New Work or Changes in Reporting? A Framework for Measuring Self-Employment Trends in Administrative Data*

Andrew Garin†, Emilie Jackson‡ and Dmitri Koustas§

July 14, 2021

Abstract

Increasingly, researchers are looking to administrative tax measures of self-employment, which show larger growth in self-employment than in mainstream surveys of the labor market. This divergence may reflect changes in real activity not well-captured by traditional surveys, but may also reflect changes in workers' propensity to report their self-employment earnings to tax authorities, potentially in response to incentives created by the tax code. Consistent with the latter explanation, we find the extensive-margin growth in reported self-employment in administrative tax records is concentrated among low-income households with children with incentives to report self-employment earnings due to tax credit phase-ins. We employ a regression discontinuity design that compares households whose children are born at the end of the tax year—and are eligible for these incentives—with households whose children are born just a few days later and face no change in reporting incentives. We find a sizable increase in reported self-employment among those who are incentivized for which the only plausible interpretation is strategic reporting. Moreover, we find these effects have grown over time, with increases being concentrated in regions where other proxies for knowledge of tax credits have grown as well. Incorporating these insights, we present a framework to adjust trends in U.S. self-employment tax filings for changes in reporting behaviors to extract true underlying trends in the composition of the workforce. By considering counterfactual scenarios, we find that between 30 and 55 percent of the growth in self-employment rates can be attributed to pure reporting changes. Once we additionally control for the aging of the workforce, we can account for most of the observed rise in self-employment.

*For discussion as part of the 2021 NBER Summer Institute – Not for broader distribution. We are grateful to Katharine Abraham, John Guyton, Lawrence Katz, Kara Leibel, Alicia Miller, Bruce Meyer, Jacob Mortenson, Michael Weber, Andy Whitten, and seminar participants at the Office of Tax Analysis at the US Treasury and National Tax Association meetings. We thank Emmanuel Saez for providing us with code used to implement the calculations in Chetty, Friedman, Saez (2013) in the restricted-access administrative data. We are grateful to the Russell Sage Foundation and the Sloan Foundation for generous support, and to Reigne Dadey for excellent research assistance. Jackson gratefully acknowledges support from the Peter G. Peterson Foundation for a post-doctoral fellowship at NBER during which part of this research was conducted. This research was authorized through the IRS SOI Joint Statistical Research Program. The researchers were granted access to tax administrative data as IRS employees through agreements under the Intergovernmental Personnel Act.

†University of Illinois Urbana-Champaign

‡NBER and Michigan State University

§University of Chicago

1 Introduction

Changes in self-employment rates are often looked to as a means to assess changes in both the structure of the labor market and the extent of entrepreneurial activity in the broader economy.[1] However, a key challenge is that self-employment is hard to measure. In fact, the share of workers reporting self-employment earnings in administrative tax records in the United States has been growing, in puzzling contrast with self-employment activity reported in workforce surveys.[2] This divergence in findings could be consistent with either of two main hypotheses: 1) government surveys are incompletely capturing changes in the nature of work, and tax records are better at picking up new forms of self-employment, and/or 2) trends documented in administrative tax data may capture other factors that do not reflect true changes in underlying work activity, such as changes in reporting/tax filing behavior.

The goal of this paper is to provide a framework for measuring self-employment using administrative data, controlling for changes in reporting/tax filing behavior that may be occurring over time. Unlike confidential surveys, individuals have incentives to be strategic about what they report on tax filings—and both those incentives and taxpayer responses may be prone to change over time. This is particularly true in the case of self-employment earnings which, unlike employment income, can be purely self-reported without any third-party verification (Slemrod, 2016). We build on key insights from the "bunching" literature (LaLumia, 2009; Saez, 2010; Chetty, Friedman, and Saez, 2013; Kuka, 2014; Mortenson and Whitten, 2020) to estimate the magnitude of the self-employment trend that is explained by changes in strategic reporting.[3]

[1] For instance, recent literature on business dynamism focuses on non-employer business statistics which include sole-proprietors (Davis, Haltiwanger, Jarmin, Krizan, Miranda, Nucci, and Sandusky, 2009; Fairlie, Miranda, and Zolas, 2019).

[2] It is now well documented that over the past two decades, self-employment rates reported in the Current Population Survey have either remained flat or have declined since 2000 (Bureau of Labor Statistics, 2018a; Katz and Krueger, 2019), while the share of tax filers reporting self-employment income on their tax returns rose dramatically between 2000 and 2014 (Abraham, Haltiwanger, Sandusky, and Spletzer, 2013; Jackson, Looney, and Ramnath, 2017; Abraham, Haltiwanger, Hou, Sandusky, and Spletzer, Forthcoming).

[3] Prior research finds that self-employment earnings are often reported at precise levels the maximize refundable tax credits like the Earned Income Tax Credit, whereas this type of "bunching" behavior is rarely observed in employer-reported wage earnings or in survey measures of self-employment earnings (LaLumia, 2009; Chetty, Friedman, and Saez, 2013; Kuka, 2014; Mortenson and Whitten, 2020). Moreover, this type

Ultimately, we provide bounded estimates on an adjusted measure of self-employment that we believe identify true changes in real work or entrepreneurship activity over the past two decades, and can be used to examine changes in economic activity and labor market composition.

We first examine the extent to which administrative trends in self-employment are plausibly driven by new kinds of app-based contract work that have grown exponentially in recent years (Harris and Krueger, 2015; Abraham, Haltiwanger, Sandusky, and Spletzer, 2018; Farrell and Greig, 2018; Katz and Krueger, 2019) or other type of "gig" work done by independent contractors. Standard survey questions about self-employment fail to capture new types of "gig" work that are nonetheless being reported on tax returns.[4] A unique feature of U.S. tax data is that this gig work is separately observable—contract work performed for firms or through online platforms is required to be separately reported by firms to the Internal Revenue Service (IRS) (Jackson, Looney, and Ramnath, 2017; Collins, Garin, Jackson, Koustas, and Payne, 2019; Lim, Miller, Risch, and Wilking, 2019). We find that while the prevalence of Online Platform Economy (OPE) earnings has grown dramatically since 2014, this cannot explain the rise in reported self-employment rates which peaked before 2014. While increases in independent contract work more generally can account for some of the rise in self-reported self-employment rates between 2000 and 2005, trends in contract work remained flat since 2005 as self-employment rates continued to climb, and year-to-year changes have essentially zero correlation with overall self-employment trends.

Second, we evaluate the extent that changes in reporting responses account for the rise in reported self-employment rates. Refundable credits like the Earned Income Tax Credit (EITC) not only create an incentive to report income close to certain refund-maximizing amounts—they also create incentives to report self-employment for some

of strategic reporting has become increasingly prevalent since 2000 (Chetty, Friedman, and Saez, 2013; Mortenson and Whitten, 2020).

[4]Although technically self-employed, contract workers may nonetheless perceive themselves to be traditional employees and report as such on surveys (Abraham, Hershbein, and Houseman, 2020; Abraham, Haltiwanger, Sandusky, and Spletzer, 2020). Further, surveys like the monthly CPS only ask about *primary* work activities and therefore may fail to capture trends in secondary or temporary self-employment captured by tax returns.

individuals who might not have otherwise. In particular tax filers with children with non-self-employment wage earnings in the EITC phase in range (below the first "kink" in the EITC schedule) face *negative* tax rates on the marginal dollar. Individuals in this situation would strictly benefit from reporting some self-employment earnings rather than none—regardless of their underlying earnings. These individuals might therefore increase their tax refunds by choosing to report informal earnings they would have otherwise not reported to the IRS (roughly a third of Americans report having some type of informal income (Bracha and Burke, 2018)), or, in some cases, by fabricating the income entirely. Just as the propensity for individuals to "bunch" at refund-maximizing kinks has grown over time, it is possible that changes in tax-code-incentivized reporting behaviors have contributed to the observed rise in the number of individuals with self-employment income.

Consistent with this story, we find that the entire increase in the propensity to report secondary self-employment income among wage employees is driven by filers in this incentivized group. While self-employment reporting rates have risen substantially among tax filers with children and wages in the phase in range, those increases are absent for those with firm-reported wages even slightly above this "kink" point where the marginal tax rate becomes positive. In contrast, we find no differential growth above or below the first EITC kink point for individuals with no kids; crucially, individuals without children never face negative marginal tax rates as the subsidy rate in the EITC phase-in region for childless adults is less than self-employment (SECA) tax rate. Notably, we find no such rise in *firm-reported* contract income receipt among those incentivized individuals with low wages and children, only a rise in self-reported self-employment. We further document *within-individual* increases in self-employment reporting in the year their first child is born—at which point they begin to face marginal tax rates. Moreover, the propensity to begin reporting self-employment after childbirth has grown secularly since 2000. Meanwhile, there is no trend in the propensity to begin receiving firm-reported independent contractor income around childbirth in all years since 2000.

A key question is whether the rise in self-reported self-employment rates—concentrated among individuals with a strong EITC incentive to report additional earned income—reflects a true increase in non-firm-facing self-employment activity or rather a greater proclivity to self-report to tax authorities that is not indicative of any change in underlying labor supply. To shed light on this question, we examine a sharp discontinuity in EITC eligibility based on the date of birth of an individual's first child motivated by earlier work by Lalumia, Sallee, and Turner (2015). An individual's EITC benefit for tax year t is calculated based on the number of children in their household during year t—a first-born child born on or before December 31 of year t would count towards the EITC calculation for that year t. By contrast, a child born only a few days later at the start of year $t+1$ would not count towards the EITC calculation in year t. Accordingly, the EITC creates sharp differences in reporting incentives for tax year t for individuals having their first child days apart at the very end of year t and those having theirs at the very beginning of year $t+1$. This motivates a regression discontinuity design comparing year-t self-employment earnings for parents with first birth right before and right after the end of tax year t. If, in the limit examining births right before and after midnight on December 31, parents face complete *ex-ante* uncertainty about the year of birth, then parents should make identical labor supply decisions in year t.

We find clear evidence of a discontinuity in the probability of reporting self-employment income around the December 31 cutoff. Our estimates indicate that having a child unexpectedly qualify towards a year's EITC calculation increases the probability of reporting self-employment income by 1.5 percentage points in years since 2010, but does not impact the probability of having freelance income reported by firms. This result does not appear to be driven by strategic timing of births to maximize tax refunds. If individuals were re-timing births to maximize EITC credits, then we should see individuals with births in December have firm-reported earnings closer to the one-child refund-maximizing level; this, however, is not the case, as firm-reported wage earnings and non-employee compensation do not vary around the threshold date. Moreover, we find that the entire self-employment effect is driven by individuals with wages in

or below the phase in range. These reporting effects have grown steadily over time. Between 2000 and 2014, the baseline effect size has grown nearly threefold from 0.5 p.p. to 1.5 p.p., and the effect conditional on having wages in the phase in range has grown over fivefold from 1 p.p. to over 5 p.p.

What explains this rise in reporting motivated by tax credits? Notably, there has not been any major expansion in the generosity of the federal EITC over the period we study.[5] However, prior research suggests that even during periods when policy remains fixed, *knowledge* of the incentives created by the EITC and related tax provisions differs across regions and spreads gradually over time (Chetty, Friedman, and Saez, 2013). This process of gradual learning could lead to a steady rise in reporting of previously unreported self-employment earnings or false reporting of self-employment income. Following Chetty, Friedman, and Saez (2013), we use sharp-bunching rates among individuals in a hold-out sample for each three-digit ZIP code in each year as an proxy for local knowledge of EITC incentives and estimate an interacted version of our RDD specification. We find that high-knowledge regions have dramatically larger effects than low-knowledge regions, and that cross-sectional differences in knowledge and within-region changes in knowledge are associated with similar increases in the reporting effect. The spread of knowledge associated with the geographic diffusion of hyper-strategic bunching behavior between 2000 and 2014 can explain a majority of the increase in reporting behavior found in our RDD design.

Finally, we use our empirical estimates to adjust the raw self-employment data for strategic reporting incentives. Our key finding is that adjusting for changes in strategic reporting behavior dramatically lowers the raw growth in tax-based measures of self-employment. We still find evidence of modest growth in self-employment, particularly in the early 2000s and in recent years with the rise of gig work mediated by online platforms, that is missing from survey data.

Our work adds to a large public economics literature that highlights that trends

[5]There are two exceptions that do affect incentives. In 2009, the EITC was increased somewhat for tax units with three or more children. In addition, the CTC was further expanded in 2008.

observed in administrative data sets may be impacted by reporting incentives, and should be interpreted with caution (Blank, Charles, and Sallee, 2009; Slemrod, 2016). Prior work has documented that this is particularly the case with self-employment earnings reported in tax data (Chetty, Friedman, Ganong, Leibel, Plumley, and Saez, 2012; Chetty, Friedman, and Saez, 2013; LaLumia, Sallee, and Turner, 2014). Our contribution is to document that these incentives appear to be changing over time in a way that spuriously drives the measured trends in self-employment. This is especially important as these raw trends are receiving increased attention amidst policy concern about the rise of gig work and other changes in the nature of work. Our new self-employment series adjusted for reporting trends, as well as our new series on gig work, should be valuable to other researchers in this area.

This paper is organized as follows: In the next section, we discuss the measurement of self-employment in administrative tax data. We also discuss important institutional details of the U.S. tax code that relate to incentives to report self-employment. Section 3 presents important motivating facts on self-employment trends. Section 4 introduces our main research design to estimate the the response of self-employment to reporting incentives. Section 5 discusses knowledge of the tax code as a driver of the change. Section 6 provides new estimates of self-employment controlling for reporting trends. We also present an application of our new data contributions, using our insights to examine the response of self-employment to the business cycle. Section 7 concludes.

2 Data and Institutional Background

In this section we describe important institutional details and our key data sources. We begin by defining self-employment, contrasting what is commonly captured in survey data and its correspondence to administrative tax data. We also discuss important institutional details of the U.S. tax schedule relevant for interpreting our results.

2.1 Self-Employment in Tax Records

Who is a self-employed worker? Self-employed workers can interact directly with consumers, such as a shopkeeper, farmer, artist, or plumber ("consumer-facing" self-employment). Self-employed workers can also engage in work for directly firms, such as a freelance writer working for a newspaper ("firm-facing" self-employment). Motivations for self-employment vary widely (Hurst and Pugsley, 2011). Someone does not need to "own" a business storefront to be self-employed. Individuals can have self-employment income from work they do as a side-business, out of their home, or even out of their car.

Self-employment is measured in different ways depending on the source. For instance, in the Census, workers are asked to self-identify, based on the job they spent the most time at, as either an employee or owner of a business. Census surveys also ask self-employed workers to distinguish between incorporated and unincorporated self-employment. The Annual Social and Economic Supplement (ASEC) to the Current Population Survey (CPS) asks individuals about earnings after expenses over the course of the previous year from any businesses they own. Reporting a positive value of earnings indicates some self-employment income over the course of the year.

The typical survey definition of self-employment undercounts secondary, informal self-employment and firm-facing self-employment (National Academies of Sciences and Medicine, 2020; Abraham, Hershbein, and Houseman, 2020; Abraham, Haltiwanger, Sandusky, and Spletzer, 2020). A key concern is misclassification of independent contracting/freelance work. Moreover, surveys are subject to limitations of how people perceive the work they do and how they would report that work in surveys. Because an independent contractor receives income from a firm, sometimes working onsite, this worker could perceive themselves as being employed by that firm, rather than being self-employed. Specialized surveys have attempted to measure the size of independent contracting. Most notably the Contingent Worker Supplement (CWS) to the CPS seeks to measure workers who are identified as independent contractors, independent consultants, or freelance workers (See Bureau of Labor Statistics, 2018b). However,

these surveys tend to be run on an ad-hoc basis. The CWS, in particular, only asks about primary work, thereby missing secondary self-employment.

A second approach to measuring self-employment comes from examining administrative tax records. However, measuring self-employment in tax records correctly depends on high degrees of tax compliance. It is well established in the public economics literature that self-employment in tax data is affected by reporting incentives and should be interpreted with caution (Chetty, Friedman, Ganong, Leibel, Plumley, and Saez, 2012; Chetty, Friedman, and Saez, 2013). Likewise, measurement of trends requires tax compliance to be constant over time. If compliance increases (decreases) measured self-employment will also increase (decrease), even if there is no change in actual work.

One key difference between wage-and-salary employees and self-employed workers comes down to who pays work-related expenses and withholds taxes. Wage and salary employees has these expenses paid by an employer (which in turn deducts them from its net proceeds) who also withholds taxes, whereas self-employed workers are responsible for both. In addition, wage workers automatically have their wage earnings transmitted to tax authorities by employers via an information return "Form W-2," thereby providing less scope for manipulation of income. A key advantage of information returns is that tax compliance is very high because of the way incentives are aligned: firms wish to deduct labor costs from their expenses, so have every incentive to report labor costs to tax authorities. Workers who receive these information returns then choose whether to file taxes, a decision which is affected by tax filing requirements and other incentives facing the taxpayer. Importantly, trends based only on W-2 information returns will not be affected by these incentives.

In contrast, self-employed workers are responsible for maintaining accurate counts of their expenses, which they can deduct from gross revenues to calculate self-employment profits. Self-employed workers are then responsible for paying the payroll taxes on their profits. These payroll taxes are 15.3 percent for much of our period of study. In the U.S., self-employed income and expenses of unincorporated businesses should be

reported on "Schedule C" of the standard "Form 1040" tax return. If a self-employed worker has net profits in excess of $433, they are required to file "Schedule SE" and pay payroll taxes on their profits.[6] In previous work, Abraham, Haltiwanger, Sandusky, and Spletzer (2020) focus on Schedule SE filers, while Jackson, Looney, and Ramnath (2017) focus on Schedule SE and Schedule C filers. In this paper, we will focus on Schedule SE because it is consistently recorded in our administrative database on an individual (as opposed to tax unit) basis.

One key advantage of U.S. tax records is that independent contractor relationships are directly observable. Payments by firms to self-employed individuals are reported on a form prepared by employers and sent to individuals and tax authorities in a similar way as wages reported on Form W-2. This allows us to separately observe trends among independent contractors. Second, this provides a measure of self-employment income that does not depend on taxpayer compliance.

Two types of information returns allow us to focus on independent contractors at these firms. The first is "Form 1099-MISC"; firms are required to report all compensation of $600 or more to self-employed independent contractors in Box 7 ("nonemployee compensation") of Form 1099-MISC. Until 2011, all "freelance" or "gig" work done for firms or for clients through intermediaries would be reported on this form.

We also pay special attention to a new and growing class of independent contract work mediated by online platforms. We refer to these arrangements—which are a subset of the broader "gig" economy"—as the "online platform economy" (OPE). We focus specifically on labor platforms on which workers directly provide services to others (for example, ridesharing or delivery) as opposed to platforms on which individuals sell goods or rent capital (for example, craft merchandise sites or homesharing).[7] One caveat is that not all OPE work is truly new work. Some forms of work in the labor

[6] Half of payroll taxes, the "employer portion," are deductible

[7] In the OPE, consumers directly interface with a digital platform technology, which matches them with contractors supplying labor and determines key parameters of the transaction. If a customer is not satisfied with the service, customer service is often handled by the corporate platform, not the worker supplying the service. Thus, although platform workers typically provide services directly to consumers, labor OPE transactions are nonetheless crucially firm-mediated—and therefore are considered independent contractors in our analysis.

OPE are clearly new economic activity, the most notable being paid ridesharing, which was largely non-existent before 2011. In other contexts, new forms of firm-mediated activity in the OPE may be supplanting informal work previously done in an informal setting, and that was unlikely to be reported to tax authorities via an information return. This is more likely the case for professional freelancers who now supply labor on the OPE. We separately identify contract income on 1099 forms reported by OPE companies in order to determine how the rise of OPE work affects the overall prevalence of contract work. We describe out method for identifying OPE-related 1099 forms in the Data Appendix.

We classify all individuals with either nonemployee compensation reported in Box 7 of a 1099-MISC or payments from one of these specific labor platforms on a 1099-K in a particular year as independent contractors in that year.[8] From the perspective of the tax code, 1099 independent contractors—those with either 1099-MISC non-employee compensation or an OPE 1099-K—are self-employed. Formally, this 1099 income, like all self-employment income, is considered active business income by the IRS. Accordingly, unless individuals become incorporated, this income should be reported to tax authorities as proceeds from a wholly-owned business on Schedule C. However, there are a number of reasons why individuals receiving a 1099 may not file their taxes as if they are self-employed. One innocent reason (albeit still running afoul of tax filing obligations) is that these individuals do not perceive themselves to be self-employed, and instead file this income as "other income" or add it to their main earnings. Other reasons include not understanding that receiving receipts over $433 mandates filing and paying self-employment taxes, even if total income falls below the standard deduction. In our subsequent analysis, we will separately examine trends for receipt of a 1099—regardless of tax filing—and tax filing of Schedule SE.

[8]We restrict to 1099-Ks with gross amounts of at least $600 in order to maintain a more consistent measure over time. See Data Appendix for more details.

2.2 Incentives in The Tax Code to Report Self-Employment Income

The tax code has incentives that could encourage certain individuals to report self-employment profits. In particular, two key credits—the Child Tax Credit (CTC) and Earned Income Tax Credit (EITC)—phase-in with higher earnings and hence subsidize earnings up to a threshold level (the site of the first "kink" where the maximum credit level is obtained). Both this threshold level and the effective subsidy rate depend on the number of children claimed on one's tax return.[9] The phase-in rate of the EITC is 7.65 percent for childless households, 34 percent for families with 1 child, 40 percent for families with two children, and 45 percent for families with three children. The CTC further increases the effective phase-in rate by 15 percentage points. The total effective tax rate combines all tax credits and tax rates, including payroll (FICA/SECA) taxes.[10]

As a result, individuals with children and earnings in the phase-in range face negative marginal tax rates, as the credit phase in rate exceeds the payroll tax rate. Families with children and with baseline earned income below the first kink point, earn a net subsidy from the federal tax system for every additional dollar of earnings reported—with additional credits provided by some states. By contrast, once earnings are above the applicable phase-in range, these incentives to report additional earnings disappear, since additional earnings are subject to the payroll tax but no marginal subsidy.[11] Similarly, households with no children always face a positive marginal tax rate after accounting for payroll taxes, hence they would always owe more in taxes from reporting any additional self-employment income.[12]

[9]The CTC applies to children under 18, and the EITC applies to children under 19.

[10]Self-employed individuals are responsible for the employer and employee portion of payroll taxes, which together are 15.3 percent. Half of payroll taxes paid (the "employer"-share) is deductible. Thus, the effective marginal tax rate on self-employment earnings is 0.153*(1-0.0765)=14.1%. At the same time, EITC and CTC credits for every dollar of self-employment income are also scaled back proportionately by the deduction.

[11]Of course, an important incentive to report self-employment income to tax authorities and pay payroll taxes is to contribute to future Social Security benefits. However, these incentives are faced by all taxpayers, not just those facing negative marginal tax rates. Moreover, the increase in Social Security benefits may be less salient to myopic consumers, and since benefits are based on the highest 35 years of earnings, earnings for young workers are likely to have little impact on future benefits.

[12]Even with the most generous state-EITC for childless individuals in the District of Columbia, they

To illustrate these incentives, Appendix Figure A.1 shows the effective tax rate on reporting self-employment income at different levels of wage earnings, taking into account the entire tax schedule. Households without children face a positive marginal tax rate on self-employment income across the wage distribution. In contrast, for families with children, additional income at amounts below the credit-maximizing kink point is taxed at a negative tax rate. Households with three children with wage income below the threshold for the minimum income to maximize the EITC ($13,870 in $2015), and above the minimum income threshold for the CTC, face a negative tax rate that implies a credit of up to 41 cents on the dollar.[13] Similarly, tax units with two children face a marginal tax rate of up to -0.37 and tax units with one child face a negative marginal tax rate of up to -0.31.

Crucially, these negative marginal tax rates create an incentive for individuals with children under 19 and wages in the credit phase-in range to report self-employment income that may not have been reported otherwise. Individuals in this "incentivized" group would be strictly better off reporting some self-employment income rather than none. Moreover, while wage earnings on W2s are directly reported to the IRS by employers, self-employment income need not be reported by any third-party, and in practice taxpayers have discretion in what they report.[14] In theory, there is a refund-maximizing earnings amount such individuals could report, which could give rise to hyper-strategic "bunching" at those income levels. Yet, in practice, there are a number of reasons why individuals may be incentivized to report additional self-employment income without reporting this "optimal" amount: Individuals could begin reporting actual income that would have otherwise not been reported; they may report the full amounts of such earnings; additionally, strategic taxpayers may also wish to reduce sus-

nonetheless face a marginal tax rate of 0 in the phase-in range.

[13]For three children, the phase in rates on the EITC and CTC are 0.45 and 0.15, respectively. After deductions, the overall marginal tax rate for a tax unit with wage earnings above $2,500 and below the standard deduction is: (1-0.765)*(-0.45-0.15+0.153)= -0.413.

[14]The reported amounts are, of course, subject to audit and individuals found to be misreporting income may face penalties. In practice, many individuals have informal sources of income that are difficult to detect when not reported (Bracha and Burke, 2018), but could potentially be reported on a tax return if desired. There is some discretion in reporting other wages and cash tips not recorded on a W-2, although in practice there is only very limited bunching occurring among non-SE filers.

picious behavior. In addition, as shown in Appendix Figure A.1, for households with 2 or more children, the marginal tax rate is essentially 0 immediately above the threshold, thereby creating neither an incentive or disincentive to report self-employment income that brings reported earnings over the threshold. We also want to note that while individuals already reporting some self-employment who have earned income in the credit phase-in range face an incentive to report additional self-employment income, such behavior would not affect the extensive margin of self-employment, which is our focus.

As we will show below, these incentives play an important role in whether households report self-employment income in practice.

2.3 Sample Construction

In our analysis, we use de-identified full-count tax records incorportating both filer-reported returns and information returns. Most of our analysis uses data from 2000-2018. While some microdata are available to us back to 1996, 2000 is when information returns are first available for 1099 independent contractors.

We determine family structure based on information in tax filings as well as links from parents to children from the Social Security Administration. These child links from the SSA provide comprehensive coverage for children born since the early 1980s who were issued a Social Security number, providing a complete picture of the number of children under 18 years of age by 2000. More information on how we clean and process the data is provided in the Data Appendix.

3 Trends

3.1 The Rising Self-Employment Share in Tax Data

In this section, we discuss trends in administrative measures of self-employment prevalence. As documented elsewhere, we show that self-employment rates are rising in ad-

ministrative tax records, in contrast with survey measures of self-employment (Jackson, Looney, and Ramnath, 2017; Abraham, Haltiwanger, Sandusky, and Spletzer, 2020). We next present an interesting new fact that self-employment trends differ when self-employment is third-party reported, as it is for gig workers and freelance contractors, compared to when this income is entirely self-reported to tax authorities. For much of our period of study, we find that growth in freelance/contractor work was substantially lower than overall self-employment growth. We show that self-employment propensities in administrative tax data have grown only for individuals with wage incomes in the phase-in region of the EITC. This finding motivates our causal research design in subsequent sections, where we isolate the reporting effect of self-employment from real changes in labor supply.

We begin by reporting the time-series for our baseline measure of self-employment in administrative tax data: the annual share of individuals in the workforce with positive net self-employment income reported on a Schedule SE.[15] In our analysis, we define the "workforce" as all individuals that have any of the following in a year: wage (W2) earnings, self-employment (Schedule SE) earnings, or 1099 non-employee compensation not reported on Schedule SE so long as the individual appears on a tax return or has other W2 income.[16]

As reported in Figure 1, the portion of the workforce with self-employment rose significantly after 2000, peaking around 2014 after increasing by around 2 percentage points. This is in sharp contrast to trends observed in labor force surveys. Appendix Figure A.2 contrasts these trends measured in tax data with comparable trends measured in the benchmark CPS-ASEC survey data based on retrospective reports of annual self-employment and wage/salary income. To extend the series further back in time and increase comparability, we use published SSA counts of W2 and Schedule SE

[15]Schedule SE filing is required if an individual's profits from self-employment are greater than $433. We focus on Schedule SE rather than Schedule C profits as our baseline because only the former is consistently available at the individual level rather than the tax-unit level in our data source throughout our period of study.

[16]When a 1099 recipient has no 1040 or W2, it is impossible to tell whether the 1099 is issued in error to someone out of the workforce, if the individual is in the workforce but not reporting correctly, or if the 1099 income was not taxable—in which case it is unclear whether or not the person was really doing "work"

recipients going back to 1987.[17] In contrast to the trend in tax data, the comparable self-employment rate in the survey data declines by about 1 percentage point over this period. Moreover, we note that the growth in self-employment in administrative records begins around 2000 where our primary microdata sample begins.

3.2 Do New Forms of Gig Work Explain the Rise?

One possibility is that this secular rise in the the share of the workforce with self-employment income heralds a fundamental change in the labor market, driven by new types of contract-based "gig" work enabled by new technologies (Katz and Krueger, 2019). The payments reported on 1099 information returns allow us to directly assess whether this is the case. As discussed above, non-employee compensation reported on 1099s is a subset of self-employment that is verified via third-party reporting from employers, typically for payments of $600 or more. One should note that not all 1099-reported labor compensation will be reported as self-employment income on Schedule SE: for instance, if someone had substantial expenses, they would not meet SE filing requirements. Nonetheless, since there is potentially discretion in both expensing and filing decisions, we present the total share of workforce participants with non-employee compensation on a 1099-MISC or a 1099-K from an online platform company in Figure 1. To investigate the contribution of new online platforms to the overall trend, we break out the portion of contract workers who have 1099-reported nonemployee compensation only from OPE platforms.

We observed in the time series plotted in Figure 1 that trends in 1099-reported contract work cannot explain the overall rise in reported self-employment on tax returns—and, in particular that the rise of new kinds of platform has little association with trends in reported self-employment. We find that 1099-based contract work did become more prevalent between 2000 and 2007, potentially contributing to the self-employment trend. However, as SE rates continued to rise between 2007 and 2014, the share with contract income remained largely constant. Notably, prior to 2014, online

[17] In this figure, the workforce is defined as all individuals with either W2 or Schedule SE income

platform work was rare and essentially nonexistent prior to 2007.

The vertical line in the figure represents a structural break in our freelance/contracting series, when new work mediated by online platforms starts to become meaningfully large as a share of the workforce. Strikingly, we find that after 2014—the time period where gig work grows most dramatically—administrative measures of self-employment are actually flat and even *decline* slightly. This finding implies that new forms of "gig" work cannot explain the observed rise in self-employment in tax data, since the SE trend had plateaued before platform work became widespread. On the whole, annual changes in the SE series and the non-employee compensation series between 2000-2018 show no statistically significant correlation.

Nonetheless, the growth in new online platform work post-2014 is substantial. This raises the question—why did larger numbers of platform workers not result in more workers reporting self-employment income? We explore this question in depth in a companion paper (Collins, Garin, Jackson, Koustas, and Payne, 2019). Crucially, that paper finds that while the number of individuals with platform earnings has grown rapidly, the vast majority of those workers only earn a small amount of money over the course of the year. The median platform worker makes less than $2,500 over a year through platform work, and, moreover, these workers typically expense over 50 percent of their gross revenues. As a result, many workers have net earnings below the Schedule SE filing threshold after deducting expenses on their Schedule C.

In addition, Collins, Garin, Jackson, Koustas, and Payne (2019) document that many platform workers with low revenues do not file a Schedule C at all, and that platform workers are less likely to file a Schedule C or SE than other types of contractors with nonemployee compensation on 1099-MISC. In 2016, 43 percent of worker with platform income reported on a 1099 did not file a schedule C at all, whereas only 22 percent of other workers with nonemployee compensation had no Schedule C. This latter observation raises the possibility that trends in non-employee compensation may diverge with trends in reported self-employment earnings because of changes in compliance or in the composition of types 1099-MISC work with different propensities to be

reported on Schedules C and SE. We directly examine this possibility in the discussion of counterfactual scenarios later in the paper.

3.3 Self-Employment Trends and Eligibility for Refundable Tax Credit

As discussed in Section 2.2, individuals with dependents on their tax return and earned income in the phase in range of the EITC and CTC have an incentive to report additional income in order to receive a larger refundable credit. In families with children and *wage* income below the top of the phase-in range, these incentives could motivate individuals to report self-employment income who otherwise have reported none.

As a preliminary exploration of the potential contribution of reporting incentives in the tax code to trends in reported self-employment, Figure 1(b) examines growth by EITC claiming and presence of children.[18] As shown in the left-hand panel, the growth in SE propensity is substantially higher for EITC recipients with children, growing from about 14 percent in 2000 to over 20 percent by 2010. In comparison, we see only modest changes in 1099 receipt and SE growth for childless filers—none of whom face negative marginal tax rates. Figure 1(c) examines growth by total earnings. The left hand panel is the SE propensities among those with total earnings less than $50,000 (in constant $2015). The right hand panel is for those with more than $50,000. We find SE rates rise most dramatically for those with less than $50,000 in total earnings, rising from about 9 percentage points in 2000 to over 12 percentage points by the 2010s, with a sharp increase during the 2007-9 recession. The time pattern for SE rates looks different for higher-income earners: SE grows between 2003-2005, falls sharply in the recession, before returning to its 2005 level. Appendix Figure A.3(a) reports additional trends by gender, showing that most of the increase occurs among women.

In theory, an incentive to report positive self-employment income rather than none only exists for individuals in families with wage earnings in the credit phase-in range.

[18] A.3(a) shows additional trends by presence of children only.

Conversely, families with combined wage earnings above the refund-maximizing amount have no incentive whatsoever to report additional self-employment income. Accordingly, we next examine more closely where in the wage distribution we see a rise in rates of self-employment reporting.

Figure 2 reports the change in self-employment propensity between 2000 and 2014, by W-2 wage earnings and presence of children.[19] The results for Schedule SE self-employment, reported in Figure 2, Panel A, are striking. We find *all* of the growth in self-employment within each wage bin comes from households with children and combined wages below the first EITC kink point, indicated by the maroon and green lines on the figure. These are precisely the households that stand to gain from reporting some small amount of self-employment rather than none. There is essentially no change in self-employment between 2000-2014 for households without children, or for households with higher levels of wage income. This raises the possibility that incentives to report self-employment income on tax returns—to either report previously unreported earnings or to fabricate income—are a factor driving growth in reports of self-employment earnings.

Panel B reports the change over time for 1099-reported self-employment. Importantly, this 1099-reported work is reported by firms, and therefore not subject to discretion by taxfilers. We see almost no change over this period for households, regardless of whether they have children. Taken together, this provides strong evidence that households are motivated by EITC incentives to report self-employment income. In the next subsection, we explore how behavior changes when a household first becomes eligible for these incentives at childbirth.

[19]The overall change in self-employment is the integral of the wage-bin propensity times the share of the population in each wage bin.

3.4 Self-Employment Reporting and Changing EITC Incentives: Event-Study Around Childbirth

As discussed above, households face a negative marginal tax rate for reporting self-employment income only if they have children. To further test the hypothesis that self-employment growth is tied to EITC incentives, we follow Chetty, Friedman, and Saez (2013) and examine how self-employment reporting changes around a woman's first childbirth, when they become eligible for a generous credit. We expound upon Chetty, Friedman, and Saez (2013) in two main ways. First, to examine the extent behavior is changing *over time*, we examine the change across different time periods. Second, we separate the rise in self-employment around childbirth into 1099-reported self-employment and self-reported work. An increase in 1099-reported work may suggest changing worker needs around childbirth draw workers into self-employment for the first time. We begin with a simple exercise, examining the raw change in self-employment at childbirth, before formalizing our analysis in an event-study framework.

We start by examining the simple raw change in self-employment in the year of childbirth. We take childbirths for all parents reported in the SSA database whether or not the child is claimed as a dependent on tax filings by that parent. Figure 3 reports the change in self-employment filing in the year of childbirth from the year before, for every cohort of first births from 1997-2018. The figure shows that the extent to which individuals begin reporting self-reported self-employment exactly when it becomes advantageous to do so has increased over this period by 0.9 percentage points, from a level of 0.9 percentage points in 1997 to 1.8 percentage points by 2014. 1099-reported work—which individuals have no discretion over reporting—differs in two key ways. First, on average, there is *no* increase in 1099-reported work in the year of childbirth. Second, there is no underlying trend in the rate of doing 1099-reported work in the year of childbirth. Appendix Figure A.4 further breaks down the trends by gender of the parent. We find that all of this increase comes from mothers: the change in self-employment in the year of childbirth among mothers has gone from 0.4

percentage points in 1997 to 2 percentage points by 2014. In contrast, 1099-reported work *decreases* in the year of childbirth for mothers; the decrease is actually slightly greater in magnitude today than in the past.

We next proceed to formalize this analysis and examine additional periods after childbirth using an event-study specification that will control for aging and business cycle effects. Our event study specification is standard and given as follows:

$$y_{it} = \sum_{k \in K} \beta_k^p I\{\text{FirstChildbirth}_i = t + k\} + \gamma^p_{a(i)\times g(i)} + \gamma^p_{t\times g(i)} + e^p_{it} \tag{1}$$

where i indexes parent, t indexes year, $a(i)$ gives the age of i. FirstChildbirth_i is i's year of first birth. $g(i)$ is the parent's gender, thus allowing for time and age effects to differ by parental gender.[20] We examine two key outcomes: having any contract/freelance work, and being an S.E. taxpayer with no contract/freelance work. We run separate regressions for different 3-year rolling windows, $p \in \{2003-2005, 2004-2006, ..., 2012-2014\}$. We exclude an indicator for the period one year prior to first birth, so that the event-time coefficients are all relative to period -1, and examine an event window of 4 years pre and post event ($k \in \{-4, ...4\} \setminus -1$). Standard errors are clustered at the individual level.

Figure 4 plots the full set of event study coefficients we estimate for two cohorts of births: 2003-2005 births and 2011-2013 births. As in the raw means, we find that 1099-reported self-employment is flat around childbirth for both cohorts. But self-reported self-employment is a different story—the propensity to self-report self-employment income increases sharply in the year of birth and by about 0.75 percentage points in subsequent years. Moreover, the magnitude of this time 0 response has grown over time: while self-reported self-employment rates grew by 0.5 percentage point after childbirth in 2003–2005, the corresponding increase was around 1.25 percentage points in 2011–2013. This contrasts with firm-reported contract work, which did not become more common after childbirth in either time period. Appendix Figure A.5 reports

[20] Accordingly, the event-study coefficients are the average of coefficients run separately for men and women, which we report in Appendix Figure A.5.

estimates separately be gender of the parent. As we found earlier, these changes over time are largely driven by mothers.

4 Labor Supply Versus Reporting

4.1 Research Design

A key question is whether the rise in self-reported self-employment rates—concentrated among individuals with a strong EITC incentive to report additional earned income—reflects a true increase in self-employment activity or rather a greater proclivity to self-report that is not indicative of any change in underlying labor supply. We find that the propensity to have self-reported self-employment income is closely tied to EITC incentives, while the propensity to have third-party-reported self-employment income is not. However, this finding might reflect real differences between self-employment that is reported by firms on 1099s and self-employment that is not. For example, finding a firm to contract for to earn extra money may be significantly more difficult than providing services to neighbors.

To shed light on this question, we examine a sharp discontinuity in EITC eligibility based on the date of birth of an individual's first child, motivated by LaLumia, Sallee, and Turner (2014). An individual's EITC benefit for tax year t is calculated based on the number of children in their household during year t—a first-born child born on or before December 31 of year t would count towards the EITC calculation for that year t. By contrast, a child born only a few days later at the start of year $t+1$ would not count towards the EITC calculation in year t. Accordingly, the EITC creates sharp differences in reporting incentives for tax year t for individuals having their first child days apart at the very end of year t and those having theirs at the very beginning of year $t+1$. The former individual faces a negative marginal tax rate on reported earned income in year t up to \$9,880 in 2015 constant dollars, whereas the latter faces a positive marginal tax rate on all earnings (after netting FICA/SECA taxes). Thus, if

both individuals had zero wage income, the individual with the birth in late December would strictly benefit from reporting a small amount of self-employment income while the latter would be worse off from doing so.

If birth dates were known with complete certainty well in advance, individuals expecting a first child near the end of tax year t could make labor supply decisions early in the year that respond optimally to the corresponding EITC incentives. However, in practice, expecting parents face substantial uncertainty about the exact date the birth will occur before the last weeks of a pregnancy. Accordingly, parents expecting a first child close to the end of the year will remain uncertain about what year the birth will occur in—and the corresponding EITC status for year t, until labor supply decisions for year t are largely sunk.

This motivates a regression discontinuity design comparing year-t self-employment earnings for parents with their first birth right before and right after the end of tax year t. If, in the limit—examining births right before and after midnight on December-31—parents face complete *ex-ante* uncertainty about the year of birth, parents should make identical labor supply decisions in year t. However, *ex post* after the children are born, the two sets of parents face different returns to self-reporting self-employment. Thus, at the limit, all differences in self-reported tax year t self-employment between parents on either side of the discontinuity will be due entirely due to differences in reporting behavior, holding underlying labor supply fixed.

The key identifying assumption is that true year-t labor supply is identical for parents on either side of the December 31 cutoff. This assumption would be violated if parents are able to schedule births around New Year's Eve to maximize their tax refunds given their *true* year-t labor supply. LaLumia, Sallee, and Turner (2014) test for such behavior and while they find some response of the timing of births to tax incentives, the effect is concentrated almost entirely among second and later births—the effect of tax incentives on first-birth timing is negligible. However, LaLumia, Sallee, and Turner (2014) also find that self-employed parents with births in December are much more likely to have earnings for that year that are bunched around the first EITC kink point

than parents with births in the following January. In our analysis, we directly address the possibility of strategic birth timing by examining discontinuities in other earnings margins around the same cutoff date. If parents are re-timing births to maximize tax credits given their sunk labor supply decisions, they should do this regardless of the source of income.

We estimate regressions of the form:

$$y_i = \alpha + \beta \mathbf{1}\{date_i \in December\} + f(date_i) + \epsilon_i \tag{2}$$

on the sample of all childbirths in December and January. In this specification, $date_i$ is the running variable—which is the child's birthdate measured as days since December 31 and ranges from -30 to 31—and $f(\cdot)$ is a cubic polynomial. The coefficient β on the indicator $\mathbf{1}\{date_i \in December\}$ is the RDD effect of having a child in tax year t instead of tax year $t+1$. The outcome y_i is a component of income tax year t for December births in year t and January births in year $t+1$.

We estimate Equation 2 on all parents with births in December of the specified tax years or the subsequent January, as identified in Social Security birth records. This sample includes both parents listed in Social Security records, regardless of whether or not that parent claims the child on their tax return. In practice, when new parents are unmarried and one parent claims the child as a dependent on their tax return, the other may not; since we cannot observe which parent is the true custodian, we consider the effects for both of the parents. This identifies an effect that is akin to an "Intent to Treat," which incorporates both the pure effect of being able to count a child towards the EITC and the probability of being the parent who claims the child.

4.2 Baseline Results on Reporting Effects

We first estimate effects pooling births around the end of each tax year from 2011 to 2018—focusing on recent years with elevated self-employment reporting rates. To validate our research design, we begin by testing for differences in *third-party* reported

income around the discontinuity at the end of the year in Panel A of Table 1. If individuals re-time births to maximize refundable tax credits, they face the same incentive to re-time their births regardless whether the source of earned income is self-reported self employment or firm-reported employment or contract income. The results in Column (1) and Appendix Figure A.7 show there is no such effect on the propensity to have employment income reported by firms. Column (2) shows there is also no impact on the level of W-2 reported wage and salary earnings in the tax unit. Moreover, the results in Column (4) show that there is no effect on the propensity to have self-employment income from independent contracting work reported by firms on 1099 forms.

However, as noted by LaLumia, Sallee, and Turner (2014), the benefits of re-timing a birth do not depend so much on the presence or amount of income *per se* as on the change in tax liability that would occur as a result. Accordingly, as a stronger test of whether parents sort across the cutoff date to maximize their EITC based on their firm-reported employment earnings, we calculate a "simulated EITC" that is the credit each individual would earn if their birth occurred in December of tax year t based on their and their spouse's (if married filing jointly) W-2 wages from year t. If individuals are sorting to maximize EITC, then wages of parents immediately to the left of the discontinuity should predict higher credits than wages of those to the right. However, the differences in simulated credits at the discontinuity, displayed in Column (3) and graphically in Panel A of Figure Figure 5, are negligible and are not statistically distinguishable from zero. Moreover, we find that individuals are equally likely to have wages in the EITC phase-in region on both sides of the cutoff. Together, these initial results support our assumptions that labor supply is fixed by the time uncertainty is resolved about when birth will occur and that births are not strategically timed based to maximize refundable credits.

Since we find no differences in wage earnings around the cutoff, we can directly quantify the difference in incentives to report SE that households face around the cutoff. Specifically, we measure the incentive to self-report self-employment income based on the effective marginal tax rate individuals would face based on their firm-

reported W-2 earnings—this marginal tax rate determines the extensive margin return to reporting a small positive amount of self-employment rather than none. In Table 1(b) we report the RDD effect on the share facing negative marginal tax rates on a first dollar of self-employment income and on the average change in the marginal tax rate. Our RDD estimates imply 25 percent of individuals with December births face negative marginal tax rates. Among these households, the average decrease in the marginal tax rate is around 30 percentage points.

Having established a large change in incentives around the cutoff, we next explore how these incentives affect the propensity to self-report self employment income. Panel B of Figure 5 graphically presents results from the regression discontinuity design. The change in self-employment status based on having any Schedule SE earnings in tax year t relative to the prior year $t-1$ is displayed in blue, alongside the change in having non-employee income reported on a 1099-MISC in tax year t relative to the prior year $t-1$, displayed in red. Each dot is the average outcome for parents with births on the corresponding date, pooled across years. The curves are the regression fits allowing for the estimated discontinuity between December 31 and January 1.

The figure shows a sharp break in the propensity to self-report self-employment income after December 31. Away from the discontinuity, the relationship between date of birth and reported self-employment is completely smooth. However, as soon as the incentive to report self-employment for tax year t disappears, the odds of self-reporting self-employment income falls dramatically. This is true both for men and for women, though the size of the discontinuity is larger for women than for men. By contrast, though the incentives are the same, there is no change in the propensity to have nonemployee compensation income reported by *firms* on a 1099 form around the same cutoff date.

Table 1(c) reports our RDD point estimates on SE filing. At the discontinuity, having a child who is eligible for EITC determination increases the probability of reporting self-employment income by 1.4 percentage points (the estimates are 1.6 percentage points for women, and 1.2 percentage points for men). Since we have shown

labor supply is held fixed at the discontinuity, and births are not strategically re-timed around the discontinuity, these results can be interpreted as pure reporting effects. Combining the extensive and intensive margin responses, the average increase in the amount of self-employment net income reported is approximately $145.

The results in Columns (2)-(3) confirm that these reporting responses are concentrated among individuals with a strict incentive to do so. Among women, we find no regression discontinuity effect on the propensity to report self-employment income in cases where combined wages exceed the first EITC kink. However, the effect for women with combined wages below the kink is large—over 5 percentage points. We find similar results for men.[21]

In the last three columns of Table 1(c), we quantify how much amount of incentive received by reporting SE, which we calculate as the increased amount of EITC actually received, above what would be expected based on wages alone (the simulated EITC amount based on W-2 wages described above), and subtracting out any increased self-employment payroll taxes paid. We find that the average incentive received is around $70–among those below the kink the average received is around $260.

In Table 1(d), we explore some of the sources of the rise in SE reporting. We find that 0.4 percentage points of the 1.4 percentage point increase in SE filing comes from individuals that would not have otherwise been in the workforce. Importantly, we find that only a small portion of the reporting effect comes from hyper-strategic "sharp-bunchers" who report within $500 of the amount of income to maximize credits.

In the final columns of Table 1(d), we examine how the types of self-employment activities reported on Schedule C differ around the threshold. Nearly one-third of the increase is concentrated among domestic/personal services, like reporting cleaning and babysitting income. Additionally, about half of the increase is among individuals who report no expenses on Schedule C, which is disproportionately larger than their baseline share of about one-third. This is suggestive that incentivized individuals

[21] We find some evidence of a self-employment reporting effect for men with combined wages above the kink level, but the point estimates are economically insignificant and are more than a full order of magnitude smaller than the effects for men with combined wages below the kink.

report earnings differently than non-incentivized individuals and may be more likely partake in self-employment with no formal record-keeping. These results support our story that taxpayers have considerable reporting discretion over this self-employment income, but do not allow us to differentiate between self-employment income that may be fabricated, or the reporting of previously unreported income.

4.3 Reporting Effects Over Time

The results above suggest that since 2010, incentives in the tax code have led individuals with new children to report self-employment income on their tax filings, independent of their underlying labor supply. We next examine whether these behaviors have evolved over time. We estimate the main RDD specification separately for birth cohorts in each year—that is, births in December of each year t and in January of the corresponding year $t + 1$.

Baseline estimates for each tax year pooling all individuals are presented in Panel A of Figure 6. We find that reporting effects increased substantially over time, growing threefold between 2000 to 2014 from an effect size of 0.5 percentage points in 2000 and 2001 to over 1.5 percentage points in 2013 and 2014. Notably, the pattern of effects across year is very similar to the pattern of self-employment increases around all childbirths (not just those near the discontinuity) over the same period, as shown in Figure 4. The magnitudes of the effect are consistent with the entire increase in the propensity to report self-employment after first childbirths stemming from changes in reporting behavior.

In Panel B of Figure 6, we contrast behavior among individuals with combined wages below the first EITC kink and those with wages above. While we find negligible effects for individuals with combined tax-unit wages above the kink in all years, we find that reporting effects have increased significantly over time for individuals with household wages in the EITC phase-in region. This effect grows approximately fivefold between 2000, when the effect size is 1 percentage point, and 2015, when the effect size

is six percentage points.[22] These results indicate that the propensity to report self-employment on tax returns *conditional* on one's true labor supply is far more responsive to incentives in the tax code in recent years than it was two decades ago.

The time path of effects in Panel A does not exactly match the time path of effects for individuals with combined wages below the kink in Panel B. This is because the size of the pooled effects in Panel A reflect both the the size of the effects for incentivized individuals in Panel B, and the share of individuals with combined wages in the EITC phase in range. While the first factor has grown nearly monotonically since 2000, the second factor reflects the aggregate distribution of wage earnings and is counter-cyclical. When joblessness rises and real wages stagnate, a larger portion of individuals will have combined wage earnings below the first EITC kink point, in turn increasing the size of the pooled effect in Panel A in recession years like 2009. Conversely, the share of individuals with combined earnings below that same kink point declines during expansions, reducing the pooled effects in Panel A in years with lower joblessness and higher real wage growth like 2017 and 2018. Appendix Figure A.8 highlights these fluctuations in the share of the workforce that falls in this incentivized group each year.

5 Knowledge of Tax Incentives and Changes in Reporting Behavior

Our analysis indicates there has been a substantial rise in tax-code-motivated reporting of self-employment income over the past two decades. This trend occurred despite no major changes in the generosity of the federal EITC over the period we study. Nonetheless, prior research suggests that even during periods when policy remains fixed, *knowledge* of the incentives created by the EITC and related tax provisions differs across regions and spreads gradually over time (Chetty, Friedman, and Saez,

[22] We also find that this increase is particularly dramatic among women with wages in the phase-in range, whose effects increased twelvefold from 0.5 percentage points in 2000 to nearly 6 percentage points in 2015.

2013). This process of gradual learning could lead to a steady rise in reporting of previously unreported self-employment earnings or false reporting of self-employment income. Learning dynamics have been shown to be particularly important in the case of the EITC. Consistent with the gradual learning story, earlier work has documented that hyper-strategic "sharp-bunching" behavior—reporting exactly the amount of self-employment income that qualifies you for the maximum tax benefit, has become more common over time (Chetty, Friedman, and Saez, 2013; Mortenson and Whitten, 2020). Chetty, Friedman, and Saez (2013) observe that this spread is partly geographic in nature: the more common "sharp-bunching" is in a locale, the more likely new mothers are to claim more generous benefits after having their first child.

Could local increases in knowledge of the tax code drive changes in strategic self-reporting of self-employment income? As a simple test, Appendix Figure A.10 compares changes in local knowledge between 2000 and 2009—proxied using the Chetty-Friedman-Saez (2013) measures of sharp-bunching at the 3-digit zipcode level—with changes in share of the local workforce that report self-employment, broken out by whether self-employment income is verified by third-party reporting by firms. Consistent with the learning hypothesis, we find that increases in local sharp-bunching rates among the self-employed strongly predict increases in the *overall* prevalence of self-reported self-employment in the local workforce—but *not* in the prevalence of firm-reported self-employment.

We next replicate Chetty, Friedman, and Saez (2013)'s measure of sharp-bunching and extend in through 2018. The resulting aggregate series is presented in Appendix Figure A.9. Consistent with Mortenson and Whitten (2020), we find that sharp-bunching has continued to increase since CFS original series ended in 2009, increasing by over 1 percentage point more to around 4 percent of the eligible EITC population with children by 2013. While the prevalence of hyper-strategic "sharp bunching" is a useful measure of knowledge, it is likely that many individuals deciding whether to report their self-income activity consider the tax benefits but do not manipulate amounts deliberately. Indeed, the magnitudes of the increases in self-reported self-employment

are too large to be explained by sharp-bunchers alone. Thus, on its own, the level of sharp-bunching does not tell us whether individuals who would report self-employment anyway are strategically manipulating self-employment income or expensing in order to end up in the bunching range (intensive-margin response), or whether knowledge of EITC incentives is leading individuals to report self-employment that they would not have otherwise reported (extensive-margin). Only changes in the extensive-margin of reporting would reflect the SE trends we documented earlier.

To assess whether changes in local knowledge might explain the evolution of the RDD estimates in Figure 6 over time, we estimate an interacted version of the main RDD specification in Equation 2 that allows the effect to vary with local knowledge, proxied by the share of sharp bunchers in each ZIP3 in each year.[23] Following Chetty, Friedman, and Saez (2013), we tabulate the sharp-bunching share in each ZIP3 in each year among individuals *without* new births in that year. We then fully interact the estimating equation:

$$\begin{aligned}\Delta y_{izt} =& \alpha + \beta BunchShare_{zt} \times \mathbf{1}\{date_i \in December\} + \delta BunchShare_{zt} \\ &+ \gamma \mathbf{1}\{date_i \in December\} + f(date_i, BunchShare_{zt}) + \epsilon_{izt} \end{aligned} \tag{3}$$

We interact each term in the polynomial with the bunching share. We estimate this specification pooling all birth cohorts in the December and subsequent January of each years 2000 through 2018.

To motivate this interacted RDD, Panel A of Figure 7 displays the reduced-form relationship between changes in self-reported self-employment status in year t for parents with first births near the end of the year, plotted against the year t sharp bunching rate among other taxpayers in the same ZIP3. For new parents with first births in December of year t and year-t wages below the first EITC kink—who would gain from reporting additional self-employment income—there is a clear relationship between local EITC knowledge and the share of new parents who begin self-reporting self-employment in-

[23] Zipcode is taken from W-2s and 1099-MISCs in the current year, or most recent of the prior two years if missing.

come. However, for parents without the incentive to report—both those with births in January of year $t+1$, or parents with December births but wages above the first EITC kink—this relationship is close to flat. Our interacted RDD formalizes this comparison, evaluated at the December 31 cutoff.

The results, presented in Table 2, suggest that local knowledge plausibly plays an important role determining the propensity to report self-employment income in response to incentives. When sharp-bunching is more common among others in the same region, new parents are significantly more likely to report self-employment income in response to incentives. Increasing the local bunching share from 0 to 2 percentage points more than doubles the magnitude of the reporting effect. We find that local information matters more for new mothers than for new fathers; moreover, new fathers' baseline behavior in low-knowledge areas is notably higher than that of new mothers. However, this knowledge proxy does not affect behaviors that are not directly incentivized by the tax code—there is no interaction effect on propensities to get third-party-reported labor income or propensities to report self-employment among those with wages above the EITC phase-in range.

These results are suggestive, but to the extent that variation in local sharp bunching across years could be driven by national trends, the estimated interaction effects could potentially be driven by other national trends besides the diffusion of knowledge. To assess whether this is the case, we examine how the relationship between self-employment reporting and local knowledge in Panel A of Figure 7 changes when only considering within-year cross-regional variation or within-region time variation. Formally, for each group (December t births and wages below the kink, t, December t births and wages above the kink, January $t+1$ births with wages below the kink), we estimate the following equation at the ZIP-year level:

$$\Delta y_{zt} = \delta_z + \delta_t + \beta BunchShare_{zt} + \epsilon_{zt} \tag{4}$$

Here, the outcome is the average change in self-reported or 1099-reported self-employment

status in year t relative to year $t-1$ for new parents in the specified group who reside in ZIP3 z in tax year t.

The results displayed in Table 3 and plotted in Panel B of Figure 7 indicate that the relationship between local knowledge and reporting behavior is similar whether isolating variation knowledge within period or within place. The SE effect in Column 1 is identical to the slope in Figure 7, by construction. Whether adding year fixed effects (isolating cross-ZIP variation), ZIP fixed effects (isolating cross-year variation), or both (isolating differential within-ZIP changes in knowledge over time, the relationship between knowledge and reporting behavior is stable. After including both ZIP and year fixed effects, we find the slope is 1. Meanwhile, limiting to within-variation, we find no relationship between changes in local knowledge and changes in reporting behavior among parents with no incentive to do so, including those with births only a month later in January. We also find no relationship between knowledge changes and propensity to have nonemployee compensation reported on 1099s among any group. Notably, the relationship between local knowledge and self-employment reporting in Column 4 is almost identical to the RDD interaction estimates of the same object for individuals below the first EITC kink in Column 6 of Table 3

To interpret the magnitude of these effects, it is useful to compare them to the aggregate bunching rates for each year presented in Figure A.9. Between 2000 and 2014, the sharp bunching rate grew by over two percentage points. Our interaction estimates in Table 2 implies that the increases in local bunching behavior comprising this aggregate increase could account for a 0.5 percentage point increase in the pooled RDD effect due to localized learning over that same period. Among individuals with wages in the EITC-phase in range, we can account for a 2 percentage point increase in the RDD reporting effect over time. This only accounts for about one-half of the observed increase in the RDD effects over time shown in Figure 6; however, this exercise only reflects localized knowledge transmission and any other knowledge spread that is not mediated by ZIP codes is not captured here. Thus, gradual transmission of knowledge through both localized and non-localized channels plausibly explains the

observed increase in reporting behavior over time.

Focusing on first childbirths may understate the full impact of increasing local knowledge of tax incentives if individuals absorb local knowledge gradually over time after their initial eligibility. In that case, greater local knowledge might impact self-employment reporting for the average eligible individual with children more than for eligible individuals immediately after their first births. For comparison, Panel C of Figure 7 replicates Panel B using the share of the total workforce reporting self-employment in each ZIP and year, rather than just individuals with first births at the end of the tax year. Similar to Panel B, Panel C shows a clear relationship between changes in local bunching rates and self-employment reporting among low-wage individuals with children, but no such relationship between bunching rates and self employment among individuals with children and wages above the credit phase-in range or individuals without children.[24] However, the increase in self-employment reporting rates among all incentivized parents, associated with one-percentage point increase in local sharp bunching rates, is four times the magnitude of the increase in self-employment reporting among incentivized parents shortly after their first birth. This suggests that spreading knowledge of tax incentives could increase self-employment reporting rates significantly more than suggested by our analysis of first births.

6 Quantification/Counterfactuals

The empirical analysis above finds that individuals with incentives to report self-employment have become increasingly likely to do so. Further, our RDD analysis finds that the entire increase in the propensity to report self-employment after first births is accounted for by a pure reporting effect with no change in underlying labor supply, and this reporting effect has grown in magnitude over time. That is, incentivized individuals are increasingly likely to report self-employment income on their tax

[24]Note that the vast majority of sharp bunchers are incentivized individuals with children, wages in the credit phase-in range, and self-employment income. However, it is not obvious how intensive-margin sharp-bunching among such individuals should be related to the propensity to report self-employment altogether.

return that they would not have in the past—these could be informal sources of income that would have otherwise been unreported or even fabricated earnings. Our objective in this section is to quantify the effect that reporting incentives have on measured growth in self-employment and present self-employment trends under an assortment of counterfactual scenarios.

First, we apply our estimates from Section 4 to adjust the observed trend in Schedule SE. The regression discontinuity estimates can be interpreted as the short-run effect of having a single child eligible towards their EITC calculation, after labor supply decisions have been sunk. We use the annual effect estimates presented in Figure 6 to estimate the share of all individuals with children under 19 who would not have reported self-employment had they not faced negative marginal tax rates, and consider how much self-employment reporting rates would have grown if the annual effect estimates presented in Figure 6 had remained at their 2000 level. Specifically, our adjustment is given as follows:

$$\frac{SE_t^*}{WF_t^*}^{RDD} = \frac{SE_t - \beta_t^{RDD,\Delta AnySE} \cdot POP_t^k}{WF_t - \beta_t^{RDD,\Delta SEOnly} \cdot POP_t^k} \tag{5}$$

where $\beta_t^{RDD,\Delta AnySE}$ and $\beta_t^{RDD,\Delta SEOnly}$ are our RDD coefficients for reporting any self-employment and only self-employment, respectively, in year t. Our RDD coefficients are at the population level and refer to the increase in self-employment among all new parents, not just those in the workforce. To convert back to a measure as a share of the workforce, we multiply these coefficients by the population with children under 19 (the main eligibility criteria for the EITC), POP_t^k. To implement this adjustment, we use population-level estimates of individuals with children under 19 in their household from the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC). We then apply our annual regression coefficients to estimate the implied number of individuals who would not have reported self-employment if the effect size remained at its 2000 level. The result of this exercise is presented as "Scenario 1" in Figure 8a. This adjustment reduces the share of the workforce reporting

self-employment by about half of one percentage point in 2014, about one-fifth of the total increase since 2000.

In our view, the RDD adjustment identifies a strict lower bound on the share of individuals with qualifying dependents who report self-employment solely due to reporting incentives. There are several reasons we think these estimates provide a lower bound. First, the RDD effect is a short-run effect identified off of individuals who have not previously been eligible for refundable credits and learn only shortly before filing their taxes that their child was born prior to the end of the tax year. It is likely that individuals become increasingly aware of the incentives provided by the tax code in the longer term, and adjust their behavior accordingly. Consistent with that hypothesis, our findings in Figure 7 Panel C suggest that the effects of spreading knowledge of the tax code on the overall incentivized workforce may be significantly larger than the effects on new parents shortly after their first births. Second, the generosity of the credit increases as individuals have subsequent children, which could lead to higher rates of credit-motivated self-employment filing among those with additional children. Finally, our analysis examines all childbirths, but in some cases children will be claimed as dependents by unmarried co-parents or other relatives, and will therefore not be eligible in the calculation of one's own EITC.

We next consider how self-employment reporting rates would have evolved if trends among individuals with incentives to report additional self-employment income paralleled trends among individuals without such an incentive. More precisely, as a second adjustment, we replace SE rates for low-wage individuals with children—who have an incentive to report SE—with the rate among comparable individuals without children who have no such incentive. We implement these comparisons within age, gender, and narrow wage bins.

Formally, the baseline self-employment rate in year t is given as follows:

$$SE_t = \sum_g \omega_{g,t} \sum_w \sum_k \zeta_{w,k,g,t} SE_{w,k,g,t} \tag{6}$$

where we denote g a gender-age cell (e.g. women in a certain age band), $k \in \{0, 1, 2+\}$ the number of children, and w the wage bin for a tax unit's total wages. In practice, we set the size of the wage bins to \$500 increments and examine age bins of [14-25, 26-55, 56-64, and 65+]. In this expression, $\omega_{g,t}$ is g's share of the overall workforce in year t, and $\zeta_{w,g,k,t}$ is the share of $\{g, k\}$ that lie in wage bin w in year t; therefore, $\sum_g \omega_{g,t} = 1$ and $\sum_w \sum_k \zeta_{w,k,g,t} = 1$. Individual are in the incentivized region of the tax schedule if they have children and wages below the first EITC kink point for k, $w^*(k)$ (i.e. $k = 1$ and $w < w^*(k = 1)$ or $k = 2+$ and $w < w^*(k = 2+)$).[25]

Our approach is to replace the actual SE rate for incentivized individuals in each year with children with those for unincentivized individuals without children in the same cell. We compute the counterfactual self employment rate as follows:

$$SE_t^* = \sum_g \omega_{g,t} \sum_k \left(\sum_{w \leq w^*(k)} \zeta^*_{\mathbf{w,g,k,t}} SE_{w,g,\mathbf{0},t} + \sum_{w > w^*(k)} \zeta^*_{\mathbf{w,g,k,t}} SE_{w,g,k,t} \right) \quad (7)$$

This expression additionally accounts for changes in the composition of the workforce implied by this exercise. Individuals with no wages require special consideration since in the absence of the incentive, some of these individuals would no longer be part of the tax workforce at all. In order to account for this, the weights must be adjusted. More precisely, to calculate the counterfactual number in the workforce in the absence of the incentive, we return to the ASEC population data. Denote $POP_{0,g,k,t}$ the population of g, k with 0 total wages in year t. We calculate a counterfactual number of such individuals with k kids who would otherwise be in the workforce in the absence of incentives as $WF^{SE*}_{0,g,k,t} = WF^{SE}_{w,g,\mathbf{0},t} / POP_{0,g,\mathbf{0},t} \cdot POP_{0,g,k,t}$, i.e. we apply the self-employment-to-population ratio for those without children, and multiply by the population with k children. (We report the implied change in the workforce in Appendix Figure A.11). We recalculate the weights of our wage bins using this counterfactual level, and denote these adjusted weights $\zeta^*_{w,g,k,t}$.

This second adjusted series is presented as "Scenario 2" in Figure 8a. Under this

[25] We further refine our definition to exclude individuals with total earnings (wages+SE) above the phase-out region, and with investment income above allowed thresholds.

counterfactual scenario, the share of the workforce observed with self-employment earnings would have been nearly a full percentage point lower in 2014—cutting the increase in those years by roughly half. By contrast, when we perform the same adjustment to the share of the workforce with independent contractor earnings reported on 1099 forms in Figure 8b, the counterfactual rate is nearly identical to the observed rate. This latter finding reflects the fact that independent contracting rates among low-wage individuals with kids evolved in parallel to rates for other individuals since 2000. The adjustment to the self-employed share eliminates most of the change between 2005 and 2012, during which time observed self-employment filings rose substantially. This difference reflects not only increasing rates of self-employment reporting among individuals with qualifying children and wages below the first EITC kink, but also an increase in the share of individuals with low wages—and hence incentives to report additional self-employment earnings—in the wake of the Great Recession.

This second approach allows for a straightforward adjustment at finer levels of geography, such as state. We simply redefine the cell to include state, where we determine state based on the recipient address on information returns. The results of our state-level adjustment are reported in Appendix Figure A.12. Panel (a) shows the raw percentage point growth in schedule SE by state. The deeper red indicates more growth. We see very dramatic growth in the South, as well as California, New York, Nevada, and Michigan. Panel (b) shows the size of our Scenario (2) adjustment by state. Many of the states that showed the most growth in Schedule SE are also places that appear to have the most incentivized growth. This figure also mirrors the geographic spread of knowledge in (Chetty, Friedman, and Saez, 2013). Finally, Panel (c) shows what our adjusted series looks like, using the same color breaks as the first map. Growth is dramatically scaled back in much of the US. Maryland and New York still have fairly dramatic growth, but otherwise our adjusted series shows much more muted growth across the country.

In the Appendix, we also report the implied changes to the *level* of the self-employment rate had there been no reporting effect in any year. These level ad-

justments report a specific scenario where incentivized people would have otherwise reported their self-employment earnings in the same way unincentivized individuals do in the absence of the incentives. However, we note that this does not necessarily correspond to a "true" self-employment rate. While it could be the case that incentivized individuals are fabricating self-employment income, it could alternatively be the case that individuals facing positive tax rates incentives under-report their self-employment income and, once incentivized, households reduce their underreporting.

6.1 Additional Quantification Exercises

In addition to accounting for incentives in our adjustments, we also consider two additional factors that may be quantitatively important for explaining observed SE trends: changes in the demographic composition of the workforce and changes in the propensity for under-reporting.

As documented in Collins, Garin, Jackson, Koustas, and Payne (2019), in all years since 2000, older workers—particularly retirement-age workers over 65—are dramatically more likely to have self-employment income and nonemployee compensation reported on 1099 forms conditional on having any labor income. Accordingly the aging of the workforce could result in higher overall self-employment rates in our baseline measure. To examine the role that the changing demographic composition of the workforce might contribute to the remaining observed rise in self-employment reporting since 2000, we calculate what the results of the Scenario 2 adjustment would be if the demographic composition of the workforce remained fixed at 2000 levels. We do this by holding constant the weights of our demographic groups at their 2000 levels:

$$SE_t^* = \sum_g \omega_{g,\mathbf{2000}} \sum_k \left(\sum_{w \leq w^*(k)} \zeta^*_{\mathbf{w},\mathbf{g},\mathbf{k},\mathbf{t}} SE_{w,g,\mathbf{0},t} + \sum_{w > w^*(k)} \zeta^*_{\mathbf{w},\mathbf{g},\mathbf{k},\mathbf{t}} SE_{w,g,k,t} \right) \quad (8)$$

This adjustment is presented as "Scenario 3" in Figure 8a. Adjusting for both demographic shifts and for differential reporting trends among incentivized individuals accounts for a large majority of the observed rise in self-employment reporting since

2000. This adjustment accounts for nearly 80% of the observed increase since 2000. Less than one half of a percentage point increase since 2000 remains unexplained.[26] Figure 8b reports the results of the same exercise of 1099-reported independent contractor income. The added demographic adjustment similarly accounts for about half of the increase in the share of the workforce with 1099-reported independent contractor income prior to the rise of the Online Platform Economy.

As a final adjustment, we assess the concern raised in Section 3 that trends in SE reporting may not line up with trends in the propensity to have non-employee income on 1099 forms due to changes in accurate reporting of non-employee compensation on Schedule C or changes in the composition of 1099 work. Changes in compliance along these lines might directly contribute to the observed rise in SE reporting in excess of changes in the prevalence of 1099-reported contract work. Note, however, that our adjustments above should already take into account changes in the propensity to report a 1099 due to the incentives we consider.

Our final adjustment accounts for such changes. To estimate changes in 1099-reporting over time, we examine the filing behavior of 1099 recipients. We run the following regression on the population of 1099 recipients who file Schedule C:

$$\mathbb{1}\{FilesSE\}_i^{t,p} = \gamma_{s\times r}^{t,p} + \epsilon_i^{t,p} \tag{9}$$

where s denotes state and r denotes a bin for the dollar value of 1099 receipts. Hence, $\gamma_{s\times r}^{t}$ report the share of individuals in a bin of non-employee compensation in a state, who file SE in year t. We run this regression separately for OPE and non-OPE 1099s (denoted p) in each year, and only among our unincentivized population, so that the phenomenon that we capture will be independent from our earlier adjustments. In a

[26] Appendix Table A.1 additionally presents trends if the age and gender composition had remained at 2000 levels given observed trends within age-by-sex cells, without any adjustment for incentives. We note that the Scenario 3 adjustment simultaneously accounting for demographic change and credit-incentivized behavior has larger impacts than adjusting for both incentivized reporting and demographics separately. This is because the tax incentives to report self-employment are concentrated among prime-age workers; as a result, once we adjust to account for the aging of the workforce, the former adjustment becomes quantitatively more important.

second step, we calculate out-of-sample predicted values in each year for individuals who do not file Schedule C. We sum these values to get the predicted share of non-C filers who should file Schedule SE. Appendix Figure A.13 reports the implied share of the workforce with non-employee compensation who are underreporting SE. We find our measure of SE underreporing is declining through 2009, thereby implying increasing rates of SE filing among 1099 recipients in these years; since then, underreporting has been increasing, reaching its highest levels in recent years amidst the rise in OPE work.

Our final SE adjustment incorporating these trends is presented as Scenario 4 in Figure 8a. After taking into account changes in 1099-reporting, early SE growth declines further. We also see that SE rates have remained relatively constant since 2014, rather than declining as a share of the workforce as we saw in the baseline. In the end, we find new online platform only makes a modest contribution to SE growth in recent years, given the small dollar amounts and high rates of expensing, as documented in Collins, Garin, Jackson, Koustas, and Payne (2019).

6.2 Application: Response of Self-Employment to the Business Cycle

To illustrate the importance of adjusting for reporting behavior, we close by examining the relationship between self-employment reported on tax returns and the state of the aggregate economy. One question studied in the self-employment literature is whether people turn to self-employment to smooth economic shocks. While unemployed and underemployed workers may seek additional opportunities, at the same time, returns to self-employment may be lower during aggregate recessions, reducing the value of this outside option. This question has mainly been studied using survey data such as the CPS, with recent attention placed on alternative work arrangements and independent contracting.[27]

[27]Parker (2009) provides a review of the literature, and Fairlie (2013) provides a more recent update examining the Great Recession. Farber (1999) and Katz and Krueger (2017) focus on alternative work as captured in the Contingent Worker Supplement (CWS) to the CPS.

It is well known that measurement error in survey data will lead to imprecise estimates; survey data may also be biased if it is not picking up changes in freelance/contract work. However, as we have highlighted above, the number of individuals facing incentives to report self-employment to increase tax refunds is also countercyclical. As shown in Appendix Figure A.8, as the wage distribution shifts down, more families have wages fall into the credit phase-in range.

Our goal is to separate out reporting-effect from real changes in self-employment in response to recessions using our adjusted series. To examine the response of self-employment to the business cycle, we follow a standard approach in this literature and exploit variation in local labor market conditions. We estimate the following specification:

$$y_{s,t} = \beta UR_{s,t} + \alpha_s + \alpha_t + e_{s,t} \tag{10}$$

where $UR_{s,t}$ is the average unemployment rate in state s in year t. We include state and year fixed effects, so our main coefficient of interest is identified off of panel variation across our time period of study, 2000-2018. The outcome $y_{s,t}$ is a measure of the state-level prevalence of self-employment based on either on raw self-employment reporting on tax returns or on the series adjusting for reporting trends in Scenario 2 above. For comparison, we also construct analogous outcomes using survey responses on the CPS-ASEC.

We first examine how self-employment evolves around the business cycle as a share of the workforce with labor income in Table 4a. We find that while the share of workers with self-employment income is significantly counter-cyclical in raw tax data, no such association exists after adjusting for reporting trends, nor does any such association exist in third-party reported 1099 data or in survey data. Column (1) shows that the baseline self-employment share of the workforce derived from tax data increases in response increases in unemployment. To put the coefficient into context, the unemployment rate increased by about 5 percentage points during the Great Recession,

so our estimated coefficient implies an increase in self-employment as a share of the workforce of (0.144*5 =) 0.72 p.p., or about (0.72/10.9=) 6.6%. In contrast, we see no response of self-employment when using our adjusted series in Column (2). This is consistent with what we find using the 1099s or using self-employment from the CPS, shown in Columns (3)-(4), respectively.

The analysis above focused on self-employment as a share of the workforce. As an alternative approach, we examine the how the prevalence of each type of work in the broader adult population varies over the business cycle in Table 4b.[28] The interpretation of our coefficients is now the log point change in employment-to-population ratio (approximately a percent change) for a percentage point increase in the unemployment rate. (Note: the denominator of the dependent variable is the same across these specifications, only the numerator is changing.)

We find that while the share of adults with wage income reported on W2s (Column 1) or contractor compensation reported on 1099s (Column 5) is significantly procyclical, the share of adults reporting self-employment income (Column 3) is constant around the business cycle. However, after adjusting for changes in reporting behavior (Column 4)—which increase self-employment reporting during recessions—we estimate that self-employment actually moves in close proportion to work reported on W2s and and 1099s. Further, the coefficients for the adjusted series are similar to effects on both employment rates and self-employment measured in the CPS-ASEC. One importance difference compared with using the CPS is that the tax outcomes are more precisely estimated.[29]

These exercises point to the importance of accounting for time-varying reporting behaviors when using self-reported self-employment measures in administrative data. Further, we believe that they demonstrate the value of our new adjusted SE series and 1099 series. By carefully considering the incentives created by the tax code, it is possible to recover useful estimates of underlying work behavior from self-reported

[28]We measure the adult population in each year and state using the CPS-ASEC.

[29]The point estimate on self-employment in the CPS-ASEC is imprecisely estimated, consistent with classical measurement error in a dependent variable.

information on tax filings.

7 Conclusions

Taken together, we highlight that new types of online platform work, or "gig" work, cannot explain increasing trends in self-employment in administrative data. Instead, we find that a substantial portion of the growth is driven by strategic reporting behavior. More precisely, we find an increase among low income individuals with kids—a group that faces negative marginal tax rates if they report self-employment income. Our regression discontinuity analysis comparing households with births in December versus January indicates that this increase is driven by changes in pure reporting behavior rather than an underlying labor response. When we consider counterfactual scenarios in which reporting behavior remained constant at 2000 level, we find that between 25 and 55 percent of the average increase in self-employment rates since 2000 can be attributed to pure reporting changes. Once we additionally adjust for the aging of the workforce and other trends in reporting, we account for nearly all the observed change in self-employment in administrative tax data.

By employing an RDD methodology comparing households with births in December versus January, of the following tax year, we find a significant increase in self-reported but not firm-reported self-employment among December births relative to January births. This well identified RDD specification leads us to believe that this is driven by changes simply in reporting behavior rather than an underlying labor response. We apply our estimates of the reporting response to provide an adjusted self-employment trend.

Our framework may be important going forward to interpret changes in self-employment as tax policy continues to evolve. For instance, in the latest relief bill the American Rescue Plan Act of 2021, the phase-in rate for childless adults will now be 15.3 percent for 2021, up from 7.65 percent, completely covering the payroll tax; thus, childless adults now face reporting incentives in states and territories that top-up the federal

EITC credit. Furthermore, our adjustments have also highlighted the impact that demographic shifts within the population over time (e.g. declining birth rates, aging population, etc.) have had on the self-employment share of the workforce. Similarly, business cycles play an important role in fluctuations in the self-employment share by increasing or decreasing the size of the group incentivized by negative marginal tax rates.

Overall, while we find that administrative data on self-employment suffer from reporting incentives, we believe that administrative data can still be highly valuable for measuring labor-market trends so long as reporting incentives are kept in mind. We show that these incentives are measurable and can be accounted for with our framework. Moreover, our work highlights that even while self-reported earnings are sensitive to reporting incentives, third-party income reported by firms provides a valuable benchmark. Despite not driving the diverging self-employment trends between survey and administrative data, increases in gig work and secondary forms of contract work are being well captured in firm-reported administrative data. To this end, our new self-employment series adjusted for reporting trends, as well as our new series on gig work, should be valuable to other researchers in this area.

References

Abraham, K., J. Haltiwanger, K. Sandusky, and J. Spletzer (2013): "Exploring Differences in Employment between Household and Establishment Data," *Journal of Labor Economics*, 31(2), S129–S172.

——— (2018): "Driving the Gig Economy," .

——— (2020): *Measuring the Gig Economy: Current Knowledge and Open Issues,*University of Chicago Press.

Abraham, K. G., J. C. Haltiwanger, C. Hou, K. Sandusky, and J. R.

Spletzer (Forthcoming): "Reconciling Survey and Administrative Measures of Self-Employment," *Journal of Labor Economics*.

Abraham, K. G., B. J. Hershbein, and S. N. Houseman (2020): "Most Self-Employed Workers Are Independent Contractors," *Policy Brief, W.E. Upjohn Institute for Employment Research*.

Blank, R. M., K. K. Charles, and J. M. Sallee (2009): "A Cautionary Tale about the Use of Administrative Data: Evidence from Age of Marriage Laws," *American Economic Journal: Applied Economics*, 1(2), 128–49.

Bracha, A., and M. Burke (2018): "How Big is the Gig?," *Working Paper*.

Bureau of Labor Statistics (2018a): "Contingent and Alternative Employment Arrangements - May 2017," USDL-18-0942.

——— (2018b): "Contingent and Alternative Employment Arrangements Technical Note," Available at https://www.bls.gov/news.release/conemp.tn.htm.

Chetty, R., J. N. Friedman, P. Ganong, K. E. Leibel, A. H. Plumley, and E. Saez (2012): "Taxpayer Response to the EITC: Evidence from IRS National Research Program," Tables available at: www.rajchetty.com/chettyfiles/eitc_nrp_tabs.pdf.

Chetty, R., J. N. Friedman, and E. Saez (2013): "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings," *American Economic Review*, 103(7), 2683–2721.

Collins, B., A. Garin, E. Jackson, D. Koustas, and M. Payne (2019): "Is Gig Work Replacing Traditional Employment? Evidence from Two Decades of Tax Returns," *SOI Working Paper*.

Davis, S. J., J. Haltiwanger, R. S. Jarmin, C. Krizan, J. Miranda, A. Nucci, and K. Sandusky (2009): *Measuring the Dynamics of Young and Small Busi-*

*nesses: Integrating the Employer and Nonemployer Universes*pp. 329–366. University of Chicago Press.

Fairlie, R. W. (2013): "Entrepreneurship, Economic Conditions, and the Great Recession," *Journal of Economics & Management Strategy*, 22(2), 207–231.

Fairlie, R. W., J. Miranda, and N. Zolas (2019): "Measuring Job Creation, Growth, and Survival among the Universe of Start-ups in the United States Using a Combined Start-up Panel Data Set," *ILR Review*, 72(5), 1262–1277.

Farber, H. S. (1999): "Alternative and Part-Time Employment Arrangements as a Response to Job Loss," *Journal of Labor Economics*, 17(4), S142–S169.

Farrell, D., and F. Greig (2018): "The Online Platform Economy in 2018: Drivers, Workers, Sellers, and Lessors," *JPMorgan Chase Institute*.

Harris, S. D., and A. B. Krueger (2015): "A Proposal for Modernizing Labor Laws for Twenty-First-Century Work: The "Independent Worker"," *The Hamilton Project Discussion Paper 2015-10*.

Hurst, E., and B. Pugsley (2011): "What Do Small Businesses Do?," *Brookings Papers on Economic Activity*, 42(2 (Fall)), 73–142.

Jackson, E., A. Looney, and S. Ramnath (2017): "The Rise of Alternative Work Arrangements: Evidence and Implications for Tax Filing and Benefit Coverage," *Office of Tax Analysis Working Paper 114*.

Katz, L., and A. Krueger (2019): "The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015," *ILR Review*, 72(2), 382–416.

Katz, L. F., and A. B. Krueger (2017): "The Role of Unemployment in the Rise in Alternative Work Arrangements," *American Economic Review*, 107(5), 388–92.

Kuka, E. (2014): "EITC and the Self-employed: Real or Reporting Effects?," *Public Finance Review*, 42(6), 691–719.

LaLumia, S. (2009): "The Earned Income Tax Credit and Reported Self-Employment Income," *National Tax Journal*, 62(2), 191–217.

LaLumia, S., J. M. Sallee, and N. Turner (2014): "New Evidence on Taxes and the Timing of Birth," *American Economic Journal: Economic Policy*, 7(2), 258–293.

Lim, K., A. Miller, M. Risch, and E. Wilking (2019): "Independent Contractors in the U.S.: New Trends from 15 years of Administrative Tax Data," *SOI Working Paper*.

Mortenson, J. A., and A. Whitten (2020): "Bunching to Maximize Tax Credits: Evidence from Kinks in the US Tax Schedule," *American Economic Journal: Economic Policy*, 12(3), 402–32.

National Academies of Sciences, E., and Medicine (2020): *Measuring Alternative Work Arrangements for Research and Policy*. The National Academies Press, Washington, DC.

Parker, S. C. (2009): "Entrepreneurship, Self-employment and the Labour Market," *The Oxford Handbook of Entrepreneurship*.

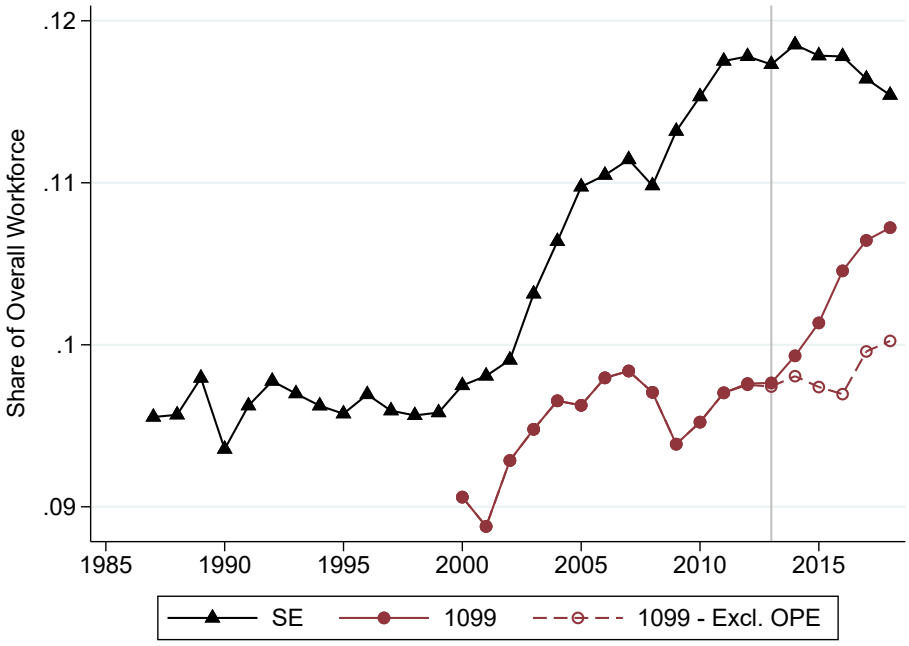
Saez, E. (2010): "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2(3), 180–212.

Slemrod, J. (2016): "The EITC and self-employment among married mothers," *National Tax Journal*, 69(4), 1003–1020.

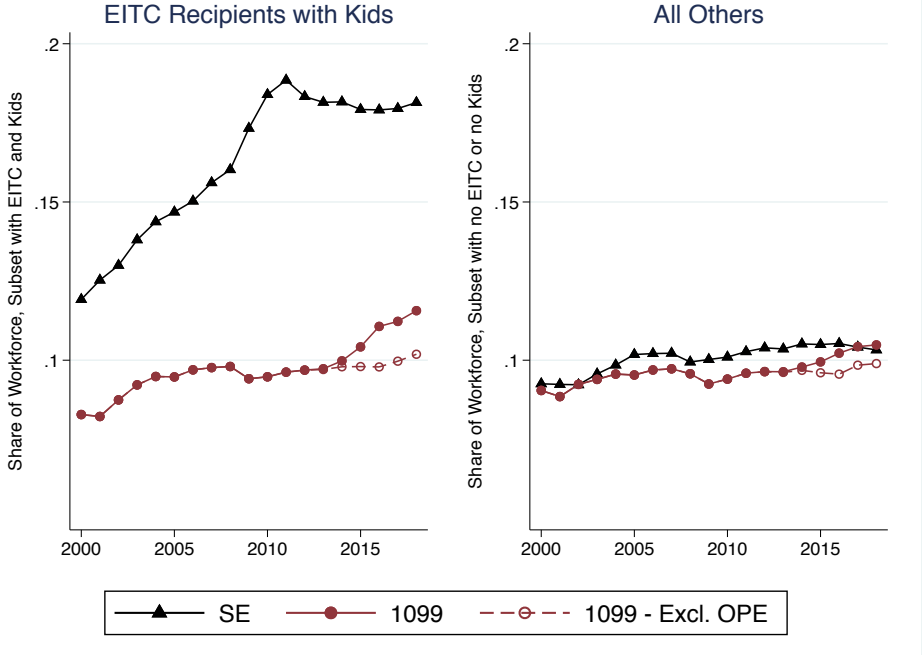
Figures

Figure 1: Share of Workforce with Self-Employment and 1099 Information Returns

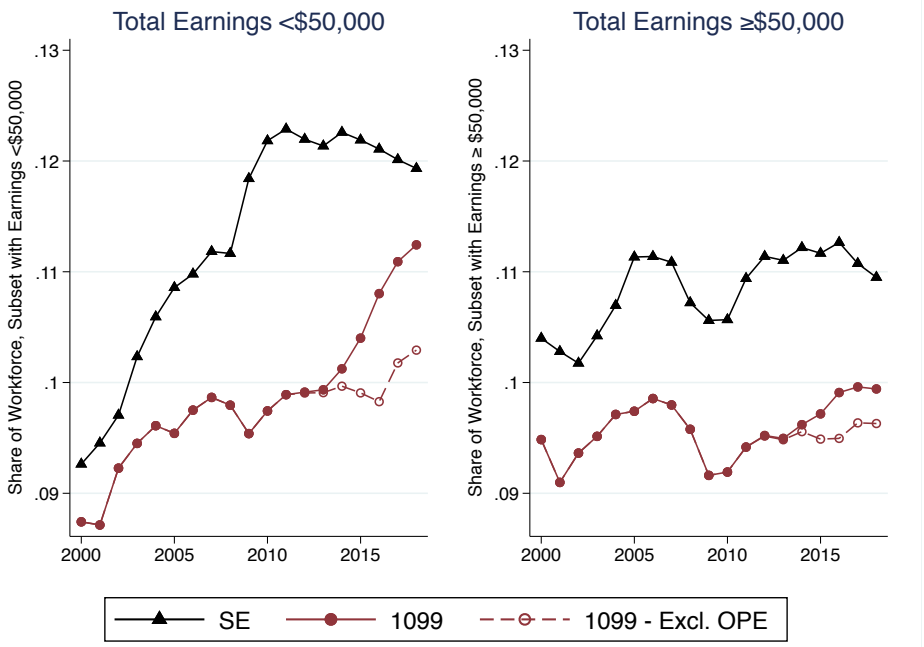
(a) Share of Overall Workforce



(b) By EITC and Kids



(c) By Total Earnings



Notes: Figure shows the share of the overall tax workforce by tax year with any SE income as filed on Schedule SE (black line) and individuals who receive a 1099 Information Return (maroon line). After the entry of OPE, we additionally distinguish the receipt of 1099 Information Returns including and excluding those received from OPE firms (dashed maroon line). In Panel (a), For years prior to 2000, SE workers and the overall workforce are drawn from annual SSA records on workers: https://www.ssa.gov/policy/docs/statcomps/supplement/2020/4b.pdf. In Panel (b), the workforce definition is restricted to EITC recipients with kids claimed on their 1040. In panel (c) Total earnings refers to the sum of wage and self-employment income by a primary tax filer and their spouse as reported on a 1040.

Figure 2: Growth Over 2000-2014 in Self-Employment, by Wage Income

(a) Any SE Filing

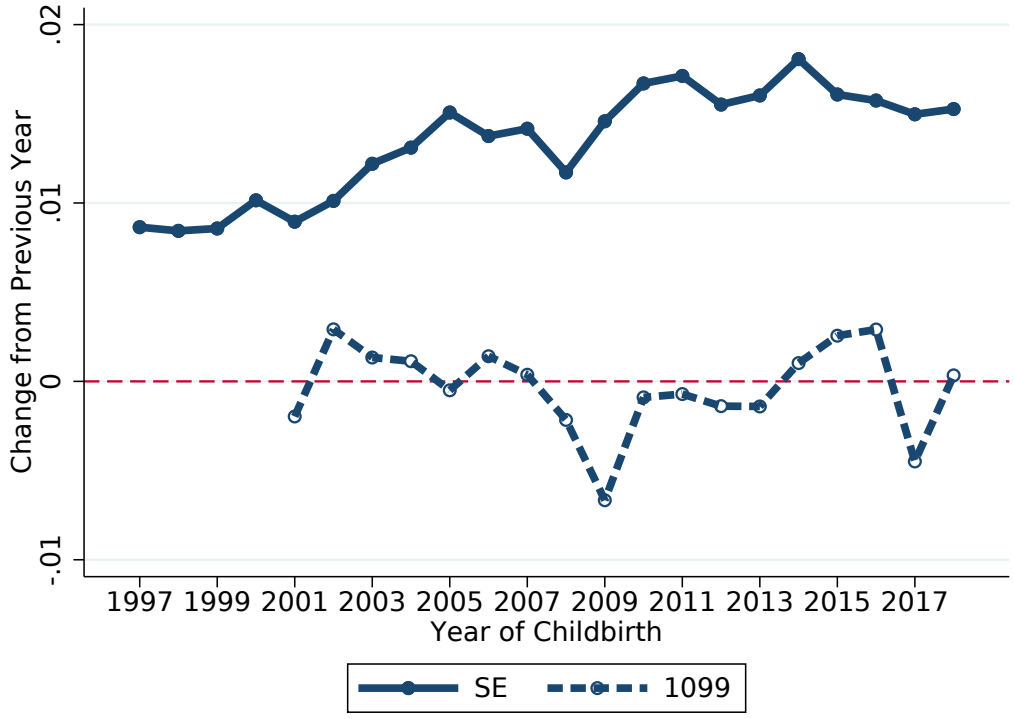


(b) Any 1099



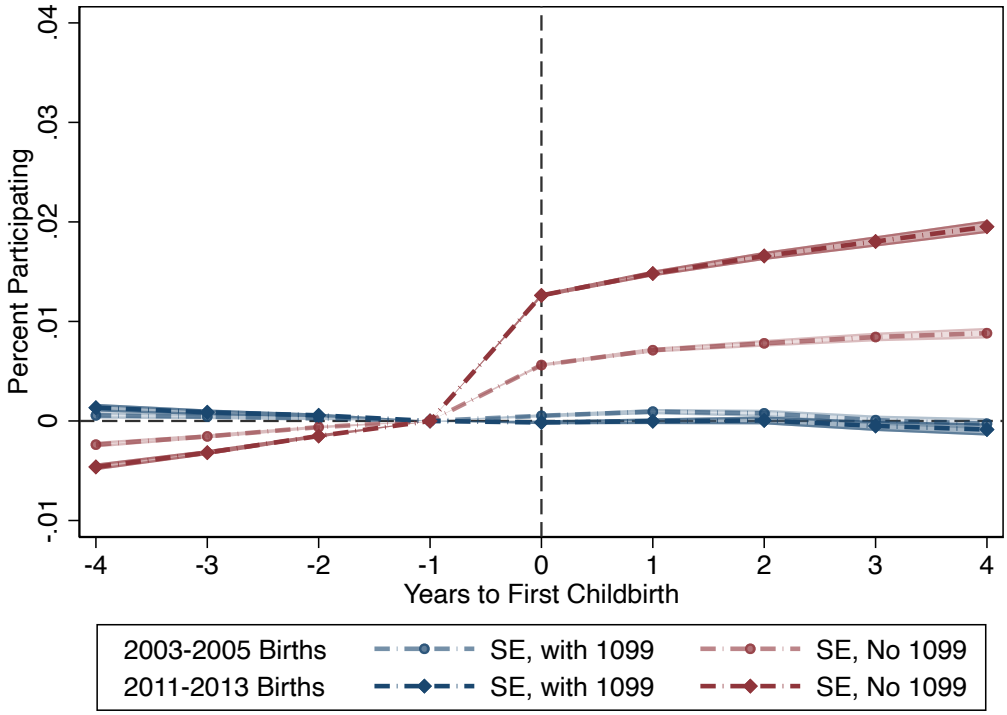
Notes: Figure shows the change in propensity to file SE (Panel A) or receive a 1099 Information Return (Panel B), for tax units with wage earnings. Wage earnings, in $500 bins, are reported on the x-axis. Change is calculated between 2000 and 2014. Wages are determined based on W2 information returns. Number of children and spouse as reported on tax return. The area to the left of the vertical lines report the earnings where reporting an additional dollar of self-employment income would face a negative marginal tax rate due to the EITC, for households with 1 (maroon line) and 2 (green line) children.

Figure 3: Change in Self-Employment Around First Childbirth, 1997-2018 Births



Notes: Figure shows the average change in propensity to file SE (solid line) or receive a 1099 Information Return (dashed line), in the year of first childbirth reported on the x-axis.

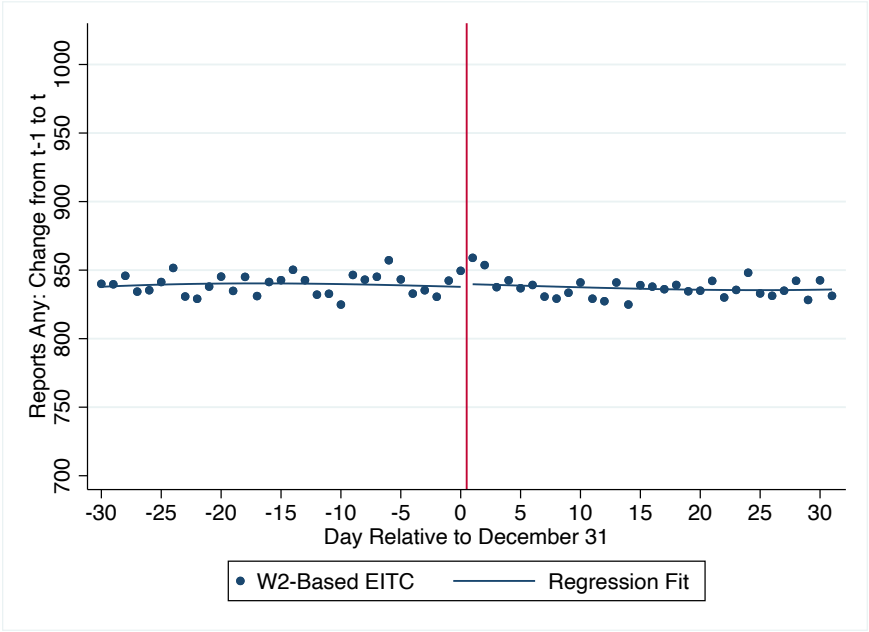
Figure 4: Childbirth Event Study Estimates, 2003-2005 births versus 2011-2013 births



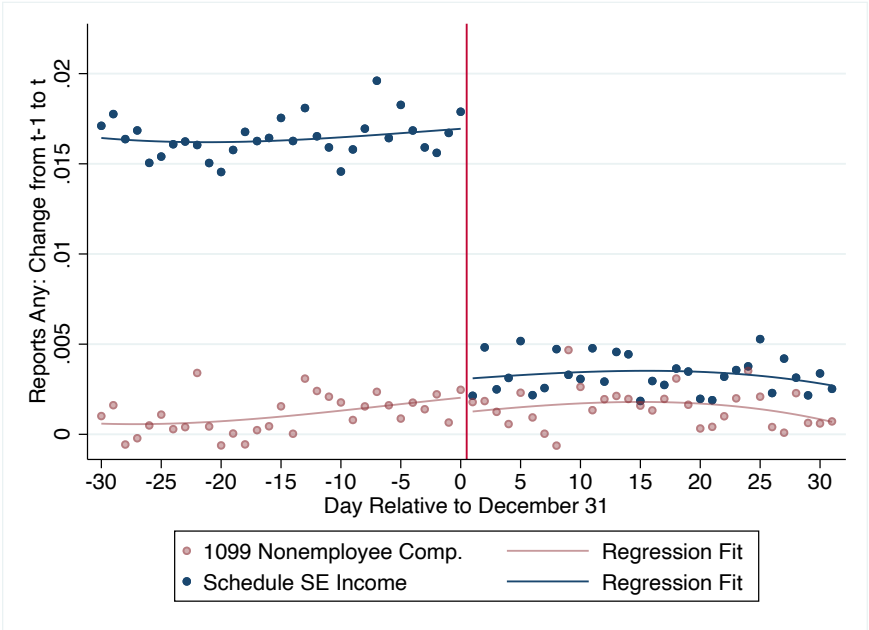
Notes: Figure plots event study coefficients for separate regressions run on the indicated time-period and for the indicated outcome. See text for more details.

Figure 5: Regression Discontinuity Design

(a) Simulated EITC Based on W2



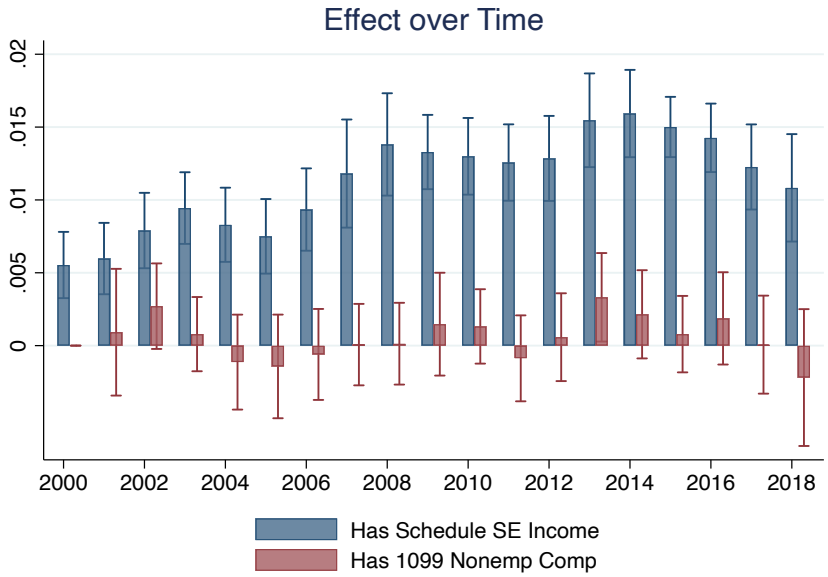
(b) Change in SE and 1099 Reporting



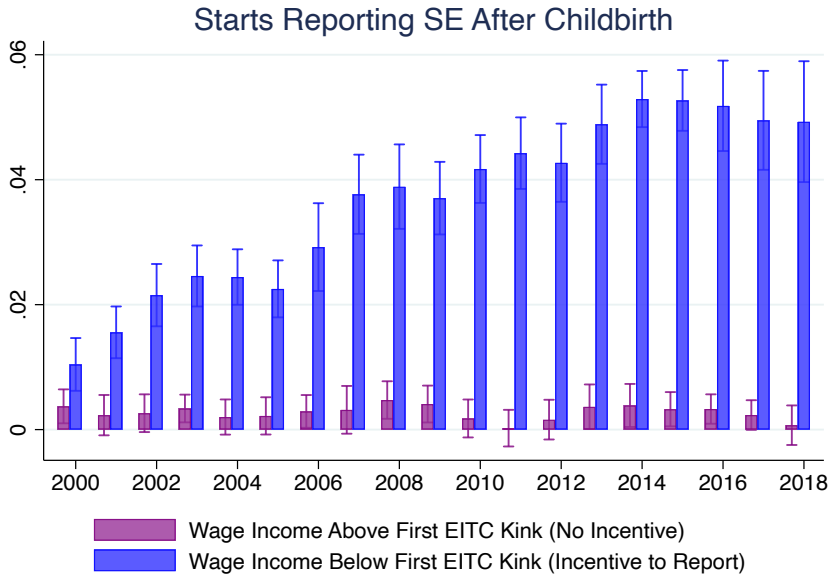
Notes: Figure graphically presents results from the regression discontinuity design, pooling births in each December 2011-2018 and each subsequent January. Panel (a) reports the change in "simulated EITC" that is the credit each individual would earn if their birth occurred in December of tax year t based on their and their spouse's (if married filing jointly) W-2 wages from year t. Panel (b) shows the change in any Schedule SE earnings in tax year t relative to the prior year $t-1$, displayed in blue, and the change in having non-employee income reported on a 1099-MISC in tax year t relative to the prior year $t-1$, displayed in red. Each dot is the average outcome for parents with births on the corresponding date, pooled across years. The curves are the regression fits allowing for the estimated discontinuity between December 31 and January 1.

Figure 6: RDD Estimates by Year

(a) Pooled Effects



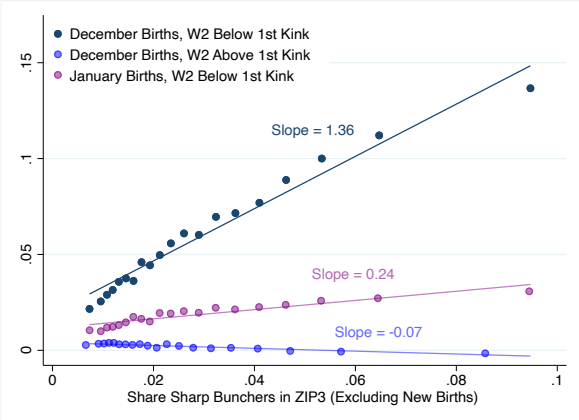
(b) Wages Above/Below 1st EITC Kink



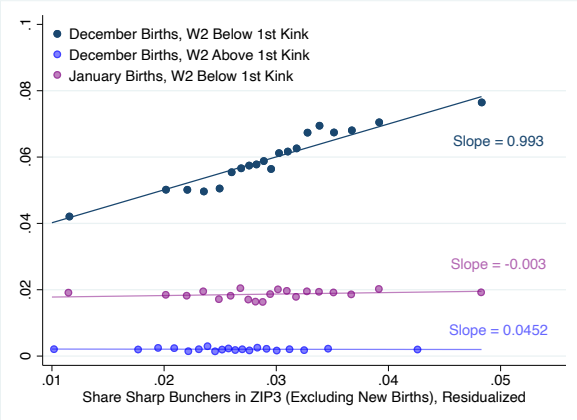
Notes: Panel A reports our baseline RDD estimates from estimating Specification 2 in the text. In Panel B, we contrast behavior among individual with combined wages below the first EITC kink and those with wages above the first EITC kink point.

Figure 7: Self-Employment Reporting and Changes in Local EITC Knowledge

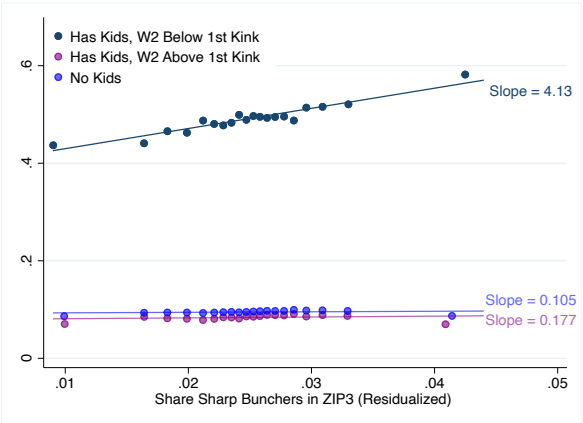
(a) Change in SE for First Births at End of Tax Year, Raw



(b) Change in SE for First Births at End of Tax Year, Within-ZIP Variation in Knowledge



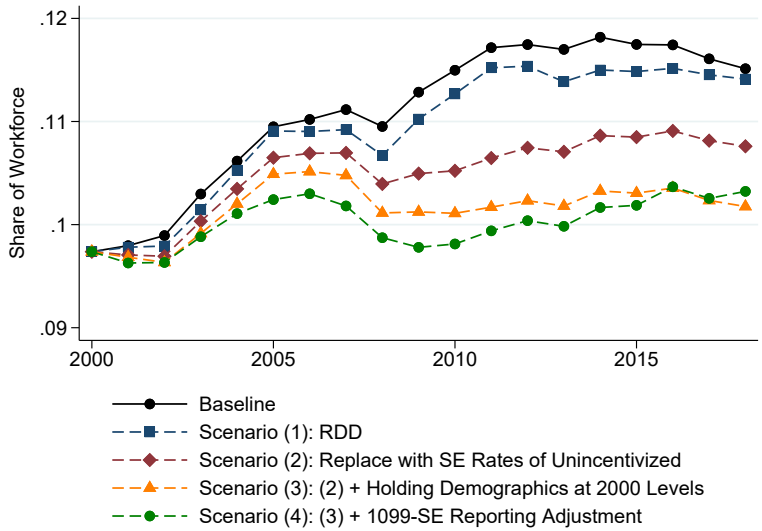
(c) Share of Total Workforce Reporting SE, Within-ZIP Variation in Knowledge



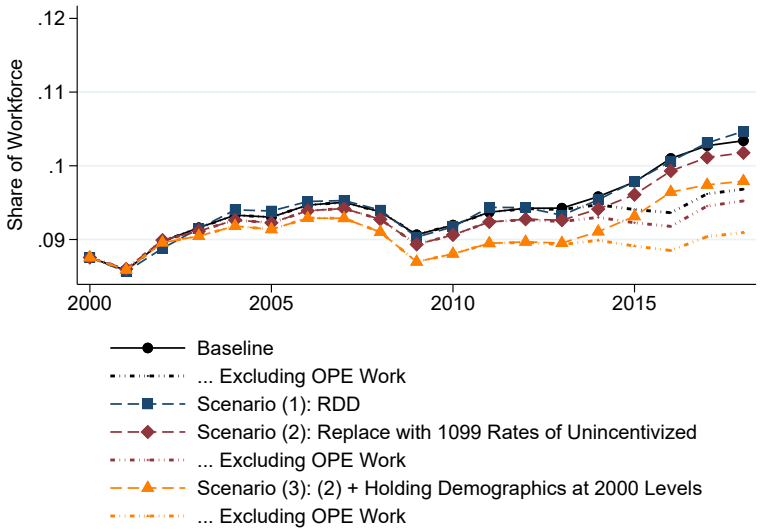
Notes: Figure (a) displays the reduced-form relationship between changes in self-reported self-employment status in year t for parents with first births near the end of the year, plotted against the year t sharp bunching rate among other taxpayers in the same ZIP3. Figure (b) displays the same relationship residualized on tax-year and ZIP3 fixed effects to control for permanent cross-sectional differences across regions and for common trends. Figure (c) plots the relationship between the share of the specified tax workforce segment reporting self-employment against the same sharp bunching rate in the above panels, residualized as in Panel B. Figures are binned scatter plots of raw data or residuals, as specified, with associated best fit lines.

Figure 8: Adjusted SE Shares Under Counterfactual Assumptions

(a) Adjusted Share of Workforce Filing Self-Employment



(b) Adjusted Share of Workforce with Nonemployee Compensation



Notes: Figure presents the baseline and implied trend in SE share of the overall tax workforce under counterfactual scenarios. Panel (a) applies these adjustments to self-employment, and Panel (b) to nonemployee compensation. "Scenario (1)" adjusts self-employment downward according using our annual RD estimates reported in Figure 5a. "Scenario (2)" replaces SE rates for individuals who have incentives to report SE with the rates among comparable individuals without this incentive. "Scenario (3)" accounts for demographic shifts over this time frame by holding the composition of the workforce constant at its 2000 level. "Scenario (4)" applies an adjustment for underreporting of independent contracting income, which is only applicable to Panel (a). Dashed lines in Panel (b) exclude the OPE.

Tables

Table 1: Baseline RDD Estimates

A: Third-Party Reported Earnings

Δ Any Wages	Δ Tax-Unit Wages	Δ W-2 Simulated EITC	Δ Any 1099	Δ **1** {Wages $< 1^{st}$ Kink}
(1)	(2)	(3)	(4)	(5)
-0.0007	30.2675	-2.1524	0.0008	-0.0077
(0.0012)	(226.3170)	(6.1099)	(0.0008)	(0.0064)
3,585,425	3,585,425	3,585,425	3,585,425	3,585,425

B: Incentives

Marginal Tax Rate < 0			Δ Marginal Tax Rate		
All	Wages $< 1^{st}$ Kink	Wages $\geq 1^{st}$ Kink	All	Wages $< 1^{st}$ Kink	Wages $\geq 1^{st}$ Kink
(1)	(2)	(3)	(4)	(5)	(6)
0.2484	0.9896	0.0002	-5.9138	-30.0722	1.9962
(0.0043)	(0.0013)	(0.0001)	(0.2424)	(0.1167)	(0.1931)
3,585,425	906,038	2,679,387	3,585,425	906,038	2,679,387

C: Self-Employment Reporting

Δ Any SE			Δ SE Amount	EITC Gain From SE ($)		
All	Wages $< 1^{st}$ Kink	Wages $\geq 1^{st}$ Kink		All	Wages $< 1^{st}$ Kink	Wages $\geq 1^{st}$ Kink
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0.0139	0.0486	0.0024	144.5037	69.7314	259.6340	6.4020
0.0008	0.0019	(0.0008)	(7.1802)	(4.2773)	(9.2667)	(4.1392)
3,585,425	906,038	2,679,387	3,585,425	3,585,425	906,038	2,679,387

D: Margins of Self-Employment

Δ Only SE	Δ Sharp Buncher	Δ SE Classified as Personal Services	Δ SE w/ Sched C Expenses
(1)	(2)	(3)	(4)
0.0041	0.0023	0.0037	0.0068
(0.0004)	(0.0003)	(0.0003)	(0.0033)
3,585,425	3,585,425	3,585,425	3,585,425

Notes: Table displays estimates from the baseline regression discontinuity design specification in Equation (2). The sample is all individuals with births in the December of each tax year t in 2011-2018 or the January immediately following tax year t. Outcomes are from year t or are changes from year t relative to the prior year $t-1$, as specified. "Wages $<$ 1st kink" subsample includes all individuals in tax-units with year t wages in the EITC phase-in region for households with one child in that year (regardless of whether their birth actually occurred in December or January); the complementary subsample includes all other individuals. Standard errors are clustered by date.

Table 2: RDD Estimates Interacted with Knowledge

	Δ Any SE	Δ Any 1099	Δ Any Wages	Δ W-2 Simulated EITC	Δ **1** {Wages $< 1^{st}$ Kink}	Δ Any SE	
						Wages $< 1^{st}$ Kink	Wages $\geq 1^{st}$ Kink
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ZIP3 Bunching	.2666	.0565	-.0178	74.36	.092	.9192	-.0162
	(.0274)	(.0702)	(.061)	(141.0)	(.0616)	(.0741)	(.0258)
RDD Main Effect	.0041	-.0013	.0007	5.701	-.0083	.0162	.003
	(.0008)	(.0045)	(.0033)	(4.696)	(.004)	(.0022)	(.0008)
N	8,015,998	8,015,998	8,015,998	8,015,998	8,015,998	1,724,421	6,291,577

Notes: Table displays estimates from the regression discontinuity specification interacted with local sharp-bunching rates in in Equation (3) in the text. The sample is all individuals with births in the December of each tax year t in 2001-2018 or the January immediately following tax year t with a valid ZIP code reported on an information return or tax filing. "Wages $<$ 1st kink" subsample includes all individuals in tax units with year t wages in the EITC phase in region for households with one child in that year (regardless of whether their birth actually occurred in December or January); the complementary subsample includes all other individuals. Specification is fully interacted with the sharp bunching rate among individuals without first births in tax year t in the individual's 3-digit ZIP code. Standard errors are clustered by date.

Table 3: The Role of Knowledge

	Births in December of Year t					Births in January of t+1
	Wages $< 1^{st}$ Kink				Wages $\geq 1^{st}$ Kink	Wages $< 1^{st}$ Kink
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Any SE Income	1.363	1.331	1.265	.9929	-.0033	.0453
	(.0783)	(.0909)	(.0896)	(.1107)	(.0214)	(.0496)
N	33,205	33,205	33,205	33,205	34,411	33,113
Δ Any 1099 Nonemployee Income	.0186	.1628	-.474	.0768	-.0161	.0935
	(.1197)	(.0456)	(.4643)	(.0786)	(.0156)	(.0619)
N	33,205	33,205	33,205	33,205	34,411	33,113
Year t FE		✓		✓	✓	✓
ZIP3 FE			✓	✓	✓	✓

Notes: Table displays estimates from regressions of average outcomes of individuals with first births at the end of each tax year—either in December or the subsequent January as specified—in each 3-digit ZIP code on the sharp bunching rate as in Equation (4). Bunching rates are calculated from individuals without first births in tax year t in the individual's 3-digit ZIP code. Observations are at the ZIP code by year level, and are weighted by the number of individuals in the specified group. Standard errors are two-way clustered by ZIP code and year.

Table 4: Response of Self-Employment to the Business Cycle Under Various Measures

(a) Share of Workforce

	Tax: Raw SE	Incentives-Adjusted	1099	ASEC: SE
	(1)	(2)	(3)	(4)
Unemployment Rate	0.144	0.0253	0.0119	-0.0151
	(0.0304)	(0.0176)	(0.0245)	(0.0514)
N	969	969	969	969
R^2	0.894	0.936	0.893	0.764
Dep. Mean	0.109	0.098	0.096	0.076
State FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

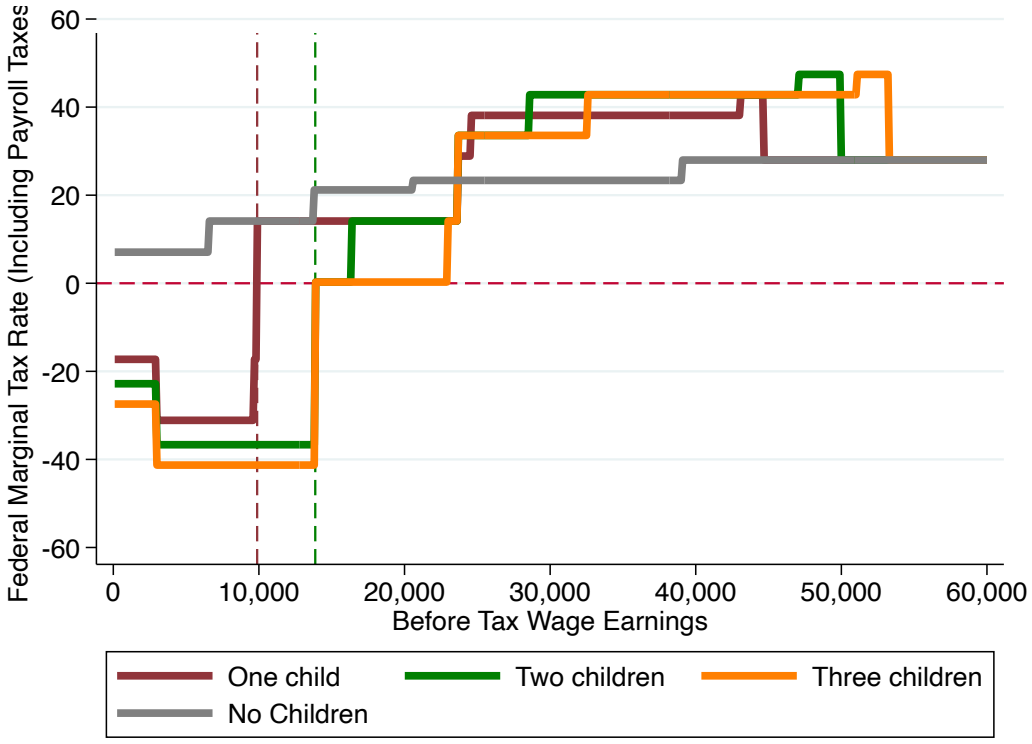
(b) Log employment-to-adult population

	Tax: Wage	ASEC: Wage	Tax: SE	Incentives-Adj	1099	ASEC: SE
	(1)	(2)	(3)	(4)	(5)	(6)
Unemployment Rate	-1.259	-0.749	0.278	-0.929	-0.950	-1.156
	(0.156)	(0.151)	(0.246)	(0.188)	(0.251)	(0.714)
N	969	969	969	969	969	969
R^2	0.919	0.904	0.924	0.958	0.914	0.805
State FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓

Notes: Table reports the results from running specification (10) in the text for the dependent variable and source indicated by table subtitles and column headers. Standard errors clustered on state in parentheses.

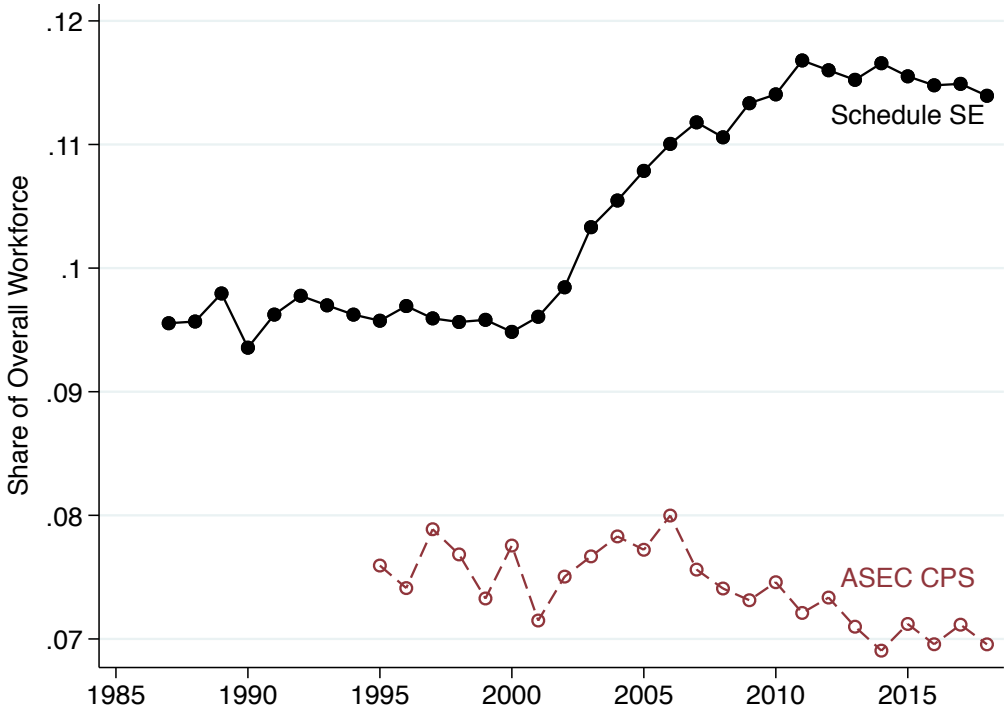
A Appendix Figures

Figure A.1: Effective Federal Marginal Tax Rate For Reporting Additional Dollar of Self-Employment Income



Notes: Figure shows the effective marginal tax rate for reporting an additional dollar of self-employment income, for a given level of before-tax wage earnings. Calculation takes into account the full tax schedule in tax year 2015 with no other credits/deductions except the EITC, CTC and standard deductions, and assumes Schedule SE payroll taxes are paid on the self-employment income and taxpayers deduct the employer-share of the payroll tax on self-employment income. Calculation assumes married filing jointly, however the marginal tax rates below the first kink point are identical for married and single parents who claim children. The area to the left of the vertical lines indicate the first kink-point of the EITC schedule, for households with 1 (maroon line) and 2 or more (green line) children.

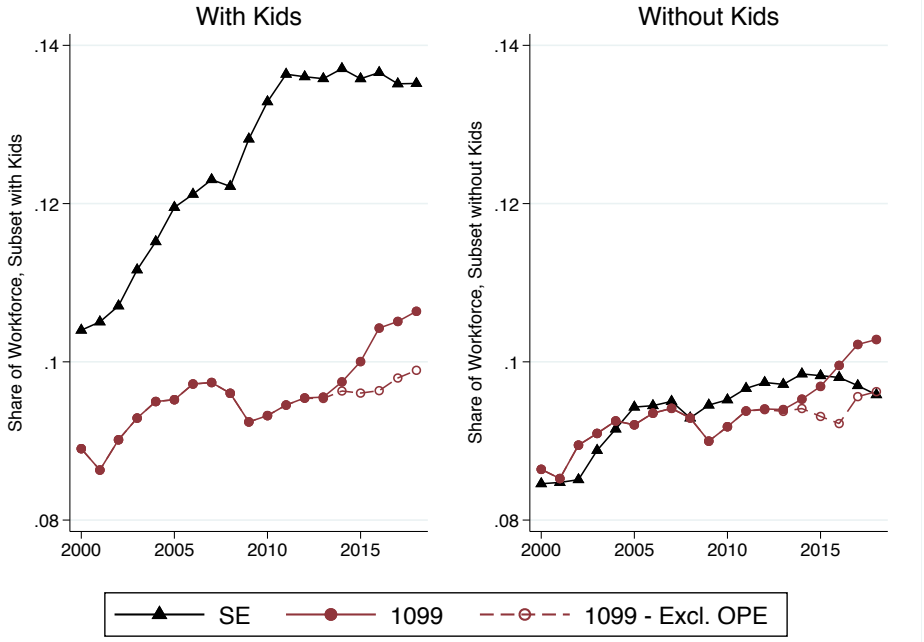
Figure A.2: Share of Workforce with Self-Employment, in Tax Returns and CPS ASEC



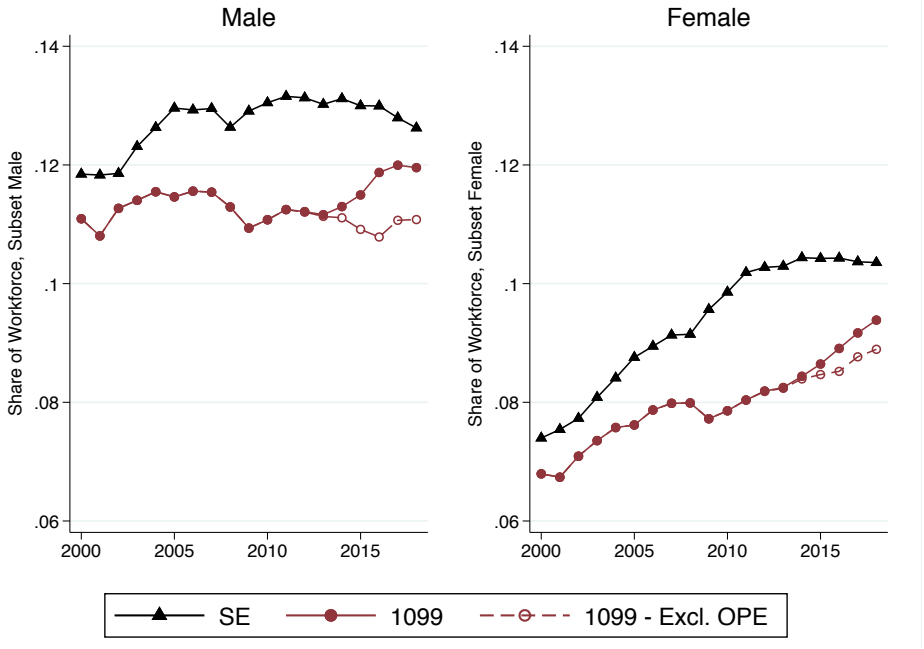
Notes: Figure shows the share of the overall tax workforce by tax year with any SE income as filed on Schedule SE (black line) and individuals who report self-employment in the CPS ASEC supplement (maroon line). For years prior to 2000, SE workers and the overall workforce are drawn from annual SSA records on workers: https://www.ssa.gov/policy/docs/statcomps/supplement/2020/4b.pdf.

Figure A.3: Share of Workforce with Self-Employment and 1099 Information Returns

(a) By Presence of Children



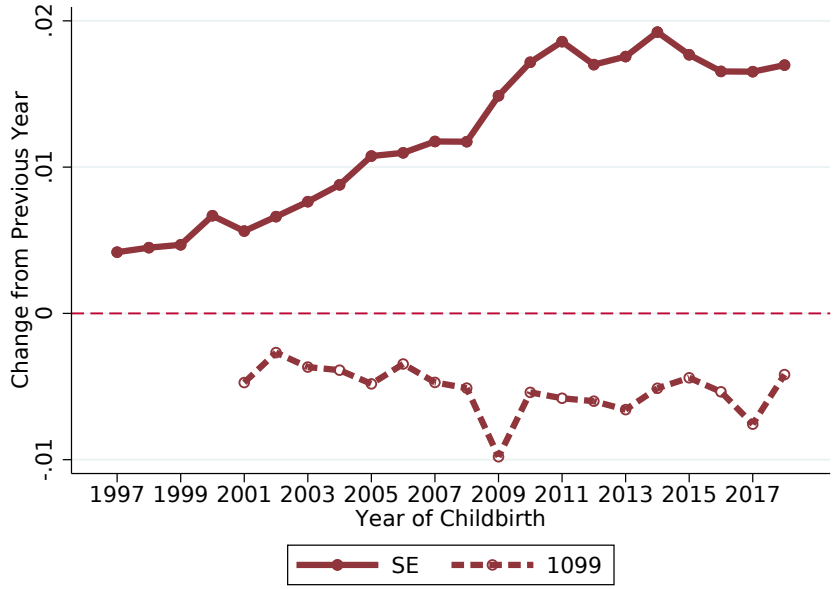
(b) By Gender



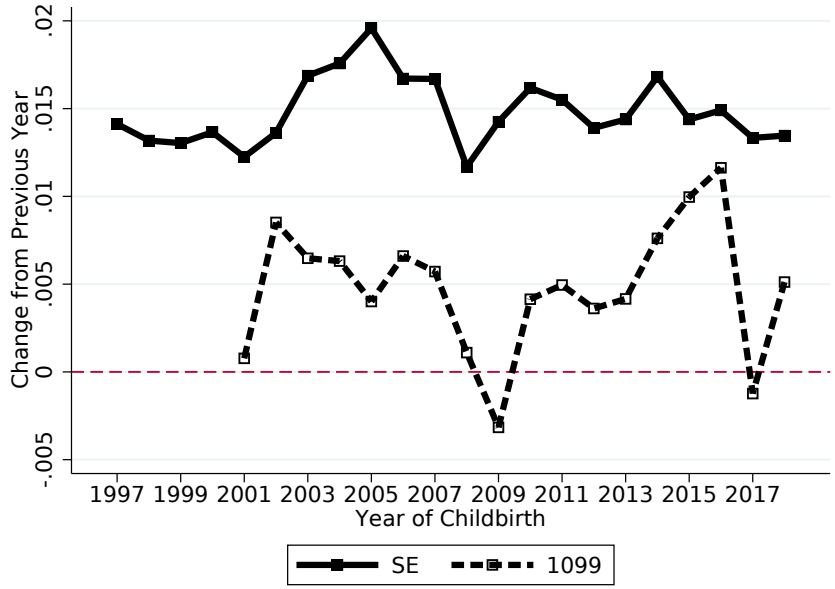
See notes for Figure 1.

Figure A.4: Change in Self-Employment Around First Childbirth, 1997-2018 Births, By Gender of Parent

(a) New Mothers



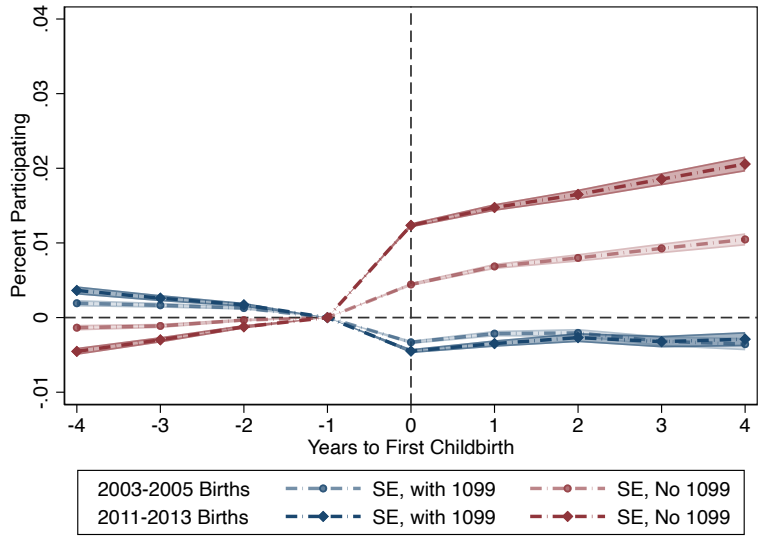
(b) New Fathers



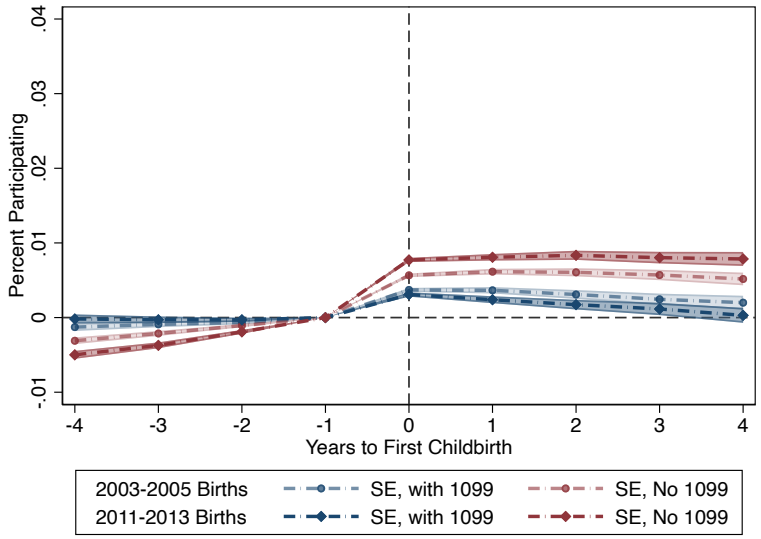
See notes for Figure 3.

Figure A.5: Childbirth Event Study Estimates, Additional Estimates

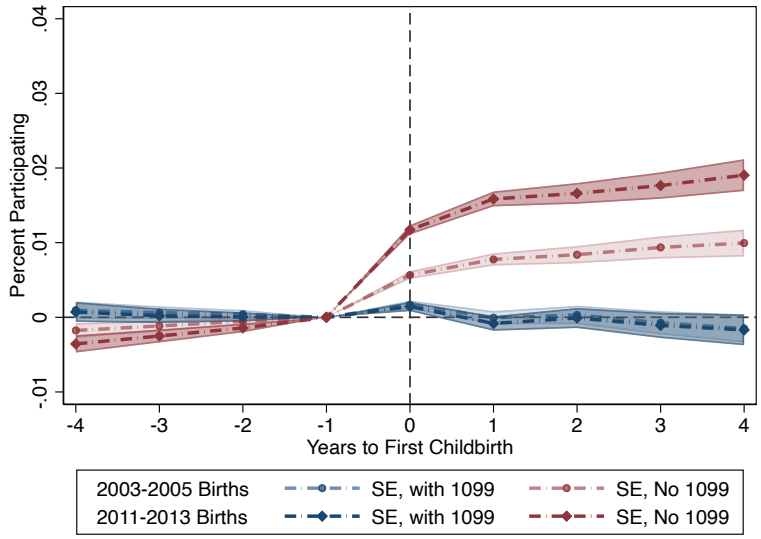
(a) Mothers Only



(b) Fathers Only



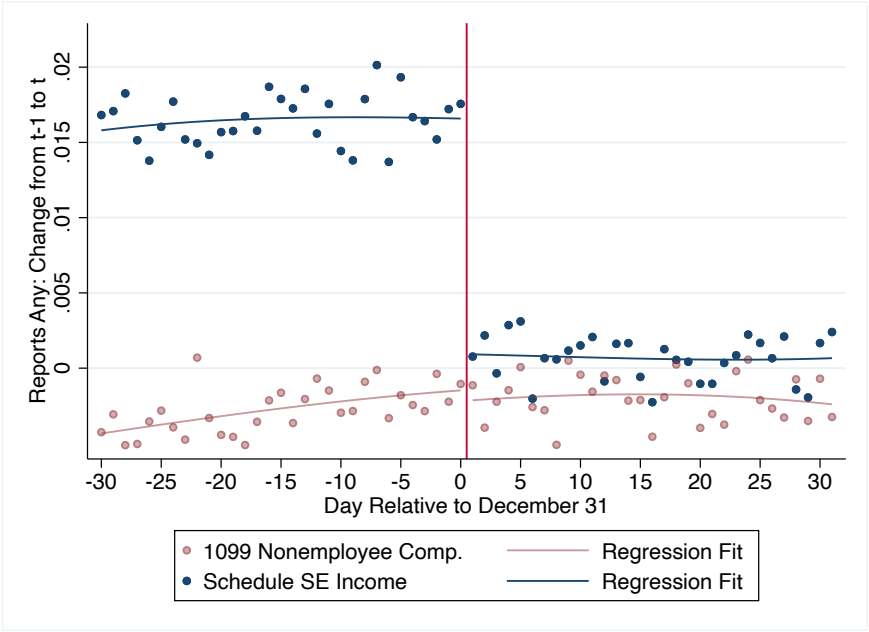
(c) All Parents, December Births Only



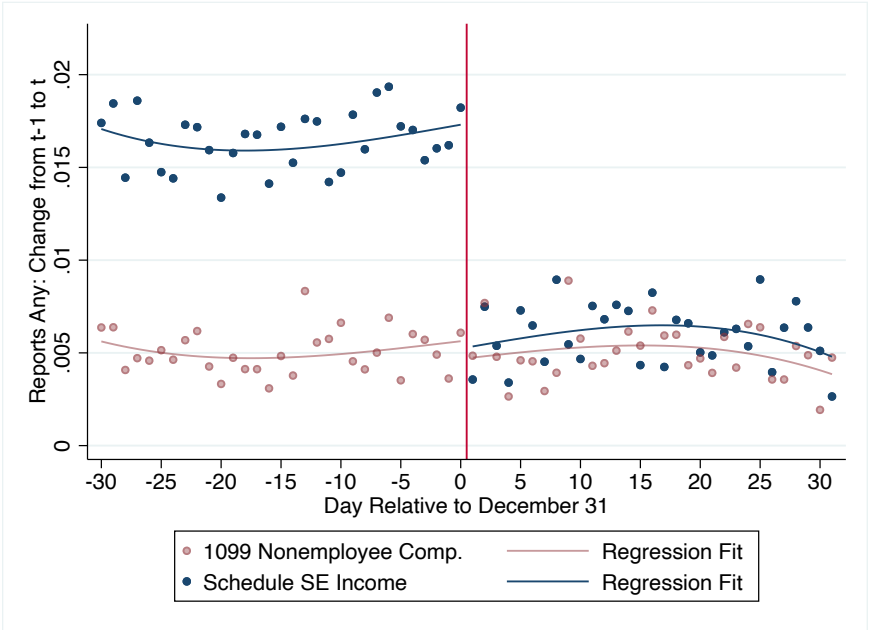
Notes: See notes for Figure 4.

Figure A.6: Regression Discontinuity Design, By Sex

(a) Women



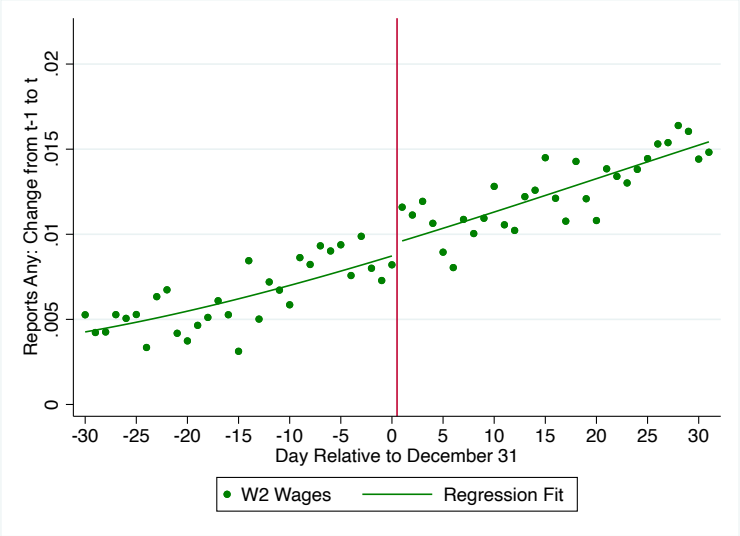
(b) Men



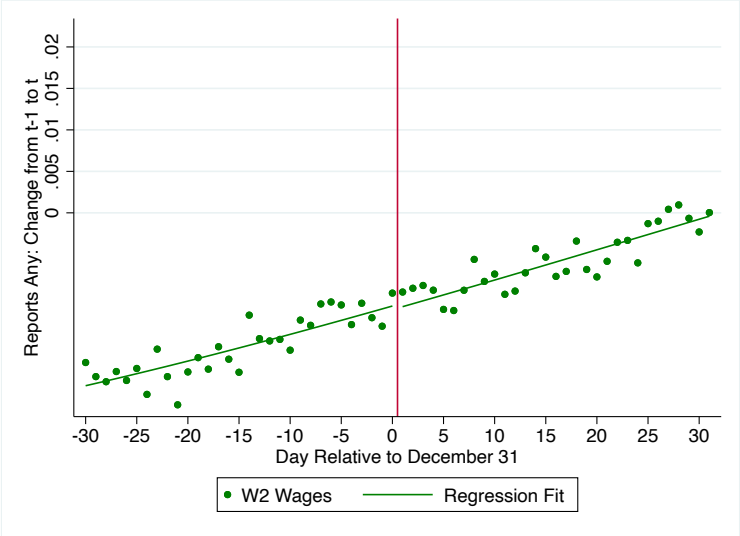
See notes for Figure 5.

Figure A.7: Regression Discontinuity Design, Wages

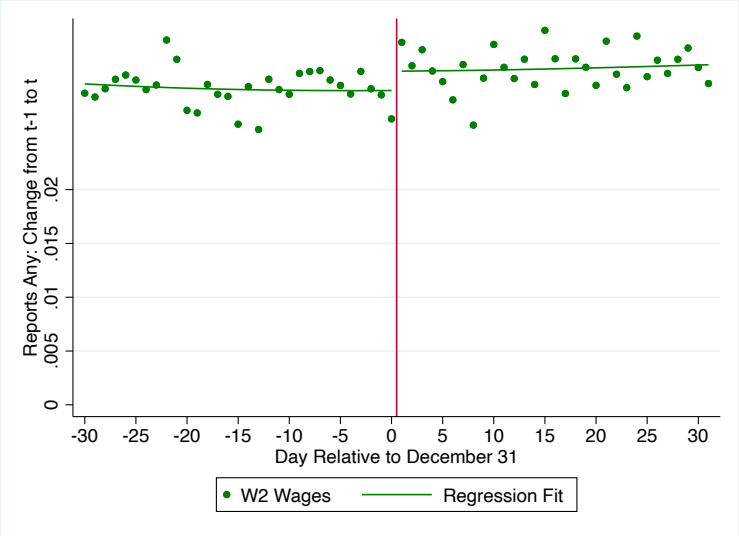
(a) All Parents



(b) Mothers Only



(c) Fathers Only



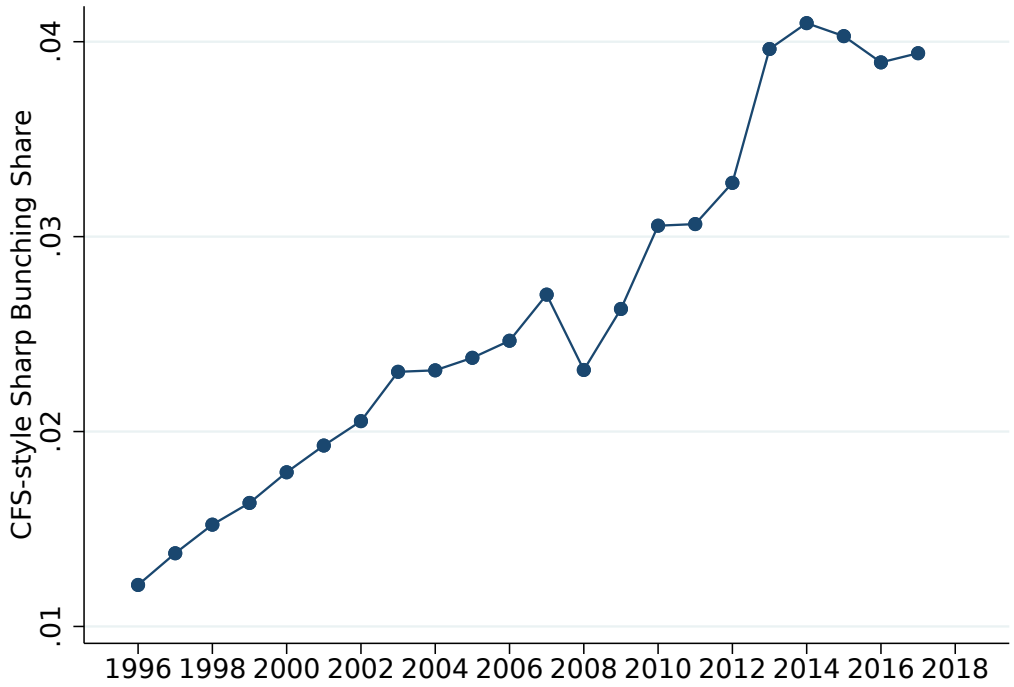
See notes for Figure 5.

Figure A.8: Share of Workforce with Incentive to Report SE



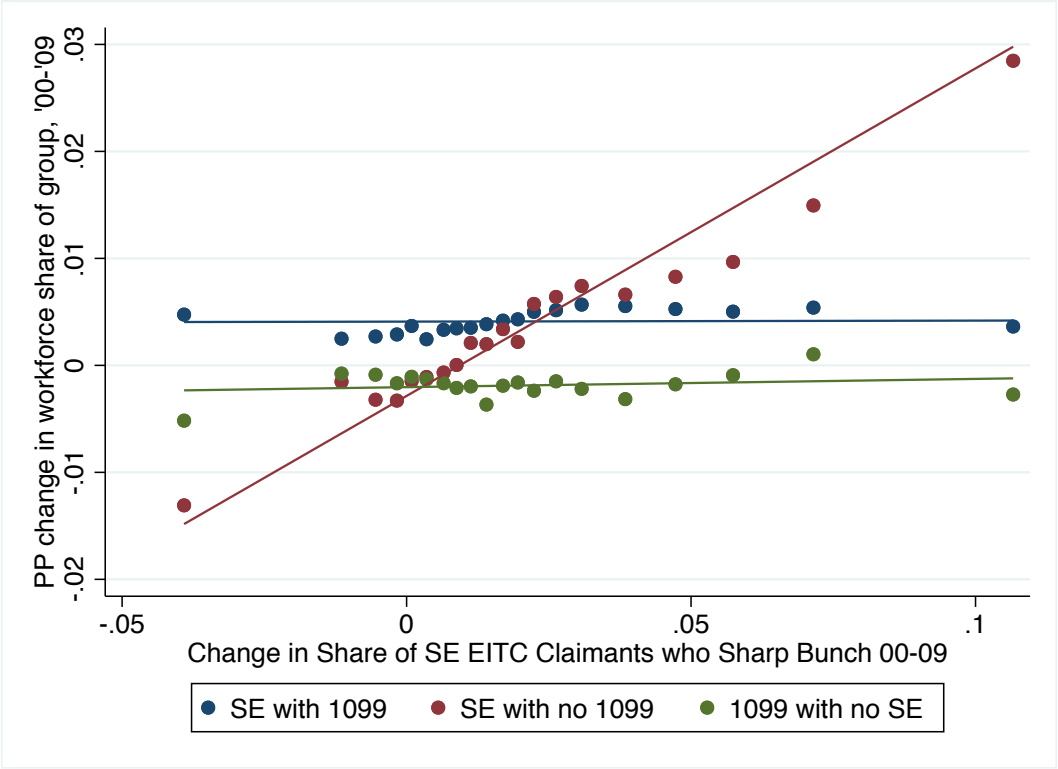
Notes: The share incentivized in each year represents the number of individuals with children and wages below the corresponding EITC kink point as a share of the tax workforce. These data measuring incentivized individuals are drawn from the CPS ASEC as described in Section 6. The tax workforce is measured in the tax data.

Figure A.9: CFS-style Sharp Bunching Share Among Eligible Taxpayers with Children, 1996-2017



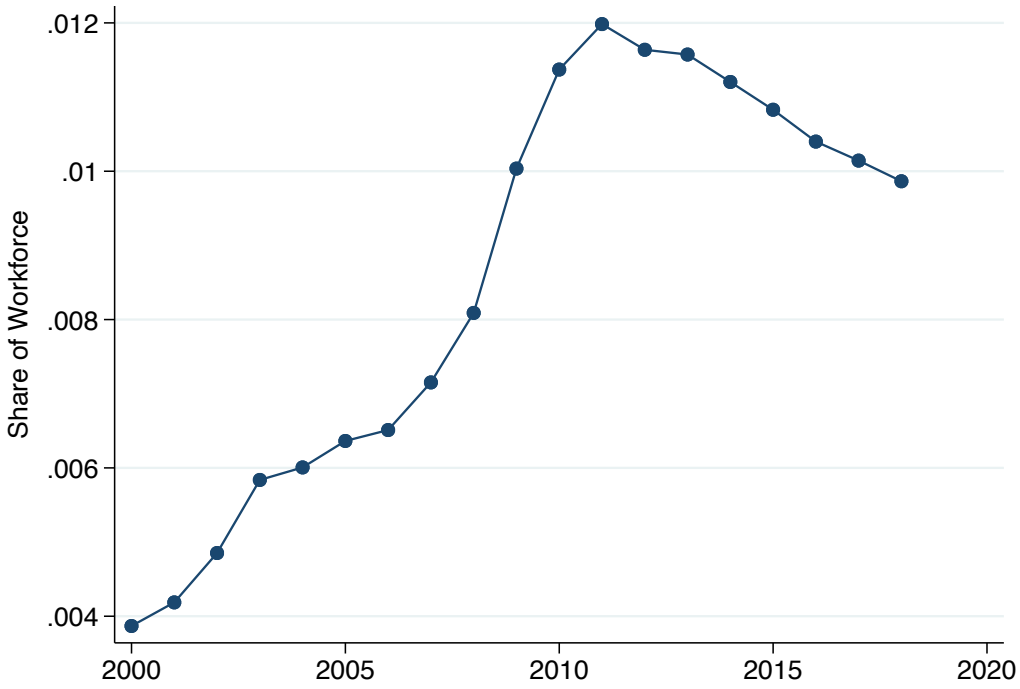
Notes: Figure plots the average of the share of tax payers who are sharp bunchers, following the methodology of Chetty, Friedman, and Saez (2013).

Figure A.10: Local SE Filing Rates and Chetty-Friedman-Saez (2013) Sharp Bunching, 2000 v. 2015



Notes: Figure displays binned scatterplots of changes in the local prevalence of different self-employment behaviors in 3-digit ZIP codes against changes in the local share of SE filers who "sharply bunch" around the applicable first EITC kink (earnings within $500 of lowest level eligible for maximum credit). Figure plots mean of each outcome variable within vingtile bins of X-axis variable as well as linear best-fit line. Outcomes are indicators for: individual files SE *and* has contractor income reported on a 1099 (blue); individual files SE *but has no* contractor income reported on a 1099 (green); and individual has contractor income reported on a 1099 but does not file SE.

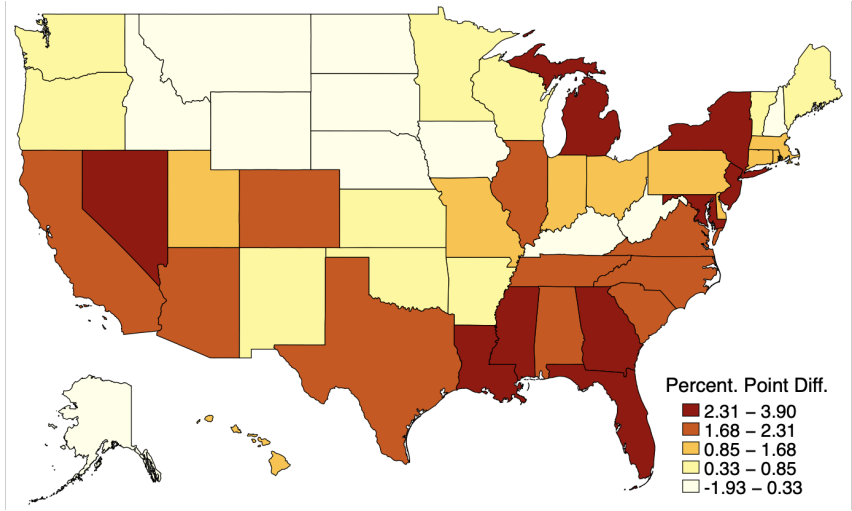
Figure A.11: Counterfactual Change in Workforce Due to Incentive



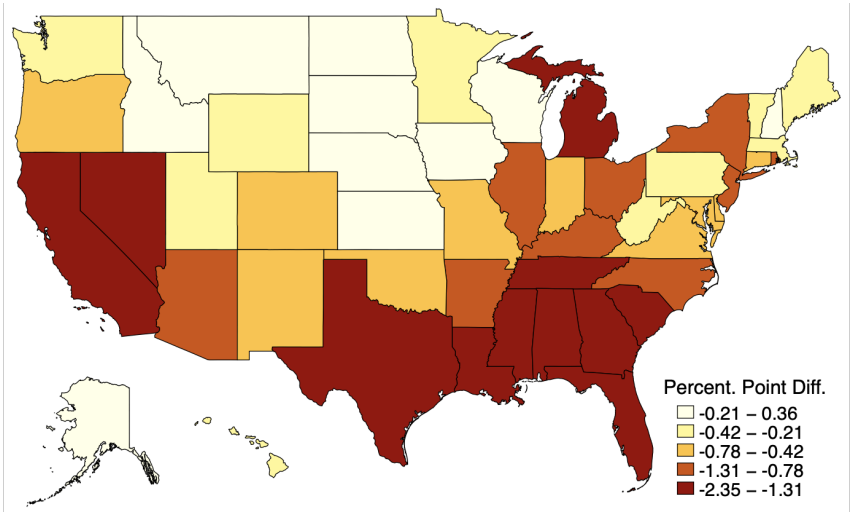
Notes: We calculate a counterfactual number of such individuals with k kids who would otherwise be in the workforce in the absence of incentives as $WF^*_{0,g,k,t} = WF_{w,g,\mathbf{0},t}/POP_{0,g,\mathbf{0},t} \cdot POP_{0,g,k,t}$, i.e. we apply the self-employment-to-population ratio for those without children, and multiply by the population with k children. Figure shows the implied share of the workforce who are in the workforce only due to the incentive, defined as $(WF_{0,g,k,t} - WF^*_{0,g,k,t})/(\sum_w WF_{w,g,k,t})$.

Figure A.12: Raw SE Filing and Application of Our Adjustment By State, 2000-2014

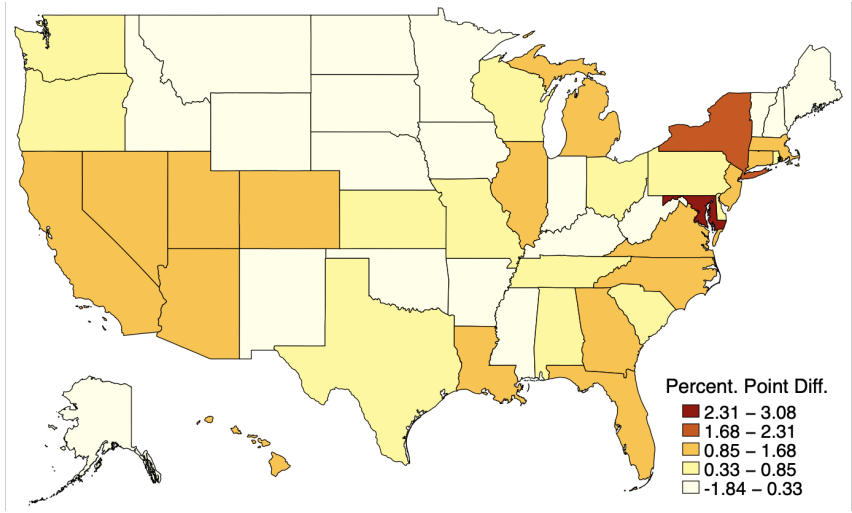
(a) Raw SE Filing Growth, by State



(b) Revision Using Scenario (2): Matched Incentives Adjustment

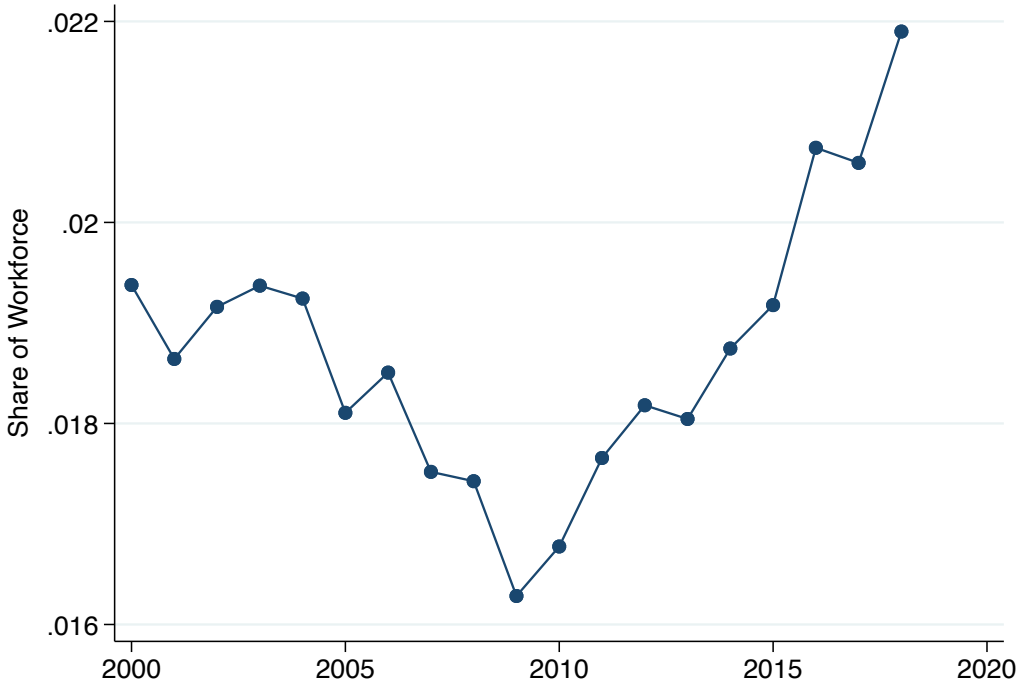


(c) Growth in Adjusted SE Series



Notes: Panel (a) reports raw percentage-point growth in SE filing by state, 2000-2014. The colors refer to quintiles of growth over time, with darker colors indicating higher levels of growth. Panel (b) reports the size of our Scenario (2) adjustment, see text for more details. Panel (c) reports our adjusted-SE series by state, using the same scale as Panel (a)

Figure A.13: Estimated Share of Workforce with 1099-Non-employee Compensation Underreporting SE



Notes: Share of workforce underreporting 1099-nonemployee compensation, estimated following the methodology described in Section 6.1.

Appendix Tables

Table A.1: Trend Adjustments: Share of Workforce with Self-Employment

	Baseline	Scenario 1 RDD Adjusted	Scenario 2 Incentive Adj	Demography Adj Only	Scenario 3 Incentives & Demography Adj	Scenario 4 Incentives & Demography & 1099-SE Reporting Adj
2000	0.0974	0.0974	0.0974	0.0974	0.0974	0.0974
2001	0.0980	0.0978	0.0971	0.0978	0.0968	0.0963
2002	0.0989	0.0979	0.0969	0.0984	0.0963	0.0963
2003	0.1030	0.1014	0.1003	0.1019	0.0991	0.0988
2004	0.1062	0.1053	0.1035	0.1049	0.1020	0.1011
2005	0.1095	0.1091	0.1065	0.1082	0.1049	0.1024
2006	0.1102	0.1090	0.1069	0.1088	0.1051	0.1030
2007	0.1112	0.1092	0.1070	0.1094	0.1048	0.1018
2008	0.1095	0.1067	0.1039	0.1072	0.1011	0.0987
2009	0.1128	0.1102	0.1049	0.1098	0.1012	0.0978
2010	0.1150	0.1127	0.1052	0.1117	0.1011	0.0981
2011	0.1172	0.1152	0.1065	0.1134	0.1017	0.0994
2012	0.1175	0.1154	0.1074	0.1135	0.1023	0.1004
2013	0.1170	0.1139	0.1071	0.1129	0.1018	0.0998
2014	0.1182	0.1150	0.1086	0.1140	0.1033	0.1017
2015	0.1175	0.1148	0.1085	0.1132	0.1031	0.1019
2016	0.1174	0.1152	0.1091	0.1131	0.1035	0.1037
2017	0.1161	0.1145	0.1081	0.1115	0.1023	0.1025
2018	0.1151	0.1141	0.1076	0.1105	0.1018	0.1032

Notes: Table reports baseline (unadjusted) share of workforce with self-employment alongside counterfactual series adjusted for reporting incentives and demographics. See text for details of our adjustment procedure.

Table A.2: Level Adjustments: Share of Workforce with Self-Employment

	Baseline	Scenario 1 RDD Adjusted	Scenario 2 Incentive Adj	Demography Adj Only	Scenario 3 Incentives & Demography Adj	Scenario 4 Incentives & Demography & 1099-SE Reporting Adj
2000	0.0974	0.0951	0.0901	0.0974	0.0901	0.1095
2001	0.0980	0.0955	0.0898	0.0978	0.0896	0.1082
2002	0.0989	0.0956	0.0897	0.0984	0.0891	0.1083
2003	0.1030	0.0991	0.0928	0.1019	0.0917	0.1111
2004	0.1062	0.1028	0.0957	0.1049	0.0944	0.1136
2005	0.1095	0.1065	0.0985	0.1082	0.0970	0.1151
2006	0.1102	0.1065	0.0989	0.1088	0.0973	0.1158
2007	0.1112	0.1067	0.0990	0.1094	0.0969	0.1145
2008	0.1095	0.1042	0.0962	0.1072	0.0936	0.1110
2009	0.1128	0.1077	0.0971	0.1098	0.0937	0.1099
2010	0.1150	0.1100	0.0973	0.1117	0.0935	0.1103
2011	0.1172	0.1125	0.0985	0.1134	0.0941	0.1117
2012	0.1175	0.1127	0.0994	0.1135	0.0947	0.1128
2013	0.1170	0.1112	0.0991	0.1129	0.0942	0.1122
2014	0.1182	0.1123	0.1005	0.1140	0.0955	0.1143
2015	0.1175	0.1122	0.1004	0.1132	0.0953	0.1145
2016	0.1174	0.1125	0.1009	0.1131	0.0958	0.1165
2017	0.1161	0.1118	0.1001	0.1115	0.0947	0.1153
2018	0.1151	0.1114	0.0995	0.1105	0.0941	0.1160

Notes: Table reports baseline (unadjusted) share of workforce with self-employment alongside counterfactual series adjusted for reporting incentives and demographics. See text for details of our adjustment procedure.

Table A.3: Level Adjustments: Share of Workforce with 1099-Reported Non-Employee Compensation

	Baseline	Scenario 1 RDD Adjusted	Scenario 2 Incentive Adj	Demography Adj Only	Scenario 3 Incentives & Demography Adj
2000	0.0906	0.0904	0.0875	0.0906	0.0876
2001	0.0887	0.0884	0.0860	0.0886	0.0859
2002	0.0928	0.0917	0.0899	0.0925	0.0896
2003	0.0947	0.0944	0.0911	0.0941	0.0905
2004	0.0965	0.0970	0.0926	0.0957	0.0918
2005	0.0962	0.0969	0.0922	0.0954	0.0914
2006	0.0979	0.0982	0.0939	0.0970	0.0929
2007	0.0983	0.0984	0.0942	0.0971	0.0929
2008	0.0970	0.0970	0.0927	0.0953	0.0910
2009	0.0938	0.0932	0.0893	0.0915	0.0870
2010	0.0951	0.0947	0.0906	0.0926	0.0880
2011	0.0969	0.0974	0.0923	0.0941	0.0895
2012	0.0975	0.0974	0.0927	0.0946	0.0897
2013	0.0975	0.0963	0.0926	0.0945	0.0895
2014	0.0991	0.0984	0.0941	0.0962	0.0911
	[0.0980]	[0.0984]	[0.0930]	[0.0950]	[0.0899]
2015	0.1012	0.1010	0.0960	0.0984	0.0931
	[0.0973]	[0.1010]	[0.0922]	[0.0943]	[0.0891]
2016	0.1045	0.1038	0.0993	0.1018	0.0964
	[0.0968]	[0.1038]	[0.0917]	[0.0938]	[0.0885]
2017	0.1063	0.1064	0.1011	0.1027	0.0974
	[0.0995]	[0.1064]	[0.0945]	[0.0956]	[0.0904]
2018	0.1070	0.1080	0.1017	0.1033	0.0979
	[0.1002]	[0.1080]	[0.0952]	[0.0961]	[0.0909]

Notes: Table reports baseline (unadjusted) share of workforce with 1099-non-reported non-employee compensation alongside counterfactual series adjusted for reporting incentives and demographics. Shares in square brackets exclude OPE work. See text for details of our adjustment procedure.

B Data Appendix

This appendix describes the technical details of our data construction where we combine data from a variety of different tax forms.

The core of our analysis draws on "masked" W2, 1099-MISC, and 1099-K information returns along with 1040 individual tax returns and associated schedules (e.g. Schedule SE). We begin with the population of individuals who appear as primary or secondary filers on a 1040 in each year. We create a record of all de-identified individuals, using masked Taxpayer Identification Numbers (TINs) appearing on these forms, attributed to either the primary filer or the attached spouse.

For all years, we merge in self-employment information for individuals and their spouses from Schedule SE. On Schedule SE (a schedule of Form 1040), individuals report all self-employment income subject to SECA taxation, so long as the total exceeds $400. This includes active income from wholly-owned businesses on Schedule C, income from partnerships on Schedule K1, and farm income on Schedule F. Importantly, SECA taxes are assessed on individuals, not income tax filing units, so Schedule SE is always identified at the individual level.

We next turn to cleaning and processing the information returns. For Form W-2, we pull all W-2s with TINs that have been validated by the IRS. We eliminate duplicate or amended returns, and we drop a small number of invalid TINs (approximately 50,000 in 2016) and TINs considered "unmatchable" (approximately 5.2 million). Both of these are small compared to the overall number of W-2s, which exceeded 240 million in 2016. We use the recipient TINs to match W-2s to our main file of individuals. Since a large number of individuals with low W-2 earnings are not required to file 1040 returns, we add all cases with valid W-2s but no 1040 to our population file.

We then merge on information from Form 1099-MISC. We pull everyone with non-zero non-employee compensation reported in Box 7. To identify the online platform economy, we begin with a list of roughly 50 large platforms based on public databases of online labor platforms, which we are able identify (along with the corresponding

EIN) in business tax returns using the unmasked firm name. Using the corresponding EIN, we then identify all 1099-MISCs in our cleaned file coming from these platforms and classify them as OPE income.

Reporting rules for intermediaries have changed over time in important ways that affect our measurement of the OPE. In 2011, a new law went into effect requiring companies that processed credit cards, electronic payments, or other transactions to report each recipient's payments on a new information return, "Form 1099-K."[30] Starting in 2012, several online intermediaries in the OPE began issuing the new Form 1099-K instead of 1099-MISC for non-employee compensation. The income paid to gig workers on OPE labor platforms is, for all practical purposes, non-employee compensation. However, 1099-Ks are also issued for income from sales that is not non-employee compensation. We therefore identify and track the labor OPE workforce over time by identifying approximately 50 important online "gig" platforms where self-employed individuals offer labor services to firms or individual clients. We then measure the total payments individuals receive from these companies that are reported on either a 1099-K or a 1099-MISC with non-employee compensation. We also explore alternative approaches to identifying OPE work, as some companies cannot be identified by this method.[31] For example, we use mentions of platform names in taxpayer-reported descriptions of business activity (line A) on Schedule C to identify additional instances of OPE work. A potentially important limitation to studying the 1099-K is that companies in the labor OPE classifying themselves as third party networks are only required to file this form if the total amount of such transactions exceeds $20,000 and the aggregate number of such transactions exceeds 200. In practice, this does not appear to impact our analysis through 2016, as we find most of the major platforms have issued 1099-Ks to all platform participants, regardless of the earnings level, in at least some years. However, beginning in 2017, more platforms begin to abide by the reporting

[30]This measure was included in The Housing and Economic Recovery Act of 2008, but did not take effect until the 2011 tax year.

[31]For some platforms that pay through the payment processor Paypal, the 1099 will be issued by Paypal, and cannot be separately tied to a company in the OPE.

thresholds, and so our measure of gig work is underestimated after 2016. In our analysis, we use Box 1 gross receipts to measure payments. We clean these forms using the same methodology described for the 1099-MISCs. We attribute 1099-K OPE payments to individuals, and add this to OPE income. We consider this income to be a part of the "1099 economy" and include it in measures of "1099 recipients" or "1099 income." So that our definition is more comparable over time, we only classify someone as an OPE worker if they receive a 1099-MISC or have 1099-K earnings of $600 or more; (Collins, Garin, Jackson, Koustas, and Payne, 2019) provides tabulations that include full counts of 1099-K workers, regardless of amount earned.

Worker characteristics Marital status and claimed dependents are defined for 1040 filers only. Marriage is determined from listing a spouse on a 1040. Dependents are determined from listing dependents (other than the spouse) on the 1040 and from a database of parent-child links maintained by the Social Security Administration. For measures of household earnings, wages and 1099 earnings are merged in for the spouse. Additional characteristics are merged in from other sources. Birth dates and gender are pulled from the DM-1 file, populated by the Social Security Administration.