OECD Social, Employment and Migration Working Papers
No. 278

New gig work or changes in reporting? Understanding self-employment trends in tax data

Andrew Garin,

Emilie Jackson,

Dmitri Koustas

https://dx.doi.org/10.1787/26465814-en

OECD

Organisation for Economic Co-operation and Development

DELSA/ELSA/WD/SEM(2022)12

Unclassified **English - Or. English**

20 September 2022

DIRECTORATE FOR EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS
EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS COMMITTEE

New Gig Work or Changes in Reporting?

Understanding Self-Employment Trends in Tax Data

OECD SOCIAL, EMPLOYMENT AND MIGRATION WORKING PAPERS No. 278

JEL classification: J21, H31, H26

Authorised for publication by Stefano Scarpetta, Director, Directorate for Employment, Labour and Social Affairs.

All Social, Employment and Migration Working Papers are now available through the OECD website at www.oecd.org/els/workingpapers .

Andrew Garin, Emilie Jackson and Dmitri Koustas.

JT03502834

OECD Social, Employment and Migration Working Papers

www.oecd.org/els/workingpapers

Acknowledgements

Garin: University of Illinois Urbana-Champaign (agarin@illinois.edu). Jackson: Michigan State University (emiliej@msu.edu). Koustas: Harris School of Public Policy, University of Chicago (email: dkoustas@uchicago.edu).

This is an updated version of the working paper available on the IRS SOI Tax Stats website:

https://www.irs.gov/pub/irs-soi/22rpnewworkchangesinreporting.pdf.

We are grateful to Katharine Abraham, Dan Black, Naomi Feldman, John Guyton, Lawrence Katz, Elira Kuka, Kara Leibel, Alicia Miller, Bruce Meyer, Jacob Mortenson, Max Risch, Michael Weber, Andy Whitten, Eleanor Wilking, and seminar participants at the Office of Tax Analysis at the US Treasury and National Tax Association meetings for helpful discussions. 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. Garin gratefully acknowledges support from the OECD Future of Work Fellowship and from the Sloan Foundation for a post-doctoral fellowship at NBER during which part of this research was conducted. 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 data as IRS employees through agreements under the Intergovernmental Personnel Act.

Abstract

Rising self-employment rates in U.S. tax data that are absent in survey data have led to speculation that tax records capture a rise in new “gig” work that surveys miss. Drawing on the universe of Internal Revenue Service (IRS) tax returns, we show that trends in firm-reported payments to “gig” and other contract workers do not explain the rise in self-employment reported to the IRS; rather, that increase is driven by self-reported earnings of individuals in the EITC phase-in range. We isolate pure reporting responses from real labor supply responses by examining births of workers’ first children around an end-of-year cutoff for credit eligibility that creates exogenous variation in tax rates at the end of the tax year *after* labor supply decisions are already sunk. We find that exposing workers with sunk labor supply to negative marginal tax rates results in large increases in their propensity to self-report self-employment—only a small minority of which leads to bunching at kink-points. Consistent with pure strategic reporting behavior, we find no impact on reporting among taxpayers with no incentive to report additional income and no effects on *firm-reported* payments of any kind. Moreover, we find these reporting responses have grown over time as knowledge of tax incentives has become widespread. Quantitatively, our results suggest that as much as 59 percent of the growth in self-employment rates, and all counter-cyclicality, can be attributed to changes in reporting behavior that are independent of changes in the nature of work. Our findings suggest caution is warranted before deferring to administrative data over survey data when measuring labor market trends.

Table of contents

Tables

Figures

1 Introduction

The emergence of new institutions and technologies over the past decade that have made it easier for firms to obtain labor from self-employed contract workers instead of employees has led to widespread speculation that the U.S. labor market is undergoing a fundamental transition towards a "gig" economy. To date, major labor force surveys like the Current Population Survey (CPS) show no increase in the self-employment rate since 2000; however, there is reason to think that such surveys might miss a shift towards a gig economy if workers do not perceive themselves as contractors or are more likely to do side jobs that are not well-captured by standard questionnaires (Katz and Krueger, 2018[1]; Abraham, Hershbein and Houseman, 2020[2]; Abraham, Hershbein and Houseman, 2020[2]). In sharp contrast to trends in the CPS, the share of individuals reporting self-employment income to the Internal Revenue Service (IRS) on their tax returns rose dramatically between 2000 and 2014.[1] We present these two measured trends side-by-side in Figure 1. Is the administrative data collected by the IRS picking up on a deep change in the labor market that major surveys currently miss?

While tax records provide a critical resource for understanding the labor market, they too come with important caveats. Unlike in confidential surveys, individuals have incentives to be strategic about what they report on tax filings—and both those incentives and reporting decisions 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, 2021[3]). A large literature has documented that self-employment reporting is highly sensitive to incentives in the tax code and there is substantial "bunching" of reported self-employment earnings at levels that maximize refundable tax credits like the Earned Income Tax Credit (EITC), even as bunching behavior is absent in survey data (LaLumia, 2021[4]; Saez, 2010[5]; Chetty, Friedman and Saez, 2013[6]; Kuka, 2013[7]; Mortenson and Whitten, 2020[8]). Moreover, this type of strategic reporting has become increasingly prevalent since 2000 (Chetty, Friedman and Saez, 2013[6]; Mortenson and Whitten, 2020[8]). These earlier findings raise the possibility that the same forces driving an increase in bunching behavior impact the reporting of self-employment work more broadly, independent of any underlying change in the labor market.

Drawing on the universe of IRS tax returns, this paper shows that changes in pure reporting behavior—given fixed labor market behaviors—contribute substantially to the observed increase in self-employment reported to the IRS. First, we draw directly on the information returns issued by firms to self-employed independent contractors (of which online-platform-based "gig" workers are a subset) and show that the rise in *taxpayer-reported* self-employment since 2000, visible in Figure 1, cannot be accounted for by the rise of new types of platform-based work in particular or by firm-reported contract work more broadly. Rather, we find the

[1] This trend in self-employment tax filings has been well-documented in earlier work (Abraham et al., 2013[23]) (Jackson, Looney and Ramnath, 2017[12]) (Abraham et al., 2021[21]) and has been interpreted as suggestive evidence of a rise in alternative work arrangements not captured in traditional surveys (Katz and Krueger, 2018[1]).

dramatic increases in taxpayer-reported self-employment are concentrated specifically among individuals in the phase-in range of refundable tax credits who have a strict incentive to report additional earnings beyond their employer-reported wages. Second, using a novel regression discontinuity design that exogenously varies exposure to negative marginal tax rates after labor supply decisions are sunk, we find that these incentives lead to increased reporting of self-employment earnings by taxpayers even with labor supply held constant. In contrast to kink-bunching designs, our design directly rules out real changes in labor supply behavior—we find that, in practice, only a small fraction of these reporting responses lead to bunching at kink points. Third, we present evidence that the magnitude of the pure reporting response to fixed tax incentives has grown over time, and the same factors that explain increased EITC bunching over time (Chetty, Friedman and Saez, 2013[6]; Mortenson and Whitten, 2020[8]) explain increased self-employment reporting rates more broadly. Finally, we show that such behaviors can quantitatively explain the majority of the increase in self-employment reporting in excess of independent contract work reported by firms.

We begin by drawing on the information returns issued to independent contractors—including app-based workers in the Online Platform Economy (OPE)—to assess the extent to which tax trends in self-employment reporting are plausibly driven by a rise in "gig" arrangements (Harris and Krueger, 2015[9]; Abraham et al., 2021[10]; Farrell, Greig and Hamoudi, 2018[11]; Katz and Krueger, 2018[1]). A unique feature of U.S. tax data is that payments to self-employed contract workers are directly observed on information returns reported by firms to the IRS. We follow previous work and break out the subset of these returns issued by OPE firms to directly measure their contribution to overall trends in contract work (Jackson, Looney and Ramnath, 2017[12]; Collins et al., 2019[13]; Lim et al., 2019[14]). We find that while the prevalence of OPE earnings has grown dramatically since 2014 this cannot explain the overall rise in reported self-employment rates which only grew prior to 2014 and have plateaued since.[2] In the full period from 2000 to 2018, year-to-year changes in firm-reported contract work have nearly zero correlation with changes in the share of workers reporting self-employment.

We then show that increases in self-employment reporting have been concentrated among individuals with a strict incentive to report additional self-employment earnings. Importantly, tax filers with children and employer-reported wage earnings in the EITC phase-in range (below the first "kink" in the EITC schedule) face *negative* tax rates on the marginal dollar—and would therefore strictly benefit from reporting some self-employment earnings, rather than none. These individuals might therefore increase their tax refunds by choosing to report informal earnings they might otherwise not have reported to the IRS, or, in some cases, by fabricating the income entirely.[3] 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. Meanwhile, such increases are absent for those with firm-reported wages even slightly above the "kink" point where the marginal tax rate becomes positive. Likewise, we find no differential growth above or below the first EITC kink point for individuals with no children who face negative marginal tax rates. Moreover, we find no such rise in the probability of having *firm-reported*

[2] By contrast, we observe no growth in firm-reported contract work in the decade between 2005 and 2014. We do find that an increased prevalence of independent contract work (outside the OPE) between 2000 and 2005 that can account for some of the increase in worker-reported self-employment; as Abraham et al (2021[21]) note, this increase in freelance work might not be reflected in survey data if workers do not perceive themselves as self-employed.

[3] While highly strategic agents should "bunch" and report total earnings at the precise level that minimizes net tax liability, in practice individuals may choose to report some self-employment to improve their tax position without bunching, especially if they have concerns about appearing suspicious. Roughly a third of Americans report having some type of informal income (Bracha and Burke, 2021[19]); it is likely that individuals have some discretion over whether or not to report such informal earnings.

contract income among incentivized individuals with low wages and children, only a rise in the propensity to *self-report* self-employment.

Do these increased propensities to self-report self-employment reflect real increases in self-employment work, or simply changes in reporting behavior given fixed labor-supply behavior? We pinpoint pure reporting behaviors holding labor supply fixed by examining a sharp discontinuity in EITC eligibility based on the date of birth of an individual's first child. An individual's EITC benefit for tax year t is calculated based on the number of children in their household during year t. Hence, a first-born child born on or before December 31 of year t would count towards the EITC calculation for that year t, while a child born only a few days later at the start of year $t + 1$ would not. This cutoff creates sharp differences in incentives when reporting income for tax year t. This motivates a regression discontinuity design (RDD) comparing year-t reported earnings for parents with births of their first children occurring right before or right after the end of tax year t.[4] Crucially, if parents with births near the discontinuity face *ex ante* uncertainty about which year the birth will occur in when deciding on year-t labor supply, then all labor supply decisions will be fully *sunk* by the time the child's eligibility for the EITC is learned. In that case, any observed differences in self-reported earnings around the cutoff reflect pure reporting adjustments holding underlying work behavior fixed.

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. By contrast, we find no such impact on the probability of having *firm-reported* contract payments or wage/salary earnings, indicating the main result reflects pure reporting behavior and not any change in labor supply.[5][5] Reassuringly, we find that the entire self-employment effect is driven by individuals with wages in or below the phase-in range who face negative marginal tax rates on the marginal dollar of self-reported self-employment earnings.

Moreover, we find that 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 more than 5 p.p. We show that changes in the generosity of the federal EITC over the period we study cannot plausibly drive the observed rise. Why, then, an increasing response over time? Prior research suggests that even during periods when tax 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[6]). 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[6]), we use sharp-bunching rates among individuals in a hold-out sample for each three-digit ZIP code in each year as a proxy for local knowledge of EITC incentives and estimate an interacted version of our RDD specification. We find that high-knowledge regions

---

[4] This design is inspired by earlier work by LaLumia, Sallee and Turner (2015[20]) who find little evidence that mothers strategically time first births around the end of the year to minimize their tax burden, but who do find that mothers with Schedule C income whose children are born in December are much more likely to bunch at the first EITC kink in that tax year than those with children born the following January. Feldman, Katuščák and Kawano (2016[25]) examine a similar discontinuity to study parent responses to children who age out of the Child Tax Credit at 17. A key difference in their setting is that their work studies changes in real labor supply behavior during the year *after* the age-out, whereas we focus on pure reporting behavior at the end of *same* tax year that eligibility changes.

[5] Unlike self-reported earnings, third-party-reported payments cannot be manipulated by individuals to avoid taxation; see, for example, Kleven et al. (2011[26]). If individuals were strategically 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. However, we find no evidence that either firm-reported wage earnings or non-employee compensation vary around the threshold date.

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. These results indicate that spreading knowledge of the tax code drove a much broader increase in self-employment reporting beyond what is reflected in kink-bunching behavior.

We conclude the paper with several exercises examining the quantitative contribution of these reporting trends to the observed overall increase in self-employment tax filings since 2000. We find that our RDD estimates alone can only explain a minority of the rise in self-employment filing; however, we argue that the RDD estimates identify a short-run response to incentives that provides a lower bound on individuals' long-run reporting response. We show that in the counterfactual scenario in which individuals with EITC-eligible children reported self-employment at the same rate as individuals without children within narrow wage-bins and demographic cells, the increase in the share of workers reporting self-employment would be reduced by over half—explaining most of the discrepancy between trends in self-reported self-employment and firm-reported payments to contractors. The remaining unexplained growth in worker-reported self-employment in the early 2000s can largely be explained by increases in firm-reported independent contracting around the same period.

Our findings ultimately offer a cautionary tale about measuring labor-market trends with administrative data (Blank, Charles and Sallee, 2009[15]; Slemrod, 2021[3]). While surveys are imperfect instruments, trends observed in administrative data may be impacted by reporting incentives and should be interpreted accordingly. In our setting, we show that such incentives give the *appearance* of a shift in the nature of work that is belied by more reliable third-party reporting by firms themselves.

More broadly, our findings and our additional descriptive analysis in Collins et al. (2019[13]) suggest that the impact of new types of online platform work on the broader labor market outside driving tasks has been negligible, despite widespread media and policy attention. We hope our analysis, our data series on the prevalence of OPE work, and our adjusted self-employment are valuable to other researchers and policymakers focused on the future of work.

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 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, which control 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

2.1 Self-Employment and Gig Work in U.S. Tax Records

In the United States, firms use W-2 information returns to report employees' wage and salary earnings to both the Internal Revenue Service and to the recipient, who then enters those earnings on their 1040 tax return. In contrast, self-employment proceeds are self-reported to the IRS by workers on their 1040 returns. In the U.S. tax system, self-employment earnings are technically active income from a "sole-proprietorship" business. Self-employed workers must keep track of their revenues, expenses, and net profits, and report those amounts on Schedule C of their 1040 return.[6] If an individual's self-employment proceeds exceed a threshold level of $433, they must file Schedule SE and pay Social Security and Medicaid ("SECA") taxes—the SECA tax amount is equivalent to the total payroll tax that would be withheld from the same level of wage and salary earnings.[7] Prior work typically measures self-employment rates in tax data by comparing annual self-employment earnings on Schedules C and SE to employment earnings on W-2 forms.[8] In our analysis, we primarily focus on Schedule SE filings, which are consistently available at the individual (rather than filing unit) level, and begin in 1996.

Self-employment activity is highly heterogeneous (Hurst and Pugsley, 2011[16]). While some self-employed workers directly serve consumers (for example, barbers and plumbers), others work for firms as non-employee independent contractors (for example, freelance IT consultants). Importantly, firms are required to report all payments to self-employed contractors in excess of $600 on 1099-MISC Box 7 (replaced by 1099-NEC in 2020). These returns are useful to us in two respects. First, they allow us to identify which self-employed workers are working for firms in "alternative" work arrangements.[9] This enables us to assess whether changes in how firms classify workers drives aggregate self-employment trends. Second, 1099s provide information on worker revenues, which is third-party reported and not sensitive to taxpayer reporting incentives. Unlike W-2s, however, 1099s report *gross revenues* and not *profits after expenses*. Hence, recipients can still influence their taxable self-employment earnings through their self-reported expenses on Schedule C.

The tax data further enable us to 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 where 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 home-sharing). OPE earnings are reported

[6] Self-employment earnings from farming are reported on Schedule F instead of Schedule C.
[7] Self-employed individuals are responsible for paying the equivalent of both 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%.
[8] For example, Abraham et al. (2017[18]) focus on Schedule SE filers, while Jackson, Looney and Ramnath (2017[12]) focus on Schedule SE and Schedule C filers.
[9] We explore this workforce in depth in earlier work (Collins et al., 2019[13]).

to the IRS either as independent contractor earnings on 1099-MISC/NEC, or, in some cases, as vendor revenues on 1099-K. We identify 1099 forms issued from OPE companies using the method in Collins et al. (2019[13]); we discuss several important measurement issues in the Data Appendix.

The self-employed workforce in tax data may differ from self-employed workers identified in survey and other data sources. In tax data, we observe total annual earnings by source. By contrast, in surveys like the Current Population Survey (CPS), workers self-identify as employed or self-employed based on their predominant activity in a reference week. "Point-in-time" measures will generally lead to lower estimates of the prevalence of self-employment work—in particular, they under-count secondary, informal self-employment and firm-facing self-employment (National Academies of Sciences, 2020[17]; Abraham, Hershbein and Houseman, 2020[2]; Abraham et al., 2017[18]). Further, even surveys that record retrospective annual earnings like the CPS Annual Social and Economic Supplement (ASEC) may nonetheless under-count self-employed workers if such workers fail to report small supplemental earnings or if independent contractors incorrectly perceive themselves as employees.

In theory, the information reported on tax returns should be more comprehensive than information on surveys, since individuals face significant penalties if they misreport their income. In reality, enforcement is imperfect and individuals have considerable scope to strategically under-report—or over-report—income to minimize their tax burden and maximize refundable credits. Thus, changes in strategic reporting behavior over time can potentially drive trends in tax data, whereas confidential surveys are often immune to such concerns.

2.2 Incentives in the Tax Code to Report Self-Employment

Certain provisions in the tax code may incentivize certain individuals to report self-employment profits. In particular, the Child Tax Credit (CTC) and Earned Income Tax Credit (EITC) phase-in with higher earnings up to a threshold level where there is a first "kink" in the EITC schedule as the maximum credit level is obtained. Both this threshold level and the effective subsidy rate depend on the number of children claimed on a tax return.[10] 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.

Individuals with children face net negative marginal tax rates in the EITC and CTC phase-in range, even after incorporating all federal taxes including payroll (FICA/SECA) taxes.[11] These subsidies are sometimes enhanced by additional credits provided by states. By contrast, the net subsidy disappears once earnings are above the applicable phase-in range, since additional earnings are subject to the payroll tax but no marginal subsidy.[12] Moreover, households with no

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

[11] Every dollar of self-employment earnings reported is subject to the 14.1 percent SECA rate (net of the deductible part of the taxes paid). In the EITC phase-in-range, the increase in Federal credits per dollar reported exceeds the increase in SECA liability, resulting in a net increase in one's refund—in this case, SECA taxes simply reduce the net refund.

[12] 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.

children *always* face a positive marginal tax rate on earnings, as the payroll tax rate exceeds the lower EITC phase-in rate.[13]

Crucially, these negative marginal tax rates create an incentive for some individuals to report self-employment income that may not have been reported otherwise. In particular, individuals with EITC-eligible children under 19 and W-2 wages below the first EITC kink are strictly better off reporting self-employment income on their 1040 return. In theory, such individuals could report self-employment earnings to maximize their net subsidy—this sort of hyper-strategic behavior would result in "bunching" at the first EITC kink point (Saez, 2010[5]).[14] In practice, though, these tax incentives may lead individuals to report additional self-employment income without going so far as to report this "optimal" amount. For example, many individuals have actual informal income (Bracha and Burke, 2021[19]) that they may choose to report in full that they may not have reported otherwise. Additionally, strategic taxpayers may avoid reporting at the refund-maximizing level if they think it might appear suspicious to auditors.[15]

Appendix Figure A.1 shows the effective tax rate on reporting a first dollar of self-employment income at different levels of W-2 wage earnings, taking into account all federal taxes. 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.[16] 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. 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 incorporating both filer-reported returns and information returns. Most of our analysis uses data from 2000-2018. While some microdata are available 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.

---

[13] Even with the most generous state-EITC for childless individuals in the District of Columbia, they nonetheless face a marginal tax rate of 0 in the phase-in range.

[14] While the first EITC kink point is the refund-maximizing level in most years, Mortenson and Whitten (2020[8]) note that temporary tax credits shift the level in some years.

[15] We note that individuals that would have reported low levels of self-employment earnings in the credit phase-in range, regardless of incentives, face an incentive to report additional self-employment income. However, their additional self-employment earnings would not impact the extensive margin of self-employment, which is our focus. We also note that individuals with W-2 earnings close to the first EITC kink point face positive marginal tax rates on any self-employment earnings that exceed the kink level, which mitigates the incentive to report additional amounts.

[16] 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: (1-0.0765)*(-0.45-0.15+0.153)=-0.413.

3 Trends in Self-Reported and Firm-Reported Data

3.1 Do New Forms of Gig Work Explain the Rise in Self-Employment Filings

The first possibility is that the rise in the share of the workforce with self-employment income demonstrates a fundamental change in the labor market, driven by new types of contract-based "gig" work (Katz and Krueger, 2018[1]). We begin by reporting the time-series for the annual share of individuals in the workforce with positive net self-employment income reported on a Schedule SE, our baseline measure of self-employment. In our analysis, we define the "workforce" as all individuals who receive wage (W-2) earnings, contract earnings (1099-MISC or OPE 1099-K), and/or self-employment (Schedule 1040-SE) earnings.[17]

As reported in Figure 2, the portion of the workforce with self-employment rose significantly (2 percentage points) after 2000, peaking in 2014. However, trends in 1099-reported contract work cannot explain the overall rise in reported self-employment on tax returns. 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.

To investigate the contribution of new online platforms to the overall trend, we break out the portion of contract workers who only have 1099-reported non-employee compensation from OPE platforms.[18] The vertical line in the figure represents a structural break in our freelance/contracting series, when online platform mediated work started to represent a meaningful share of the workforce. Strikingly, we find that where gig work grows most dramatically (2014 onwards), administrative measures of self-employment are flat or *decline* slightly. This implies that new forms of "gig" work cannot explain the rise in self-employment in tax data. On the whole, annual changes in the SE series and the non-employee compensation series between 2000-2018 show no statistically significant correlation.

Nonetheless, there is substantial growth in new online platform work after 2014. This raises the question: why did larger numbers of platform workers not result in more workers reporting self-employment income? The vast majority of new platform workers only earn a small amount of money over the course of the year, and typically expense over half 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. Additionally, platform workers are less likely to file a Schedule C or SE than other types of contractors with non-employee compensation on a 1099-MISC.[19]

[17] Following Collins et al. (2019[13]), we only include individuals with contract earnings when they receive a W-2 or file a tax return.
[18] Note that in 2017 and 2018, we impute growth in OPE 1099-Ks using state-level data from Massachusetts and Vermont, both of which introduced a state 1099-K with a lower $600 filing threshold. As discussed in Collins et al. (2019[13]), prior to 2017, although the federal filing requirement was $20,000, in practice this was not binding.
[19] We examine filing and expensing behavior among contractors in further detail in a companion paper (Collins et al., 2019[13]).

Therefore, trends in non-employee compensation may diverge from trends in reported self-employment earnings because of changes in compliance, or changes in the composition of types of 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.

3.2 Self-Employment Reporting and Eligibility for Refundable Credits

We turn to our second hypothesis and examine trends in relation to incentives. 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 would have reported none.

As shown in Figure A.2a, 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.[20] In comparison, we see only modest changes in 1099 receipt and SE growth for childless filers, who do not face negative marginal tax rates. Figure A.2c examines growth by total earnings. We find SE rates rise most dramatically for those with less than $50,000 in total earnings (in constant $2015), 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 earners with more than $50,000: SE grows between 2003-2005, and falls sharply in the recession, before returning to its 2005 level. Figure A.2d reports additional trends by gender, showing that most of the increase occurs among women.

In theory, only families with wage earnings in the credit phase-in range are incentivized to report positive 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 3 reports the change in self-employment propensity between 2000 and 2014, by W2 wage earnings and the presence of children.[21] The results for Schedule SE self-employment, reported in Figure 3, Panel A, are striking. We find dramatic growth in self-employment among 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 gain from reporting some small amount of self-employment rather than none. The change in self-employment between 2000-2014 for households without children, or for households with higher levels of wage income is much smaller. 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 tax filers. We see much smaller changes over this period for households, regardless of whether they have children.

We further show in Annex C that the changes in self-employment reported on 1040-SE occurs precisely at the time of one's first childbirth, when they first face reporting incentives, and continues to grow in subsequent years. Moreover, we show that this increase in self-employment at the time of childbirth has grown steadily since 2000. However, we find no change

[20] Figure 2b shows additional trends by presence of children only.

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

in *firm-reported* contract income around childbirth. While suggestive that households are motivated by EITC incentives to report self-employment income, these findings could also reflect real increases in types of self-employment activity not reported on 1099 forms. The need to distinguish between real labor supply changes and pure reporting behavior motivates our research design in the next section.

4 Isolating Pure Reporting Behavior

4.1 Regression Discontinuity Design

A key question is whether the increases in self-reported self-employment rates documented above reflect 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. To isolate reporting behavior, we examine a sharp discontinuity in EITC eligibility based on the date of birth of an individual's first child. 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, creating an incentive for low-wage parents to report additional self-employment income on tax day. 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 and their parents would have no such incentive. Because exact birthdates are difficult to precisely forecast far in advance, parents expecting a first child close to the end of the year are 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 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, they 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. Accordingly, all differences in self-reported tax year t self-employment between parents on either side of the discontinuity will be due entirely to differences in reporting behavior.

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.[22] In our analysis, we test for strategic birth timing by examining discontinuities in *firm-reported* earnings, which should exist if individuals are sorting. Additionally, we plot the distribution of births around the end of tax years 2011-2018 (corresponding to our benchmark sample below) in Appendix Figure A.3 and find that the distribution of births is mostly smooth, with expected decreases around the Christmas and New Year's holidays. To ensure that selective avoidance of births during the New Year's holiday does not influence our results, we

[22] Early work by Dickert-Conlin and Chandra (1999[27]) documented a correlation between tax rates and birth timing. More recent work by LaLumia, Sallee and Turner (2015[20]) tests for such behavior using administrative tax records 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—meanwhile, the effect of tax incentives on first-birth timing is negligible. However, LaLumia, Sallee and Turner (2015[20]) 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, consistent with some difference in reporting behavior depending on the year of birth. Our design tests for broader reporting responses beyond highly-strategic bunching behavior.

omit a three-day bandwidth “donut hole” from our baseline analysis sample and examine robustness to the inclusion of the donut hole and alternate bandwidths.

We estimate regressions of the form:

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

on the sample of childbirths in December and January. In this specification, $date_i$ is the running variable—which is the child’s birth date measured as days since December 31— and outcome y_i is measured in tax year t. In our baseline analysis, we include births in a fifteen-day bandwidth around midnight on New Year’s Eve, omitting births in the three-day bandwidth “donut hole” around the start of the new year in case of any potential shifting around the holiday, and let $f(\cdot)$ be a linear function allowing the slope to change around the cutoff.[23] We also examine alternative bandwidths and polynomial specifications as robustness checks. 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 $t + 1$. We estimate Equation 1 on all parents with first births in the specified window as identified in Social Security birth records. This sample includes known parents listed in Social Security records, regardless of whether or not that parent claims the child on their tax return.[24]

4.2 Baseline Results on Reporting Effects

We first estimate effects by pooling births all around the end of tax years 2011 to 2018—focusing on recent years with elevated self-employment reporting rates. We begin by testing for differences in *third-party* reported income around the discontinuity at the end of the year in Table 1. The results in Column (1) show no effect on the propensity to have employment income reported by firms on a W-2 form. Importantly, the results in Column (2) also show no effect on the propensity to have self-employment income from independent contracting work reported by firms on 1099 forms. Further, Columns (3) and (4) show that there is also no impact on the *level* of W-2 reported wage and salary earnings in the tax unit.

As noted by LaLumia, Sallee and Turner (2015[20]), the benefits of strategically 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. Thus, 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” similar to Chetty, Friedman and Saez (2013[6]) 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 their EITC, then the wages of parents immediately to the left of the discontinuity should predict higher credits than the wages of those to the right. However, the differences in simulated credits at the discontinuity, displayed in Column (5) of Table 1 and graphically in Panel A of Figure 4, are negligible and statistically indistinguishable from zero. Moreover, the results in Column (4) show that individuals are equally likely to have wages in the EITC phase-in region on both sides of the cutoff. These results support the claim that labor supply is fixed around the cutoff.

[23] We use robust standard errors rather than clustering standard annexs by date. Kolesár and Rothe (2018[28]) show that SEs clustered by a discrete running variable “do not guard against model misspecification, and [...] have poor coverage properties”.
[24] We examine all parents with known births. However some births in the data are missing identifiers for one or both parents. Missing identifiers are more common for fathers than for mothers in the birth data.

Accordingly, we can measure the difference in individuals' incentives to report SE income around the cutoff given their wage earnings. We measure the incentive to self-report self-employment income using the effective federal marginal tax rate (inclusive of SECA taxes) that individuals would face on a first dollar of self-employment income beyond their firm-reported W-2 earnings, given the year their child was actually born.[25] Panel B of Figure 4 shows that 25 percent of individuals with December births—but none with January births—face negative marginal tax rates on reported self-employed. We report point estimates separately for individuals with W-2 wages above and below the first EITC kink point in the first two columns of Appendix Table B.1. We find that this change in reporting incentives occurs only in those with wages below the first EITC kink, among whom a December birth virtually always makes the MTR negative (and 30 percentage points lower on net).

While these reporting incentives exist in theory, the returns on reporting can only be actualized if an individual files a tax return and claims their child as an eligible dependent. We examine effects on claiming behavior in Columns 3-6 of Appendix Table B.1 and within narrower wage bins in Figure A.4. Notably, a first birth only increases their propensity to claim a child on their return by about 60 percentage point.[26] This is partly because individuals are not able to claim children as dependents if they are claimed by another filer (e.g. another parent on a different return or a grandparent), and partly because some individuals already claimed children as dependents prior to the first birth identified in the SSA data.[27] This implies our first-birth RDD identifies an "intent to treat" effect rather than the direct effect of having a claimable child. Hence, the true effect of having an eligible child is larger by a factor of 1/0.6 = 1.7. In contrast, the propensity to report children around a first birth is markedly lower among individuals with wages below the first EITC kink. This low claiming rate also reflects much lower filing rates among low earners—likely including many individuals who do not file despite being eligible for a net refund.[28]

Panel C of Figure 4 presents our main results on self-employment reporting. There is a clear break in the propensity to self-report self-employment income after December 31. Notably, the relationship between date of birth and reported self-employment is completely flat away from the discontinuity, further suggesting that the result is not driven by local sorting around the cutoff date. By contrast, though the incentives are the same. There is no change in the propensity to have 1099 non-employee compensation (reported by *firms*) around the same cutoff date in Panel D.

Panel A of Table 2 reports baseline RDD estimates of the response of self-employment reporting to reporting incentives. The estimates in Column 1 indicate that, at the discontinuity, having a qualifying child for EITC determination increases the probability of reporting self-employment income by 1.34 percentage points (over a baseline rate of 7.19 percentage points). Given our prior results, these estimates can be interpreted as pure reporting effects. The result in Column (2) confirms that there is no "placebo" effect on self-employment in the tax year prior

[25] We calculate this for all individuals regardless of filing status based solely on tax-unit-level W-2 reported earnings and the date of birth of the child using TAXSIM.

[26] Note that there is an effect on 1040 filing for individuals with wages slightly above the first EITC kink point, despite there being *self-employment* reporting incentives or reporting effects for those individuals. This is expected—these individuals qualify for a refundable EITC benefit given their wages so long as their children were born during the tax year, but must file a 1040 return in order to claim that benefit.

[27] This would occur, for example, if an individual had married another person with a child from a previous marriage. This also could occur due to incorrect identification of first births due to missing SSA data on one's first child's parentage.

[28] A leading reason for low EITC take-up is that many eligible individuals do not file a tax return. For instance, in tax year 2016, only single filers with wages above $10,350 are legally required to file a 1040 ($20,700 for married filers). Recent work by Goldin et al. (2022[29]) finds that simply informing EITC-eligible individuals about free tax preparation services increases their propensity to file a 1040 and benefit from the EITC.

to tax year *t*.[29] We explore the robustness of our baseline estimates to alternative RDD specifications in Appendix Figure A.5. The estimates are not sensitive to the inclusion of a "donut-hole" and are highly stable across a wide range of bandwidths. We also find that our results are robust to including a quadratic polynomial in the running variable with slope breaks around the cutoff, though overfitting the quadratic on a small number of days results in unstable point estimates in specifications with smaller bandwidths.

Importantly, the result in Column (5) of Table 2 implies that while some individuals who begin reporting self-employment due to the incentive "bunch" their earnings within $500 of the first EITC kink, this sort of "sharp-bunching" behavior accounts for less than one-fifth of the baseline SE reporting effect. In addition, although we found no effects of reporting incentives on the propensity to have contract income reported on a 1099, the result in Column (6) implies that over a quarter of the baseline self-employment reporting effect comes from individuals who have 1099 contract income but who nonetheless would not have reported any self-employment earnings on their 1040.

The results in Panels B and C of Table 2 confirm that these reporting responses are concentrated among individuals with a strict incentive to do so. As predicted, we find *no* effect of self-employment reporting for individuals with wages above the EITC phase-in range. By contrast, we find that the presence of incentives increases the share reporting self-employment earnings by 4.6 percentage points (over a base rate of 10.8 percentage points) among individuals with W-2 wages below the first EITC kink (including individuals with no W-2 wages). In Figure 5, we estimate effects separately for individuals within $2000 bins of year-*t* tax unit W-2 wages measured in constant 2015 dollars. The results in the figure show that the reporting effect diminishes as W-2 wages approach the first EITC kink and disappears completely above the kink, at which point individuals face strictly positive marginal tax rates on any reported self-employment earnings.

In Columns (7) and (8), we find that individuals with wages below the first EITC kink report $509 more in self-employment income ($647 more including their spouse) if their child is born in December. How much do these individuals gain from reporting those self-employment proceeds? To quantify the tax benefit from reporting self-employment income, we calculate each individual's tax burden first using their own and their spouse's wages and self-employment earnings, and then based on wages alone. We then take the difference to measure the net reduction in one's net tax burden (inclusive of refundable credits).[30] For individuals without EITC-eligible children, this "gain" is negative, as individuals pay higher taxes as a result of reporting additional earned income. However, the results in Column (9) show that the realized gains from reporting self-employment increase for those with wages below the first EITC kink and EITC-eligible children. On average, these individuals receive net subsidies for the self-employment reported on their 1040.[31] Dividing the average increase in the propensity SE below the kink (4.6 percentage points), by the average benefit ($448.5), implies a SE reporting response of 0.01 percentage point for every $1 in incentive.

[29] The magnitude of the reporting effect in Column (1) is remarkably similar to the post-2010 observed increase in self-employment around first births in our event-study analysis in Annex C, suggesting that the those changes likely reflect pure reporting behavior given fixed labor supply.

[30] In this exercise, we use TAXSIM to estimate the federal tax (including FICA/SECA taxes) owed on combined tax unit W-2 and Schedule SE earnings, or W-2 earning alone, given each individual's marital status and number of claimed dependents. For computational reasons, we do not incorporate additional information on deductions or other types of income.

[31] The effect for individuals with wages above the EITC phase-in range in Column (9) is negative despite there being no self-employment reporting effect in Column (1)—this simply reflects the bite of the EITC phase-out for individuals with qualifying children.

We note that the regression discontinuity design only identifies the *short-run* reporting response to one's initial, and unexpected, exposure to negative marginal tax rates. However, continued exposure to those same incentives over time might lead to larger shifts in reporting behavior as individuals become aware of the incentives and how to optimally respond to them. One piece of suggestive evidence comes from examining differences in behavior in the *subsequent* tax year $t + 1$, when *both* children born in December of tax year t and those born in January of t + 1 count towards EITC and CTC determination in tax year $t + 1$. In year $t + 1$, individuals on both sides of the year t cutoff face *identical* incentives in the following year $t + 1$. However, if there is learning-by-doing—that is, if a prior year of eligibility better prepare ones to maximize their refund in the future—then accumulated experience could lead individuals with December births in year t may still respond more than individuals with January births in $t + 1$.[32] Consistent with continued learning, we find statistically significant reporting effects in year t+1 in Column (4) of Table 2. The effect size is considerable, amounting to about one-quarter of the initial year-t reporting effect.[33] This finding suggests that the short-run effects identified in the baseline specification understate the longer-run impact of tax-code incentives on individuals' self-employment reporting.

4.3 Reporting Effects over Time

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 $t + 1$. Figure 6 contrasts the effects 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 latter effect grows approximately *fivefold* between 2000, when the effect size is 1 percentage point, and 2015, when the effect size is six percentage points. Yet, for this same group, we find no effect on having 1099-reported non-employee compensation in all years. These results indicate that the propensity to report self-employment *conditional* on one's true labor supply is far more responsive to incentives in the tax code in recent years than two decades ago.

4.4 More Over-reporting or Less Under-reporting?

Our results imply that individuals with an incentive to report self-employment income are increasingly likely to do so irrespective of any changes in their work. One potential explanation is that individuals who are not self-employed are more likely to fabricate self-employment income on their tax returns to maximize refunds, artificially inflating the self-employment rate. An alternative explanation is that individuals with actual self-employment income that would have otherwise gone unreported to the IRS are now incentivized to report that income on a tax

32 This comparison identifies learning that increases with years of eligibility (conditional on number of years as a parent) similar to Ramnath and Tong (2017[30]). This comparison does not capture general learning that occurs the longer one has spent as a parent, since the RDD holds this fixed.

33 Unlike our baseline design, which examines reporting behavior after period-t labor supply is sunk, the analysis of outcomes in period t+1 does not hold labor supply fixed. Accordingly, the effect in period t+1 could reflect further increases in reporting behavior or actual increases in self-employment. However, we find no effect on firm-reported 1099 payments in period $t + 1$ (not tabulated)—this is suggestive evidence that the period-t+1 increase in self-employment is a pure reporting effect. We also find negligible effects on reporting in subsequent year, implying that the behavior of the January-birth group fully converges to that of the December-birth group after 2–3 years of eligibility.

return.[34] Our baseline findings are therefore consistent with *either* a rise in over-reporting or a decline in under-reporting.

To shed light on this matter, we examine the audits of a representative stratified random sample of 1040 filers conducted in tax years 2001 and 2006–2014 as part of the IRS's National Research Program (NRP) Individual Income Tax Reporting Compliance Studies. Using data from these audits, we measure how often *individuals* on 1040 returns inaccurately report having self-employment earnings or inaccurately report having *no* self-employment earnings. We plot the propensity of each type of audit result in Figure A.7, separately for individuals with and without an incentive to report self-employment.[35] We find, first, that while audits consistently find that 2–2.5 percent of tax filers individuals have self-employment income but report none, that rate is constant over time and nearly identical for individuals with and without EITC reporting incentives (if incentives were driving behavior, we would expect under-reporting to fall among incentivized individuals). By contrast, we find that while almost no one incorrectly reported self-employment income in 2001, the share doing so has risen substantially over subsequent years, but *only* among individuals with a tax incentive to do so.

These findings suggest that our baseline results do, to some extent, reflect a rising propensity to report self-employment income despite having none. However, the rising rates of such behavior in Figure A.7 can only explain a one percentage point increase in reported self-employment among incentivized individuals, just one-fifth of the rise in reporting among this group over the same period implied by our RDD estimates in Panel B of Figure 6.Indeed, the NRP studies are not necessarily always able or intended to detect a decline in under-reporting of self-employment income. In particular the studies we examine focus on individuals who file a 1040 return, yet it is likely that many individuals with self-employment that is not reported to the IRS do not file a 1040 at all. More generally, NRP audits may detect many types of informal self-employment income when they go unreported—though individuals could legitimately report such income to claim an EITC refund if they chose to do so. Thus, while we do find evidence of a rise in over-reporting of self-employment income, we cannot rule out simultaneous declines in under-reporting.

[34] For example, Bracha and Burke (2021[19]) report that nearly one in three workers have some type of informal income—a rate that far exceeds the share of workers reporting self-employment income to the IRS. Individuals with a tax incentive to report self-employment income may choose to *truthfully* report such informal income that might otherwise be omitted on a tax return

[35] We classify individuals on the basis of their firm-reported W-2 income and the number of eligible children determined by the audit.

[35] We classify individuals on the basis of their firm-reported W-2 income and the number of eligible children determined by the audit.

5 Mechanisms Driving Changes over Time

Our analysis indicates a substantial rise in tax-code-motivated reporting of self-employment income over the past two decades. One potential explanation for the growth in reporting effects over time is increased generosity of tax credits that might amplify the incentives to report. To test this explanation, we examine how the impact on post-wage MTRs has evolved over time by estimating the specifications in Column (5) of Table B.1 for each year 2000–2018. The results, plotted in A.6, show that the only major change in self-employment reporting incentives for families with one child occurred in 2009 when the Child Tax Credit was expanded, which lowered the federal MTR for low-wage families from about -25% to about -30%. This result holds even when incorporating state EITC expansions in our analysis in Panel B. While this expansion may have impacted reporting behavior, we observe that the vast majority of the increase in reporting effects apparent in Figure 6 occurs from 2000 to 2007, prior to the announcement of the 2009 reform. Accordingly, it appears that changes in policy cannot account for the observed increase in self-employment reporting effects observed over this period.

Even during periods when policy remains fixed, *knowledge* of the incentives created by the EITC and related tax provisions can differ across regions and spread gradually over time (Chetty, Friedman and Saez, 2013[6]). This process of gradual spread of awareness of incentives in the tax code 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[6]; Mortenson and Whitten, 2020[8]). Chetty, Friedman and Saez (2013[6]) 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. While the prevalence of hyper-strategic "sharp bunching" is a useful measure of knowledge, knowledge of the tax code can lead to increases in self-employment reporting that are not associated with bunching behavior. Indeed, our results in Table 2 imply that sharp bunching behavior accounts for only a small fraction of reporting responses to tax incentives.

To directly 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 1 that allows the effect to vary with local knowledge, proxied by the share of sharp bunchers in each ZIP3 in each year.[36] To avoid any mechanical dependence, we tabulate the

[36] We replicate the sharp-bunching measure from Chetty, Friedman and Saez (2013[6]) and extend it through 2018. This measure is defined as the share of 1040 filers with children in a 3-digit ZIP code and year with AGI under $50,000 in adjusted 2010 dollars who report self-employment earnings and total earned income within $1,000 of the first EITC kink for that year. The resulting aggregate series is presented in Appendix Figure A.8. ZIP codes for individuals in the RDD sample are taken from W-2s and 1099-MISCs in the current year, or most recent of the prior two years if missing.

sharp-bunching share in each ZIP3 in each year following Chetty, Friedman and Saez (2013[6]) only among individuals *without* new births in the December at the end of the tax year or in the following January. We then fully interact the estimating equation:

$$\begin{aligned}\Delta y_{izmt} =& \alpha + \beta BunchShare_{zmt} \times \mathbf{1}\{date_i \in December\} + \delta BunchShare_{zmt} \\ &+ \gamma \mathbf{1}\{date_i \in December\} + f(date_i, BunchShare_{zmt}) \\ &+ \zeta_{zm} + \phi_{zt} + \epsilon_{izmt}\end{aligned} \quad (2)$$

Here, t denotes the RDD cohort (the tax year corresponding to the year-end December-January pair) and m denotes the calendar month. We interact all terms, including the function of the running variable, with the bunching share for each ZIP3 and tax year. Since nationwide average bunching rates increase systematically over time, we include ZIP-by-calendar-month and cohort-by-calendar month fixed effects in some specifications to isolate only *within-ZIP* variation in bunching rates. The fixed effects absorb the RDD main effect (η) in each ZIP and each cohort, but not the interaction effect (β). We estimate this specification on all observations in our sample with non-missing ZIP codes, pooling all birth cohorts in the December and subsequent January of years 2000 through 2018 to capture the full evolution of knowledge since 2000.

The results, presented in Table 3, 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 parents with older children in the same region, new parents are significantly more likely to report self-employment income in response to changing incentives around their first birth. The results in Column (1) indicate that reporting effects in regions with low knowledge are only a small fraction of the baseline estimates in Table 2, but reporting effects grow significantly with greater local knowledge of tax incentives. The coefficients on the interaction terms are nearly the same in magnitude when including ZIP-month and cohort-month fixed effects. This indicates that the interaction term estimates reflect differential increases in knowledge within regions rather than aggregate time trends. The results in Columns (3) and (4) show that as knowledge spreads, new parents exposed to incentives are more likely to bunch at the first EITC kink, but increases in that type of bunching behavior accounts for less than one-third of the overall increase in self-employment behavior. Reassuringly, there is no RDD interaction effect for individuals with wages above the first EITC kink in Panel C, or any interaction effects on the propensity to have of 1099-reported payments or to sort around the cutoff based on the EITC value of wages in Columns (5)-(8).

The magnitude of the interaction term estimates in Table 2 are large enough to explain most of the increase in RDD effects over time in Figure 6 . The estimates in Columns (1) and (2) imply that an increase in knowledge corresponding to a 3 percentage point rise in sharp bunching—approximately the national increase observed in Appendix Figure A.8—should increase the RDD effect among incentivized individuals by 2.5–3 percentage points. Accordingly, local knowledge spread can account for the majority of the roughly 4 percentage point rise in the RDD estimates from 2000 to 2018 in Panel B of Figure 6. However, this exercise only reflects *localized* knowledge transmission; any other knowledge spread that is not mediated by ZIP codes (e.g. broad-based dissemination via the internet or broader social networks) would not be captured by our analysis. 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, as we found above. In that case, the spread in local knowledge might have a greater impact on self-employment reporting for the average eligible individual with children more than captured by short-run effects after their first birth. For comparison, Appendix Table B.2 estimates panel regressions of ZIP-3 self-employment reporting rates on the ZIP3 bunching measure used in the interacted RDD specification.37 For comparability to the estimates in Column (2) of Table 3, we include ZIP and year fixed effects and weight by population. Whereas a 1 percentage point increase in bunching increased self-employment reporting rates among incentivized individuals by less than 1 percentage point in our interacted RDD analysis, we find in Appendix Table B.1 that the same increase in bunching is associated with 3.6 percentage point higher self-employment reporting rates among all incentivized individuals with children. While the ZIP-level regressions cannot isolate reporting effects from labor supply effects as the RDD estimates do, we nonetheless find no effect on the propensity to have 1099-reported contract income and only negligible impacts on self-employment reporting by individuals not facing reporting incentives (with high wages or no children), consistent with pure reporting effects. This suggests that spreading local knowledge could increase overall self-employment reporting rates among incentivized individuals over time 3–4 times as much as implied by the growth of our RDD estimates in Figure 6.

[37] In this analysis, ZIP-3 self-employment rates are defined as a share of all individuals in the workforce (among whom ZIP codes are always observed), not as a share of all individuals with children, as we do not observe ZIP codes for all individuals with children. For comparison, 93% of the RDD sample with non-missing ZIP codes in Table 3 are in the workforce by our definition.

6 Quantification and Counterfactuals

6.1 Accounting for Changes in Reporting Behavior

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 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. In this section we 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 and tax unit W-2 wages below the first EITC kink who would not have reported self-employment had they not faced negative marginal tax rates.[38] We then consider how much self-employment reporting rates would have grown by using the annual effect estimates presented in Figure 6 to account for those only reporting self-employment due to incentives. Specifically, our adjustment is given as follows:

$$\frac{SE_t^*}{WF_t^*}^{RDD} = \frac{SE_t - \beta_t^{RDD,\Delta AnySE,BTK} \cdot POP_t^{k,BTK}}{WF_t - \beta_t^{RDD,\Delta SEOnly,BTK} \cdot POP_t^{k,BTK}} \tag{3}$$

where $\beta_t^{RDD,\ \Delta AnySE,BTK}$ and $\beta_t^{RDD,\ \Delta SEOnly,\ BTK}$ are our RDD coefficients for reporting any self-employment and only self-employment, respectively, among those below the kink, 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) and who have wages below the EITC kink, $POP_t^{k,BTK}$, and also adjust the denominator to remove individuals only in our workforce measure due to the incentive.

[38] To reduce noise, we impose the effect on individuals with W-2 wages above the kink is zero in all our years; this is consistent with our estimates in Figure 6.

To implement this adjustment, we use population-level estimates of individuals with children under 19 in their household from the Social Security birth records. We then apply our annual regression coefficients to estimate the implied number of individuals who would not have reported self-employment if they had not been incentivized to do so. The result of this exercise is presented as "Scenario 1" in Figure 7a. This adjustment reduces the share of the workforce reporting self-employment by about half of one percentage point in 2014, about one quarter of the total increase since 2000.

The evidence presented above indicates that 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. Our results in Column 4 of Table 2 indicate 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 Appendix Table B.2 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.

To account for how changes in awareness of incentives may have impacted reporting behavior more broadly, 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.[39]

Under this second counterfactual scenario, presented as "Scenario 2" in Panel A of Figure 7a, 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. 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

[39] Formally, 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:

$$SE_t^* = \sum_g \omega_{g,t} \sum_k \left(\sum_{w \le w^*(k)} \zeta^*_{\mathbf{w},\mathbf{g},\mathbf{k},\mathbf{t}} SE_{w,g,\mathbf{0},t} + \sum_{w > w^*(k)} \zeta^*_{\mathbf{w},\mathbf{g},\mathbf{k},\mathbf{t}} SE_{w,g,k,t} \right) \quad (4)$$

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, *w* the wage bin for a tax unit's total wages,[41] and $\omega_{g,t}$ is *g*'s share of the overall workforce in year *t*. Individuals 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 $w < w^*(k = 2+)$). We further refine our definition to exclude individuals with total earnings (wages+SE) above the phase-out region. 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. To calculate the counterfactual number in the workforce in the absence of the incentive, 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). 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.10. We recalculate the weights of our wage bins using this counterfactual level, and denote these adjusted weights $\zeta^*_{w,g,k,t}$.

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.

Crucially, when we perform the same adjustments to the share of the workforce with third-party-reported independent contractor earnings on 1099 forms in Figure 7b there is essentially no impact on the series. This provides important validation for the use of individuals without children within narrow wage bins as a counterfactual for incentivized reporters with children. In particular, this implies that *firm-reported* independent contracting activity among low-wage individuals with kids evolved in parallel to rates for other individuals since 2000; only *self-reported* earnings trended differently. Moreover, we this counterfactual adjustment dramatically reduces the divergence in trends between self-reported self-employment and firm-reported non-employee compensation. After the adjustment, a majority of the remaining increase in self-employment reported on schedule SE can be accounted for the increased prevalence of 1099-reported contractor payments between 2000 and 2005. As noted by Abraham et al. (2021[21]), it is plausible that this early trend in independent contracting is not reflected in surveys like the CPS to the extent that such workers do not perceive themselves as self-employed.

In Appendix Table B.3, we report what the implied changes to the measured *level* of the self-employment rate would be if individuals were entirely unaware of tax incentives to report self-employment in all years. These level adjustments 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.[40]

6.2 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 propensity for under-reporting and changes in the demographic composition of the workforce.

First, we assess the concern raised in Section 3 that trends in SE income 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 SE adjustment incorporating these trends is described in Annex D and presented in Column (4) of Table B.3 and as “Scenario 3” in Figure D.2a. We find our measure of SE underreporting 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. In the end, however, 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, which we take at face value for this exercise.

[40] However, we note that this does not necessarily correspond to a “true” self-employment rate. As discussed above in Section 4.4, although 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.

As documented in Collins et al. (2019[13]), older workers—particularly retirement-age workers over 65—are dramatically more likely to have self-employment income and non-employee compensation reported on 1099 forms conditional on having any labor income, in all years since 2000. 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. Further details are described in Annex D and this adjustment is presented as Column (5) in Table B.3 and "Scenario 4" in Figure D.2b. 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 75% of the observed increase since 2000. Less than one half of a percentage point increase since 2000 remains unexplained.

6.3 Application: Reporting Over the Business Cycle

Our second adjustment 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.11. Panel (a) shows the raw percentage point growth in schedule SE by state. The deeper red indicates more growth. We see 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[6]). Finally, Panel C shows what our adjusted series under Scenario 2 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.

To illustrate the importance of adjusting for reporting behavior, we close by examining the state-level relationship between self-employment reported on tax returns and macroeconomic conditions. 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.[41]

Tax data has potential benefits for studying this question. 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, the reporting issues highlighted above may become more severe during economic downturns, as the number of individuals facing incentives to report self-employment to increase tax refunds is also countercyclical. As shown in Appendix Figure A.9, as the wage distribution shifts down, more families have wages fall into the credit phase-in range.

[41] Parker (2008[31]) provides a review of the literature, and Fairlie (2013[32]) provides a more recent update examining the Great Recession. Farber (2015[33]) and Katz and Krueger (2017[24]) focus on alternative work as captured in the Contingent Worker Supplement (CWS) to the CPS.

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:

$$y_{s,t} = \beta UR_{s,t} + \alpha_s + \alpha_t + e_{s,t} \tag{5}$$

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.

The results, reported in Table 4a, show that the share of workers with self-employment income is significantly countercyclical in raw tax 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), 1099s (Column 3) or using self-employment from the CPS (Column 4).

As an alternative approach, we examine how the prevalence of each type of work in the broader adult population varies over the business cycle in Table 4b. The interpretation of our coefficients becomes the log point change in the employment-to-population ratio (approximately a percent change) for a percentage point increase in the unemployment rate.[42] (The denominator of the dependent variable is the same across these specifications, only the numerator is changing). We find that while the number of adults with wage income reported on W-2s (Column 1) or contractor compensation reported on 1099s (Column 5) is significantly pro-cyclical, 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 W-2s 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.[43]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 information on tax filings.

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

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

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 U.S. tax data.

Our results caution against trusting trends in administrative data over trends in survey data by default. Nonetheless, administrative records like tax data can be powerful tools for measuring labor-market trends so long as reporting incentives are kept in mind. This paper shows that these incentives are measurable and can be accounted for. 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 tax data, increases in gig work and secondary forms of contract work are being well captured in firm-reported tax 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. et al. (2021), "Driving the Gig Economy", https://www.aeaweb.org/aea/2022conference/program/pdf/14036_paper_DDZKbDhn.pdf?display (accessed on 28 June 2022). [10]

Abraham, K. et al. (2021), "Reconciling survey and administrative measures of self-employment", *Journal of Labor Economics*, Vol. 39/4, pp. 825-860, https://doi.org/10.1086/712187/SUPPL_FILE/20079DATA.ZIP. [21]

Abraham, K. et al. (2017), "Measuring the Gig Economy: Current Knowledge and Open Issues", in Corrado, C. et al. (eds.), *Measuring and Accounting for Innovation in the 21st Century*, University of Chicago Press, http://www.nber.org/chapters/c13887 (accessed on 6 June 2018). [18]

Abraham, K. et al. (2013), "Exploring differences in employment between household and establishment data", *Journal of Labor Economics*, Vol. 31/2 PART2, pp. S129-S172, https://doi.org/10.1086/669062/0. [23]

Abraham, K., B. Hershbein and S. Houseman (2020), *Most Self-Employed Workers Are Independent Contractors Most Self-Employed Workers Are Independent Contractors*, W.E. Upjohn Institute for Employment Research , Kalamazoo, MI, https://doi.org/10.17848/pb2020-24. [2]

Blank, R., K. Charles and J. Sallee (2009), "A Cautionary Tale about the Use of Administrative Data: Evidence from Age of Marriage Laws", *American Economic Journal: Applied Economics*, Vol. 1/2, p. 149, https://doi.org/10.1257/APP.1.2.128. [15]

Bracha, A. and M. Burke (2021), "How Big is the Gig? The Extensive Margin, The Intensive Margin, and The Hidden Margin", *Labour Economics*, Vol. 69, https://doi.org/10.1016/J.LABECO.2021.101974. [19]

Chetty, R., J. Friedman and E. Saez (2013), "Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings", *American Economic Review*, Vol. 103/7, pp. 2683-2721, https://doi.org/10.1257/aer.103.7.2683. [6]

Collins, B. et al. (2019), "Is Gig Work Replacing Traditional Employment? Evidence from Two Decades of Tax Returns", *SOI Working Paper*, Internal Revenue Service. [13]

Dickert-Conlin, S. and A. Chandra (1999), "Taxes and the timing of births", *Journal of Political Economy*, Vol. 107/1, pp. 161-177, https://doi.org/10.1086/250054/0. [27]

Fairlie, R. (2013), “Entrepreneurship, Economic Conditions, and the Great Recession”, *Journal of Economics & Management Strategy*, Vol. 22/2, pp. 207-231, https://doi.org/10.1111/JEMS.12017. [32]

Farber, H. (2015), “Alternative and Part-Time Employment Arrangements as a Response to Job Loss”, *https://doi.org/10.1086/209946*, Vol. 17/4, https://doi.org/10.1086/209946. [33]

Farrell, D., F. Greig and A. Hamoudi (2018), *The Online Platform Economy in 2018: Drivers, Workers, Sellers, and Lessors*, JPMorgan Chase Institute, http://www.jpmorganchaseinstitute.com (accessed on 27 June 2022). [11]

Feldman, N., P. Katuščák and L. Kawano (2016), “Taxpayer Confusion: Evidence from the Child Tax Credit”, *American Economic Review*, Vol. 106/3, pp. 807-35, https://doi.org/10.1257/AER.20131189. [25]

Goldin, J. et al. (2022), “Tax filing and take-up: Experimental evidence on tax preparation outreach and benefit claiming”, *Journal of Public Economics*, Vol. 206, https://doi.org/10.1016/J.JPUBECO.2021.104550. [29]

Harris, S. and A. Krueger (2015), “A Proposal for Modernizing Labor Laws for Twenty-First-Century Work: The "Independent Worker"”, *Discussion Papers*, No. 2015-10, The Hamilton Project, http://www.hamiltonproject.org/assets/files/modernizing_labor_laws_for_twenty_first_century_work_krueger_harris.pdf (accessed on 31 May 2018). [9]

Hurst, E. and B. Pugsley (2011), “What Do Small Businesses Do?”, *Brookings Papers on Economic Activity*, Vol. 42/2, pp. 73-142, https://www.brookings.edu/wp-content/uploads/2011/09/2011b_bpea_hurst.pdf (accessed on 27 June 2022). [16]

Jackson, E., A. Looney and S. Ramnath (2017), “The Rise of Alternative Work Arrangements: Evidence and Implications for Tax Filing and Benefit Coverage”, *Working Papers*, No. 114, Office of Tax Analysis, https://www.treasury.gov/resource-center/tax-policy/tax-analysis/Documents/WP-114.pdf (accessed on 31 May 2018). [12]

Katz, L. and A. Krueger (2018), “The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015:”, *ILR Review*, Vol. 72/2, pp. 382-416, https://doi.org/10.1177/0019793918820008. [1]

Katz, L. and A. Krueger (2017), “The Role of Unemployment in the Rise in Alternative Work Arrangements”, *American Economic Review*, Vol. 107/5, p. 392, https://doi.org/10.1257/AER.P20171092. [24]

Katz, L. and A. Krueger (2016), “The Rise and Nature of Alternative Work Arrangements in the United States”, *NBER Working Paper Series*, No. 22667, NBER, http://www.nber.org/papers/w22667.pdf (accessed on 31 May 2018). [22]

Kleven, H. et al. (2011), “Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark”, *Econometrica*, Vol. 79/3, pp. 651-692, https://doi.org/10.3982/ECTA9113. [26]

Kolesár, M. and C. Rothe (2018), “Inference in Regression Discontinuity Designs with a Discrete Running Variable”, *American Economic Review*, Vol. 108/8, pp. 2277-2304, https://doi.org/10.1257/AER.20160945. [28]

Kuka, E. (2013), “EITC and the Self-employed: Real or Reporting Effects?”, *Public Finance Review*, Vol. 42/6, pp. 691-719, https://doi.org/10.1177/1091142113496130. [7]

LaLumia, S. (2021), “The Earned Income Tax Credit and Reported Self-Employment Income”, *National Tax Journal*, Vol. 62/2, pp. 191-217, https://doi.org/10.17310/NTJ.2009.2.01. [4]

LaLumia, S., J. Sallee and N. Turner (2015), “New Evidence on Taxes and the Timing of Birth”, *American Economic Journal: Economic Policy*, Vol. 7/2, p. 293, https://doi.org/10.1257/POL.20130243. [20]

Lim, K. et al. (2019), “Independent Contractors in the U.S.: New Trends from 15 years of Administrative Tax Data”, *SOI Working Paper*, Internal Revenue Service. [14]

Mortenson, J. and A. Whitten (2020), “Bunching to Maximize Tax Credits: Evidence from Kinks in the US Tax Schedule”, *American Economic Journal: Economic Policy*, Vol. 12/3, p. 432, https://doi.org/10.1257/POL.20180054. [8]

National Academies of Sciences, E. (2020), *Measuring Alternative Work Arrangements for Research and Policy*, National Academies Press, https://doi.org/10.17226/25822. [17]

Parker, S. (2008), *Entrepreneurship, Self-employment and the Labour Market*, Oxford University Press, https://doi.org/10.1093/OXFORDHB/9780199546992.003.0016. [31]

Ramnath, S. and P. Tong (2017), “The Persistent Reduction in Poverty from Filing a Tax Return”, *American Economic Journal: Economic Policy*, Vol. 9/4, p. 394, https://doi.org/10.1257/POL.20150486. [30]

Saez, E. (2010), “Do Taxpayers Bunch at Kink Points?”, *American Economic Journal: Economic Policy*, Vol. 2/3, pp. 180-212, https://doi.org/10.1257/POL.2.3.180. [5]

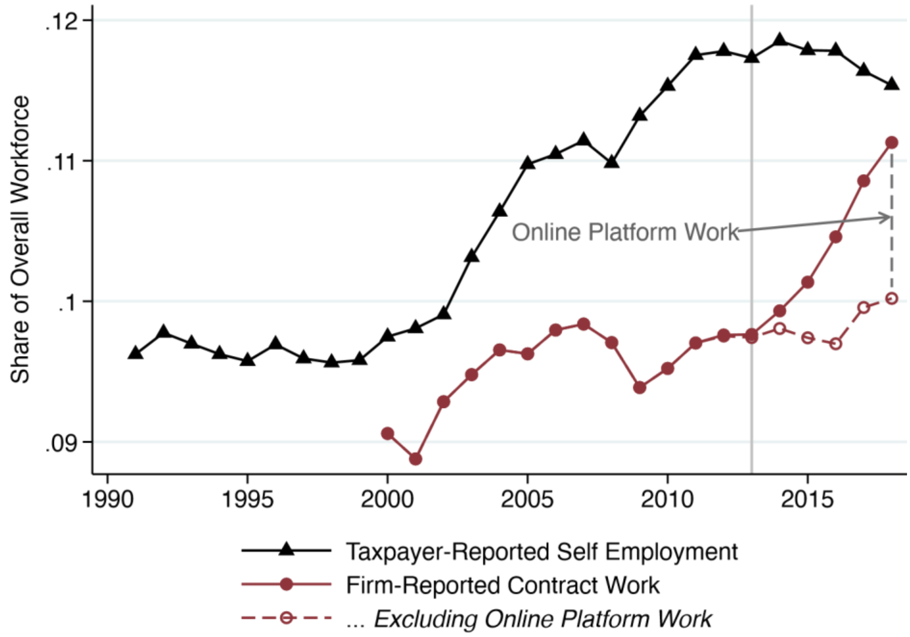
Slemrod, J. (2021), “Caveats to the Research Use of Tax-Return Administrative Data”, *National Tax Journal*, Vol. 69/4, pp. 1003-1020, https://doi.org/10.17310/NTJ.2016.4.13. [3]

Figures

Figure 1. Share of Workforce with Self-Employment in Tax Returns and March CPS

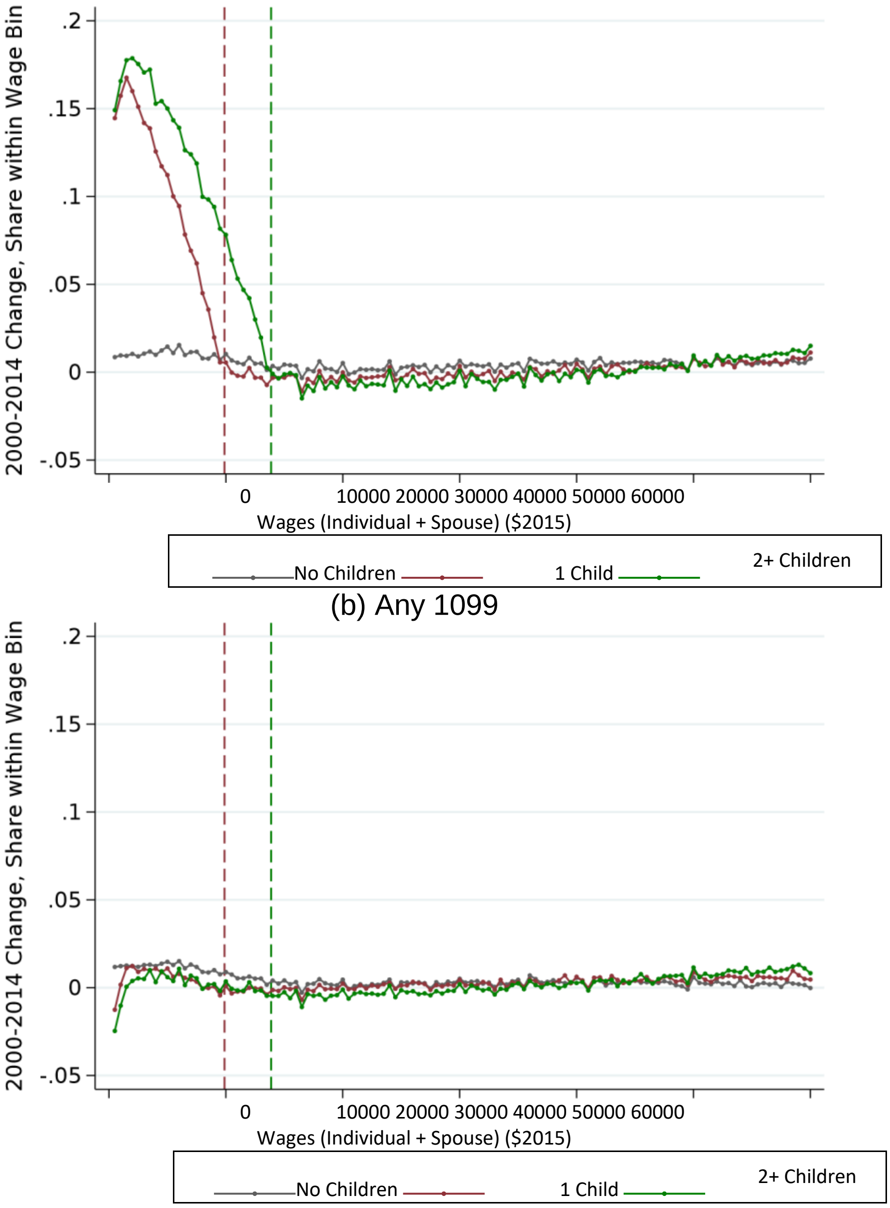
Note: Figure compares the share of the workforce—defined as all individuals with any employment earnings or self-employment earnings during the year—reporting positive self-employment earnings in tax records and in the March supplement to the Current Population Survey. The black line reports this share in tax data, where the workforce is defined as all individuals with positive earnings on either a W-2 return or on Schedule SE during the year, and individuals are classified as self-employed if they report any positive earnings on a Schedule SE. To extend the series before 1999, we draw on tabulate Social Security Administration records available at https://www.ssa.gov/policy/docs/statcomps/supplement/2020/4b.pdf. The maroon line reports the share in the March CPS; for comparability to the tax data, the workforce is defined as all individuals reporting positive wage/salary earnings or self-employment earnings for the specified year (reported retrospectively in the following March), and individuals are classified as self-employed if they report any self-employment income for the year.

Figure 2. Self-Reported and Firm-Reported Self-Employment Earnings



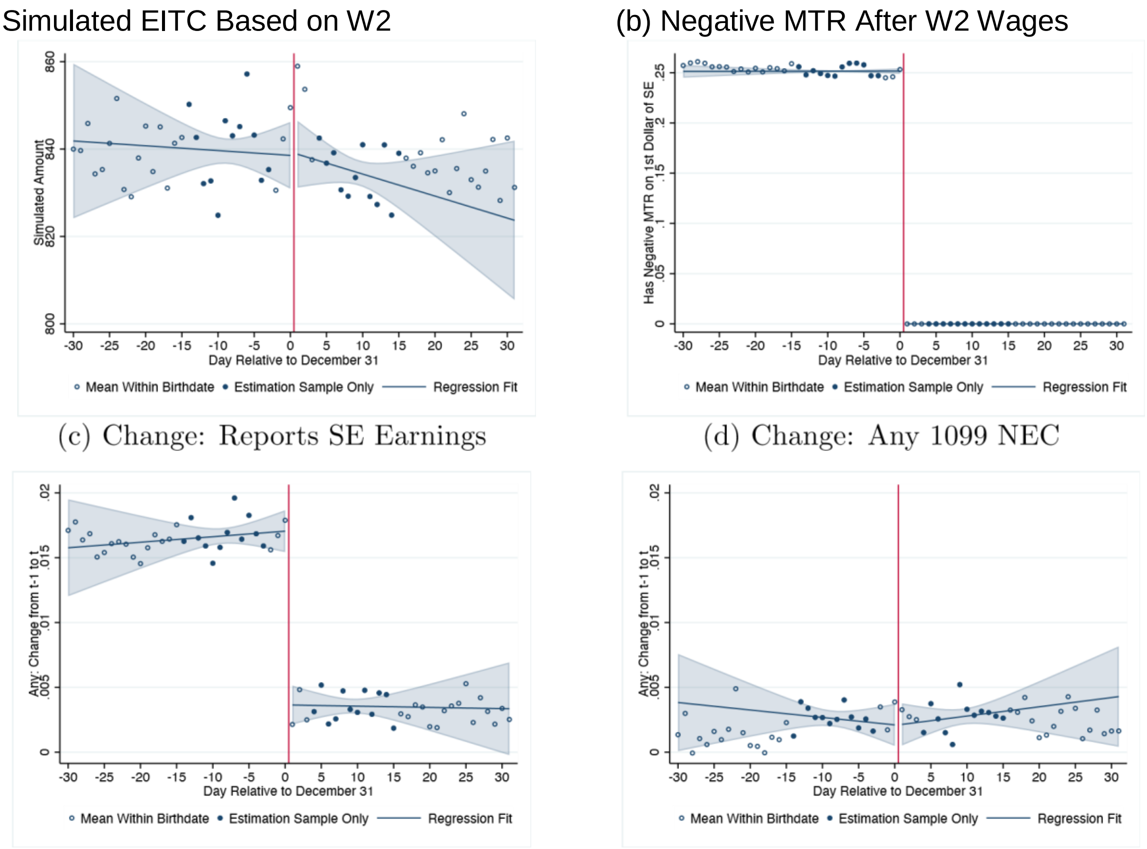
Note: Figure shows the share of individuals in the workforce reporting self-employment income on Form 1040 Schedule SE (same series as in Figure 1) in each year (black line) and the share with firm-reported nonemployee labor compensation exceeding $600 on a 1099 Information Return (maroon line). In both series, the workforce is defined as all individuals with positive earnings on either a W-2 return or on Schedule SE during the year. Following the method in Collins, Garin, Jackson, Koustas, and Payne (2019), we separately break out the subset of independent contractors whose 1099-reported payments come exclusively from online platform economy firms. See Annex E for additional details on data construction. For years prior to 2000, SE workers and the overall workforce are drawn from tabulated SSA records available at https://www.ssa.gov/policy/docs/statcomps/supplement/2020/4b.pdf.

Figure 3. Growth in Reported Self-Employment 2000-2014, by Wage Income



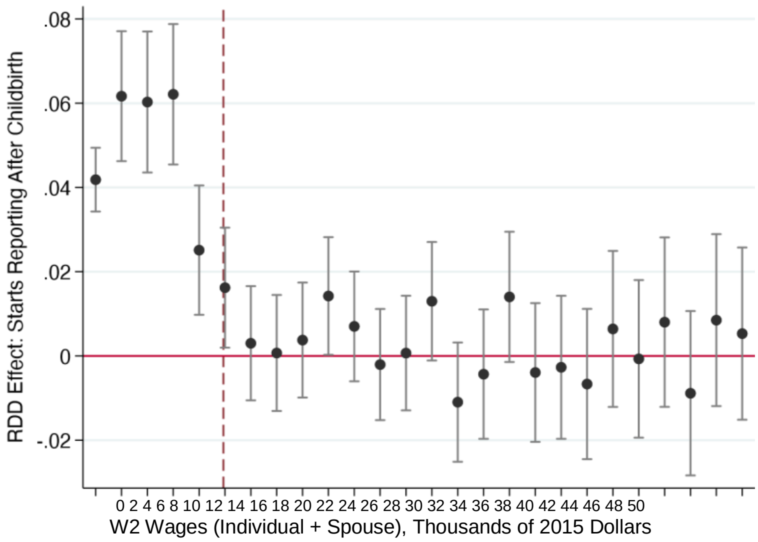
Note: Figure shows the change in propensity to file SE (Panel A) or receive a 1099 Information Return (Panel B), for tax units with wage earnings. Wage earnings, in $500 bins, are reported on the x-axis. Change is calculated between 2000 and 2014. Wages are determined based on W2 information returns. 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 4. Regression Discontinuity Design



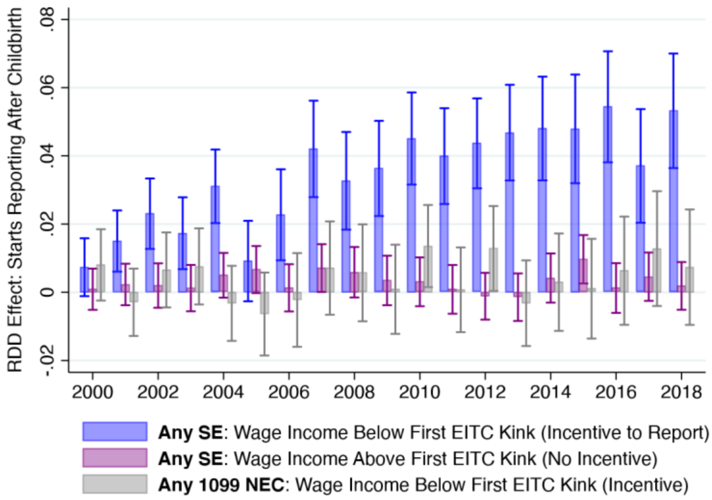
Note: Figure graphically presents results from the baseline regression discontinuity design specification in Equation (1), pooling births in each December 2011-2018 and each subsequent January. Panel (a) examines the change in “simulated EITC” that is the credit each individual would earn if their birth 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) examines whether one would face a negative federal marginal tax rate (including SECA taxes) on a first dollar of self-employment earnings beyond one’s W-2 reported wage/salary earnings and those of any spouse in year t, given the year their child was actually born. Panel (c) examines the change in whether the one reports any Schedule SE earnings in tax year t relative to the prior year t − 1. Panel (d) examines the change in having non-employee income reported on a 1099-MISC in tax year t relative to the prior year t − 1. Each dot is the average outcome for parents with births on the corresponding date, pooled across years. The solid dots are the calendar dates used in the main estimation window. The line segments are the regression fits allowing for the estimated discontinuity between December 31 and January 1, with 95 percent confidence bands from robust standard errors displayed.

Figure 5. RDD Estimates by Tax Unit W-2 Wage Earnings



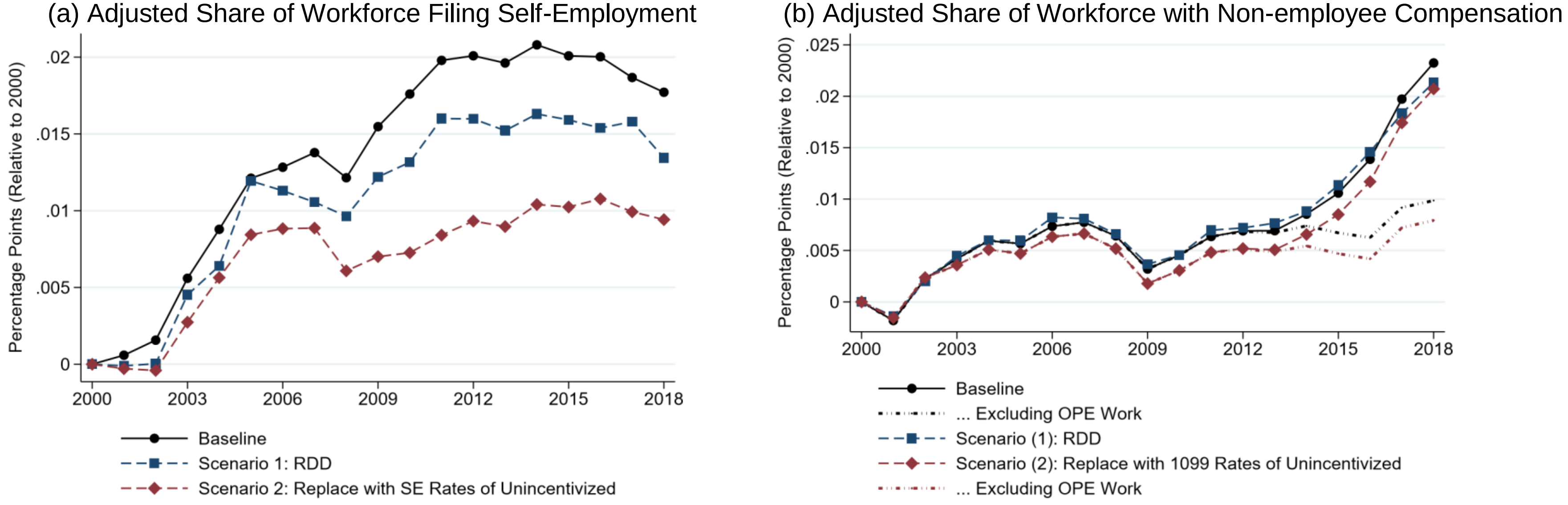
Note: Figure presents results from the regression discontinuity design specification in Equation (1) pooling births in each December 2011-2018 and each subsequent January, estimated separately for individuals within $2000 bins of year-t tax unit (self plus spouse) W2 wages, measured in constant 2015 Dollars. The dashed maroon line is the amount where the first EITC kink occurs for families with one child based on the 2015 schedule.

Figure 6. RDD Estimates by Year



Note: Figure reports our baseline RDD estimates from estimating Specification 1 in the text within individual cohorts. We report separate estimates for individuals with combined wages below the first EITC kink and those with wages above the first EITC kink point. Years correspond to the tax year t, at the end of which the births occur in the corresponding December or January.

Figure 7. Adjusted SE Shares Under Counterfactual Assumptions



Note: Figure presents the percentage point change in the share of the workforce with self-employment income or 1099-reported contract payments relative to 2000, along with counterfactual increases in each series under alternative scenarios. Panel (a) applies these adjustments to self-employment, and Panel (b) to non-employee compensation. “Scenario (1)” adjusts self-employment downward using our annual RDD estimates reported in Figure 6, under the counterfactual assumption that the RDD effects remained constant in all years. “Scenario (2)” reports how overall rates would have evolved if trends for individuals who have incentives to report self-employment followed those rates among comparable individuals without such incentives. Dashed lines in Panel (b) exclude the OPE.

Tables

Table 1. RDD Estimates: Third-Party Reported Earnings

	Δ Any Wages	Δ Any 1099-NEC	Tax Unit W-2 Wages	Tax Unit Wages < 1st Kink	Sim. 1-Child EIC W-2 Only
	(1)	(2)	(3)	(4)	(5)
December Births					
Coeff	0.003	0.00003	46.1	-0.0009	-0.8
	(0.002)	(0.001)	(297.5)	(0.002)	(5.8)
N	1382740	1382740	1382740	1382740	1382740
DV Mean Level, Jan Births	0.8	0.08	55244.5	0.3	834.5

Note: Table displays estimates from the baseline regression discontinuity design specification in Equation (1) on third-party reported earnings. The sample is all individuals with births in the last fifteen days of December of each tax year t in 2011-2018 or the first fifteen days of January immediately following tax year t, omitting births within three days of the start of the new year. Outcomes are from year t or are changes from year t relative to the prior year $t-1$, as specified. Simulated 1-Child EIC Levels are calculated as the EIC amount one would receive if their first child had been born in year t (irrespective of when the true birth occurred) given only one's W-2 reported wage/salary earnings and those of any spouse reported on a 1040. We report mean year t levels of each dependent variable for individuals with first births in January of $t+1$. Robust standard errors are displayed in parentheses.

Table 2. RDD Estimates: Main Effects on Self-Employment Reporting

	∆Any SE	AnySEEarnings			∆AnySE& Sharp Buncher	∆Has1099NEC & Reports SE	∆Individual SE Earnings	∆TaxUnit SE Earnings	TaxBenefitfrom Reporting SE
		Tax Year t-1	Tax Year t	Tax Year t+1					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. All Parents									
Coeff	0.0134**	-0.00101	0.0124**	0.00395**	0.00254**	0.00388**	130.8**	153.6**	90.62**
	(0.00118)	(0.00120)	(0.00127)	(0.00135)	(0.000326)	(0.000950)	(11.57)	(19.53)	(9.372)
N	1382740	1382740	1382740	1382740	1382740	1382740	1382740	1382740	1382740
DV Mean Level, Jan Births	0.0719	0.0683	0.0719	0.0878	0.00203	0.0406	855.4	1770.5	-417.7
Panel B. With Wages < 1st EITC Kink									
Coeff	0.0460**	-0.00139	0.0446**	0.0146**	0.00939**	0.0128**	509.0**	646.5**	448.5**
	(0.00272)	(0.00270)	(0.00315)	(0.00335)	(0.00121)	(0.00197)	(30.10)	(43.90)	(15.40)
N	349240	349240	349240	349240	349240	349240	349240	349240	349240
DV Mean Level, Jan Births	0.108	0.0900	0.108	0.143	0.00804	0.0503	1403.1	2319.4	-404.7
Panel C. With Wages ≥1st EITC Kink									
Coeff	0.00240	-0.000849	0.00155	0.000433	0.000229	0.000858	3.108	-12.85	-30.45**
	(0.00128)	(0.00131)	(0.00131)	(0.00140)	(0.000152)	(0.00108)	(11.59)	(21.44)	(11.37)
N	1033500	1033500	1033500	1033500	1033500	1033500	1033500	1033500	1033500
DV Mean Level, Jan Births	0.0597	0.0610	0.0597	0.0692	0	0.0374	670.6	1585.3	-422.1

Note: Table displays estimates from the baseline regression discontinuity design specification in Equation (1) on third-party reported earnings. The sample is all individuals with births in the last fifteen days of December of each tax year t in 2011-2018 or the first fifteen days of January immediately following tax year t, omitting births within three days of the start of the new year. “Wages < 1st kink” subsample includes all individuals in tax units (self plus spouse if filing a 1040 jointly) with year t wages in the EITC phase-in region for households with one child in that year (irrespective of whether their birth actually occurred in December or January); the complementary subsample includes all other individuals. Outcomes are from year t or are changes from year t relative to the prior year t− 1, as specified. ”Sharp bunchers” are individuals with earning income within $500 of the level where the first EITC kink occurs. ”Net Tax Benefit from Reporting SE” is the increase in net taxes one would pay if both they and their spouse (if present on a 1040) did not report their self-employment earnings to the IRS. We report mean levels of each dependent variable for individuals with first births in January of t + 1 in each subsample; levels are from year t except in columns 2 and 4, which report year t− 1 and t + 1 means, respectively. Robust standard errors are displayed in parentheses.

Table 3. RDD Estimates: Interactions with Knowledge Spread

	Δ Any 1040-SE Income		Δ Any 1040-SE & Sharp Buncher		Δ Any 1099-NEC		Simulated W2-Only One Child EIC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. All Parents								
December Birth	0.00250*	-	-0.000716*	-	0.000982	-	-0.162	-
	(0.00126)	-	(0.000323)	-	(0.00138)	-	(6.197)	-
December Birth × ZIP Bunching	0.337**	0.303**	0.0998**	0.0831**	0.0120	-0.0250	-8.682	44.61
	(0.0453)	(0.0597)	(0.0146)	(0.0183)	(0.0477)	(0.0633)	(198.0)	(266.4)
N	3094399	3094387	3094399	3094387	3094399	3094387	3094399	3094387
Panel B. With Wages < 1st EITC Kink								
December Birth	0.0134**	-	-0.000445	-	0.00335	-	8.664	-
	(0.00336)	-	(0.00130)	-	(0.00340)	-	(11.71)	-
December Birth × ZIP Bunching	0.999**	0.868**	0.311**	0.254**	0.0788	-0.0616	-221.5	344.4
	(0.123)	(0.143)	(0.0526)	(0.0665)	(0.112)	(0.144)	(339.5)	(467.4)
N	665093	665074	665093	665074	665093	665074	665093	665074
Panel C. With Wages ≥1st EITC Kink								
December Birth	0.00151	-	-0.000294	-	0.000666	-	-1.938	-
	(0.00132)	-	(0.000151)	-	(0.00149)	-	(7.200)	-
December Birth × ZIP Bunching	0.0557	0.0506	0.0156*	0.0135	-0.0199	-0.0191	65.15	-108.6
	(0.0457)	(0.0611)	(0.00654)	(0.00800)	(0.0514)	(0.0694)	(238.5)	(320.1)
N	2429306	2429295	2429306	2429295	2429306	2429295	2429306	2429295
Zip × Month FE		X		X		X		X
Cohort × Month FE		X		X		X		X

Note: Table displays estimates from the regression discontinuity specification in Equation (2) in the text including interactions with local sharp-bunching rates at the 3-Digit ZIP code level for each year. Bunching rates are calculated omitting individuals in the RDD sample following Chetty, Friedman, and Saez (2013). 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 in year t. The bunching variable is interacted with the discontinuity and the running-variable slope terms. Columns 2, 4, 6, and 8 include year t (cohort) by birth calendar month and ZIP3 by birth calendar month fixed effects so that all bunching variation in the interaction term comes from within-ZIP changes over time; these FEs absorb the RDD main effect. Robust standard errors are displayed in parentheses. See notes to 2 for additional details.

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	X	X	X	X
Year FE	X	X	X	X

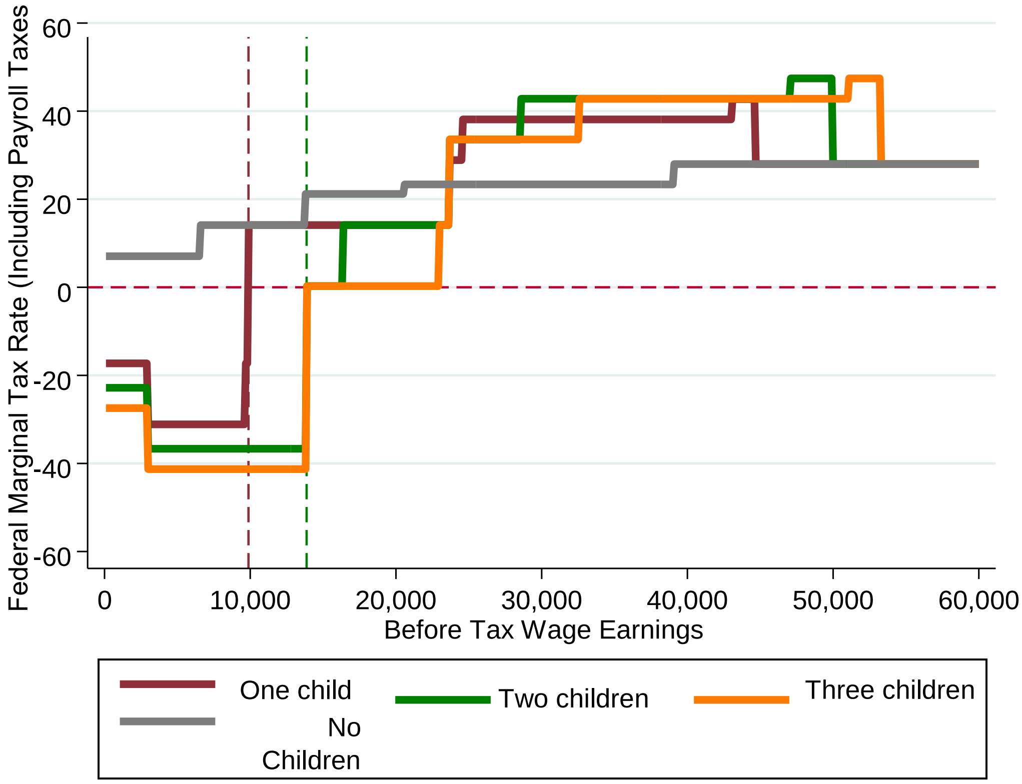
(b) Log employment-to-adult population

	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	X	X	X	X	X	X
Year FE	X	X	X	X	X	X

Note: Table reports the results from running specification (5) in the text for the dependent variable and source indicated by table subtitles and column headers. Standard errors clustered on state are displayed in parentheses.

Annex A. Appendix Figures

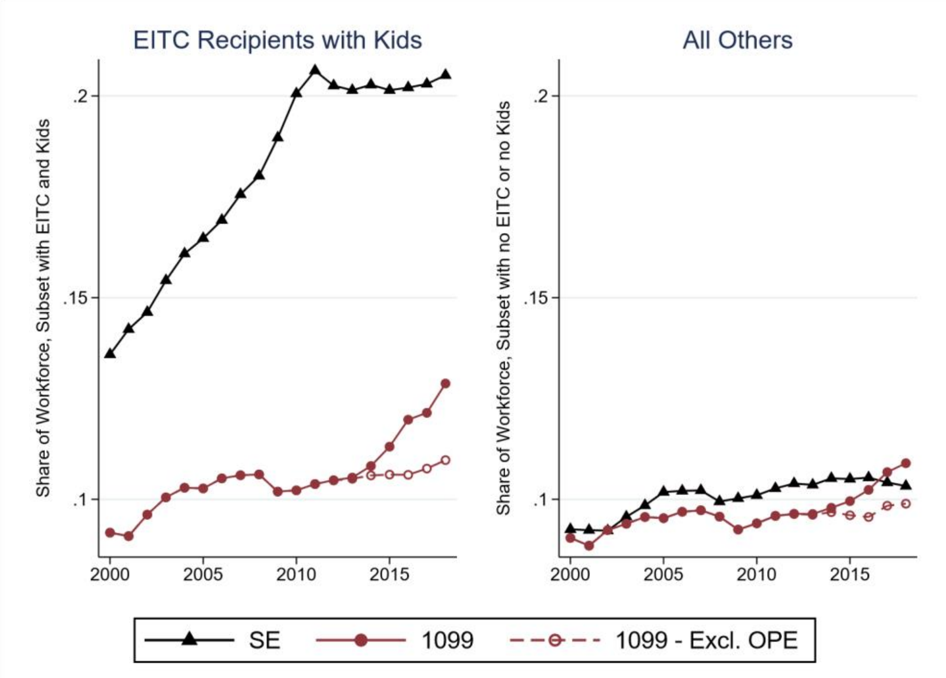
Figure A.1. Effective Federal Marginal Tax Rate for Reporting Additional Dollar of Self-Employment Income



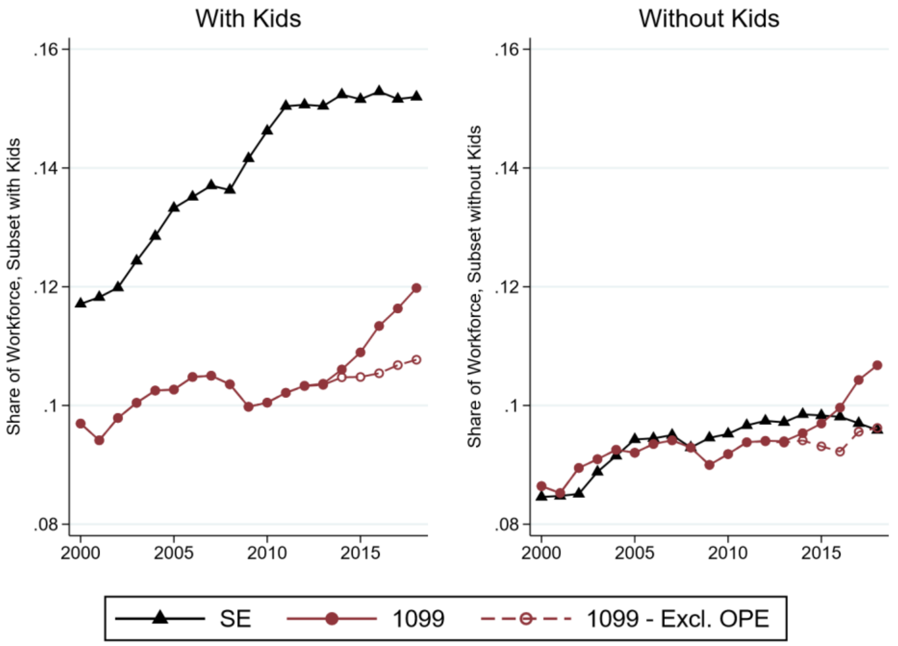
Note: Figure shows the effective marginal tax rate for reporting an additional dollar of self-employment income, for a given level of before-tax wage earnings. Calculation takes into account the full tax schedule in tax year 2015 with no other credits/deductions except the EITC, CTC and standard deductions, 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 and 1099 Information Returns

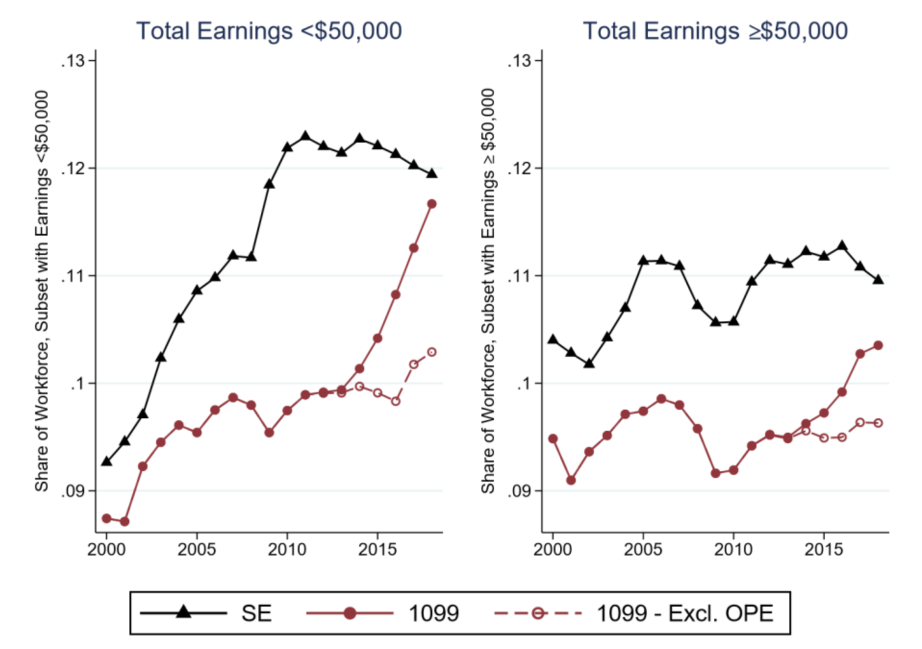
(a) By EITC and Kids



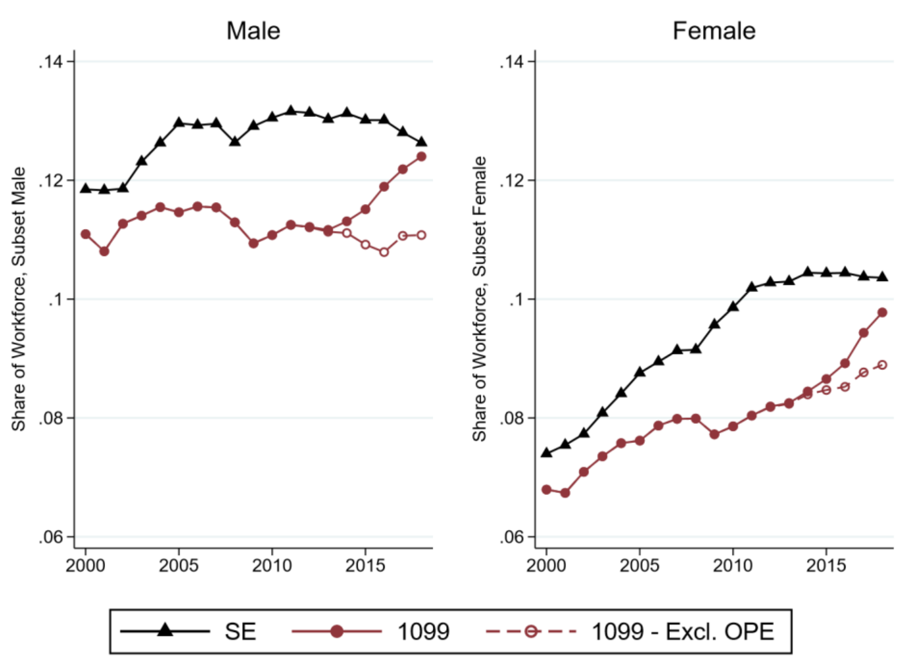
(b) By Presence of Children



(c) By Total Earnings

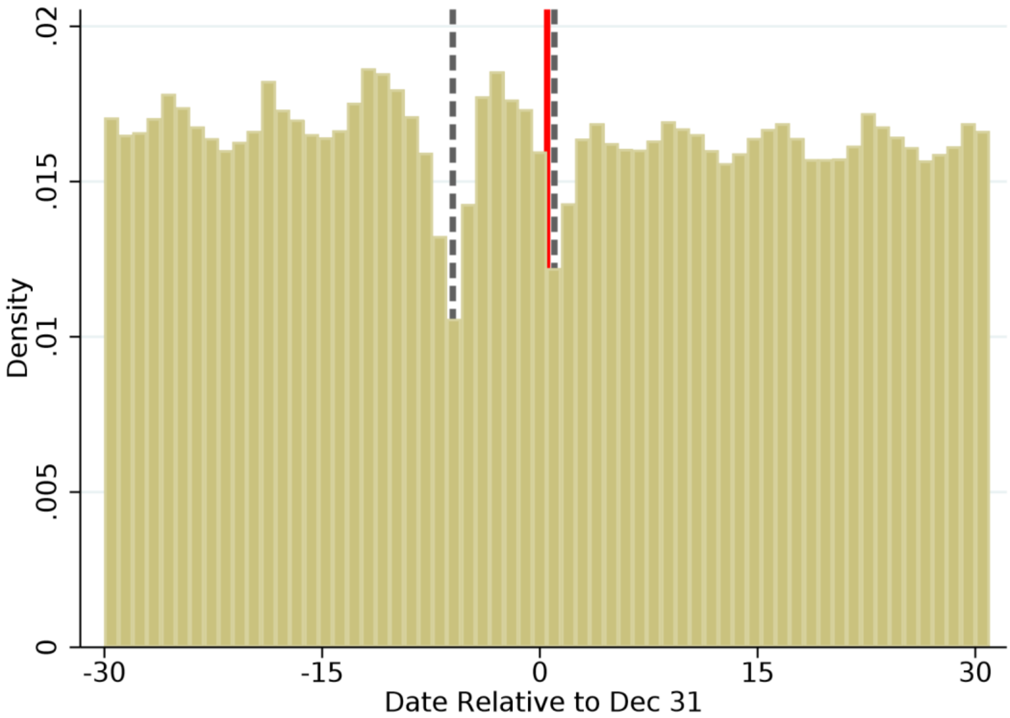


(d) By Gender



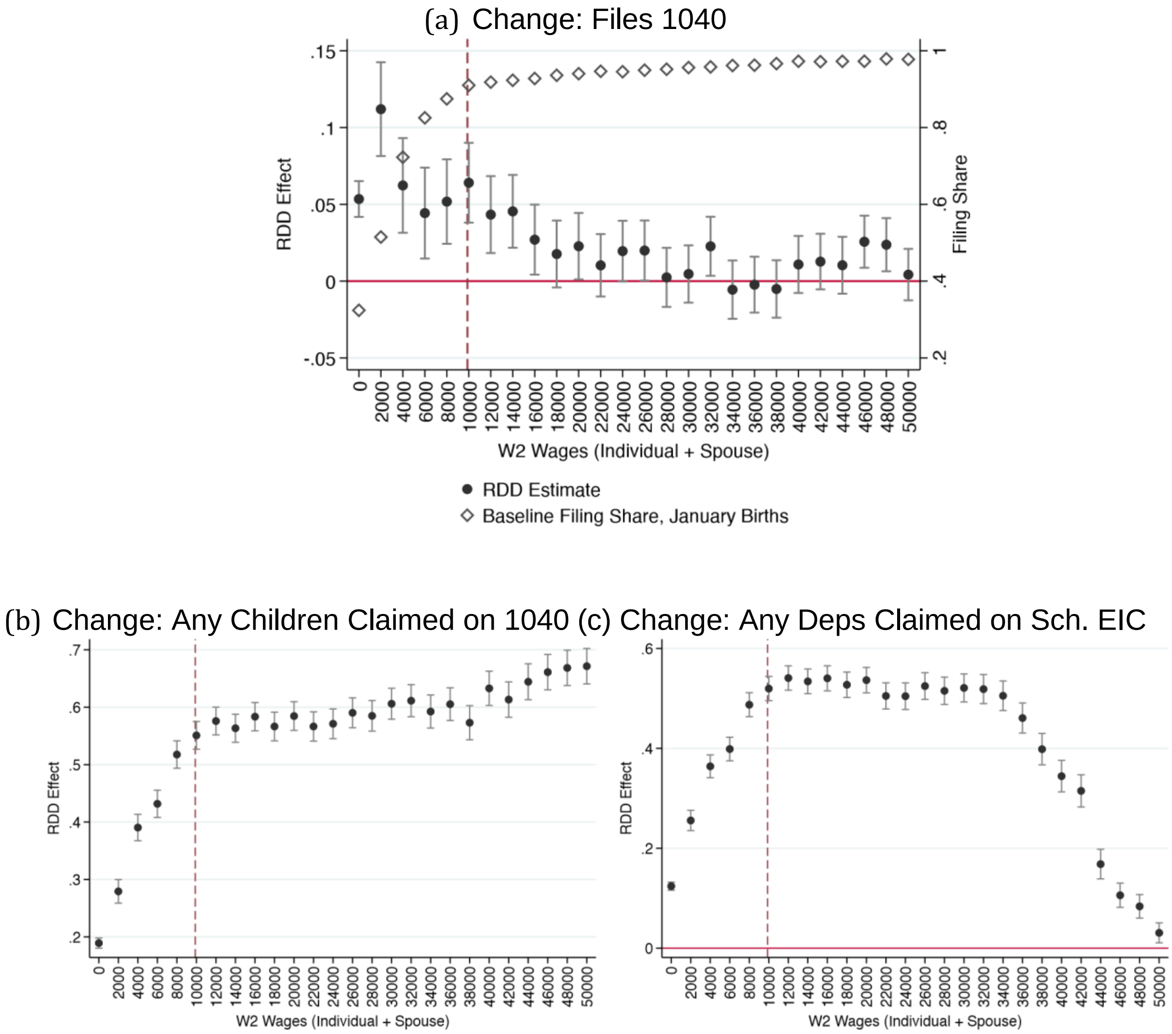
Note: Figure shows the share of the overall tax workforce by tax year with any SE income as filed on Schedule SE (black line) and individuals who receive a 1099 Information Return (maroon line). After the entry of OPE, we additionally distinguish the receipt of 1099 Information Returns including and excluding those received from OPE firms (dashed maroon line). In Panel (a), the workforce definition is split on EITC recipients with kids claimed on their 1040. In panel (b), we split by presence of kids 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. In Panel (d), we split by gender.

Figure A.3. Distribution of First Births Around End of Tax Years 2011-2018



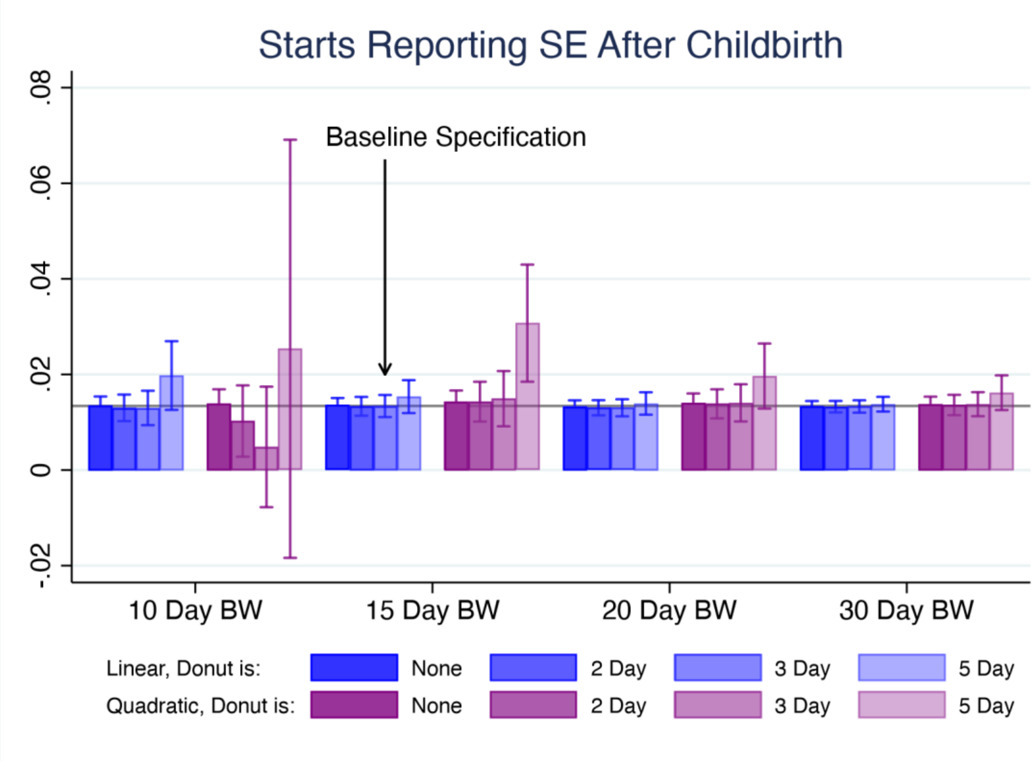
Note: Histogram reports distribution of all first births in December of each tax year 2011-2018 or the following January in our SSA sample (corresponding to the sample in our baseline analysis). The solid red line denotes the end of tax year t and the dashed grey lines correspond to the Federal holidays on Christmas day (December 25) and New Year's day (January 1).

Figure A.4. RDD Filing Effects by Tax Unit W2 Wage Earnings



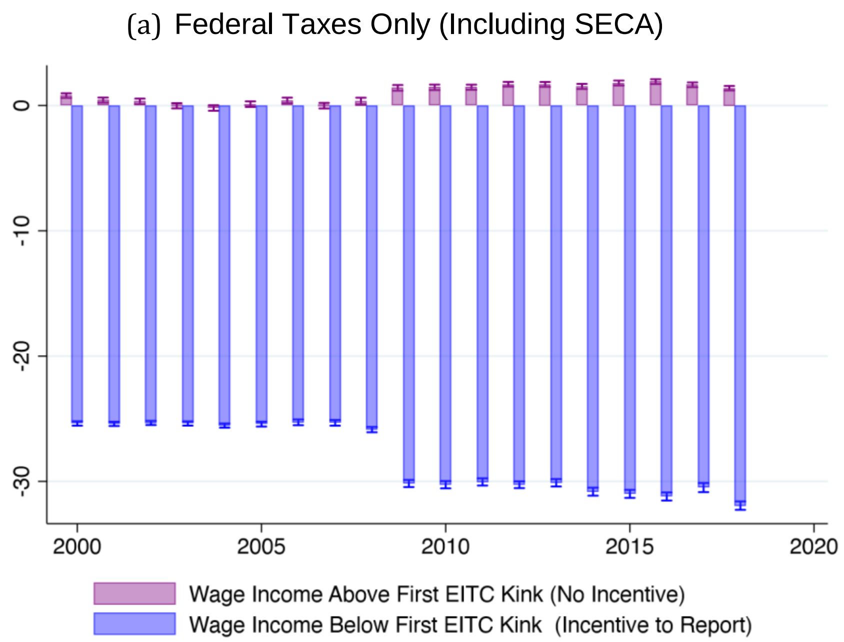
Note: Figure presents results from the baseline regression discontinuity design specification in Equation (1) pooling births in each December 2011-2018 and each subsequent January, estimated separately for individuals within $2000 bins of year-t tax unit (self plus spouse) W2 wages, measured in constant 2015 Dollars. The dashed maroon line is the earnings amount where the first EITC kink occurs for families with one child based on the 2015 schedule.

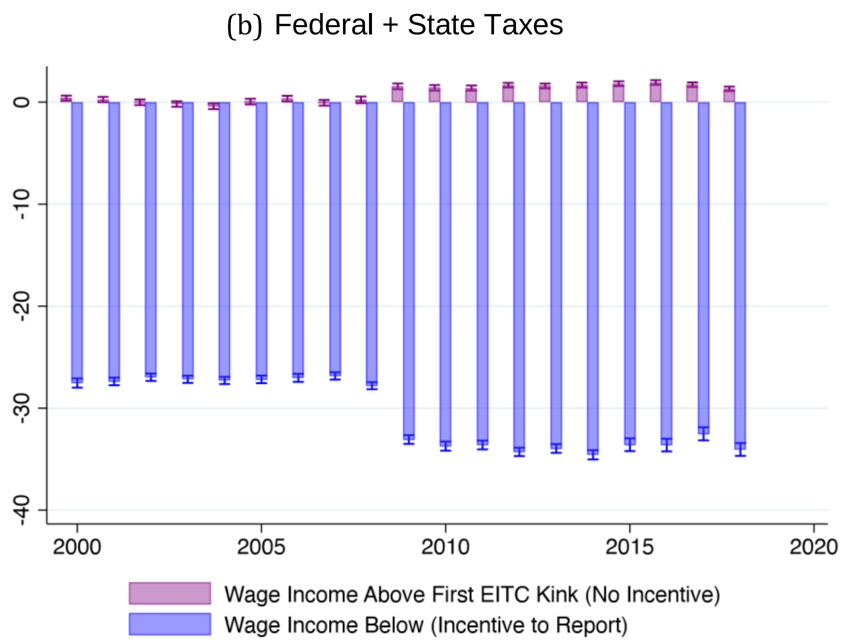
Figure A.5. Robustness of Main RDD Effects



Note: Figure displays main regression discontinuity effects on the change in whether the one reports any Schedule SE earnings in tax year t relative to the prior year t−1 from Column 1 in Table 2 under alternative specifications. Donut hole widths are bandwidths omitted from the regression sample. Quadratic specifications allow slopes to differ across the threshold. The horizontal black line corresponds to the size of the benchmark estimate in Table 2.

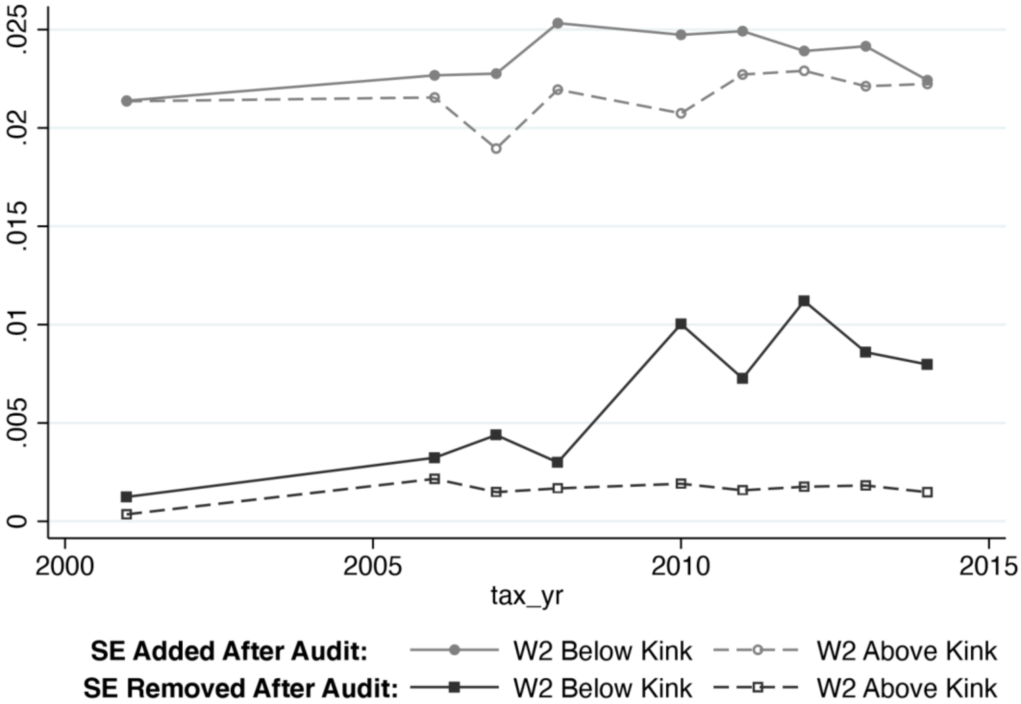
Figure A.6. RDD Effects on MTRs After Wages by Year





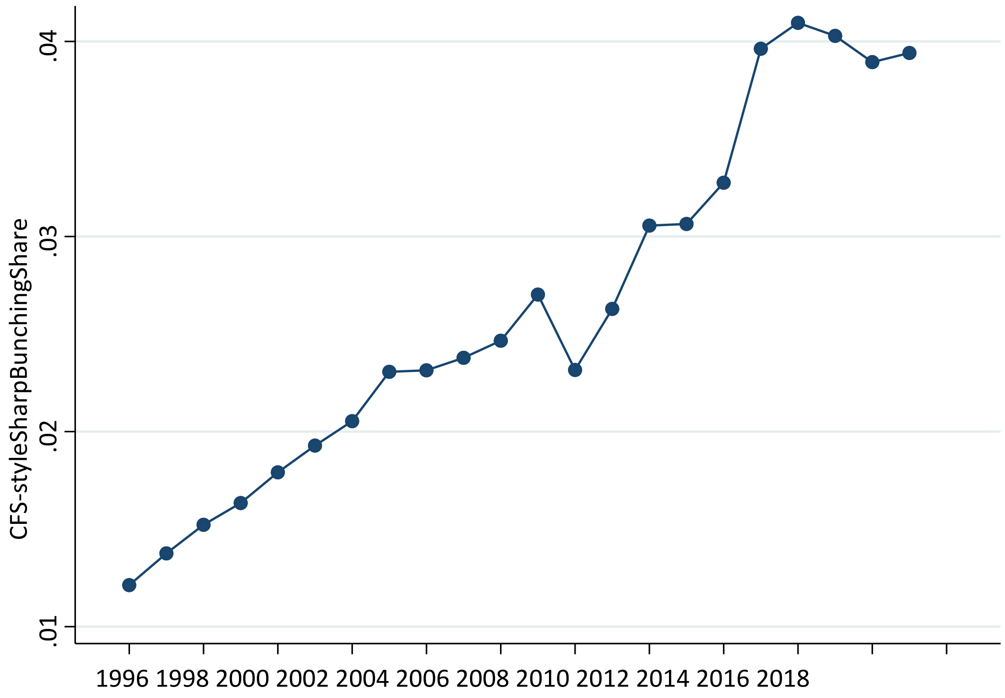
Note: Figure reports our baseline RDD estimates from estimating Equation 1 in the text within individual cohorts. Years correspond to the tax year t, at the end of which the births occur in the corresponding December or January. Outcomes are marginal tax rates on a first dollar of self-employment earnings, conditional on own and spouse's W2 wage earnings, calculated using TAXSIM.

Figure A.7. Changes in Self-Employment Status after NRP Audits



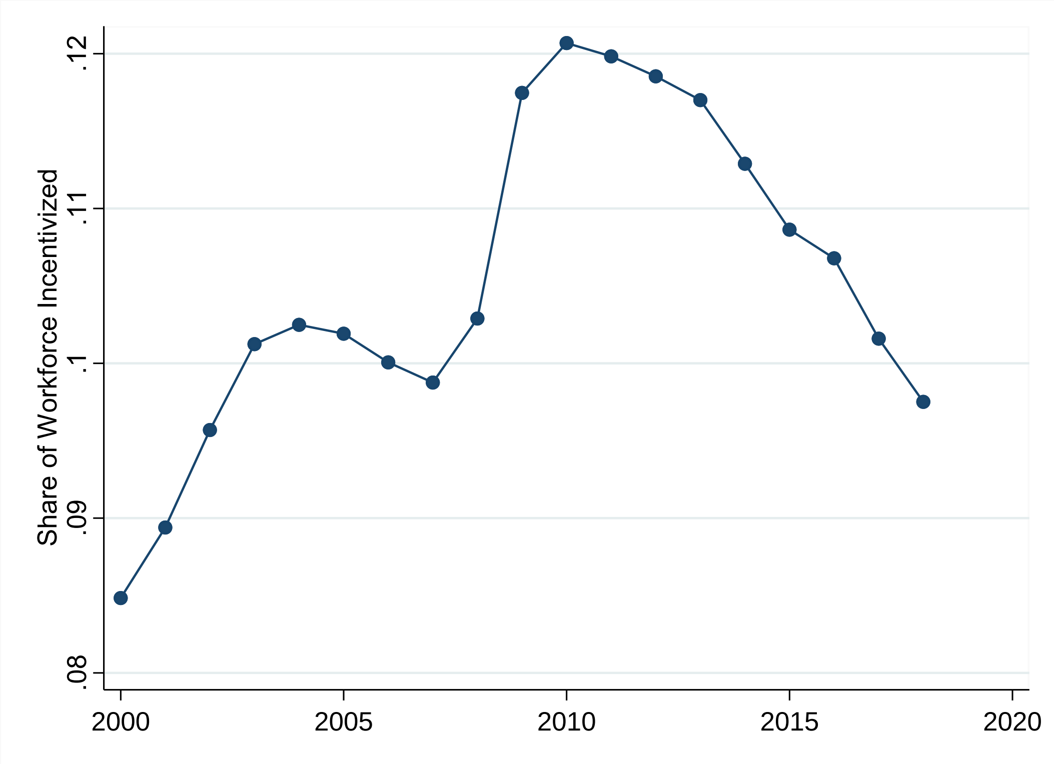
Note: Figure displays results of audits of a representative stratified random sample of 1040 filers conducted in tax years 2001 and 2006–2014 as part of the IRS's National Research Program (NRP) Individual Income Tax Reporting Compliance Studies. Using sampling weights for representativeness, the figure plots the share of individuals with 1040 returns who are found to have incorrectly not reported self-employment income on Schedule SE when they should have, and the share of individuals found to have reported positive self-employment income on Schedule SE when they actually should have reported none. Each propensity is calculated separately for individuals with and for individuals without an incentive to report self-employment. Individuals are classified based on their firm-reported W2 income and the number of eligible children determined by the audit.

Figure A.8. CFS-style Sharp Bunching Share Among Eligible Taxpayers with Children, 1996-2017



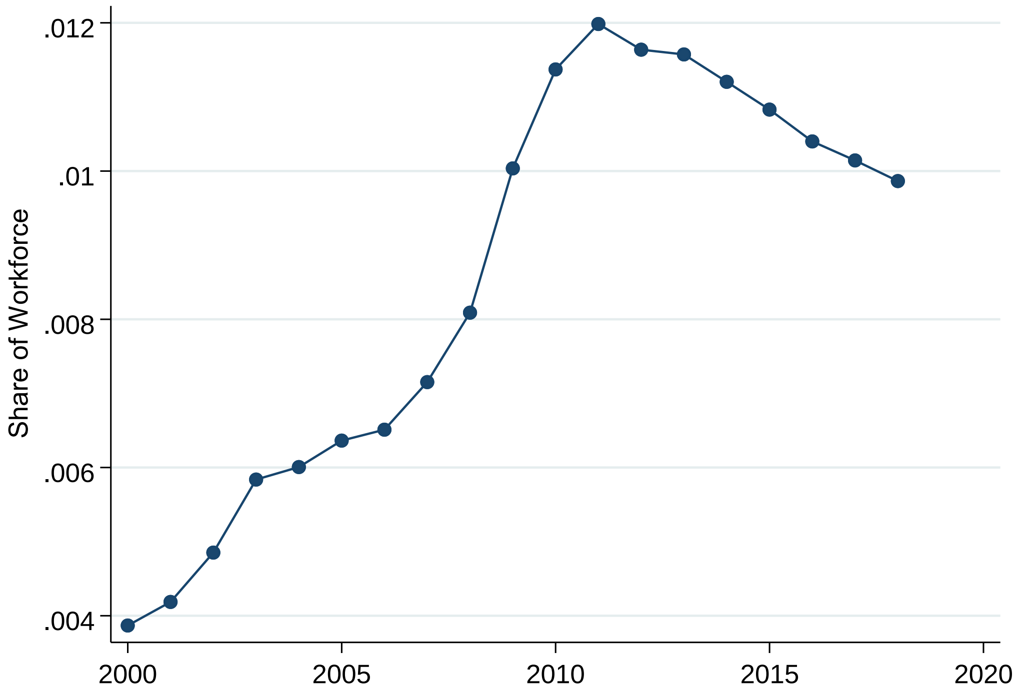
Note: Figure plots the average of the share of tax payers who are sharp bunchers, following the methodology of Chetty, Friedman, and Saez (2013).

Figure A.9. Share of Workforce with Incentive to Report SE



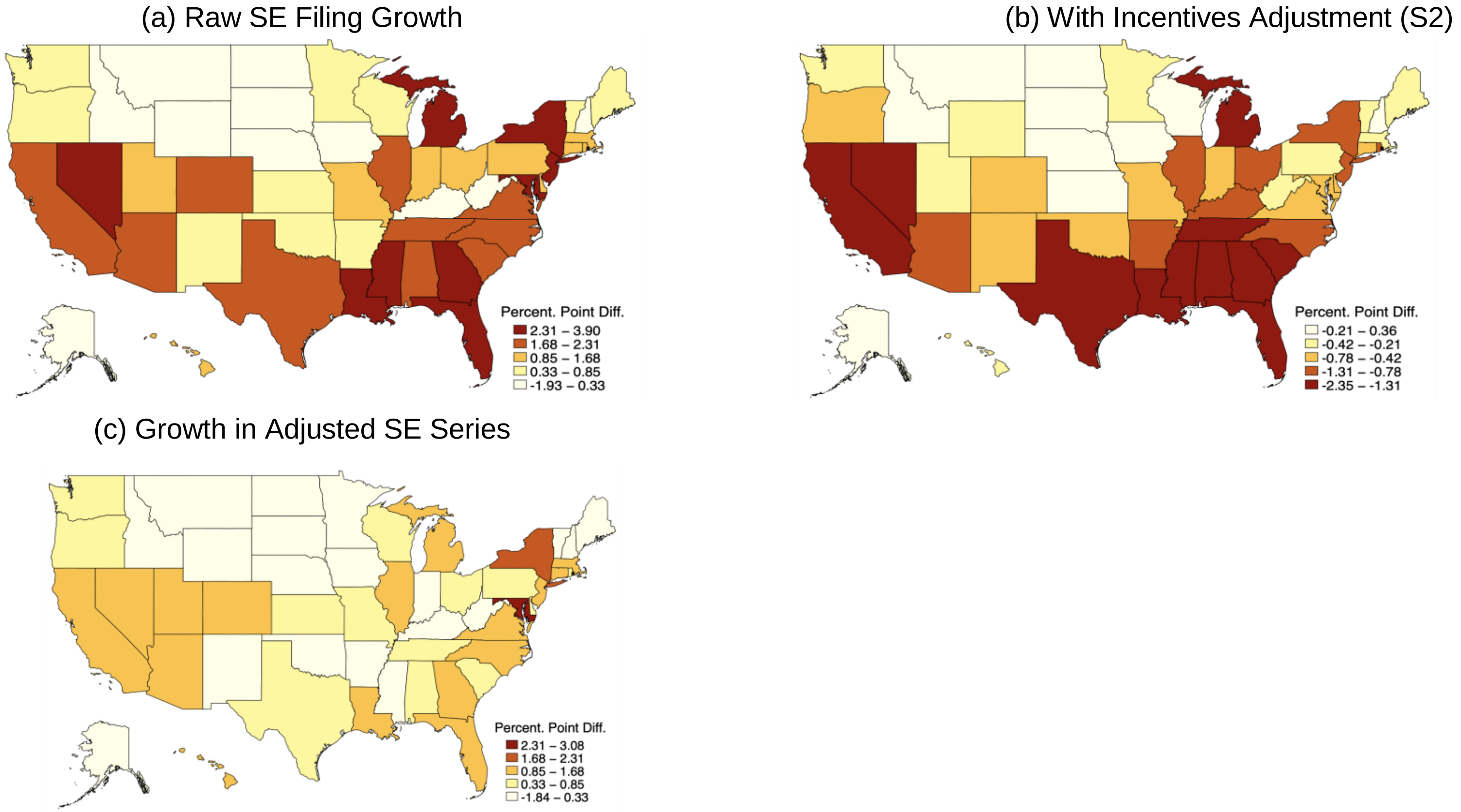
Note: The share incentivized in each year represents the number of individuals with children and wages below the corresponding EITC kink point as a share of the tax workforce.

Figure A.10. Counterfactual Change in Workforce Due to Incentive



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

Figure A.11. Raw SE Filing and Application of Our Adjustment By State, 2000-2014



Note: 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).

Annex B. Appendix Tables

Table B.1. RDD Estimates: "First-Stage" Effects on Filing Status and Reporting Incentives

	MTR After Wages	Has Neg MTR After Wages	Δ Any 1040	Δ Any Children	Δ Any EITC Dependents
	(1)	(2)	(3)	(4)	(5)
Panel A. All Parents					
Coeff	-6.534**	0.251**	0.0221**	0.598**	0.225**
	(0.0826)	(0.00138)	(0.00153)	(0.00180)	(0.00162)
N	1382740	1382740	1382740	1382740	1382740
DV Mean Level, Jan Births	22.72	0	0.860	0.131	0.0758
Panel B. With Wages < 1st EITC Kink					
Coeff	-30.69**	0.991**	0.0614**	0.299**	0.252**
	(0.0565)	(0.000634)	(0.00472)	(0.00366)	(0.00354)
N	349240	349240	349240	349240	349240
DV Mean Level, Jan Births	8.370	0	0.528	0.105	0.0912
Panel C. With Wages ≥1st EITC Kink					
Coeff	1.619**	0.000393**	0.00875**	0.699**	0.215**
	(0.0352)	(0.0000662)	(0.00129)	(0.00192)	(0.00181)
N	1033500	1033500	1033500	1033500	1033500
DV Mean Level, Jan Births	27.56	0	0.973	0.140	0.0707

Note: Table displays estimates from the baseline regression discontinuity design specification in Equation (1) on third-party reported earnings. The sample is all individuals with births in the last fifteen days of December of each tax year t in 2011-2018 or the first fifteen days of January immediately following tax year t, omitting births within three days of the start of the new year. "Wages < 1st kink" subsample includes all individuals in tax units (self plus spouse if filing a 1040 jointly) with year t wages in the EITC phase-in region for households with one child in that year (irrespective of whether their birth actually occurred in December or January); the complementary subsample includes all other individuals. Outcomes are from year t or are changes from year t relative to the prior year t−1, as specified. Marginal tax rates (MTRs) after wages are calculated as the federal marginal tax rate on the first dollar of self-employment earnings (including SECA taxes) beyond one's W-2 reported wage/salary earnings and those of any spouse reported on a 1040, given the year their child was actually born. We report mean year t levels of each dependent variable for individuals with first births in January of t + 1 in each subsample. Robust standard errors are displayed in parentheses.

Table B.2. Panel Relationship Between ZIP Bunching and SE Reporting

	Individuals With Children & Wages Below Kink	Individuals With Children & Wages Above Kink	Individuals Without Children
	(1)	(2)	(3)
Outcome: Workforce Share with SE			
ZIP Bunching Share	3.636**	0.123**	0.113**
	(0.290)	(0.0267)	(0.0325)
N	15709	15744	15782
Outcome: Workforce Share with 1099 NEC			
ZIP Bunching Share	-0.00210	0.0674**	0.0797**
	(0.148)	(0.0179)	(0.0228)
N	15709	15744	15782
Zip FE	X	X	X
Year FE	X	X	X

Note: Panels display estimates of panel regressions of self-employment rates and non-employee compensation reported on 1099-MISC within each specified workforce segments on the year-by-ZIP3 bunching measures calculated as in Chetty, Friedman, and Saez (2013). Sample is all individuals in the tax workforce 2000–2018, collapsed to the ZIP-year-subgroup level. Regressions are weighted by the workforce population in each cell. Standard errors are clustered by year and ZIP3.

Table B.3. Counterfactual Self-Employment Rates

	Baseline	Scenario 1 RDD Adjusted	Scenario 2 Incentivized = Unincentivized	Scenario 3 Scenario 2 & 1099-to-SE Adj	Scenario 4 Scenario 2 & Fixed Demog.
2000	0.0974	0.0969	0.0901	0.0901	0.1095
2001	0.0980	0.0968	0.0898	0.0896	0.1084
2002	0.0989	0.0969	0.0897	0.0891	0.1088
2003	0.1030	0.1014	0.0928	0.0917	0.1122
2004	0.1062	0.1033	0.0957	0.0944	0.1150
2005	0.1095	0.1088	0.0985	0.0970	0.1166
2006	0.1102	0.1082	0.0989	0.0973	0.1174
2007	0.1112	0.1074	0.0990	0.0969	0.1165
2008	0.1095	0.1065	0.0962	0.0936	0.1136
2009	0.1128	0.1091	0.0971	0.0937	0.1134
2010	0.1150	0.1100	0.0973	0.0935	0.1141
2011	0.1172	0.1129	0.0985	0.0941	0.1161
2012	0.1175	0.1129	0.0994	0.0947	0.1176
2013	0.1170	0.1121	0.0991	0.0942	0.1171
2014	0.1182	0.1132	0.1005	0.0955	0.1192
2015	0.1175	0.1128	0.1003	0.0953	0.1195
2016	0.1174	0.1123	0.1008	0.0957	0.1216
2017	0.1160	0.1127	0.1000	0.0946	0.1206
2018	0.1151	0.1103	0.0995	0.0941	0.1214

Note: Table reports baseline (unadjusted) share of workforce with self-employment earnings alongside counterfactual series adjusted for shifts in reporting behavior and demographic change "Scenario (1)" examines how self-employment would have evolved in the absence of any reporting incentives captured in our RDD estimates; specifically, it reports the counterfactual replacing our RDD estimates reported in Figure 6 with zero in all years. The adjustment in "Scenario (2)" replaces SE rates for individuals with incentives to report SE with the rates among comparable individuals without this incentive in each year. "Scenario (3)" applies an adjustment accounting for changes in the propensity of *unincentivized* individuals with 1099 contract income to report their self-employment proceeds on a 1040 over time. "Scenario (4)" accounts for demographic shifts over this time frame by holding the gender and age composition of the workforce constant at its 2000 levels. See text for further details.

Table B.4. Level Adjustments: Share of Workforce with 1099-Reported Non-Employee Compensation

	Baseline	Scenario 1 RDD Adjusted	Scenario 2 Incentivized = Unincentivized	Scenario 4 Scenario 2 & Fixed Demog.
2000	0.0906	0.0903	0.0875	0.0876
2001	0.0887	0.0889	0.0860	0.0859
2002	0.0928	0.0923	0.0899	0.0896
2003	0.0947	0.0947	0.0911	0.0905
2004	0.0965	0.0963	0.0926	0.0918
2005	0.0962	0.0962	0.0922	0.0914
2006	0.0979	0.0985	0.0939	0.0929
2007	0.0983	0.0984	0.0942	0.0929
2008	0.0970	0.0969	0.0927	0.0910
2009	0.0938	0.0939	0.0893	0.0870
2010	0.0951	0.0948	0.0906	0.0880
2011	0.0969	0.0972	0.0923	0.0895
2012	0.0975	0.0975	0.0927	0.0897
	[0.0974]	[0.0975]	[0.0927]	[0.0896]
2013	0.0975	0.0979	0.0926	0.0895
	[0.0973]	[0.0979]	[0.0924]	[0.0893]
2014	0.0991	0.0991	0.0941	0.0911
	[0.0980]	[0.0991]	[0.0930]	[0.0899]
2015	0.1011	0.1016	0.0960	0.0931
	[0.0973]	[0.1016]	[0.0922]	[0.0891]
2016	0.1044	0.1048	0.0992	0.0964
	[0.0968]	[0.1048]	[0.0917]	[0.0885]
2017	0.1063	0.1065	0.1011	0.0974
	[0.0995]	[0.1065]	[0.0945]	[0.0904]
2018	0.1070	0.1075	0.1017	0.0979
	[0.1002]	[0.1075]	[0.0952]	[0.0909]

Note: 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. “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)” is not relevant for adjustments to 1099-reported non-employee compensation and thus is not included. “Scenario (4)” accounts for demographic shifts over this time frame by holding the composition of the workforce constant at its 2000 levels

Annex C. Self-Employment Reporting and Changing EITC Incentives: Event Study Around Childbirth

As discussed in the main text, only households with children face a negative marginal tax rate for reporting self-employment income. To further test the hypothesis that self-employment growth is tied to EITC incentives, we follow Chetty, Friedman and Saez (2013[6]) and examine how self-employment reporting changes around a person's first childbirth, when they become eligible for a generous credit. We expand upon Chetty, Friedman and Saez (2013[6]) in two main ways. First, to investigate 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 C.1a 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 C.2 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:

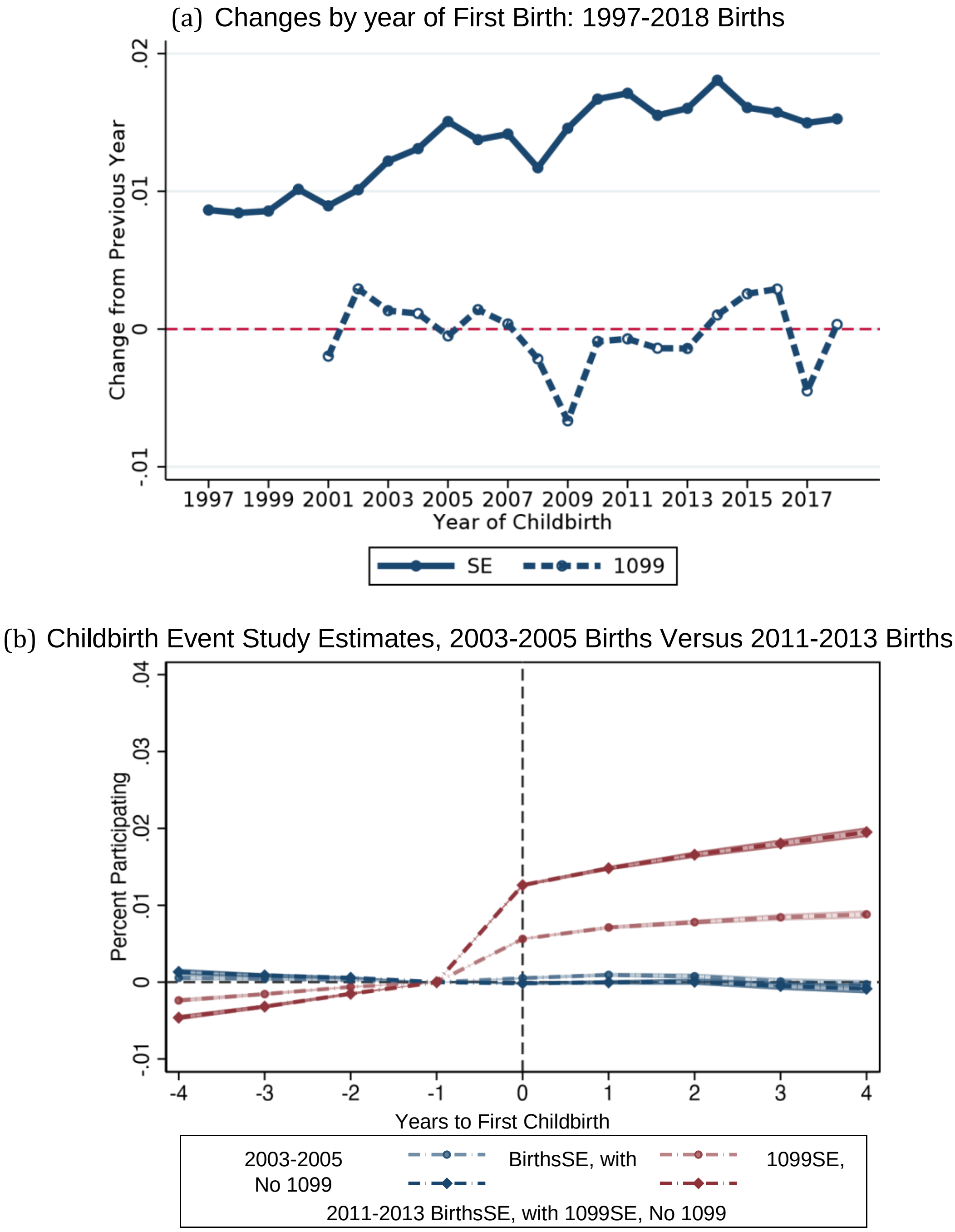
$$y_{it} = \sum_{k\in K} \beta_k^p I\{\text{FirstChildbirth}^i = t+k\} + \gamma^p_{a(i)\times g(i)} + \gamma^p_{t\times g(i)} + e^p_{it} \qquad (6)$$

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.[44] 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 C.1b 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 C.3 reports estimates separately by gender of the parent. As we found earlier, these changes over time are largely driven by mothers.

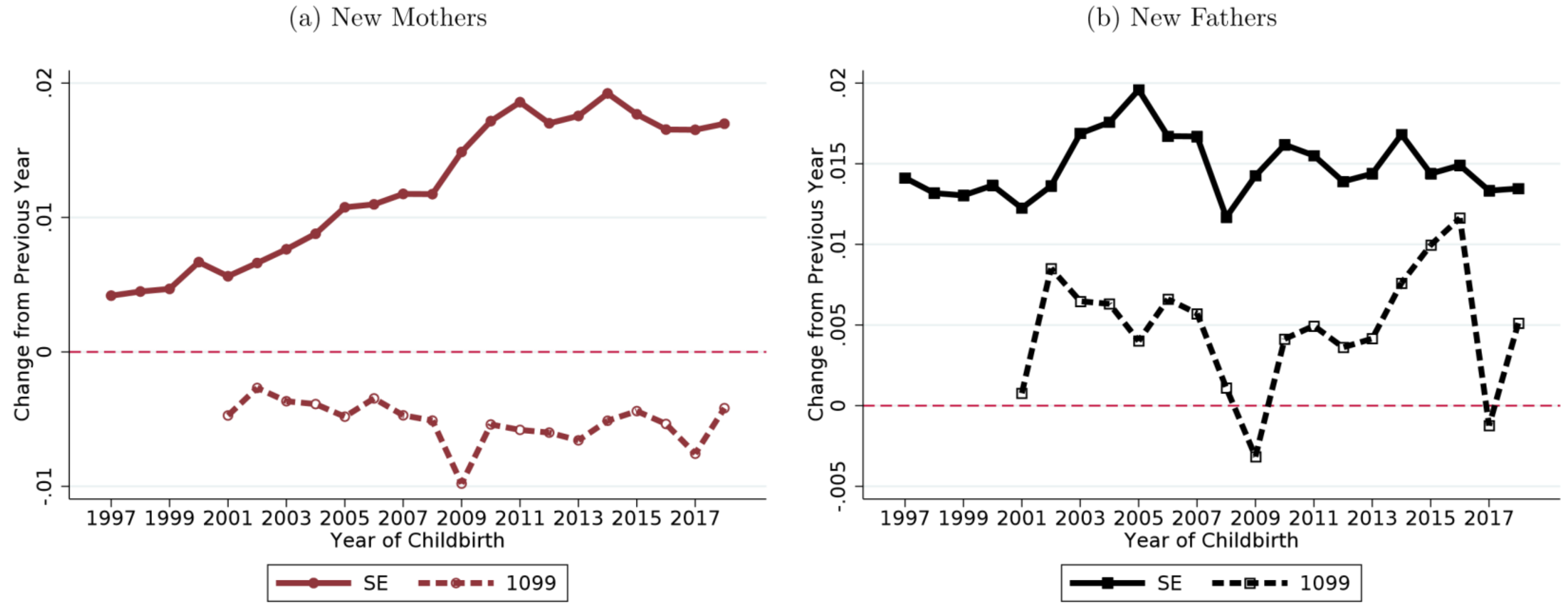
[44] Accordingly, the event-study coefficients are the average of coefficients run separately for men and women, which we report in Appendix Figure C.3.

Figure C.1. Change in Self-Employment Around First Childbirth, 1997-2018 Births



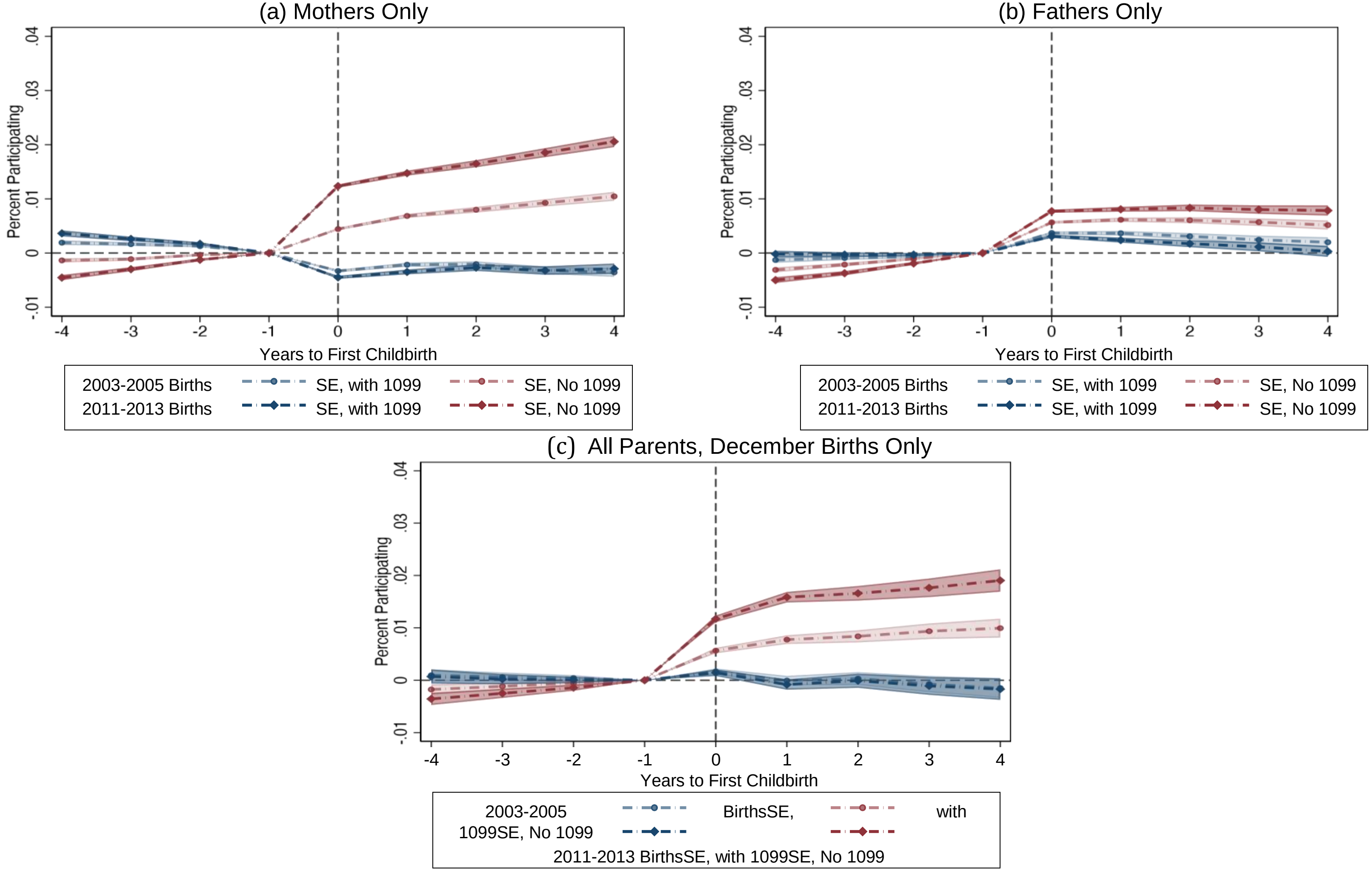
Note: Panel A shows the average change in propensity to file SE (solid line) or receive a 1099 Information Return (dashed line), in the year of first childbirth reported on the x-axis. Panel B plots event study coefficients for separate regressions run on the indicated time-period and for the indicated outcome. See text for more details.

Figure C.2. Change in Self-Employment Around First Childbirth, 1997-2018 Births, By Gender of Parent



See notes for Figure B.1a.

Figure C.3. Childbirth Event Study Estimates, Additional Estimates



Note: See notes for Figure C.1b.

Annex D. Additional Counterfactual Exercises

This Appendix describes the details of two additional counterfactual exercises.

First, 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:

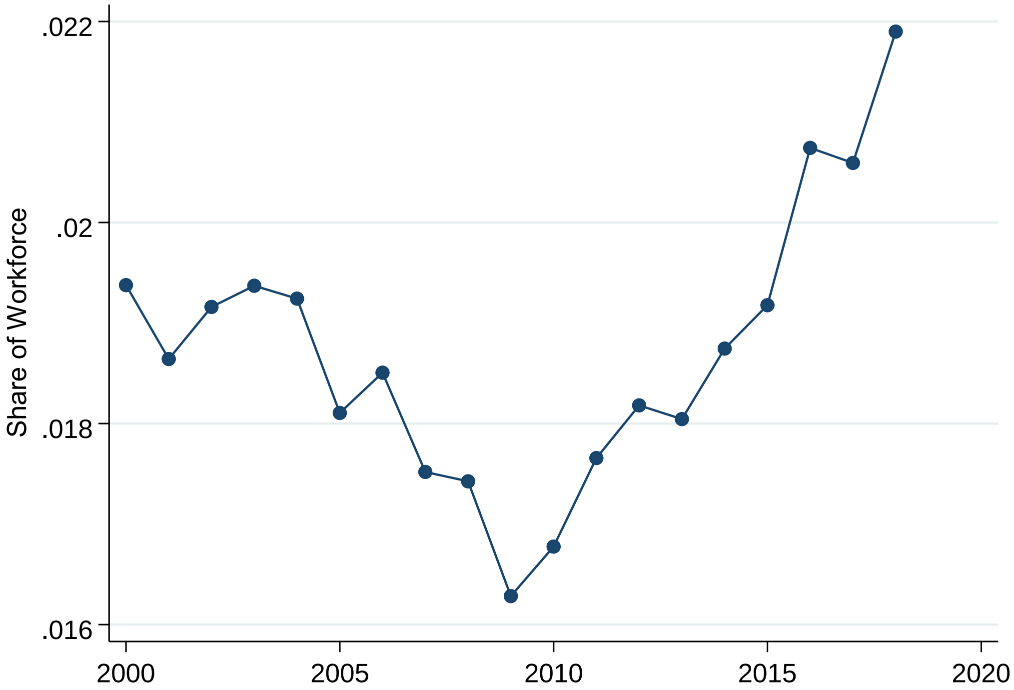
$$\mathbb{1}\{FilesSE\}_i^{t,p} = \gamma_{s\times r}^{t,p} + \epsilon_i^{t,p} \tag{7}$$

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 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 D.1 reports the implied share of the workforce with non-employee compensation who are underreporting SE.

To account for demographic changes over time, we hold constant the weights of our demographic groups, $\omega_{g,t}$ at their 2000 levels, i.e.

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

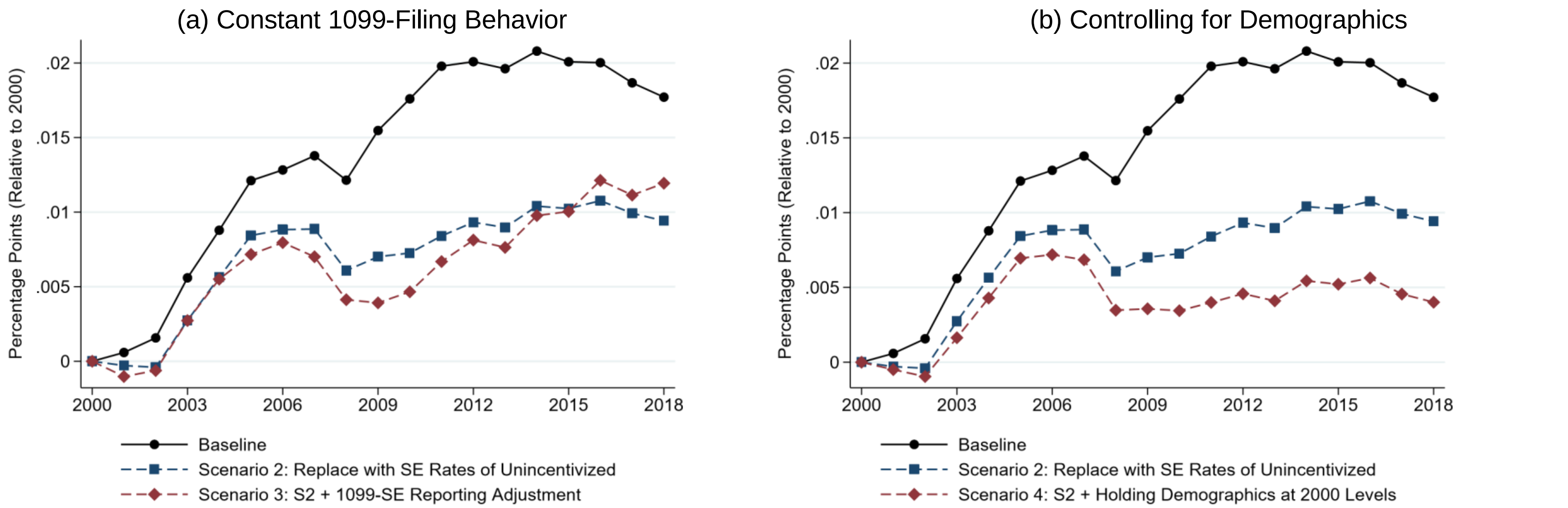
Figure D.1. Estimated Share of Workforce with 1099-Non-employee Compensation Underreporting SE



Note: Share of workforce with self-employment contract payments reported on 1099 forms but not reported on their 1040-SE, estimated using on individuals with no self-employment reporting incentives following the methodology described in Section 6.2 (which accounts for expensing behavior).

Figure D.2. Adjusted SE Shares Under Additional Counterfactual Assumptions

Figure C.2:



Note: Figure presents the baseline and implied trend in SE share of the overall tax workforce under counterfactual scenarios. In panel (a), “Scenario (3)” applies an adjustment accounting for changes in the propensity of unincentivized individuals with 1099 contract income to report their self-employment proceeds on a 1040 over time. In panel (b), “Scenario (4)” accounts for demographic shifts over this time frame by holding the composition of the workforce constant at its 2000 levels.

Annex E. 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 de-identified, or “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 use the list of roughly 50 large labor platforms from Collins, Garin, Jackson, Koustas, and Payne (2019) that are mentioned in public databases than can be identified in the tax data (along with the corresponding EIN) using the unmasked firm name. Using the corresponding masked 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.”[45] Starting in 2012, several online intermediaries in the OPE began issuing the new Form 1099-K instead of 1099-MISC for nonemployee 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 also identify and track the 1099-Ks issued by the approximately 50 important online “gig” platforms where self-employed individuals offer labor services to firms or individual clients mentioned above. 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

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

companies cannot be identified by this method.[46] 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 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.

[46] 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.