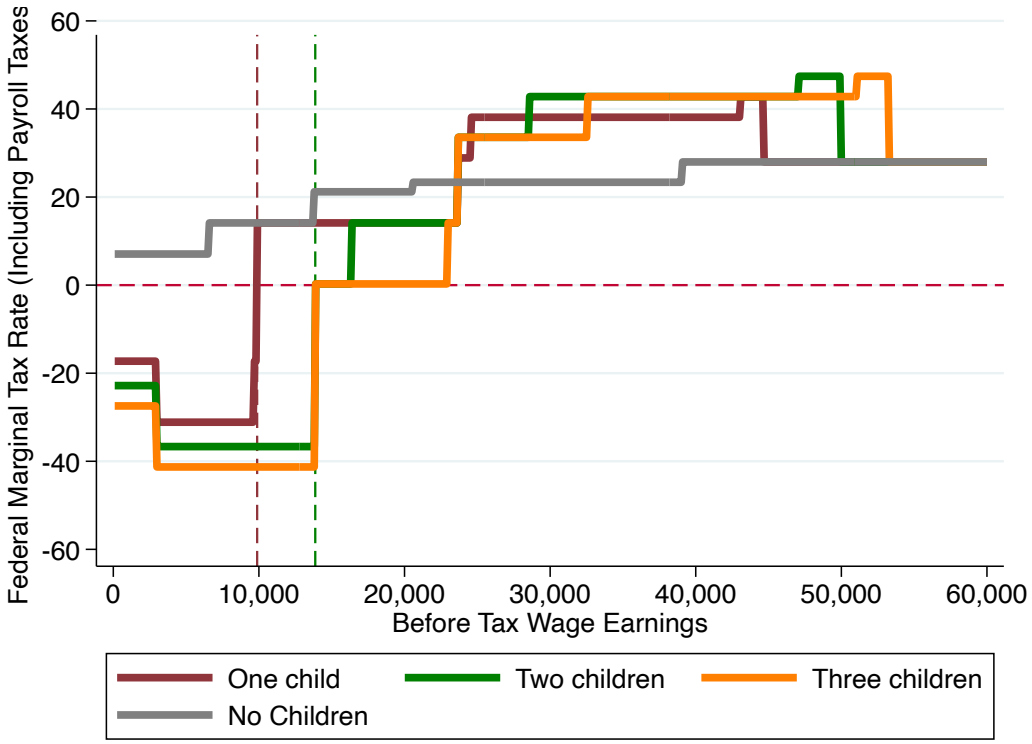
Online Appendix For: New Gig Work or Changes in Reporting? Understanding Self-Employment Trends in Tax Data

Andrew Garin, Emilie Jackson and Dmitri Koustas

June 20, 2024

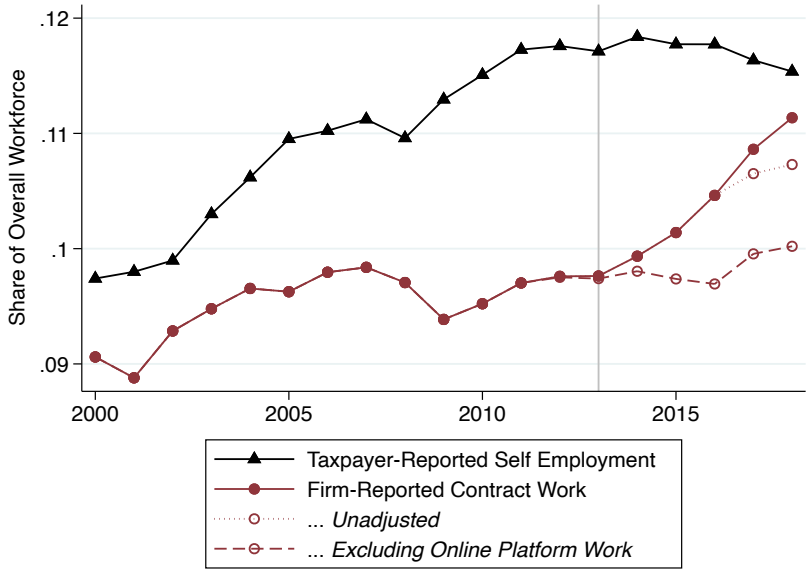
A Appendix Figures

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



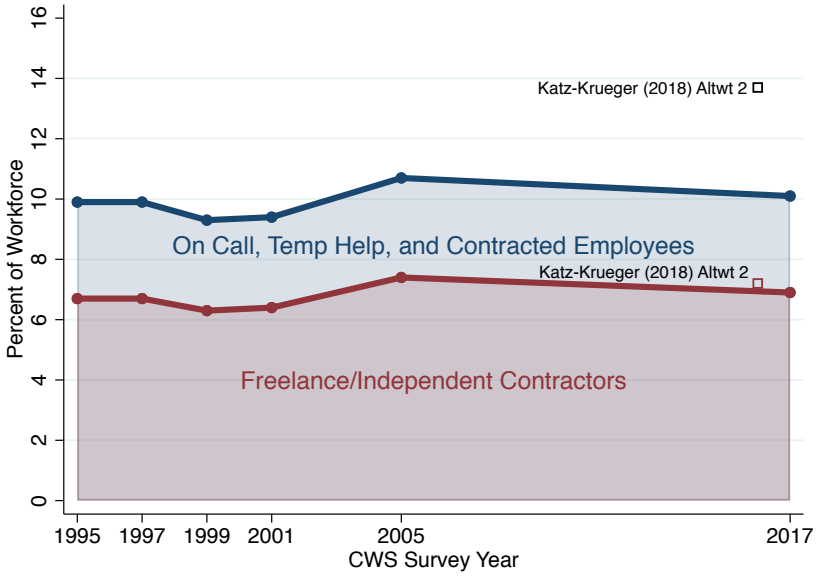
Notes: Figure shows the effective marginal tax rate for reporting an additional dollar of self-employment income, for a given level of before-tax wage earnings. Calculation takes into account the full tax schedule in tax year 2015 with no other credits/deductions except the EITC, CTC and standard deductions,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

Figure A.2: Firm-Reported Self-Employment Earnings, Raw Data and 1099-K Imputation



Note: After 2016, several online platform companies reported payments to gig workers on form 1099-K, adhering to the higher $20,000 reporting threshold for that form. In the solid red line, the 2017 and 2018 levels of total 1099-reported work reflect an imputation of how trends in online platform work would have evolved if reporting thresholds had remained constant described in **?** and **?**; this series reproduces the corresponding series in Figure 4 of **?**. This imputation is based on trends in state-level 1099-K data in Massachusetts and Vermont with a lower $600 threshold in those states; see **?** for details on the construction of the imputed series. While the imputed series approximates the number of workers engages in such work, it overstates the number of these workers for whom a 1099 return was actually filed. These raw data are shown in the "Unadjusted" series. See additional notes for Figure **??**.

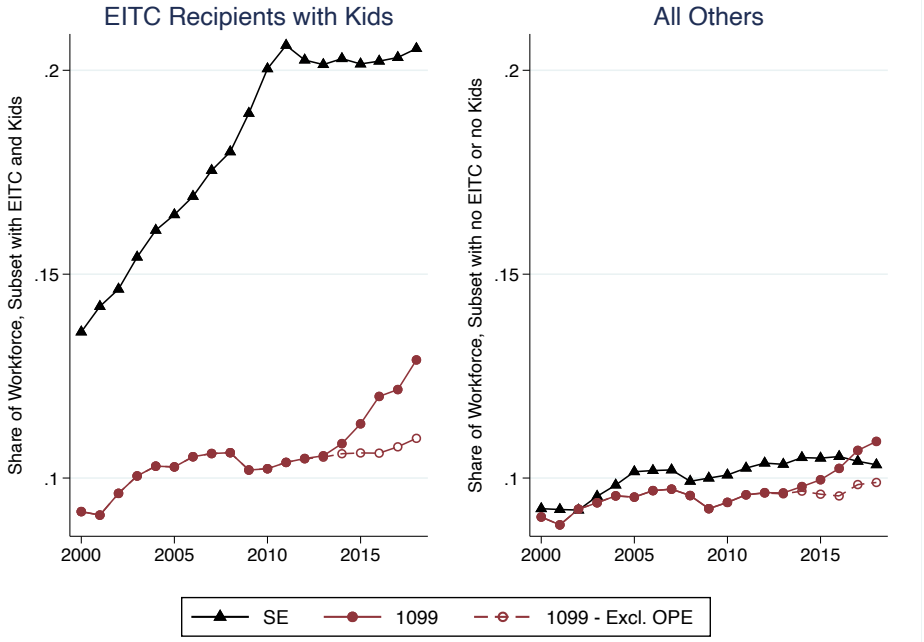
Figure A.3: Survey-Based Measures of Independent Contracting



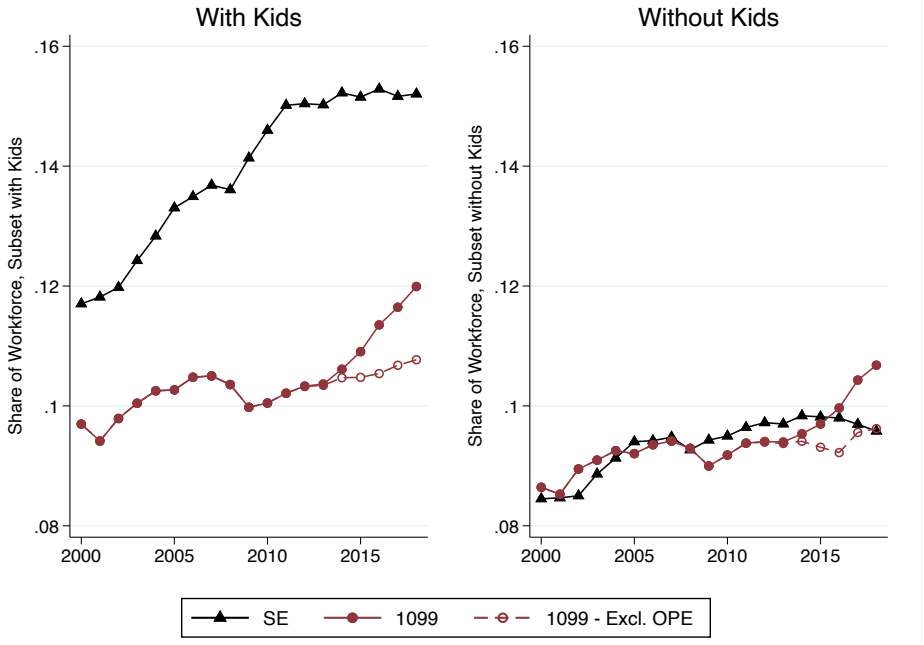
Source: Contingent Worker Supplement (1995, 1997, 1999, 2001, 2005, 2017), and **?**. We report **?**'s "Altwt. 2" estimates, an alternative weighting of the trends originally presented in **?** that downweight multiple job holders to match the rate of multi-job holding in the October 2015 CPS.

Figure A.4: Share of Workforce with Self-Employment and 1099 Information Returns

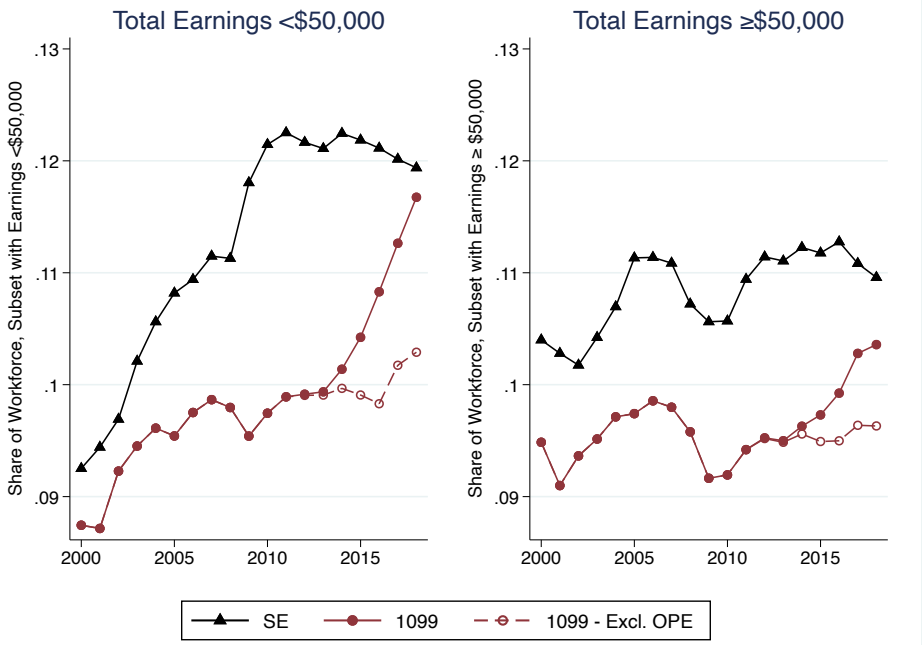
(a) By EITC and Kids



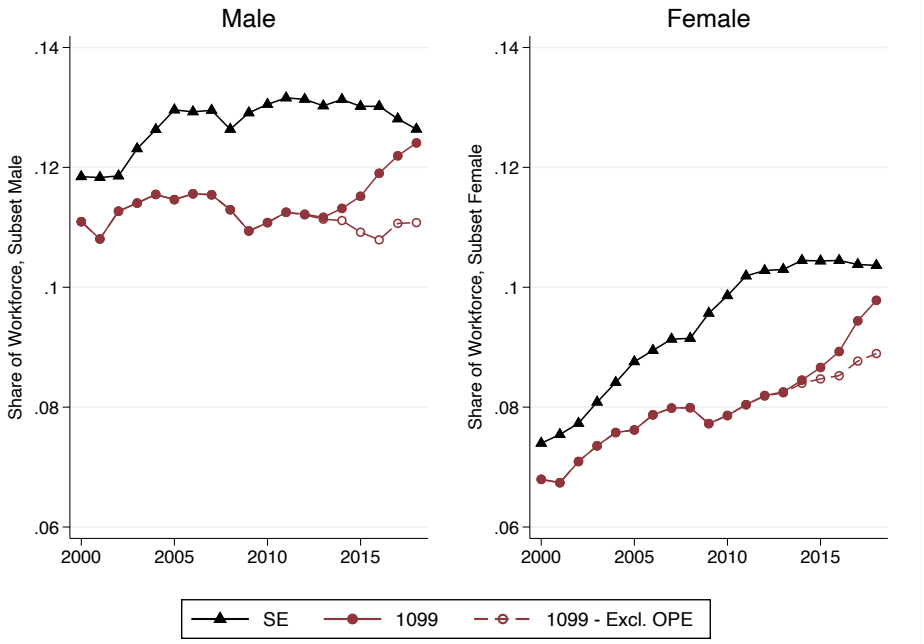
(b) By Presence of Children



(c) By Total Earnings

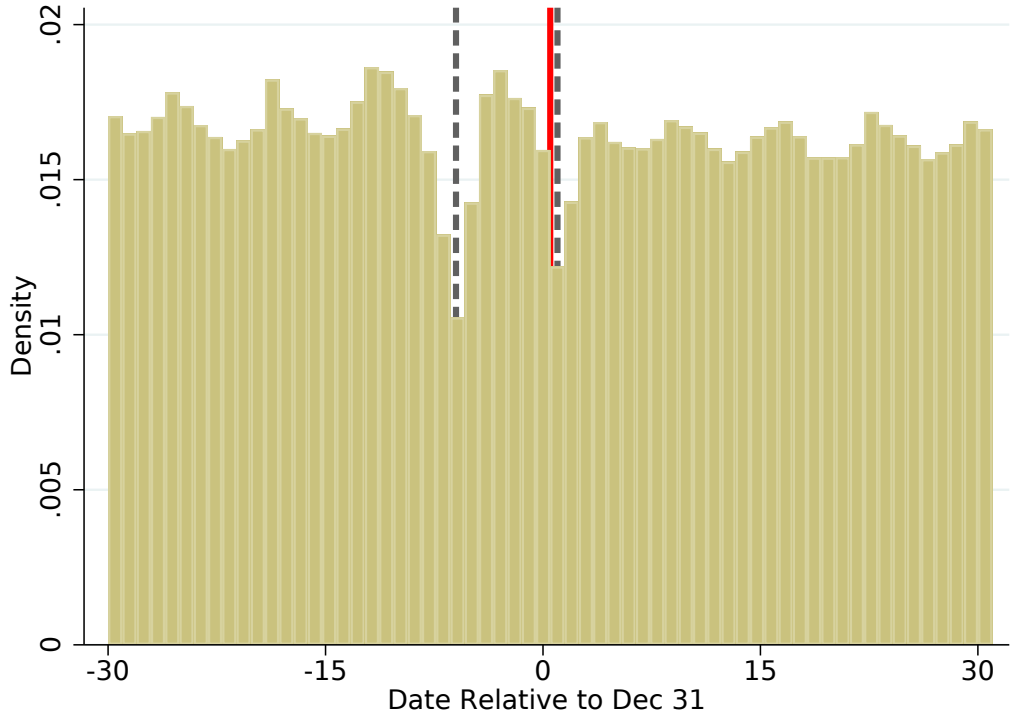


(d) By Gender



Notes: Figure shows the share of the overall tax workforce by tax year with any SE income as filed on Schedule SE (black line) and individuals who receive a 1099 Information Return (maroon line). After the entry of OPE, we additionally distinguish the receipt of 1099 Information Returns including and excluding those received from OPE firms (dashed maroon line). In Panel (a), 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.

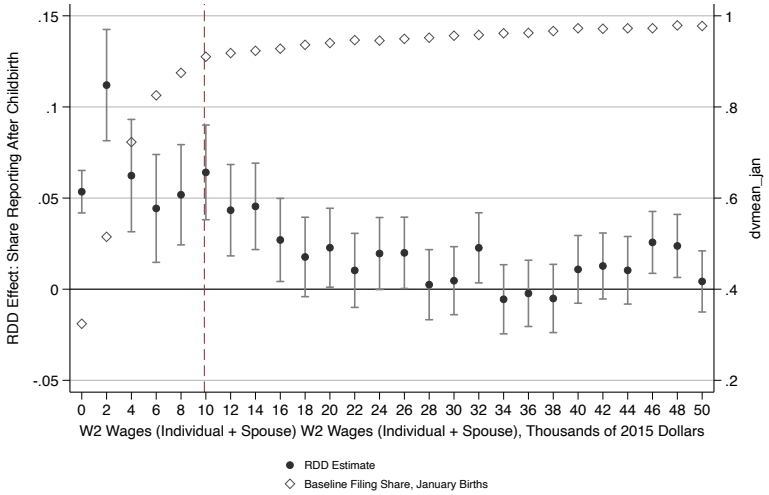
Figure A.5: Distribution of First Births Around End of Tax Years 2011-2018



Notes: 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).

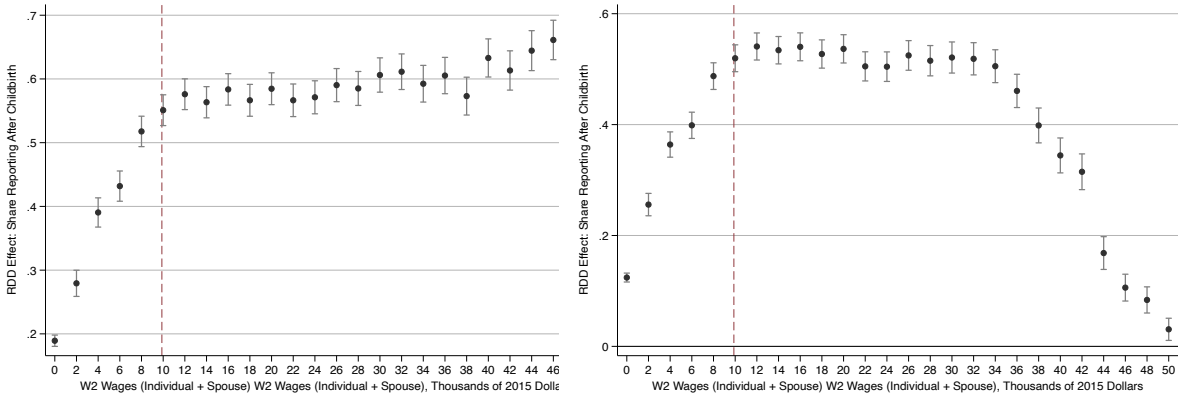
Figure A.6: RDD Filing Effects by Tax Unit W2 Wage Earnings

(a) Change: Files 1040



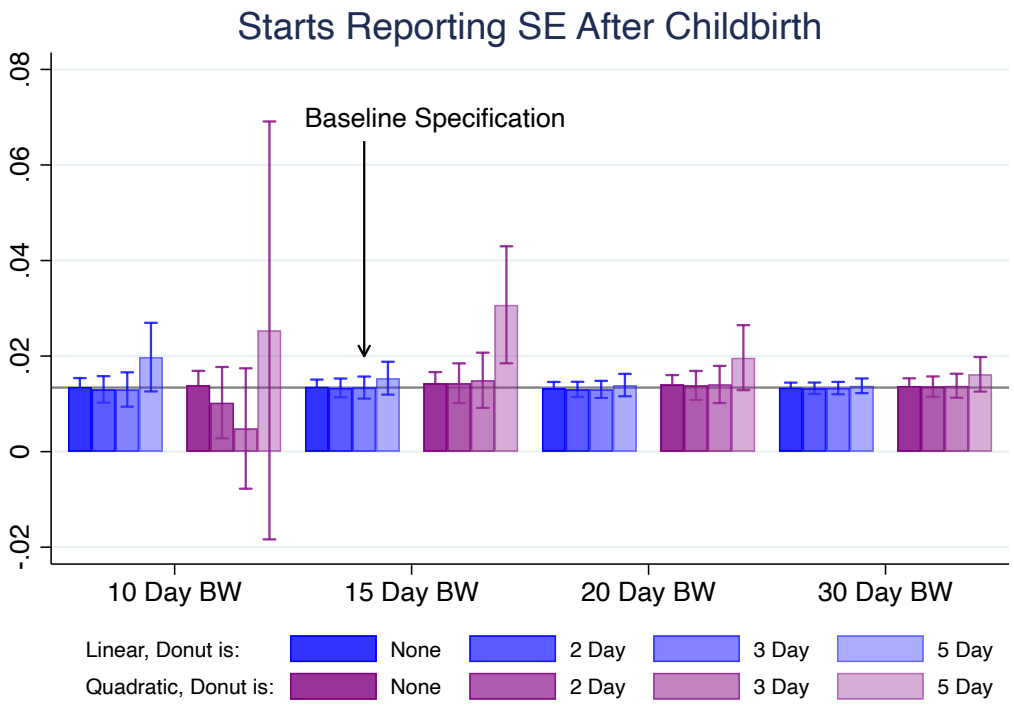
(b) Change: Any Children Claimed on 1040 EITC

(c) Change: Any Deps. Claimed on Sch.



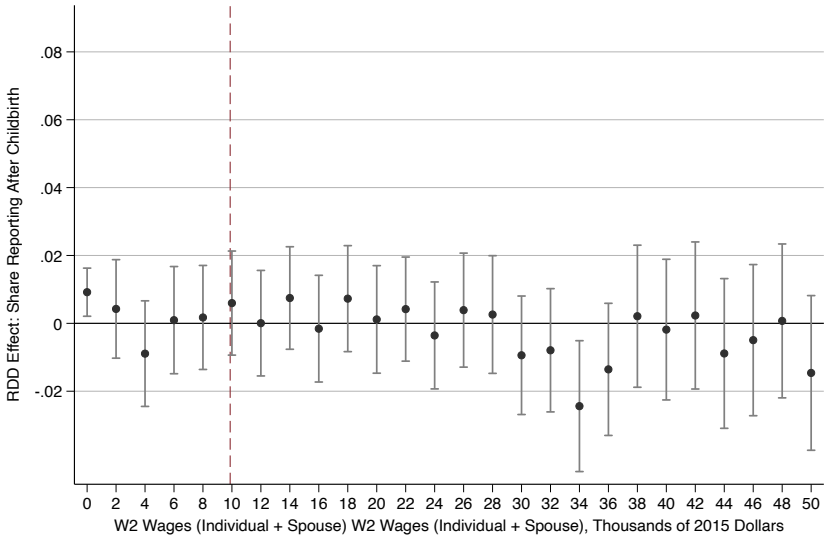
Notes: Figure presents results from the baseline regression discontinuity design specification in Equation (**??**) pooling births in each December 2011-2018 and each subsequent January, estimated separately for individuals within $2000 bins of year-$t$ tax unit (self plus spouse) W2 wages, measured 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.7: Robustness of Main RDD Effects



Notes: 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 **??** 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 **??**.

Figure A.8: RDD Effects on 1099 Earnings by Tax Unit W-2 Wages

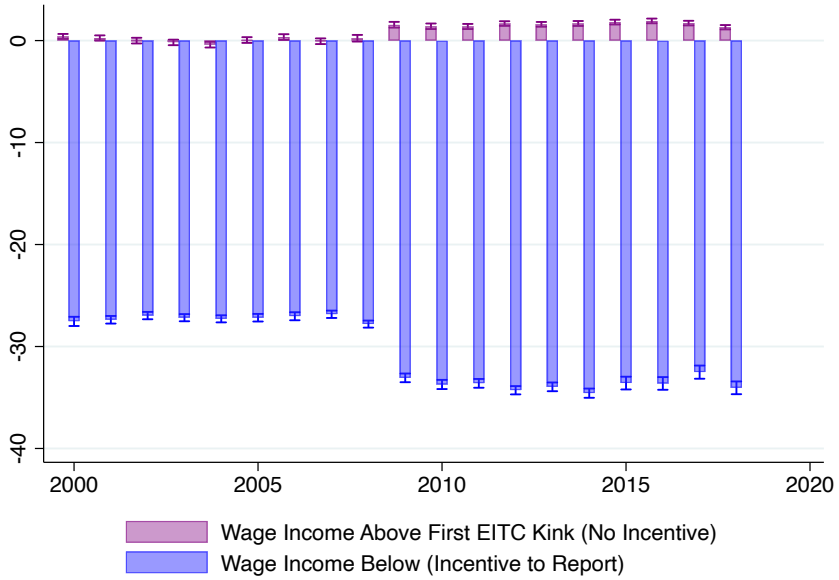


Notes: Figure replicates Figure **??** using the change in having non-employee income reported on a 1099-MISC in tax year t relative to the prior year $t-1$ as the outcome. The dashed maroon line is the amount where the first EITC kink occurs for families with one child based on the 2015 schedule.

Figure A.9: RDD Effects on MTRs After Wages by Year

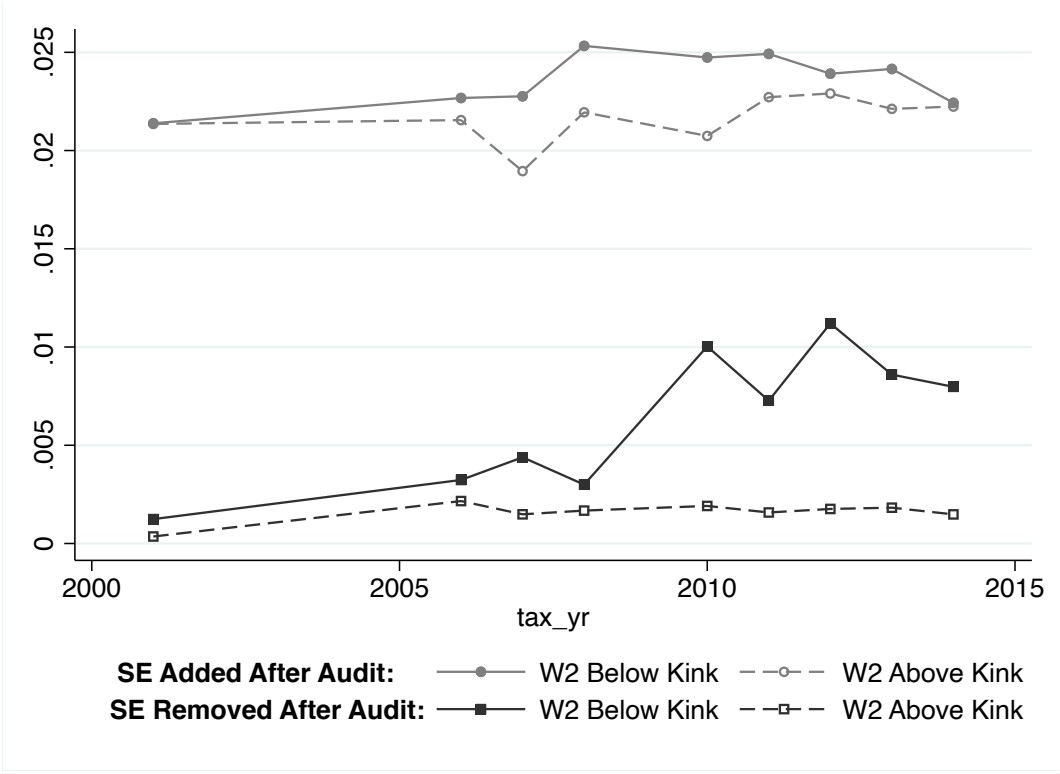
(a) Federal Taxes Only (Inlcuding SECA)

(b) Federal + State Taxes



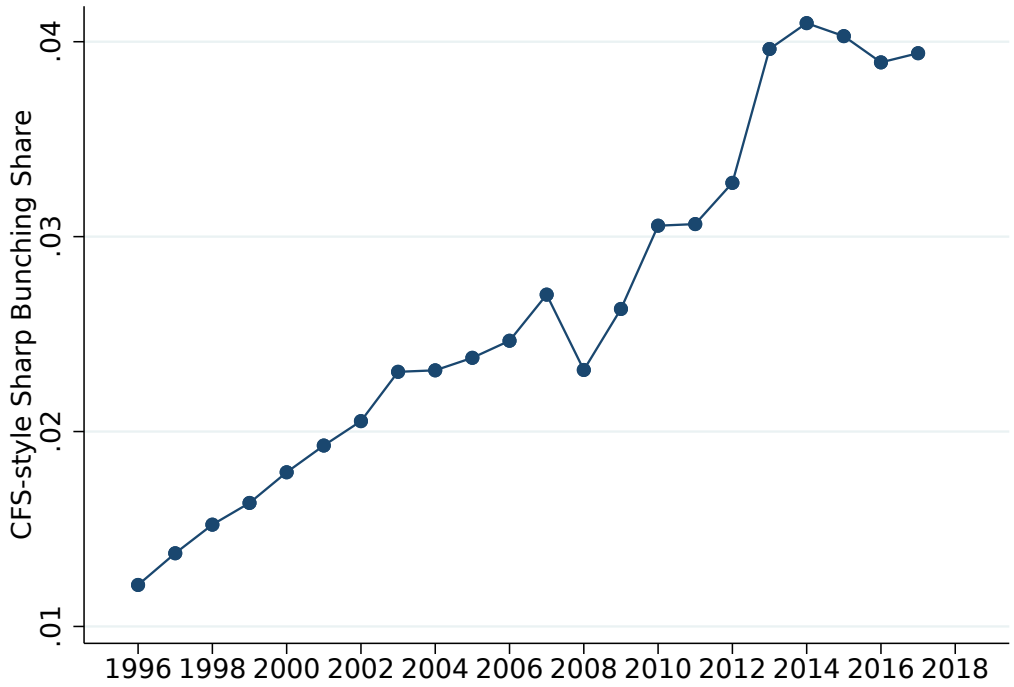
Notes: Figure reports our baseline RDD estimates from estimating Equation **??** in the text within individual cohorts. Years correspond to the tax year t, at the end of which the births occur in the corresponding December or January. Outcomes are marginal tax rates on a first dollar of self employment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

Figure A.10: Changes in Self-Employment Status after NRP Audits



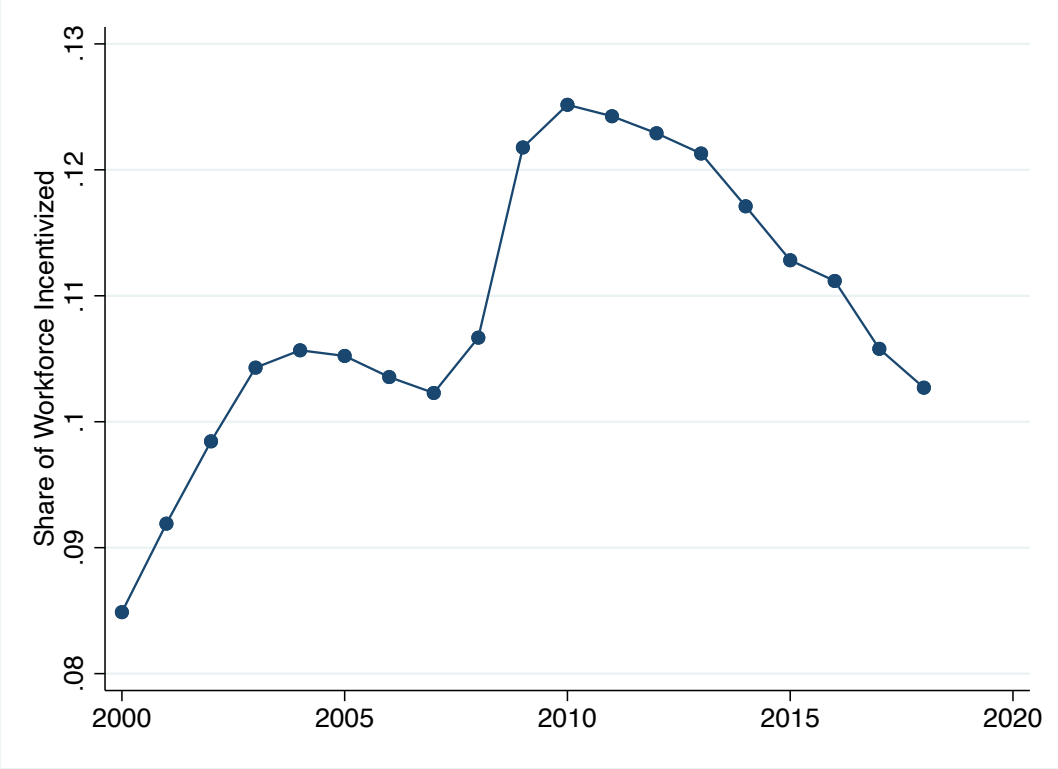
Notes: Figure displays results of audits of a representative stratified random sample of 1040 filers conducted in tax years 2001 and 2006–2014 as part of the IRS's National Research Program (NRP) Individual Income Tax Reporting Compliance Studies. Using sampling weights for representativeness, the figure 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.11: CFS-style Sharp Bunching Share Among Eligible Taxpayers with Children, 1996-2017



Notes: Figure plots the average of the share of tax payers who are sharp bunchers, following the methodology of **?**.

Figure A.12: Share of Workforce with Incentive to Report SE



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

Appendix Tables

Table A.1: RDD Estimates: 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 $< 1^{st}$ *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 $\geq 1^{st}$ *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

Notes: Table displays estimates from the baseline regression discontinuity design specification in Equation (**??**) 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 A.2: RDD Estimates: Lead and Lag Effects

	Any SE Earnings			Any 1099 Earnings		
	Tax Year t-1	Tax Year t	Tax Year t+1	Tax Year t-1	Tax Year t	Tax Year t+1
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All Parents						
Coeff	-0.00101	0.0124**	0.00395**	-0.000625	-0.000592	-0.000792
	(0.00120)	(0.00127)	(0.00135)	(0.00124)	(0.00126)	(0.00124)
N	1382740	1382740	1382740	1382740	1382740	1382740
DV Mean Level, Jan Births	0.0683	0.0719	0.0878	0.0734	0.0761	0.0738
Panel B. With Wages $< 1^{st}$ *EITC Kink*						
Coeff	-0.00139	0.0446**	0.0146**	-0.00429	0.000536	-0.000281
	(0.00270)	(0.00315)	(0.00335)	(0.00258)	(0.00281)	(0.00272)
N	349240	349240	349240	349240	349240	349240
DV Mean Level, Jan Births	0.0900	0.108	0.143	0.0813	0.0973	0.0908
Panel C. With Wages $\geq 1^{st}$ *EITC Kink*						
Coeff	-0.000849	0.00155	0.000433	0.000626	-0.000940	-0.000937
	(0.00131)	(0.00131)	(0.00140)	(0.00141)	(0.00139)	(0.00139)
N	1033500	1033500	1033500	1033500	1033500	1033500
DV Mean Level, Jan Births	0.0610	0.0597	0.0692	0.0707	0.0689	0.0680

Notes: Table displays estimates from the baseline regression discontinuity design specification in Equation (**??**) 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 years $t-1$, t, and $t+1$, as specified. We report mean levels of each dependent variable in years $t-1$, t, and $t+1$ for individuals with first births in January of $t+1$ in each subsample. Robust standard errors are displayed in parentheses.

Table A.3: 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	✓	✓	✓
Year FE	✓	✓	✓

Notes: 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 **?**. 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 A.4: Workforce Share Reporting Self-Employment on Schedule SE in Counterfactual Scenarios

	Baseline	Scenario 1 RDD Adjusted	Scenario 2 Incentivized = Unincentivized
2000	0.0974	0.0974	0.0974
2001	0.0980	0.0973	0.0976
2002	0.0990	0.0975	0.0982
2003	0.1030	0.1019	0.1019
2004	0.1062	0.1038	0.1045
2005	0.1095	0.1093	0.1074
2006	0.1102	0.1086	0.1076
2007	0.1112	0.1078	0.1080
2008	0.1096	0.1069	0.1055
2009	0.1129	0.1096	0.1076
2010	0.1151	0.1105	0.1083
2011	0.1173	0.1133	0.1092
2012	0.1176	0.1133	0.1097
2013	0.1171	0.1126	0.1091
2014	0.1184	0.1137	0.1100
2015	0.1177	0.1133	0.1096
2016	0.1177	0.1128	0.1096
2017	0.1164	0.1132	0.1084
2018	0.1154	0.1108	0.1074

Notes: 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 **??** 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. See text for further details.

Table A.5: Workforce Share with 1099-Reported Non-Employee Compensation in Counterfactual Scenarios

	Baseline	Scenario 1 RDD Adjusted	Scenario 2 Incentivized = Unincentivized
2000	0.0906	0.0906	0.0906
2001	0.0888	0.0892	0.0894
2002	0.0929	0.0927	0.0937
2003	0.0948	0.0951	0.0950
2004	0.0965	0.0967	0.0964
2005	0.0963	0.0966	0.0959
2006	0.0980	0.0989	0.0976
2007	0.0984	0.0988	0.0982
2008	0.0971	0.0973	0.0968
2009	0.0939	0.0944	0.0940
2010	0.0952	0.0953	0.0955
2011	0.0970	0.0977	0.0971
2012	0.0976	0.0979	0.0973
	[0.0975]	[0.0978]	[0.0972]
2013	0.0976	0.0984	0.0970
	[0.0974]	[0.0982]	[0.0968]
2014	0.0993	0.0996	0.0984
	[0.0980]	[0.0983]	[0.0971]
2015	0.1014	0.1021	0.1001
	[0.0974]	[0.0981]	[0.0961]
2016	0.1046	0.1054	0.1030
	[0.0969]	[0.0977]	[0.0954]
2017	0.1086	0.1091	0.1069
	[0.0996]	[0.1000]	[0.0981]
2018	0.1114	0.1121	0.1096
	[0.1002]	[0.1009]	[0.0986]

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

B 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 **?** and examine how self-employment reporting changes around a person's first childbirth, when they become eligible for a generous credit. We expand upon **?** 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 B.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 B.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:

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

where i indexes parent, t indexes year, $a(i)$ gives the age of i. FirstChildbirth$_i$ is i's year of first birth. $g(i)$ is the parent's gender, thus allowing for time and age effects to differ by parental gender.[1] 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, $p \in \{2003-2005, 2004-2006, ..., 2012-2014\}$.

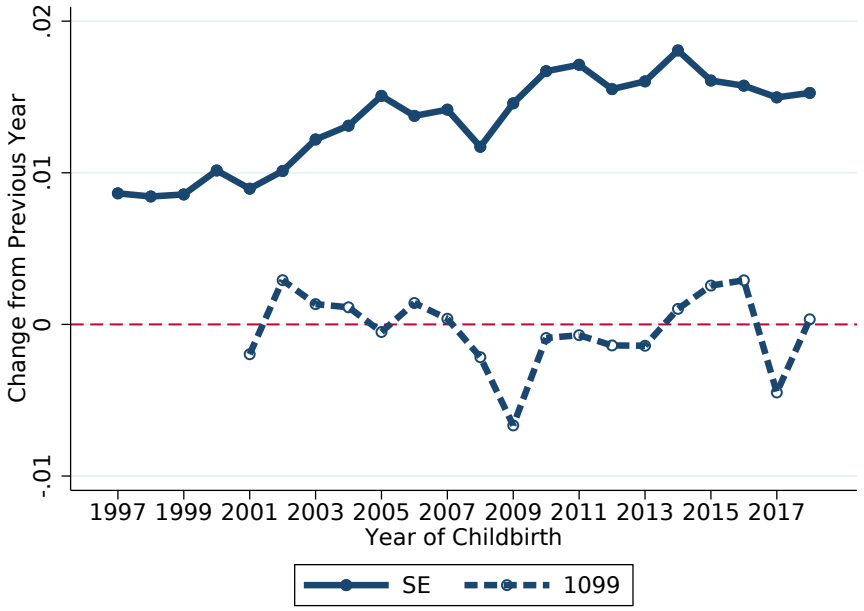
We exclude an indicator for the period one year prior to first birth, so that the event-time coefficients are all relative to period -1, and examine an event window of 4 years pre and post event ($k \in \{-4, ...4\} \setminus -1$). Standard errors are clustered at the individual level.

Figure B.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 B.3 reports estimates separately by gender of the parent. As we found earlier, these changes over time are largely driven by mothers.

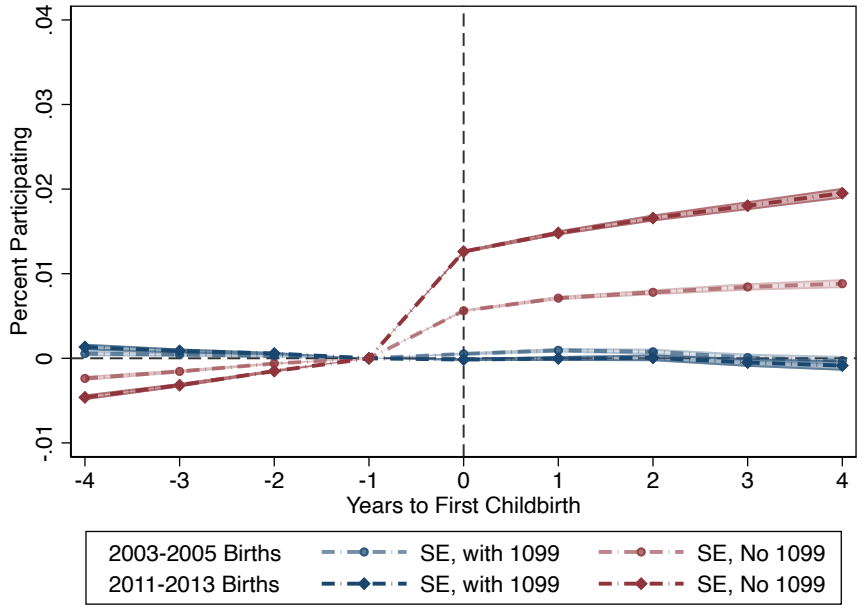
[1] Accordingly, the event-study coefficients are the average of coefficients run separately for men and women, which we report in Appendix Figure B.3.

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

(a) Changes by year of First Birth: 1997-2018 Births



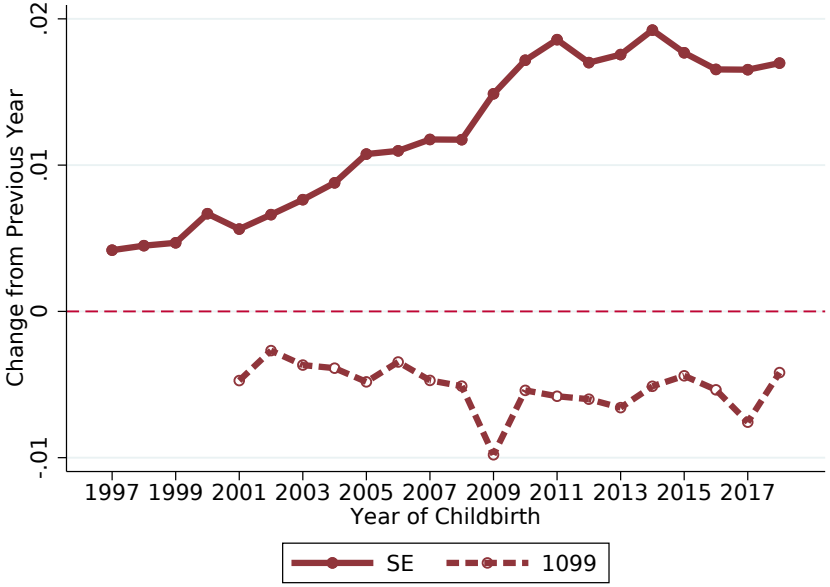
(b) Childbirth Event Study Estimates, 2003-2005 Births Versus 2011-2013 Births



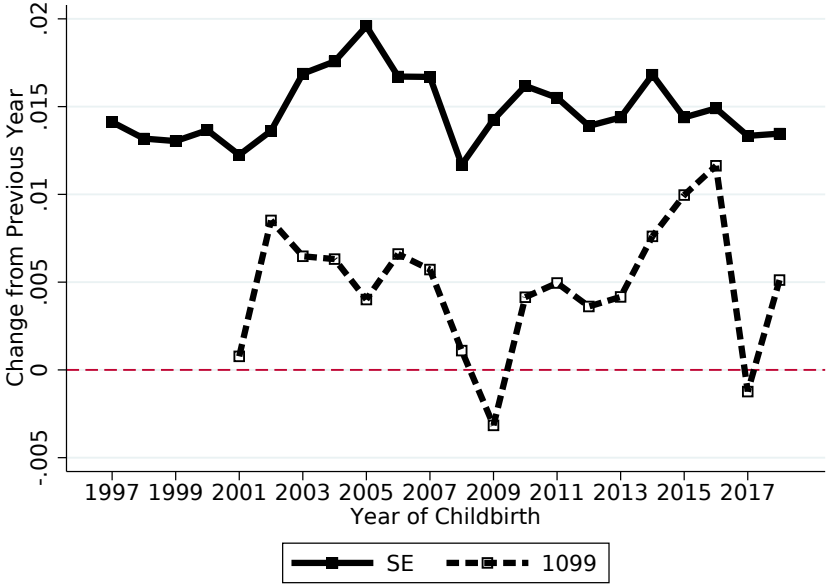
Notes: Panel A shows the average change in propensity to file SE (solid line) or receive a 1099 Information Return (dashed line), in the year of first childbirth reported on the x-axis. Panel B plots event study coefficients for separate regressions run on the indicated time-period and for the indicated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

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

(a) New Mothers



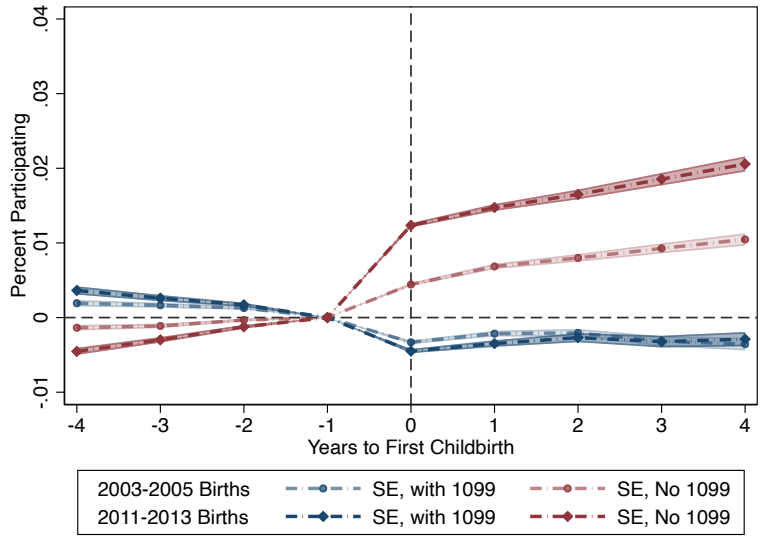
(b) New Fathers



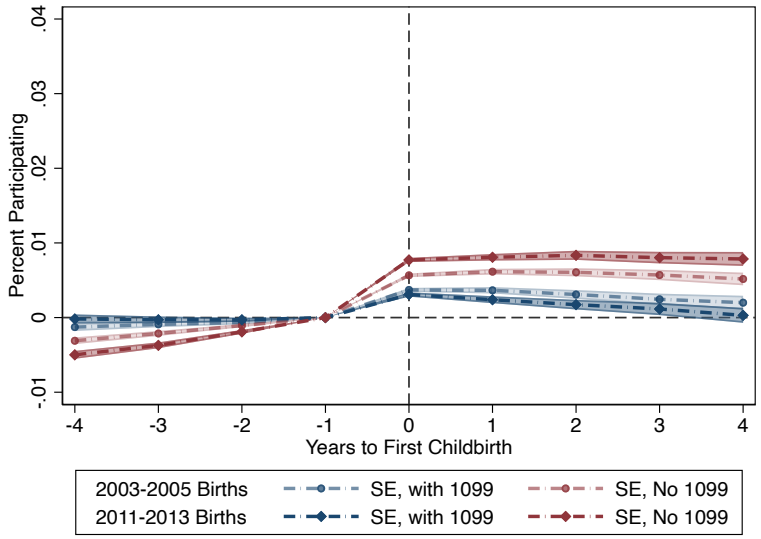
See notes for Figure B.1a.

Figure B.3: Childbirth Event Study Estimates, Additional Estimates

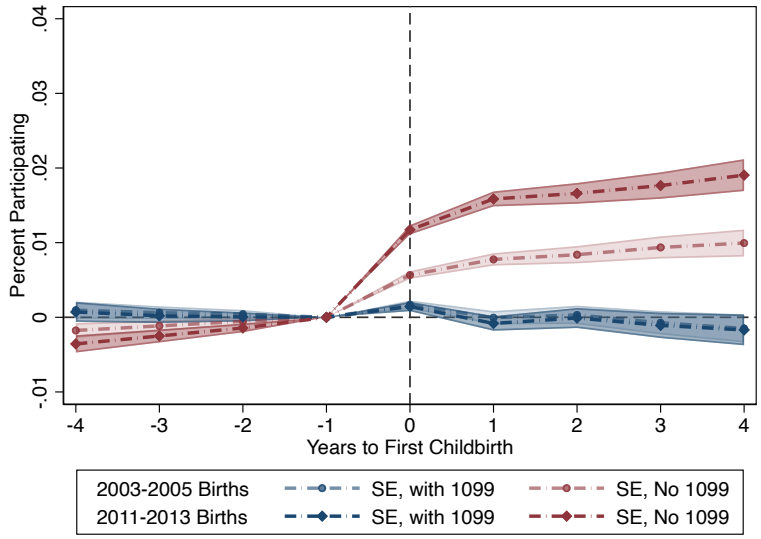
(a) Mothers Only



(b) Fathers Only



(c) All Parents, December Births Only



Notes: See notes for Figure B.1b.

C 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 **?** 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."[2] Starting in 2012, several online

[2]This measure was included in The Housing and Economic Recovery Act of 2008, but did not take effect

intermediaries in the OPE began issuing the new Form 1099-K instead of 1099-MISC for non-employee compensation. The income paid to gig workers on OPE labor platforms is, for all practical purposes, non-employee compensation. However, 1099-Ks are also issued for income from sales that is not non-employee compensation. We therefore 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 companies cannot be identified by this method.[3] 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; (**?**) 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,

until the 2011 tax year.

[3]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.

populated by the Social Security Administration.

D Proofs of Propositions

D.1 Proof of Propositon 4.1

If non-complier types are not eligible for or not aware of the credit ($E_i = 0$ or $I_i = 0$, respectively), or if they are eligible and aware but have wage earnings in excess of the top of the phase-in range ($E_i = 1$, $I_i = 1$, and $w_i > \hat{r}$) then they expect all self-employment profits to be taxed at a positive marginal tax rate and report $z_i = 0$. By contrast, if $E_i = 1$, $I_i = 1$, and $w_i < \hat{r}$, then there is a net benefit from reporting positive self-employment earnings up to $z_i = \hat{r} - w_i$.

In principle, it is optimal for all non-complier types with $E_i = 1$, $I_i = 1$, and $w_i < \hat{r}$ to report *exactly* $z_i = \hat{r} - w_i$. One can relax this sharp condition without fully modelling reporting decisions as in **?** by considering that the government likely knows this and suspects individuals who report exactly $z_i = \hat{r} - w_i$ are reporting fraudulently. One might then suppose that individuals, knowing the government to behave this way, choose some amount $z_i \in (0, \hat{r} - w_i]$ with the amount depending on personal attitudes towards detection risk.

D.2 Proof of Propositon 4.2

Individuals without children or with wage earnings above the refund-maximizing amount ($BTK_i = 0$) report self-employment honestly with probability $1 - \theta$ and report zero self employment earnings with probability θ, such that

$$\rho^0 = (1 - \theta)\sigma^0 + \theta \times 0 = (1 - \theta)\sigma^0$$

which is strictly below the true unemployment rate.

For individuals with $BTK_i = 1$, their behavior depends both on their compliance type N_i and their information type I_i. All honest types report honestly. Non-complier types perceive a strictly positive marginal tax rate and therefore report zero self-employment earnings when they are unaware of the credit (I_i=0). Non-complier types with $BTK_i = 1$ who are aware of the credit *always* report positive self-employment earnings. Accordingly

$$\rho^1 = (1 - \theta)\sigma^1 + \theta\lambda \times 1 + \theta(1 - \lambda) \times 0 = (1 - \theta)\sigma^1 + \theta\lambda$$

In a low information environment, the self-reported self-employment rate among these individuals will be below the true self-employment rate. However, with high degrees of informa-

tion—specifically, if the share of non-compliant filers who know about the credit exceeds the share with actual self-employment profits—the reported self-employment rate can exceed the true rate within this group.

Since the overall shares can be expressed as $\sigma^{tot} = \sigma^0(1-\kappa)+\sigma^1\kappa$ and $\rho^{tot} = \rho^0(1-\kappa)+\rho^1\kappa$, the above results imply that

$$\begin{aligned} \rho^{tot} &= (1-\kappa)\rho^0 + \kappa\rho^1 \\ &= (1-\kappa)[(1-\theta)\sigma^0] + \kappa[(1-\theta)\sigma^1 + \theta\lambda] \\ &= (1-\theta)\sigma^{tot} + \kappa\theta\lambda \end{aligned}$$

thereby completing the proof.