment Income, by Counterfactual Distribution Estimation Method, Subsidies Assigned According to Number of Filers

Year	Filers	Children	(1) Baseline	(2) Linear	(3) Quadratic	(4) Cubic	(5) 2012	(6) 2013
2014	1	0	0.730†***	0.816†***	0.625†***	0.625†***	0.802†***	0.600†***
			(0.107)	(0.134)	(0.139)	(0.139)	(0.151)	(0.099)
2014	2	0	0.704†***	0.246†***	0.514†***	0.514†***	0.728†***	0.533†***
			(0.097)	(0.038)	(0.083)	(0.083)	(0.100)	(0.076)
2014	2	1	0.641†***	0.769†***	1.142†***	1.143†***	0.859†***	0.486†***
			(0.143)	(0.232)	(0.265)	(0.265)	(0.231)	(0.086)
2014	2	2	0.997†***	-0.383†	-0.042	-0.041	0.693†	0.590†**
			(0.333)	(0.549)	(0.241)	(0.241)	(0.828)	(0.250)
2014	2	3	0.879†***	-0.416†	-0.512†	-0.512†	0.866†	0.430†**
			(0.264)	(0.415)	(0.786)	(0.787)	(0.545)	(0.212)
2015	1	0	0.897†***	1.037†***	0.784†***	0.783†***	0.935†***	0.754†***
			(0.118)	(0.149)	(0.170)	(0.171)	(0.129)	(0.118)
2015	2	0	0.888†***	0.259†***	0.602†***	0.602†***	0.982†***	0.761†***
			(0.122)	(0.035)	(0.091)	(0.091)	(0.136)	(0.110)
2015	2	1	0.683†***	1.124†***	1.183†***	1.186†***	0.849†***	0.542†***
			(0.144)	(0.317)	(0.272)	(0.273)	(0.248)	(0.105)
2015	2	2	1.061†*	-0.283†	-0.024	-0.024	0.906†	0.692†
			(0.546)	(0.322)	(0.344)	(0.345)	(1.431)	(0.422)
2015	2	3	0.797†*	0.263	0.151	0.152	1.274†	1.253†***
			(0.453)	(0.243)	(0.469)	(0.470)	(1.090)	(0.486)

Source: Authors' calculations, IRS Form 1040 data, 2000 through 2015. Bunching estimates are based on 2,173,586 total returns filed in non-expansion states with income between 50 and 150 percent of the federal poverty level (FPL) that report self-employment income in 2014 and 2,208,635 such returns in 2015.
Note: Estimates are produced using counterfactual distributions for each group generated via the technique indicated at the top of the row. The DD column provides the main estimates for reference. The linear, quadratic, and cubic columns estimate counterfactual distributions using within-bin regressions of the share of returns on a time trend of the given order. The 2012 (2013) column uses the the baseline difference-in-differences approach to identify the bunching region but uses the 2012 (2013) empirical distribution as the counterfactual distribution. Elasticities are estimated separately for groups defined by year, presence of self-employment income, number of filers, and number of children (topcoded at 4). Groups listed exhibited significant bunching behavior that implied statistically significant elasticities of reported income in both 2014 and 2015 in the baseline approach. Standard errors reported here are the standard deviation of 500 bootstrap estimates of the elasticities in question. Elasticities are calculated using subsidy values that assume all children receive health insurance coverage through CHIP. The baseline post-tax income measure incorporates federal and state taxes and credits, calculated using the National Bureau of Economic Research's TAXSIM program. Significant bunching at the 10 percent confidence level or greater is indicated by †. Asterisks indicate statistically significant elasticities at the 10 percent (*), 5 percent (**), and 1 percent (***) levels. In groups that do not exhibit statistically significant bunching, elasticities are calculated using the 95-105 percent FPL income range.

Table 5: Elasticities of Taxable Income Under Alternative Baseline Taxation and Subsidy Assignment Assumptions

			(1)	(2)	(3)	(4)
			Subsidy Assignment:			
			Number of Filers		Household Size	
Year	Filers	Children	Post-Tax	Pre-Tax	Post-Tax	Pre-Tax
2014	1	0	0.730	0.791	0.730	0.791
			(0.107)	(0.116)	(0.107)	(0.116)
2014	2	0	0.704	0.746	0.704	0.746
			(0.097)	(0.103)	(0.097)	(0.103)
2014	2	1	0.641	0.562	0.507	0.444
			(0.143)	(0.125)	(0.113)	(0.099)
2014	2	2	0.997	0.810	0.649	0.528
			(0.333)	(0.271)	(0.217)	(0.176)
2014	2	3	0.879	0.724	0.483	0.398
			(0.264)	(0.217)	(0.145)	(0.119)
2015	1	0	0.897	0.970	0.897	0.970
			(0.118)	(0.128)	(0.118)	(0.128)
2015	2	0	0.888	0.942	0.888	0.942
			(0.122)	(0.129)	(0.122)	(0.129)
2015	2	1	0.683	0.599	0.540	0.474
			(0.144)	(0.126)	(0.114)	(0.010)
2015	2	2	1.061	0.863	0.691	0.562
			(0.546)	(0.444)	(0.355)	(0.289)
2015	2	3	0.797	0.657	0.439	0.362
			(0.453)	(0.374)	(0.250)	(0.206)

Source: Authors' calculations, IRS Form 1040 data, 2000 through 2015. Bunching estimates are based on 2,173,586 total returns filed in non-expansion states with income between 50 and 150 percent of the federal poverty level that report self-employment income in 2014 and 2,208,635 such returns in 2015.
Note: Estimates are produced using the same methods of identifying bunching and estimating counterfactual distributions used in the baseline analysis. Groups listed exhibited significant bunching behavior that implied statistically significant elasticities of reported income in both 2014 and 2015. Different approaches to subsidy assignment assume children get health insurance through CHIP (number of filers) or, counterfactually, purchased using the premium tax credit (household size). Pre-tax elasticities use modified adjusted gross income as the baseline resource measure, while post-tax elasticities incorporate federal and state taxes and credits, calculated using the National Bureau of Economic Research's TAXSIM program. The baseline estimates reported in Table 1 assign subsidies based on number of filers and use the post-tax income measure; these estimates are reproduced in the first column of this table. Standard errors reported here are the standard deviation of 500 bootstrap estimates of the elasticities in question. Bunching and statistical significance indicators are omitted because they are consistent across columns and identical to those reported in Table 1.

Appendix A Adjustments to ACS Sample Weights

Due to variation in the quality of the Personally Identifiable Information (PII) provided across data sources, and across respondents within data sources, CARRA's PVS system is not able to match every individual record to a PIK, especially in surveys like the ACS in which respondents' social security numbers are not provided. On average, the match rate between ACS records and PIKs is high (around 91 percent each year for 2012-2015), but there are differences in match rates across demographic groups that suggest PIKs are not missing at random. For example, 93-94 percent of white respondents receive PIKs in each year of ACS data used here, while only about 88 percent of black respondents and 80-81 percent of Hispanic respondents receive PIKs. There are also smaller differences in match rates across age, sex, and education groups.

Given that the observable characteristics of the people who are assigned PIKs differ from those of people who are not assigned PIKs, one might be concerned that their unobservable characteristics also differ in a way that could bias estimates that are based on matched ACS-tax data. In order to account for these differences between ACS respondents with and without PIKs, we adjust the ACS sample weights using the inverse of each observation's predicted probability of receiving a PIK, based on logistic regressions of PIK-receipt on measures of age, race, sex, and education. Specifically, we estimate

$$HasPIK_i = \sum_{j=1}^{9} \alpha_j Age_i^j + \sum_{j=1}^{8} \beta_j Race_i^j + \sum_{j=1}^{3} \gamma_j Sex_i^j + \sum_{j=1}^{8} \delta_j Educ_i^j + \varepsilon_i \tag{A1}$$

where $HasPIK_i$ is an indicator for presence of a PIK, and the other variables are indicators for membership in various age, race, sex, and education categories. The age categories are 0-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+, and missing. The race/ethnicity categories are white, black, Native American, Asian, Native Hawaiian/Pacific Islander, other, Hispanic, and missing, with Hispanics of any race included only in the Hispanic category. The sex categories are male, female, and missing. The education categories are less than high school, high school, GED, some college but no degree, associate's degree, bachelor's degree, graduate degree, and missing. We estimate this regression separately for each year. Coefficients are then used to predict π, the probability that each individual is assigned a PIK. We then create adjusted weights by multiplying each individual's ACS sample weight by $1/\pi$.

Appendix B Difference-in-Discontinuities Details and Tables

The figures discussed in Section 5.3 suggest that there is some differential change in health insurance coverage above the premium tax credit eligibility threshold after the credit became available but little change in labor market outcomes. It can be difficult, however, to assess the magnitude of these changes visually. In order to provide some clarity on that front, we present descriptive estimates of the changes in health insurance coverage and labor market outcomes near the threshold.

The estimates are produced using a difference-in-discontinuities strategy similar to Grembi et al. (2012). As the name suggests, the difference-in-discontinuities framework combines aspects of regression discontinuity (RD) analysis with aspects of difference-in-differences analysis.[38] Conceptually, the approach essentially controls for breaks in outcomes of interest that occur at the policy-relevant threshold but which predate the policy in question by comparing regression discontinuity estimates at that threshold from before and after the introduction of the policy of interest. This analysis uses matched ACS-tax return data for people without self-employment income from 2012-2015. In order to focus on the people most likely to be affected by the credit, we exclude people below the age of 26 (who may receive health insurance coverage through their parents' plans) and above the age of 64 (who are eligible for Medicare). We also exclude people who report zero or negative income, as well as those whose income exceeds 2,000 percent of the federal poverty level (FPL). As in the figures above, we use MAGI from the previous year as the running variable in this analysis.

We take two approaches to difference-in-discontinuities analysis. The first, which we will refer to as the flexible functional form (FFF) approach, uses the full dataset and involves fitting flexible polynomials on both sides of the premium tax credit eligibility threshold before and after the credits became available. We estimate the equation

$$\begin{aligned} L_{ist} = &\sum_{k=0}^{p}\left(\delta_k Y_i^{\star k}\right) + Elig_i \sum_{k=0}^{p}\left(\gamma_k Y_i^{\star k}\right) + After_t\left[\sum_{k=0}^{p}\left(\alpha_k Y_i^{\star k}\right) + Elig_i \sum_{k=0}^{p}\left(\beta_k Y_i^{\star k}\right)\right] + \\ &\lambda X_i + \eta_s + \theta_t + \varepsilon_{ist} \end{aligned} \tag{B1}$$

where L_{ist} is a measure of health insurance coverage or labor supply for person i in state s at time t. $Y_i^\star$ is a measure of income relative to poverty that has been re-centered around the poverty line, that is, $Y_i^\star = Y_i - 100$. $Elig_i$ indicates that $Y_i^\star < 0$, and $After_t$ indicates that an observation is from 2014 or 2015. The order of the polynomials is given by p, and here we set $p = 7. X_i$ contains controls for age, sex, race, education, and family structure. The difference-in-discontinuities estimate is given by β_0.

[38] For more details on RD estimation, see Imbens and Lemieux (2008), Van der Klaauw (2008), and Lee and Lemieux (2010).

The other approach focuses on the region close to the premium tax credit eligibility threshold. Instead of high-order polynomials estimated over a wide income range, this approach uses local linear regressions (LLR) within a narrow window surrounding the threshold. We estimate the equation

$$\begin{aligned} L_{ist} = &\delta_0 + \delta_1 Y_i^\star + Elig_i\left(\gamma_0 + \gamma_1 Y_i^\star\right) + After_t\left[\left(\alpha_0 + \alpha_1 Y_i^\star\right) + Elig_i\left(\beta_0 + \beta_1 Y_i^\star\right)\right] + \\ &\lambda X_i + \eta_s + \theta_t + \varepsilon_{ist} \end{aligned} \tag{B2}$$

where variables are defined and indexed as above. We use individuals whose income is within 30 percent FPL of the eligibility threshold in our baseline analysis. The parameter of interest is again β_0.

In addition to the demographic controls included in these regressions, the difference-in-discontinuities design also controls for possible discontinuities in outcomes of interest that occur at the premium tax credit eligibility threshold but predate the credit's introduction. Such discontinuities could arise from the fact that he eligibility threshold is also the FPL. It is conceivable that social pressure or stigma surrounding the poverty designation, or other factors unobservable to the econometrician, could lead individuals to sort themselves onto one side of the FPL or the other, giving rise to discontinuities in the conditional mean functions for some outcomes at that point in the income distribution. Thus, the estimates presented in the tables below reflect changes in the listed outcomes that occur at the eligibility threshold due to factors other than demographics and pre-ACA sorting around the FPL.

Table B1: Mean Characteristics of Individuals Near the Premium Tax Credit Eligibility Threshold

	2012-13			2014-15			
	Below FPL	Above FPL	Difference	Below FPL	Above FPL	Difference	Difference in Differences
Age	40.100	40.170	0.070	39.940	40.090	0.150	0.080
	(0.214)	(0.251)	(0.330)	(0.205)	(0.275)	(0.343)	(0.476)
Female	0.617	0.610	-0.007	0.619	0.605	-0.014	-0.007
	(0.006)	(0.006)	(0.008)	(0.007)	(0.009)	(0.011)	(0.014)
White	0.416	0.407	-0.009	0.398	0.423	0.025	0.034
	(0.056)	(0.054)	(0.078)	(0.049)	(0.053)	(0.073)	(0.107)
Black	0.275	0.271	-0.004	0.296	0.278	-0.018	-0.014
	(0.045)	(0.046)	(0.064)	(0.044)	(0.043)	(0.061)	(0.089)
Native American	0.012	0.0115	-0.0005	0.0105	0.0105	0.0000	0.0005
	(0.0044)	(0.0041)	(0.0060)	(0.0035)	(0.0035)	(0.0049)	(0.0077)
Asian	0.0285	0.0293	0.0008	0.0299	0.0281	-0.0018	-0.0026
	(0.0039)	(0.0040)	(0.0056)	(0.0033)	(0.0032)	(0.0045)	(0.0072)
Native HI/PI	0.00162	0.00131	-0.00031	0.00305	0.00135	-0.00170	-0.00139
	(0.00053)	(0.00054)	(0.00076)	(0.00069)	(0.00044)	(0.00082)	(0.00112)
Other Race	0.00203	0.00204	0.00001	0.00187	0.00122	-0.00065	-0.00066
	(0.00053)	(0.00043)	(0.00068)	(0.00072)	(0.00029)	(0.00078)	(0.00103)
Hispanic	0.265	0.278	0.013	0.261	0.258	-0.003	-0.016
	(0.090)	(0.086)	(0.125)	(0.078)	(0.084)	(0.115)	(0.169)
Less than HS	0.218	0.208	-0.010	0.204	0.188	-0.016	-0.006
	(0.026)	(0.024)	(0.036)	(0.015)	(0.022)	(0.026)	(0.044)
High School	0.272	0.280	0.008	0.277	0.282	0.005	-0.003
	(0.011)	(0.011)	(0.015)	(0.007)	(0.008)	(0.010)	(0.018)
GED	0.066	0.062	-0.004	0.061	0.068	0.007	0.011
	(0.004)	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)	(0.006)
Some College	0.255	0.257	0.002	0.259	0.251	-0.008	-0.010
	(0.008)	(0.008)	(0.011)	(0.007)	(0.006)	(0.010)	(0.015)
Assoc. Degree	0.073	0.080	0.007	0.085	0.085	0.000	-0.007
	(0.009)	(0.008)	(0.013)	(0.004)	(0.009)	(0.010)	(0.016)
Bach. Degree	0.088	0.088	-0.000	0.085	0.098	0.013	0.013
	(0.007)	(0.006)	(0.009)	(0.005)	(0.006)	(0.008)	(0.012)
Grad. Degree	0.028	0.025	-0.003	0.029	0.028	-0.001	0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)
Married	0.340	0.359	0.019	0.339	0.344	0.005	-0.014
	(0.016)	(0.015)	(0.022)	(0.013)	(0.008)	(0.015)	(0.027)
Observations	1,356,264	1,409,132		1,354,112	1,427,855		

Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2011 through 2015.
Note: Means are estimated using people with previous-year income from 1 percent to 10 percent below or above the federal poverty level. Standard errors are clustered at the state level.

Table B2: Difference-in-Discontinuities Estimates, Flexible Functional Form

	Any Health Insurance	Directly Purchased Health Insurance	Employer Sponsored Health Insurance
β_0	0.0140	0.0173*	-0.00440
	(0.0113)	(0.00908)	(0.0117)
Observations	2,248,866	2,248,866	2,248,866
R-squared	0.181	0.016	0.198
Outcome Mean Near Threshold	0.663	0.114	0.391

	Employed	Unemployed	Not in Labor Force	Employed 3+ Quarters	Employed Full Year	Employed Full Time	Employed FTFY	Hours Worked per Week	Weeks Worked per Year	Hours Worked per Year
β_0	-0.0150*	0.0105*	0.00447	-0.0108*	-0.0115	0.00709	0.00670	0.187	-0.0958	12.43
	(0.00755)	(0.00571)	(0.00592)	(0.00542)	(0.00945)	(0.0101)	(0.0101)	(0.193)	(0.349)	(13.43)
Observations	2,248,866	2,248,866	2,248,866	2,248,866	2,248,866	2,248,866	2,248,866	1,760,400	1,866,877	1,760,400
R-squared	0.111	0.020	0.119	0.116	0.109	0.151	0.144	0.111	0.056	0.124
Outcome Mean Near Threshold	0.759	0.0531	0.188	0.689	0.622	0.566	0.487	37.50	47.31	1,794

Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2011 through 2015.
Note: Estimation is based on fitting seventh-order polynomials in previous-year income above and below the premium tax credit eligibility threshold before and after it was introduced. All regressions include controls for age, race, sex, education, and family structure, as well as state and year fixed effects. Standard errors are clustered at the state level. Regressions are estimated using ACS sample weights that have been adjusted for differences in match probabilities across groups. Full time is defined as at least 35 hours per week. Full year is defined as at least 50 weeks per year. Weeks per year is reported in bins in ACS. For this analysis, we use the midpoint of each bin. Analysis of hours and weeks outcomes excludes individuals who report not working. Hours per year is the product of hours per week and weeks per year. Analysis includes all matched ACS-tax observations from 2012-2015 for people between the ages of 26 and 64 who are not self-employed, live in non-expansion states, and have income greater than zero but not exceeding 2,000 percent federal poverty level (FPL). Outcome means near the premium tax credit eligibility threshold are calculated using 2014 and 2015 ACS respondents with previous-year income in the 90-99 percent FPL income bins. Asterisks indicate statistical significance at the 10 percent (*), 5 percent (**), and 1 percent (***) levels.

Table B3: Difference-in-Discontinuities Estimates, Local Linear Regression

	Any Health Insurance	Directly Purchased Health Insurance	Employer Sponsored Health Insurance
β_0	0.0132	0.0148*	-0.00232
	(0.0111)	(0.00729)	(0.0136)
Observations	199,515	199,515	199,515
R-squared	0.082	0.043	0.067
Outcome Mean Near Threshold	0.663	0.114	0.391

	Employed	Unemployed	Not in Labor Force	Employed 3+ Quarters	Employed Full Year	Employed Full Time	Employed FTFY	Hours Worked per Week	Weeks Worked per Year	Hours Worked per Year
β_0	-0.0175*	0.00555	0.0120	-0.0122*	-0.0104	0.00751	0.0101	0.313	-0.0608	15.14
	(0.00980)	(0.00439)	(0.00803)	(0.00639)	(0.0106)	(0.0120)	(0.0122)	(0.217)	(0.337)	(14.31)
Observations	199,515	199,515	199,515	199,515	199,515	199,515	199,515	142,445	157,337	142,445
R-squared	0.088	0.013	0.106	0.078	0.069	0.099	0.090	0.071	0.026	0.063
Outcome Mean Near Threshold	0.759	0.0531	0.188	0.689	0.622	0.566	0.487	37.50	47.31	1,794

Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2011 through 2015.
Note: Estimation is based on fitting linear regressions above and below the premium tax credit eligibility threshold before and after it was introduced. All regressions include controls for age, race, sex, education, and family structure, as well as state and year fixed effects. Standard errors are clustered at the state level. Regressions are estimated using ACS sample weights that have been adjusted for differences in match probabilities across groups. Full time is defined as at least 35 hours per week. Full year is defined as at least 50 weeks per year. Weeks per year is reported in bins in ACS. For this analysis, we use the midpoint of each bin. Analysis of hours and weeks outcomes excludes individuals who report not working. Hours per year is the product of hours per week and weeks per year. Analysis includes all matched ACS-tax observations from 2012-2015 for people between the ages of 26 and 64 who are not self-employed, live in non-expansion states, and have previous-year income within an amount equivalent to 30 percent of the federal poverty level (FPL) of the premium tax credit eligibility threshold. Outcome means near the premium tax credit eligibility threshold are calculated using 2014 and 2015 ACS respondents with previous-year income in the 90-99 percent FPL income bins. Asterisks indicate statistical significance at the 10 percent (*), 5 percent (**), and 1 percent (***) levels.

Table B4: Difference-in-Discontinuities Estimates, Annual Labor Market Outcomes, No 2014 Data

	Employed 3+ Quarters	Employed Full Year	Employed Full Time	Employed FTFY	Hours Worked per Week	Weeks Worked per Year	Hours Worked per Year
	Flexible Functional Form, p=7						
β_0	-0.0110	-0.0163	0.00621	-0.00191	0.147	-0.207	5.235
	(0.00904)	(0.0126)	(0.0173)	(0.0168)	(0.277)	(0.379)	(16.64)
Observations	1,686,494	1,686,494	1,686,494	1,686,494	1,317,977	1,398,652	1,317,977
R-squared	0.117	0.110	0.152	0.145	0.110	0.057	0.124
	Local Linear Regression, Bandwidth = 30% FPL						
β_0	-0.0118	-0.0120	0.000374	0.00123	0.297	-0.0594	19.01
	(0.00969)	(0.0130)	(0.0165)	(0.0158)	(0.222)	(0.416)	(18.12)
Observations	149,568	149,568	149,568	149,568	106,257	117,555	106,257
R-squared	0.079	0.069	0.100	0.091	0.070	0.027	0.063

Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2011 through 2015.
Note: In the top panel, estimation is based on fitting seventh-order polynomials in previous-year income above and below the premium tax credit eligibility threshold before and after it was introduced. In the bottom panel, estimation is based on fitting linear regressions within an amount equivalent to 30 percent of the federal poverty level (FPL) of above and below the premium tax credit eligibility threshold before and after it was introduced. All regressions include controls for age, race, sex, education, and family structure, as well as state and year fixed effects. They also exclude data from the 2014 ACS, since the the outcomes in question incorporate at least some pre-ACA period for the vast majority of respondents. Standard errors are clustered at the state level. Regressions are estimated using ACS sample weights that have been adjusted for differences in match probabilities across groups. Full time is defined as at least 35 hours per week. Full year is defined as at least 50 weeks per year. Weeks per year is reported in bins in ACS. For this analysis, we use the midpoint of each bin. Analysis of hours and weeks outcomes excludes individuals who report not working. Hours per year is the product of hours per week and weeks per year. Analysis includes matched ACS-tax observations from 2012-2015 for people between the ages of 26 and 64 who are not self-employed and live in non-expansion states. Asterisks indicate statistical significance at the 10 percent (*), 5 percent (**), and 1 percent (***) levels.

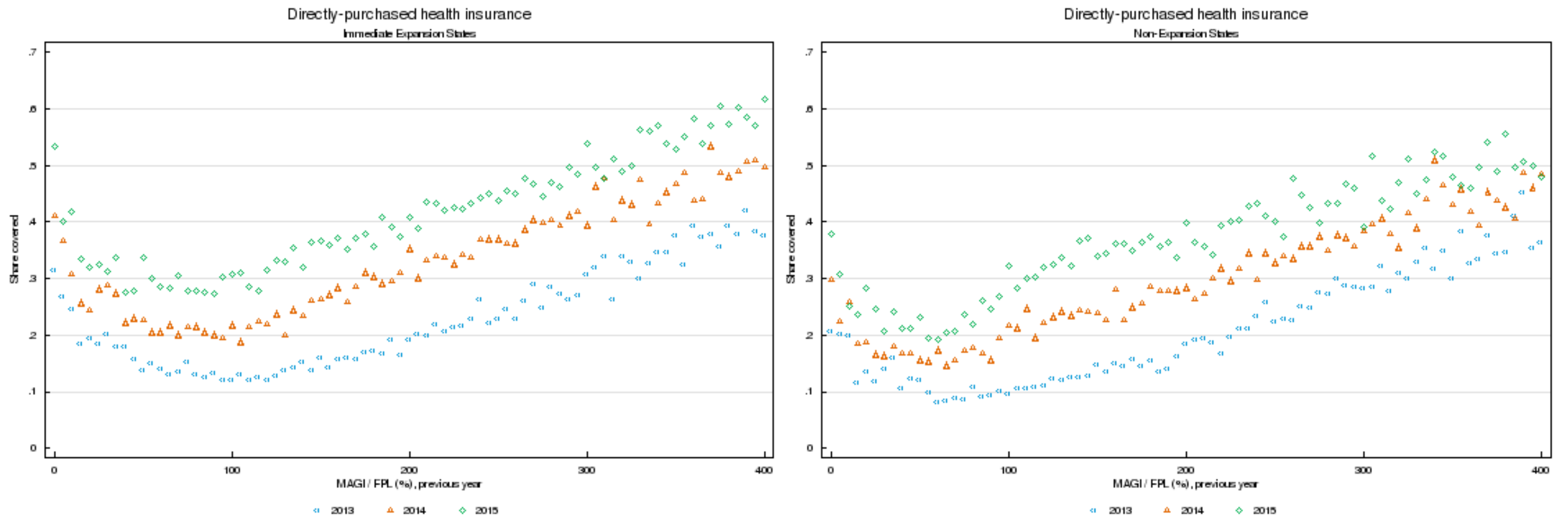
Table B5: Difference-in-Discontinuities Estimates of Effects on Weeks and Hours Measures, with Imputed Zeroes

	(1)	(2)	(3)	(4)	(5)	(6)
	Flexible Functional Form			Local Linear Regression		
	Hours Worked per Week	Weeks Worked per Year	Hours Worked per Year	Hours Worked per Week	Weeks Worked per Year	Hours Worked per Year
β_0	-0.244	-0.362	-13.30	-0.466	-0.534	-18.07
	(0.523)	(0.395)	(23.51)	(0.594)	(0.499)	(25.87)
Observations	1,686,494	1,686,494	1,686,494	149,568	149,568	149,568
R-squared	0.159	0.127	0.166	0.102	0.099	0.101
Outcome Mean Near Threshold	28.47	37.61	1,362	28.47	37.61	1,362

Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2011 through 2015.
Note: In columns 1-3, estimation is based on fitting seventh-order polynomials in previous-year income above and below the premium tax credit eligibility threshold before and after it was introduced. In columns 4-6, estimation is based on fitting linear regressions within an amount equivalent to 30 percent of the federal poverty level (FPL) of above and below the premium tax credit eligibility threshold before and after it was introduced. All regressions include controls for age, race, sex, education, and family structure, as well as state and year fixed effects. Standard errors are clustered at the state level. Regressions are estimated using ACS sample weights. Full time is defined as at least 35 hours per week. Full year is defined as at least 50 weeks per year. Weeks per year is reported in bins in ACS. For this analysis, we use the midpoint of each bin. Zeroes are assigned to missing values of hours per week and weeks per year. Hours per year is the product of hours per week and weeks per year. Analysis includes matched ACS-tax observations from 2012-2015 for people between the ages of 26 and 64 who are not self-employed and live in non-expansion states. Outcome means near the premium tax credit eligibility threshold are calculated using 2014 and 2015 ACS respondents with previous-year income in the 90-99 percent FPL income bins. Asterisks indicate statistical significance at the 10 percent (*), 5 percent (**), and 1 percent (***) levels.

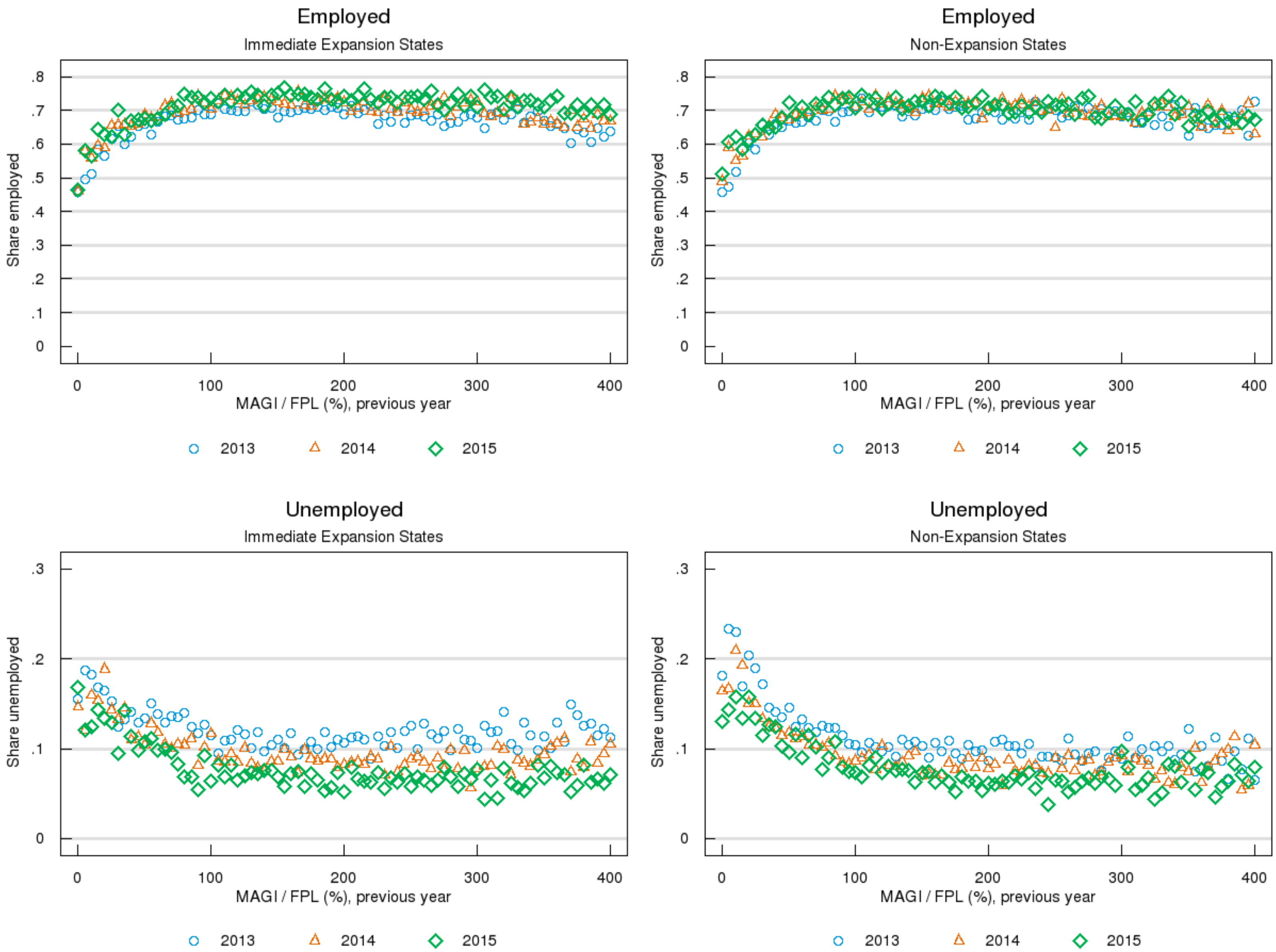
Appendix C Workers without Other Insurance, Including Ages 19-25

Figure C1: Health Insurance Coverage among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status

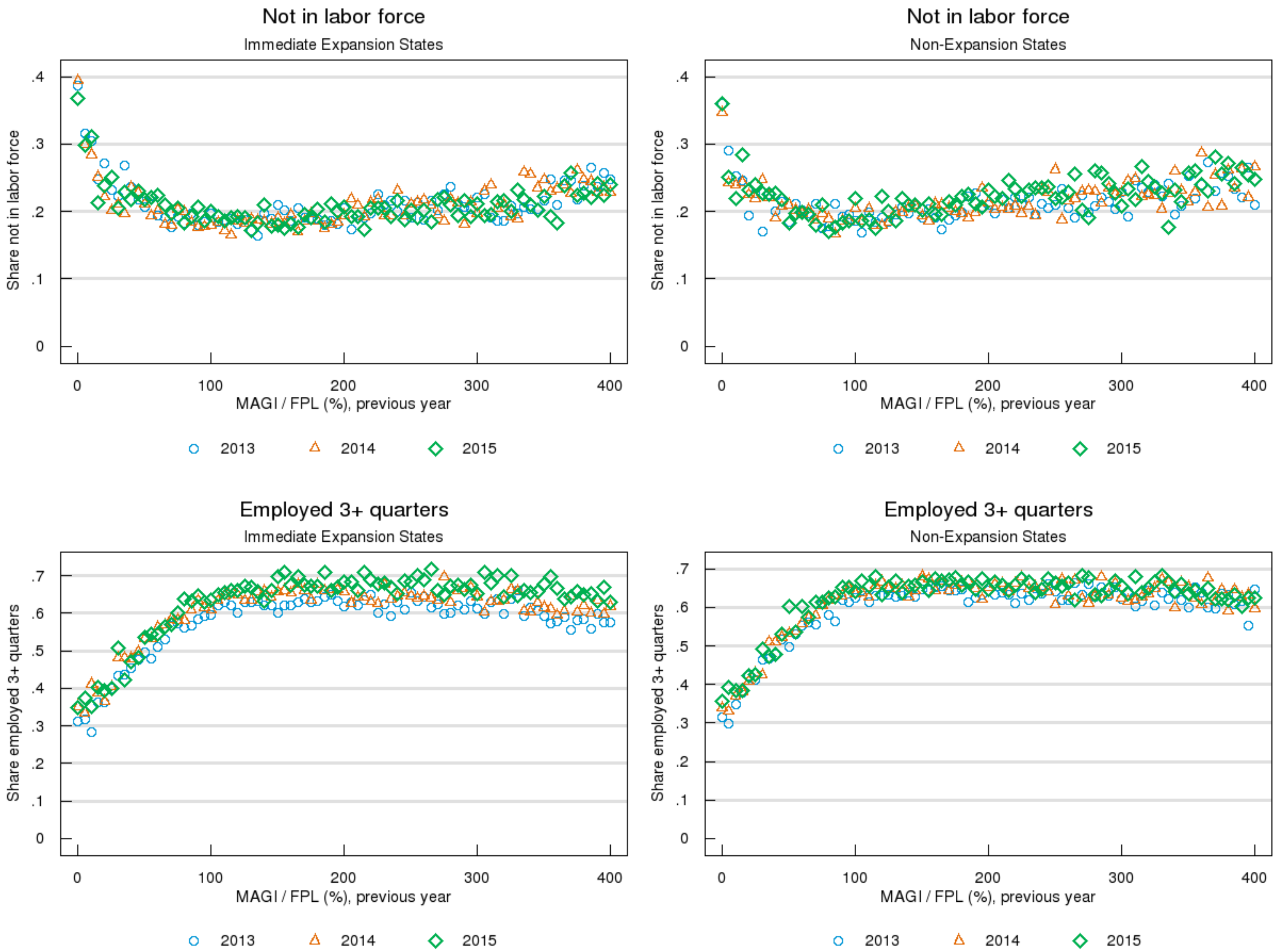


Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2011 through 2015. Estimates in the range shown are based on 480,439 observations.
Note: Sample includes individuals aged 19 to 64 without self-employment income who do not report being covered by a type of insurance that would disqualify them from receiving the premium tax credit. Income is obtained from the tax return on which each individual appeared for the year prior to responding to the ACS. Each point represent the mean outcome within an income bin that is 5 percent of the federal poverty level in width, plotted at the minimum value contained in the bin.

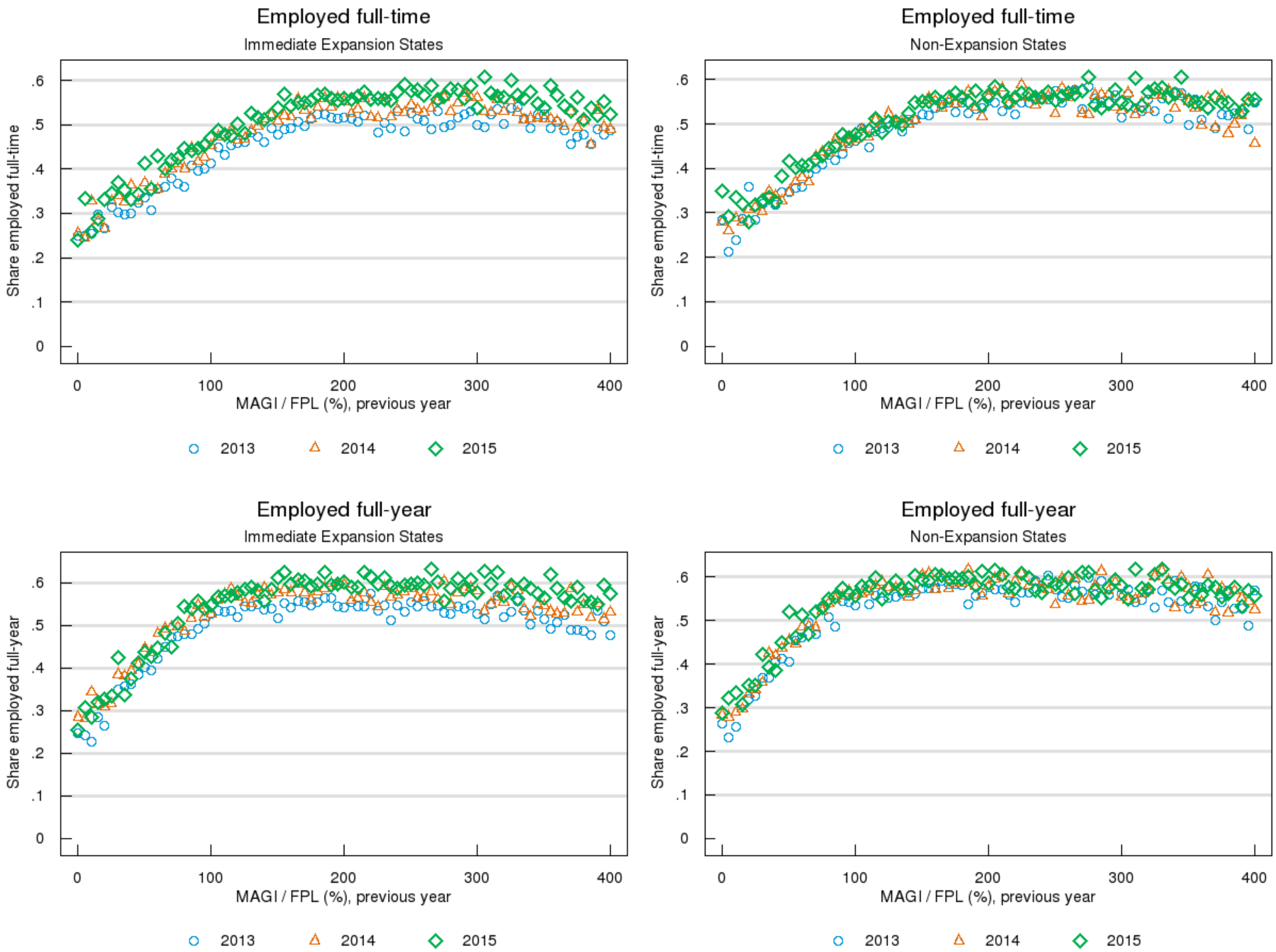
Figure C2: Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



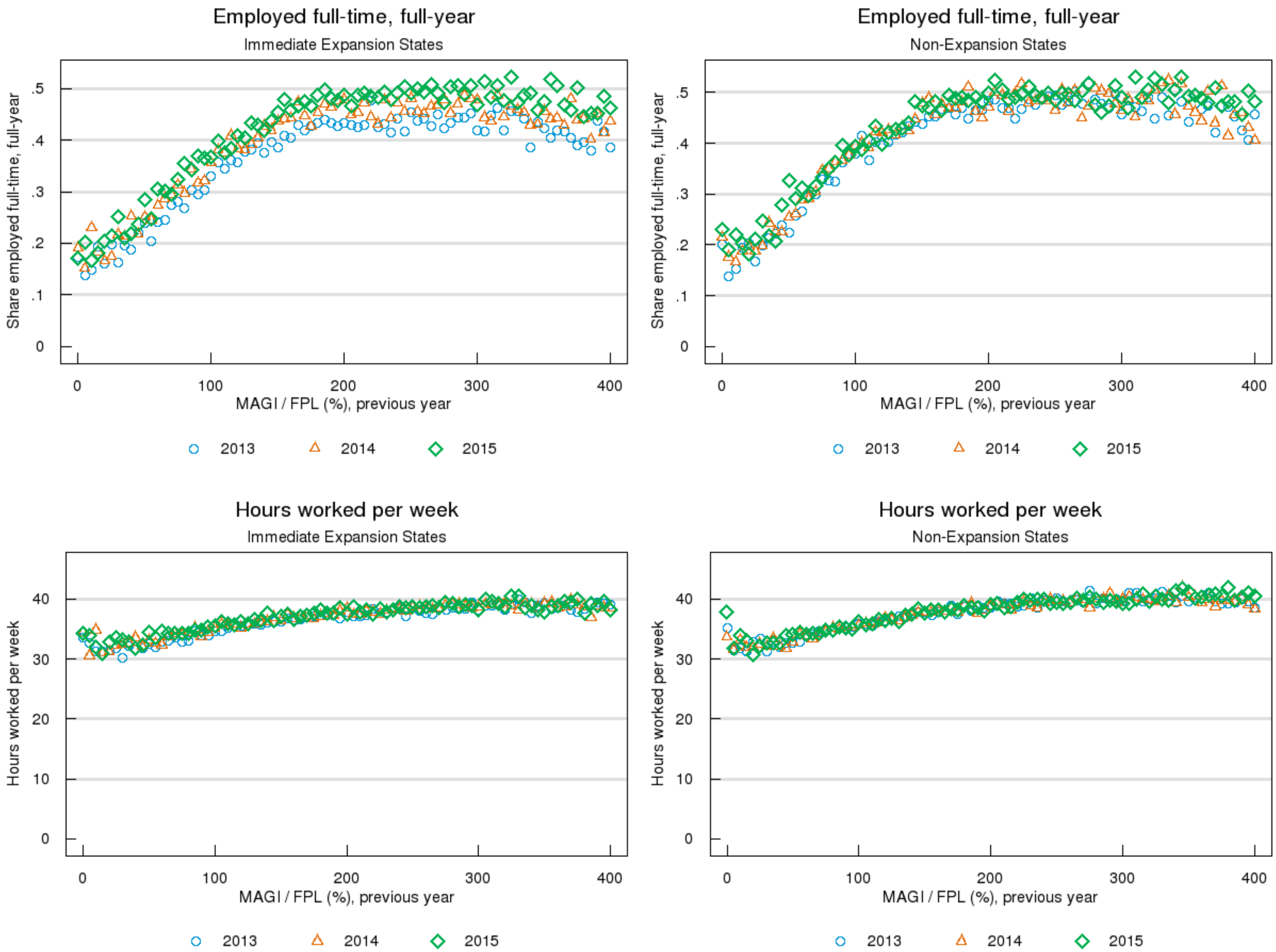
Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



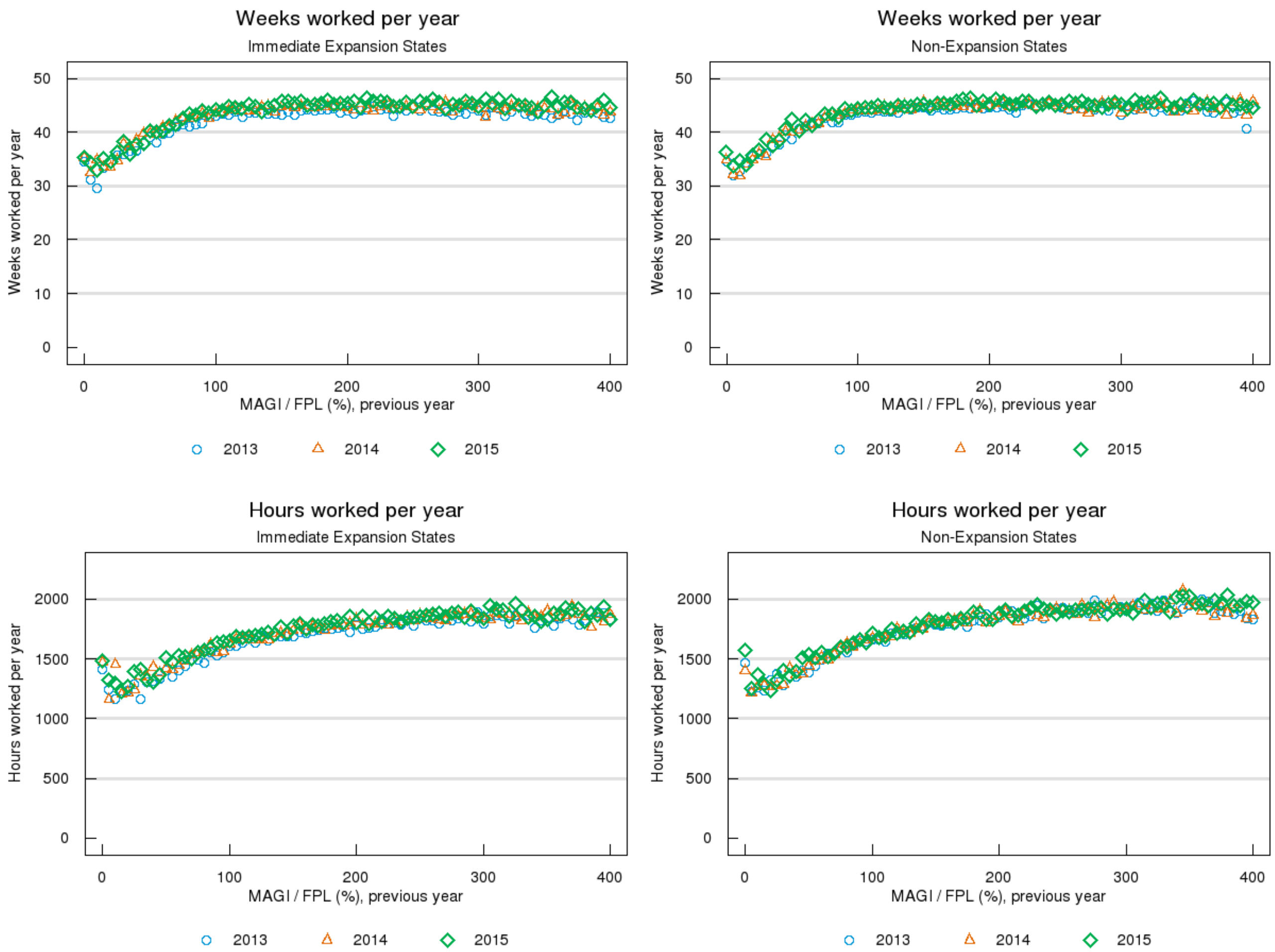
Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



Continued – Labor Market Outcomes among Wage and Salary Workers, by Income, Year, and Medicaid Expansion Status



Source: Authors' calculations, linked ACS-IRS Form 1040 data, 2011 through 2015. Estimates in the range shown are based on 480,439 observations.
Note: Sample includes individuals aged 19 to 64 without self-employment income who do not report being covered by a type of insurance that would disqualify them from receiving the premium tax credit. Income is obtained from the tax return on which each individual appeared for the year prior to responding to the ACS. Full time is defined as at least 35 hours per week. Full year is defined as at least 50 weeks per year. Weeks per year is reported in bins in ACS. For this analysis, we use the midpoint of each bin. Analysis of hours and weeks outcomes excludes individuals who report not working. Hours per year is the product of hours per week and weeks per year. Each point represent the mean outcome within an income bin that is 5 percent of the federal poverty level in width, plotted at the minimum value contained in the bin.