



## Session 4. Future Directions in Tax Administration

**Moderator:**

***Stephanie Parrish***  
*IRS, Online Services*

**Who is Minding the Nanny Tax?**

***Brian Erard***  
*B. Erard & Associates*

**Can IRS Move Paper Filers to Free Assisted Tax Preparation?**

***Rizwan Javaid***  
*IRS, RAAS*

**Creating a Fully Synthetic Individual Income Tax Return Data File and Validation Server**

***Len Burman***  
*Urban-Brookings Tax Policy Center*

**Discussant:**

***Laura Kawano***  
*University of Michigan*

# Who is Minding the Nanny Tax?

Brian Erard

B. Erard & Associates



# What is the Nanny Tax?

- When households employ domestic help such as a nanny, senior caregiver, housekeeper, health aide, cook, or gardener, they are required to file Form 1040 Schedule H to report payroll taxes (Social Security, Medicare, FUTA).
- They also have state payroll responsibilities, including state unemployment insurance (SUTA), and, potentially, workers' comp, disability insurance, and/or other state taxes.
- Domestic workers, of course, are responsible for paying income tax on their earnings.
- They also need to make quarterly estimated tax payments unless their employer voluntarily withholds income tax.

# Why is the Nanny Tax important?

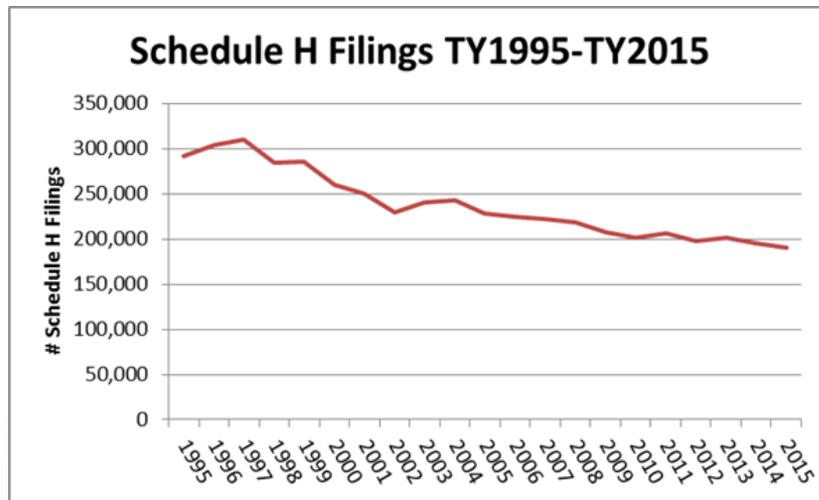
- It is the pathway for domestic workers to receive employment-related benefits, such as Social Security, Medicare, and Unemployment Insurance.
- Being paid over the table allows workers to establish credit for renting an apartment, buying a car, securing a mortgage, or obtaining a personal loan.
- Important social interactions:
  - Domestic workers are relatively vulnerable, while the households that hire them are typically relatively well off.
- Household workplace is an area where all of the labor market and social issues surround illegal immigration come to a head.
- Nanny tax evasion results in an inefficient allocation of resources and tax revenue losses.

# Brief history of federal nanny taxes

- Prior to 1995, household employers had to file a quarterly return (Form 942) to report FICA and a separate annual return (Form 940) to report FUTA.
- In response to the “Nannygate” scandal that prevented Zoë Baird from becoming Attorney General, Congress attempted to simplify the employer requirements to reduce burden and improve compliance.
  - Employers file all federal payroll taxes with their individual income tax return on a new Form (Schedule H).
  - \$50 filing threshold for Form 942 replaced by a higher indexed threshold (currently \$2,100).
  - “Babysitter exemption” for those under age 18.

# Did simplification work?

- Immediate **drop** of 40% in Nanny Tax filings!
- Declining trend in filings over time.
- Ongoing parade of “Nannygate” cases.



# 2 Possible explanations for declining trend in filings

- Decline in direct employment of domestic workers.
  - Shift to hiring businesses to provide household services
    - Merry Maids
    - Child care centers; child care in the provider's home
- Increased noncompliance.

# Estimating the Nanny Tax gap

- From CPS-ASEC, an estimated 734,000 workers with a longest-job as a domestic employee in 2015; 637,000 of them with earnings above the filing threshold.
  - But only 191,000 Schedule H returns filed in that year.
- This does not count many workers who moonlighted or worked for a portion of the year as a domestic employee.
- To account for these additional workers, rely on monthly CPS.

# CPS monthly data

- Average number of domestic employees across the January – December 2015 monthly CPS files is 790,000.
- Outgoing rotation group of monthly survey provides information on second job held (another 69,000 domestic employees).
- So, monthly data suggest an average of 859,000 domestic employees, compared to original ASEC figure of 734,000.
- But, even this figure is too low, because it is based on employment at a point in time.

# Exploiting panel nature of monthly CPS

- Instead of number of individuals who worked as a domestic employee during the reference week in a monthly survey, want to know how many were domestic employees over at least some portion of the year.
- Monthly CPS data has 8 outgoing rotation groups that are interviewed in (up to) four consecutive months in the current year and then again in (up to) four consecutive months the following year.
- By matching individuals across monthly surveys, we can see if they worked as household employees during any of the interview months.

# Number of months as domestic employee over up to 4 interview months

Number of Monthly Reports of Domestic Employment	Weighted Percentage Frequency (A)	Percentage of Months Correctly Assigned as a Domestic Worker (B)	(A) Times (B)
1	33.34%	25%	8.3%
2	16.96%	50%	8.5%
3	12.82%	75%	9.6%
4	36.88%	100%	36.9%
<b>Total</b>	<b>100.00%</b>		<b>63.3%</b>

# Extending the analysis to 8 interview months over 16 month period

- Only about 42% of individuals who worked as a domestic employee for one or more of the 8 interview months would show up in a given monthly survey as a domestic employee in their main job.
- So, the actual number who worked for a household over at least a portion of this period is about 2.4 times as high as the monthly point-in-time measure.

# Additional adjustments

- Survey undercount foreign-born residents
- Remove those with earnings below Schedule H filing threshold.
- Account for the fact that many domestic workers have multiple employers over the course of the year, and each of these employers has a Schedule H filing requirement.

# Estimated Schedule H filing compliance rate of only 5.3%!

Occupation	Frequency	Percentage Frequency	Average # Employers	Total # Employers
Maid/Housekeeper	834,111	45.8%	3	2,502,332
Child Caregiver	549,214	30.1%	1.5	823,821
Personal Home Care Aide	298,698	16.4%	1.1	328,568
Nurse/Therapist/Health Aide	32,818	1.8%	1.1	36,100
Cook/Food Prep. Worker	13,857	0.8%	1.1	15,242
Other Miscellaneous	93,734	5.1%	1.1	103,108
Total	1,822,432	100.0%	2.1	3,809,171
<b>Actual Number of Sch. H Filings</b>	<b>190,852</b>			
<b>Required Number of Sch. H Filings</b>	<b>3,619,000</b>			
<b>Sch. H Filing Compliance Rate</b>	<b>5.3%</b>			

# Estimated Payroll and Income Tax Gap

\$3.3 to \$5.7 billion in unpaid payroll and income taxes (compliance rate of between 22% and 32%).

# How to change the culture of noncompliance?

- Education and outreach to address Nanny Tax misconceptions, such as:
  - Nannies are “independent contractors” .
  - Nanny matching services are responsible for payroll tax collection and remittance.
  - You can avoid Nanny Tax requirements by having an employee sign a contract that states you are not responsible for the tax.
  - “Don’t ask, don’t tell” policy absolves employer from legal and tax issues when hiring undocumented immigrants.
- Improve ambiguous publications and provide an online decision tool to help households and domestic workers clarify whether there is an employer-employee relationship.
  - Establish a rebuttable presumption of employer-employee relationship.
- Coordinate with states to make the compliance process easier and less burdensome.
- Significantly step up enforcement and improve case selection and audit techniques.
  - Consider an amnesty in advance of enforcement surge.
- Consider requiring household employer withholding.



**June 20, 2018**  
**IRS – TPC Research Conference**

## **Can IRS Move Paper Filers to Assisted Tax Preparation?**

*Rizwan Javaid | Internal Revenue Service*  
*Brenda Schafer | Internal Revenue Service*  
*Jacob Goldin | Stanford University*  
*Adam Isen | Department of the Treasury*  
*Tatiana Homonoff | New York University*

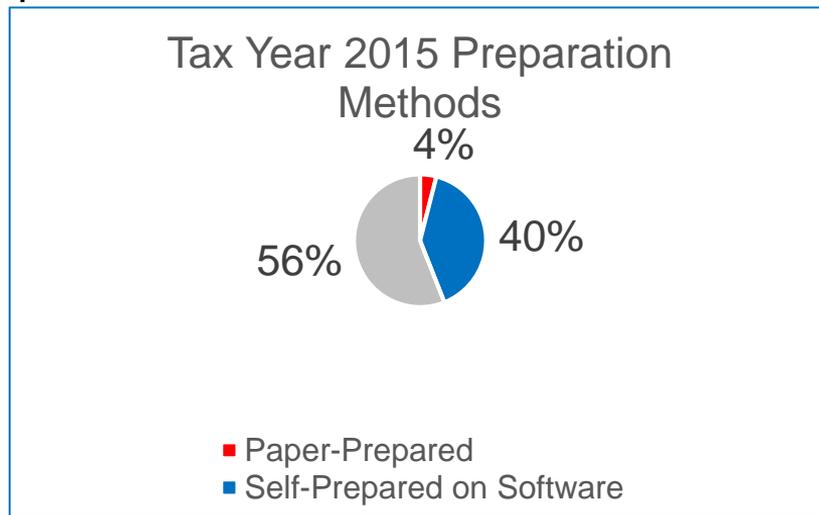
DISCLAIMER: The views and opinions presented in this presentation reflect those of the authors. They do not necessarily reflect the views or the official position of the Internal Revenue Service or the Department of the Treasury.

# Agenda

- Background
- Why the Choice of How to Prepare a Return Matters
- Tax Season 2017 Pilot Design
- Implementation and Data Issues
- Experimental Analysis
- Results
- Further Work

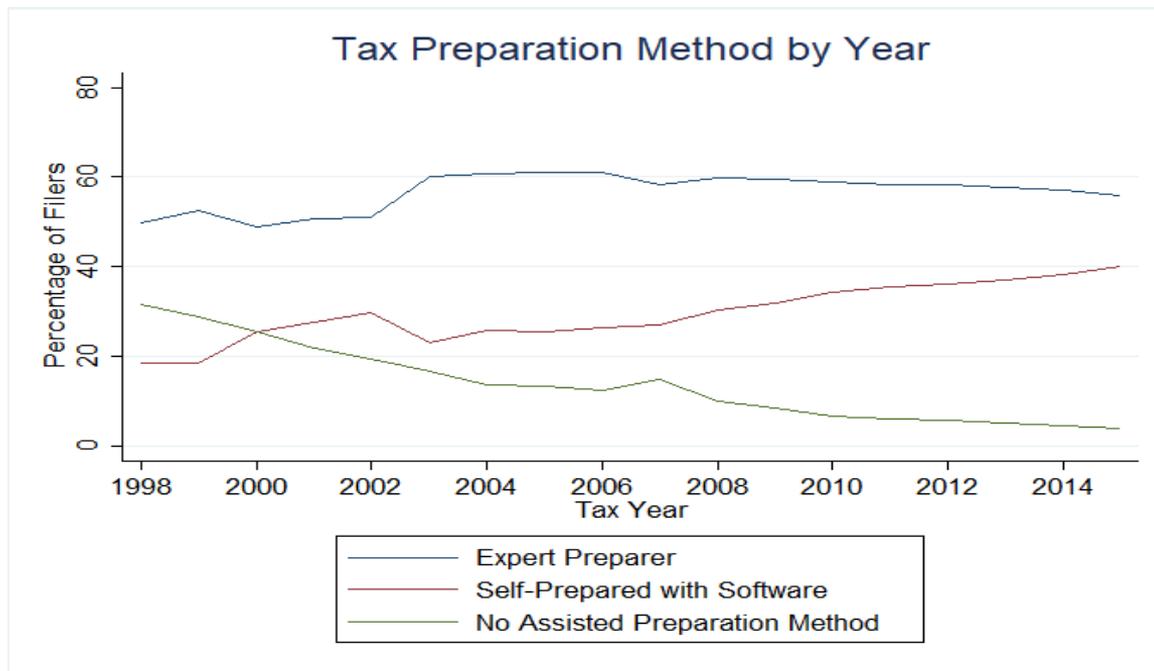
## Background: Tax Preparation Method Choice

- Most taxpayers use an assisted method of tax preparation
  - Self-prepared using software
  - Third-party preparer (e.g., paid professional, VITA)
- Approximately 6 million taxpayers prepared their returns on paper in TY2015



Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2015 Returns. Data extracted November 2017.

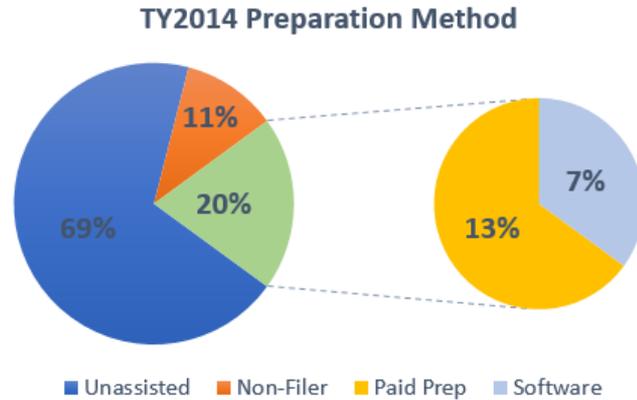
## Background: Tax Preparation Method Trend



Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. Data extracted November 2017.

- Use of a third-party preparer is fairly stable
- Taxpayers who self-prepared on paper in the past seem to be switching to self-preparation on software

# Background: Prior Years Preparation Method



<b>Preparation Methods for TY2012-TY2014</b>	
Self-Prepared Exclusively (all 3 years)	0.50
Used Assisted Prep Method ( $\geq 1$ year)	0.34
Software	0.22
In-Person Preparer	0.15
Didn't File ( $\geq 1$ year)	0.23

Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2012-TY2014 Returns. Data extracted November 2017.

## Background: Free Assisted Tax Preparation

### VITA Program (In-Person Assistance)

- Income must be  $\leq$  \$54,000 (for 2016)
- 67% of paper filers meet the income test

### Free File and MyFreeTaxes Programs

- Income must be  $\leq$  \$64,000 (for 2016)
- 72% of paper filers meet the income test

**Fact:** Only 2% – 3% of apparently-eligible taxpayers use a free assisted method to prepare their returns.

### Potential Barriers

- Lack of awareness (brief “marketing” period, advertising and outreach efforts may not be reaching eligible taxpayers)
- Stigma of going to a VITA site
- No VITA site near them, limited hours, too busy, language barriers
- Lack of internet connectivity
- No perceived need or benefit (easy return, preparer quality concerns)
- Privacy, security concerns
- Inertia

# Why the Choice of How to Prepare a Return Matters

## Taxpayer Compliance Concerns

- Selecting and obtaining forms and instructions may be difficult
  - No access to the Internet or a printer
  - Cost to travel to a site to pick up forms
- Tax law is complex
  - Taxpayer may not read or understand form instructions
  - Confusion may cause missed deductions or credits
- Taxpayer may be unaware of allowed deductions or credits
  - Example: EITC take-up estimates are 92% for assisted methods) and 85% for paper-prepared returns
- Math errors delay refunds
  - 0.5% for assisted vs 16.1% for unassisted

## Tax Administration Concerns

- Cost to process electronic return (~\$0.25) vs. paper (~\$4.00)
- Generally higher error rates than assisted preparation

# Tax Season 2017 Outreach Pilot

## Outreach Pilot Plan

- Send one postcard in January 2017 with information about free assisted tax preparation options
- Message variations promoting VITA, Free File, MyFreeTaxes, and addresses for closest VITA sites to taxpayer
- Five treatment groups with ~80,000 taxpayers per group
- Treated group received one postcard, control group received no postcard

## Sample Population

- TY2015 paper filers whose AGI were within the VITA eligibility ranges
- Restrict sample to taxpayers within 30 miles of at least two Tax Season 2017 VITA site

## Final Sample

- N=2,041,537 | Treated=627,268 (31%)
- Stratified random sampling
- Oversampled apparently-EITC eligible taxpayers who did not claim EITC on their TY2015 return

## Measure

- TY2016 Return Preparation Method
- TY2016 Return Filing Method

# All Treatments: Address Side of Postcard

Social norms  
messaging



Internal Revenue Service  
c/o Westat  
1600 Research Blvd. RW2634  
Rockville, MD 20850-3129  
RETURN SERVICE REQUESTED

IRS

2 out of 3 people  
qualify for **FREE** tax  
preparation services!

PRESORTED  
FIRST-CLASS MAIL  
POSTAGE & FEES PAID  
IRS  
PERMIT NO. G-48

112200000051

ATTN: ANITA VENTIMACCHIATO  
1 STARBUCKS BLVD  
SEATTLE, WA 98101-1234

# Treatment 1: VITA Addresses and Free File

Here are 2 easy options for **FREE** Tax Preparation

## Option 1



### File Online

Free tax software helps you prepare your return from home.

- People with income below \$64,000 qualify
- Prepare your return from home
- Free tax software

Step-by-step directions at: [www.irs.gov/FreeFile](http://www.irs.gov/FreeFile)

## Option 2



### File In-Person

IRS-certified volunteers prepare your taxes for free at a

Volunteer Income Tax Assistance site

- Most people with income below \$54,000 qualify

Please bring with you:

- Photo Identification
- Social Security Card(s)
- Tax Documents (e.g., W-2, 1099-MISC)

Find locations at: [www.irs.gov/vita](http://www.irs.gov/vita)  
or call **800-906-9887**

#### Locations near you:

Community Activities Center  
305 S Dogwood Dr  
Harrisonburg, VA 22801-1926  
"Opens 2/1/2017; SUN: 9:00AM - 12:00PM; TUE: 8:30AM - 11:30AM, 12:30PM - 3:30PM; SAT: 9:00AM - 12:00PM"

James Madison University VITA  
Lake View-Bluestone Drive  
Harrisonburg, VA 22807-0000  
Opens 1/28/2017; FRI: 4:00PM - 8:00PM

Publication 5270 (12-2016) Catalog Number 69330F Department of the Treasury Internal Revenue Service [www.irs.gov](http://www.irs.gov)

# Treatment 2: VITA Info and Free File

Here are 2 easy options for **FREE** Tax Preparation

## Option 1



### File Online

Free tax software helps you prepare your return from home.

- People with income below \$64,000 qualify
- Prepare your return from home
- Free tax software

Step-by-step directions at: [www.irs.gov/FreeFile](http://www.irs.gov/FreeFile)

## Option 2



### File In-Person

IRS-certified volunteers prepare your taxes for free at a Volunteer Income Tax Assistance site

- Most people with income below \$54,000 qualify

Please bring with you:

- Photo Identification
- Social Security Card(s)
- Tax Documents (e.g., W-2, 1099-MISC)

Find locations at: [www.irs.gov/vita](http://www.irs.gov/vita) or call **800-906-9887**

Publication 5270-A (12-2016) Catalog Number 69332B Department of the Treasury Internal Revenue Service [www.irs.gov](http://www.irs.gov)

## Treatment 3: VITA Addresses Only

Here is an easy option for **FREE** Tax Preparation



File In-Person

IRS-certified volunteers prepare your taxes for free at a Volunteer Income Tax Assistance site

- > Most people with income below \$54,000 qualify

**Locations near you:**

Buker Community Center  
22 Armory St  
Augusta, ME 04330-1234  
Appointment Only  
(207) 582-3053

Millinocket Timber Cruisers Snow  
Mobile Club  
Northern Timber Cruise Trail  
Millinocket, ME 04462-1234  
Opens 2/1/2017; TUE: 10:00AM -  
3:00PM; WED: 10:00AM -  
3:00PM

Please bring with you:

- > Photo Identification
- > Social Security Card(s)
- > Tax Documents  
(e.g., W-2, 1099-MISC)

Find locations at: [www.irs.gov/vita](http://www.irs.gov/vita) or call **800-906-9887**

Publication 5270-B (12-2016) Catalog Number 69333M Department of the Treasury Internal Revenue Service [www.irs.gov](http://www.irs.gov)

## Treatment 4: Free File Only

Here is an easy option for **FREE** Tax Preparation



File Online

- People with income below \$64,000 qualify
- Prepare your return from home
- Free tax software

Step-by-step directions at: [www.irs.gov/FreeFile](http://www.irs.gov/FreeFile)

IRS Free-File is a partnership between the IRS and the Free File Alliance, a group of industry-leading commercial tax preparation companies that have agreed to provide free commercial online tax preparation and electronic filing. IRS does not endorse any individual Free File Alliance company or retain any taxpayer information entered on the Free File Site.

Publication 5270-C (12-2016) Catalog Number 69334X Department of the Treasury Internal Revenue Service [www.irs.gov](http://www.irs.gov)

## Treatment 5: MyFreeTaxes Only

*Here is an easy option for **FREE** Tax Preparation*



**File Online**

- People with income below \$64,000 qualify
- Prepare your return from home
- Free tax software

Step-by-step directions at: [www.MyFreeTaxes.com](http://www.MyFreeTaxes.com)

MyFreeTaxes is a free online tax preparation service provided by the United Way (a non-profit organization).  
IRS does not endorse MyFreeTaxes or retain any taxpayer information entered on the MyFreeTaxes site.

Publication 5270-D (12-2016) Catalog Number 69336T Department of the Treasury Internal Revenue Service [www.irs.gov](http://www.irs.gov)

## Implementation and Data Issues

### Printer Errors

- Treatment 1: 50% didn't receive; 50% received two copies
- Treatments 1 and 3: Addresses/hours sometimes misprinted
- Poor print quality (Many undelivered postcards had significant smearing, bends, and tears)

A second wave of Treatments 2, 4, and 5, was sent to an additional sample in March 2017 who met the following conditions:

- Must not have already filed 2016 return
- Filed 2015 return in second half of filing season

**Missing “V” Code Returns Data** – Data coding error for most of the tax season. It is likely that 2M software-prepared paper returns appear to be paper-prepared returns

**Free File and MyFreeTaxes Returns Data** – Identified as software/assisted prep, but not by software type

## Experimental Analysis

$$y_i = \beta \text{Postcard}_i + \delta X_i + \varepsilon_i$$

$$\text{Postcard}_i = \beta \text{Treated}_i + \delta' X_i + \varepsilon'_i$$

$y$  denotes TY2016 return filing method

$\text{Postcard}$  denotes receipt of treatment message

$X$  contains indicators and interactions for

- Wave 2 eligibility
- Over-sampled EITC non-claimant group
- Mishandled treatment group (Treatment 1 from January)

Weighted to account for over-sampling of groups into treatment

## Preliminary Result: Increase in Use of Assisted Preparation

Postcards including VITA and FF information led to:

- Modest increase in assisted preparation
  - Results may have shown slightly higher impact if not for “V” Code returns data issue
- Significant increase in Software and VITA
- Significant decrease in use of Paid Preparer
- Significant decrease in likelihood of not filing at all

	Any Assisted	Software	VITA	Paid Preparer	Non-Filer
T2: VITA Info & FF	0.629 *** (0.125)	0.558 *** (0.103)	0.203 ** (0.032)	-0.136 * (0.079)	-0.188 ** (0.095)
Control Mean	21.70%	13.10%	1.00%	7.50%	12.20%
N	2,041,537				

Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2016 Returns. Data extracted November 2017.

Standard errors reported in parenthesis.

\*90% significance | \*\*95% significance | \*\*\*99% significance

## Preliminary Result: Switch to an Assisted Method Impact of Message Variation

	Preparation Method Switch			
	Software	VITA	Paid Preparer	Any Assisted
T2: VITA Info & FF	0.581 *** (0.114)	0.214 *** (0.035)	-0.199 * (0.088)	0.606 *** (0.139)
T4,T5: Online Prep Only	0.570 *** (0.091)	0.022 (0.027)	-0.132 (0.071)	0.464 *** (0.112)
T3: VITA Addresses	0.200 (0.136)	0.563 *** (0.049)	-0.018 (0.108)	0.730 *** (0.168)
Control Mean	13.10%	1.00%	7.50%	21.70%

Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2016 Returns. Data extracted November 2017.

Standard errors reported in parenthesis.

\*90% significance | \*\*95% significance | \*\*\*99% significance

- Software assisted preparation experienced the same effect whether only information about free online preparation was provided or both VITA and software information was provided
- The combination of VITA and FF resulted in a significant increase for VITA
- The individuals who only received information about VITA were almost 60% more likely to use VITA than the control group

## Preliminary Result: Impact of Providing VITA Addresses

Providing information on how to find a VITA site was more effective at increasing usage than only providing general VITA information

- About 1% of the control group switched to a VITA site
- Smudges and address errors may have reduced impact of the outreach
- Providing VITA addresses resulted in the largest increase when taxpayers were living less than 2 miles from a VITA site

	VITA Office Location			
	Any Distance	<10 Miles	<5 Miles	<2 Miles
T2: VITA Info & FF	0.255 *** (0.062)	0.272 *** (0.283)	0.283 *** (0.082)	0.236 ** (0.118)
T1: VITA Addresses & FF	0.397 *** (0.065)	0.420 *** (0.072)	0.481 *** (0.087)	0.591 *** (0.134)
Control Mean	0.94%	0.97%	1.03%	1.11%
N	994,421	832,373	622,304	294,901

*Treatment 1 sample restricted to individuals who received the treatment*

*Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2016 Returns. Data extracted November 2017.*

*Standard errors reported in parenthesis.*

*\*90% significance | \*\*95% significance | \*\*\*99% significance*

## Preliminary Result: Effects of Timing

- The early mailing was more effective in moving taxpayers to an assisted tax preparation method
- The taxpayers who received the March mailing and switched to an assisted method were more likely to use a paid preparer

	Any Assisted	Software	VITA	Paid Prep	Nonfiler
Treated (Early Intervention)	0.512 *** (0.101)	0.574 *** (0.082)	0.086 *** (0.025)	-0.154 ** (0.064)	-0.121 (0.074)
Treated (Late Intervention)	0.569 (0.057)	0.519 (0.113)	0.055 (0.033)	0.005 * (0.850)	-0.196 (0.107)
Control	21.70%	13.10%	1.00%	7.50%	12.20%

*Treatment 1 sample restricted to individuals who received the treatment*

*Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2016 Returns. Data extracted November 2017.*

*Standard errors reported in parenthesis.*

*\*90% significance | \*\*95% significance | \*\*\*99% significance*

## Preliminary Result: Treatment Effects on Sample by Age

	Any Assisted	Software	VITA	Paid Prep	Unassisted	Non-Filer
Any Treatment (Age <=30)	0.614 *** (0.147)	0.739 *** (0.131)	0.074 *** (0.023)	-0.191 ** (0.095)	-0.280 (0.156)	-0.335 *** (0.107)
Any Treatment (Age >30)	0.554 *** (0.080)	0.383 *** (0.061)	0.194 ** (0.025)	-0.022 (0.052)	-0.535 *** (0.095)	-0.019 (0.066)
Control Mean	21.70%	13.10%	1.00%	7.50%	66.00%	12.20%

Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2016 Returns. Data extracted November 2017.

Standard errors reported in parenthesis.

\*90% significance | \*\*95% significance | \*\*\*99% significance

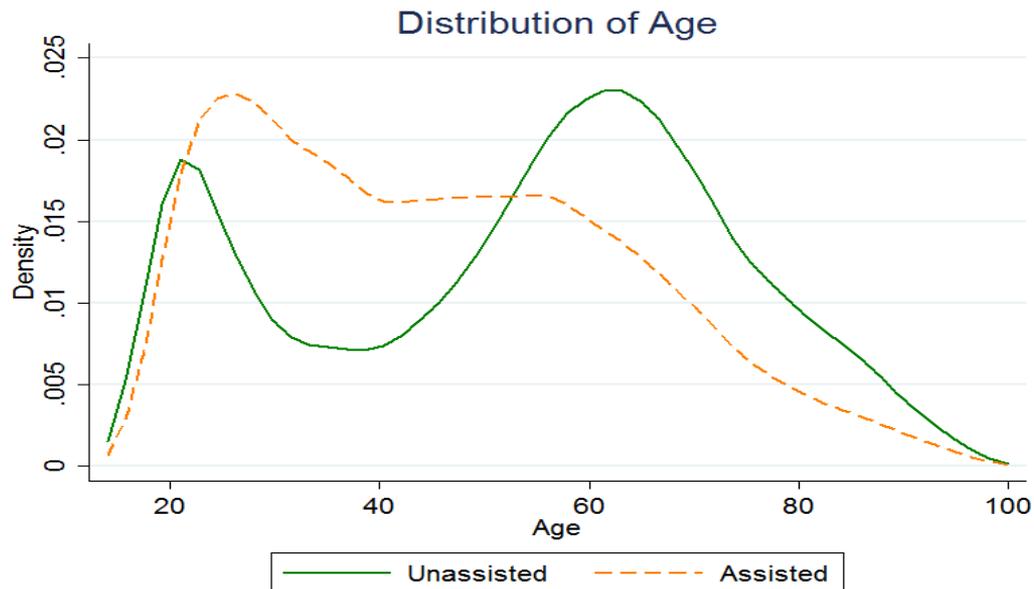
- Both age groups showed significant increases in assisted tax preparation
- Younger taxpayers were more likely to use software
- Older taxpayers were more likely to use VITA
- Older taxpayers had a significant decrease in unassisted preparation
- Younger taxpayers had a significant decrease in non-filing

## Further Work

- Continue data analysis
  - Root cause identification
  - Treatment persistence – Review TY2017 preparation and filing methods on same set of taxpayers with no additional outreach
  - Research whether the nonfilers were required to file
- Consider testing focused outreach to segments with the largest positive impact
- Consider messaging around reduction of math errors as a benefit of using an assisted method
- Extend research to nonfilers

# Appendix

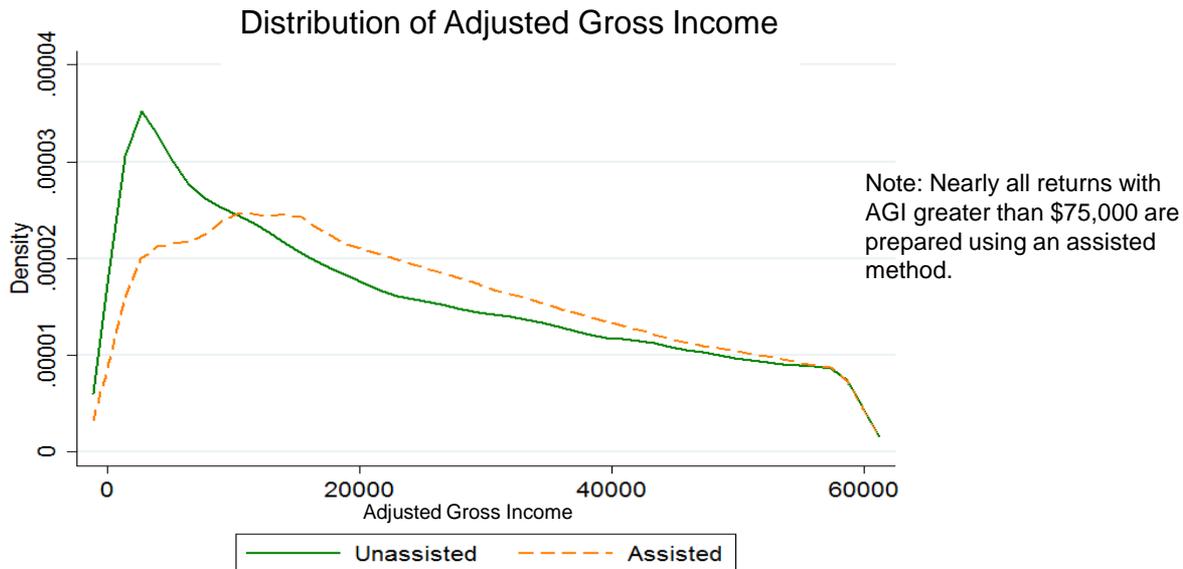
## Background: Tax Preparation Method by Age



Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2015 Returns. Data extracted November 2017.

- Assisted tax preparation choice is representative of the taxpayer population
- Self preparation on paper is more prevalent among taxpayers under age 30 or over age 60

# Background: TY2015 Preparation Method by AGI



Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2015 Returns. Data extracted November 2017.

Potential reasons why lower-income taxpayers may choose paper preparation:

- Simple tax situation (e.g., W-2 only, no dependents, no itemized deductions)
- Cost (real or perceived) of paid preparation or software
- Distrust of online filing
- Habit

## Related Literature

Descriptive work on filing methods and tax return characteristics

- Guyton et al. (2003)
- Marcuss et al. (2013)
- Davis-Smith et al. (2016)

Effect of tax preparation method on taxpaying behavior

- Paid preparers: Kopczuk & Pop-Eleches (2007)
- Online software: Gunter (2016)

“Nudge” behavior change literature

- Benefit take-up: Bhargava & Manoli (2015); Guyton et al. (2015); Manoli & Turner (2016); Chetty & Saez (2013)
- Compliance: Hallsworth et al. (2017); Chirco et al. (2017); Meiselman (2017)

# Tax Return Characteristics by Filing Method

	<b>Assisted</b>	<b>Unassisted</b>
<b>Age</b>		
Age (mean)	46	54
Age (median)	44	57
<b>Income</b>		
Household Income (mean)	\$58,871	\$47,290
Household Income (median)	\$36,338	\$31,588
Any Business Income	12.3%	8.2%
<b>Family Size</b>		
Joint Filer	35.9%	33.3%
Persons on Return	1.93	1.44

Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2015 Returns. Data extracted November 2017.

# Treatment Group Balance

	Coefficient on Treated (p-value)	F-stat to reject balance across treatment groups
Household Income (2015)	8.78 (0.74)	0.16 (0.99)
EITC Amount (2015)	0.16 (0.85)	1.53 (0.14)
Persons on Return (2015)	0.01 (0.35)	0.79 (0.61)
Math Error (2015)	-0.00 (0.93)	0.85 (0.55)
EITC Notice (2015)	0.00 (0.65)	1.15 (0.33)
W-2 Wages (2016)	-44.35 (0.15)	0.49 (0.86)

Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2015 Returns. Data extracted November 2017.

## Tax Claiming Behavior by Filing Method

	<b>Assisted</b>	<b>Unassisted</b>
<b>Tax Benefits</b>		
Itemizer Rate	30.0%	21.2%
EITC Claim Rate	19.5%	6.9%
<b>Mistakes</b>		
Math Error	0.5%	16.1%
Missed EITC Notice	0.1%	2.9%

*Source: IRS, Compliance Data Warehouse. Individual Return Transaction File. TY2015 Returns. Data extracted November 2017.*

# Summary of Experimental Variation

Treatment Group		Online Info	In-Person Info	VITA Addresses	Included in Wave 1	Included in Wave 2	Alternate Online Info
(1)	Software + VITA	X	X		X	X	
(2)	Software + VITA w/ Addresses	X	X	X	X		
(3)	VITA Only		X	X	X		
(4)	Software Only	X			X	X	
(5)	Alternate Software	X			X	X	X

# Estimating the Effects of Tax Reform on Compliance Burden

June 22, 2017

Daniel L. Berger, Eric Toder, Victoria Bryant, John Guyton and Patrick Langetieg  
IRS – TPC Conference



**TAX POLICY CENTER**  
URBAN INSTITUTE & BROOKINGS INSTITUTION

- Compliance costs are one part of the resource cost of taxation, these costs reflect the social cost imposed by taxes
- Slemrod (2005)
  - Compliance costs are predominately time and out of pocket expenses
  - These costs include record keeping, preparation, learning about new forms / laws, lawyers, accountants, software etc.
- What can be done to lower compliance costs?

- TPC has recently built a version of the Individual Taxpayer Burden Model (ITBM) used by IRS RAAS into TPC's microsimulation model
- IRS developed an adapted version of the model to work specifically with the SOI Public Use File (PUF)
- This model allowed TPC to analyze baseline compliance costs and changes in compliance costs associated with reform plans

- Rational taxpayer cost-minimization framework
  - Decreasing marginal costs with income
  - Time / money trade off based on productivity
- Calibrated to observe behavior
- Used in conjunction with tax calculator
- Compliance Cost Factors
  - Economic Activity
  - Tax preparation method
  - Complexity of taxpayer's reporting requirements

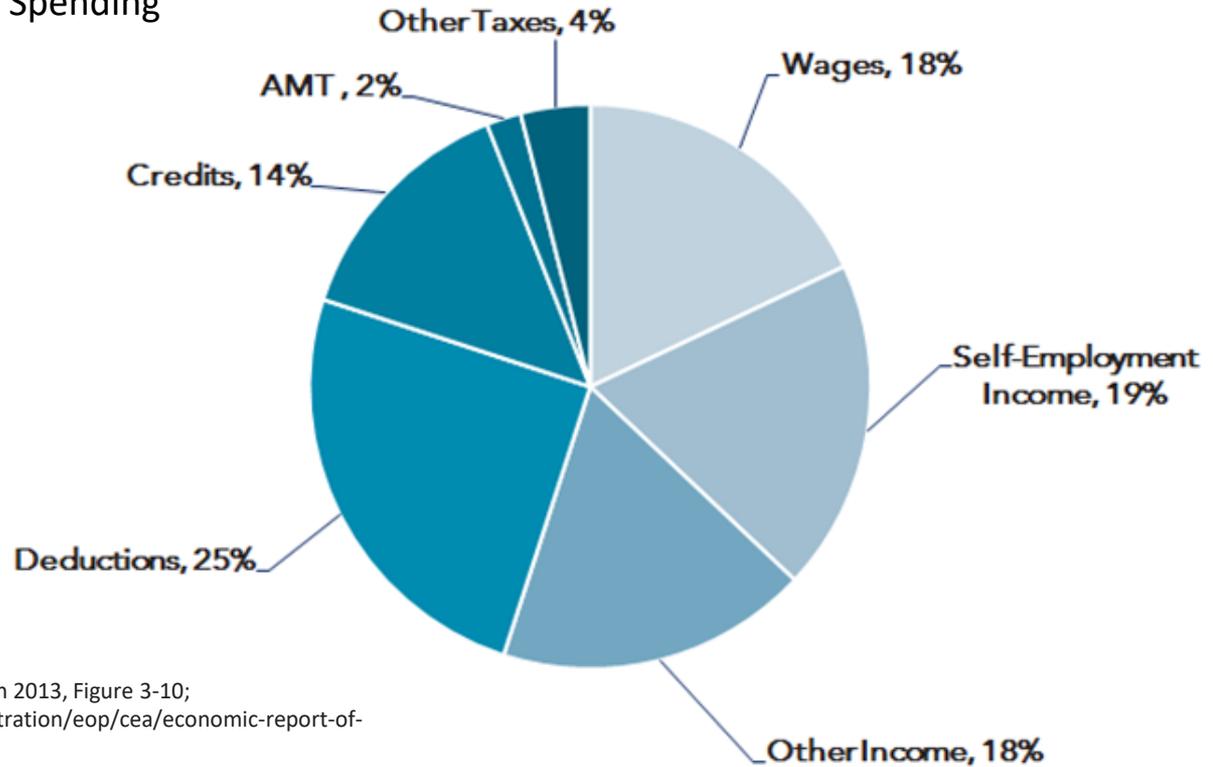
- Capture the degree to which reporting requirements demand additional recordkeeping
- Examples of the categories of increasing difficulty
  - **Low:** wages, interest, dividends
  - **Medium:** EITC, itemized deductions, business income
  - **High:** AMT credits, AMT taxable income, rental depreciation,

- Coefficients include preparation method, complexity categories, tax return line counts and modified positive income (MPI)
- The TPC adapted model is stratified by filing status
- Complexity category coefficients are slightly higher in adapted model
- The model was calibrated to meet aggregate totals, which may have implications for distributional estimates

# Allocation of IRS Model Individual Taxpayer Compliance Cost, 2010



**FIGURE 1**  
Composition of Discretionary Spending  
Percent of Total

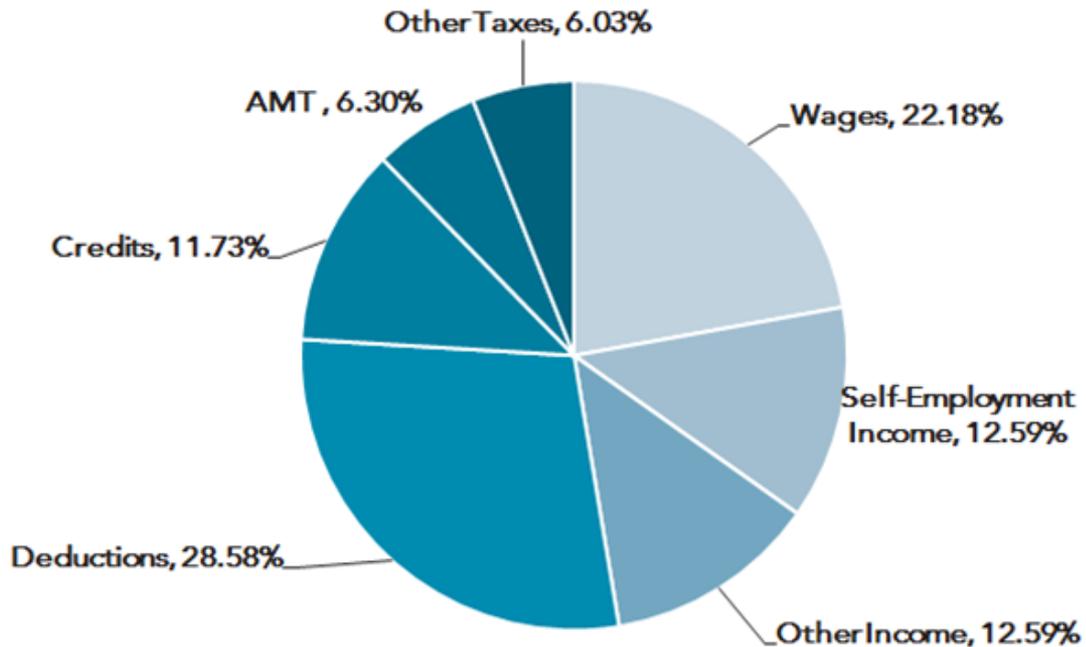


Source: Economic Report of the President, March 2013, Figure 3-10;  
<https://obamawhitehouse.archives.gov/administration/eop/cea/economic-report-of-the-President/2013>

FIGURE 2

## Composition of Discretionary Spending

Percent of Total



Source: Urban Brookings Tax Policy Center Microsimulation Model (version 0217-1)

# Baseline Compliance Burden Estimates



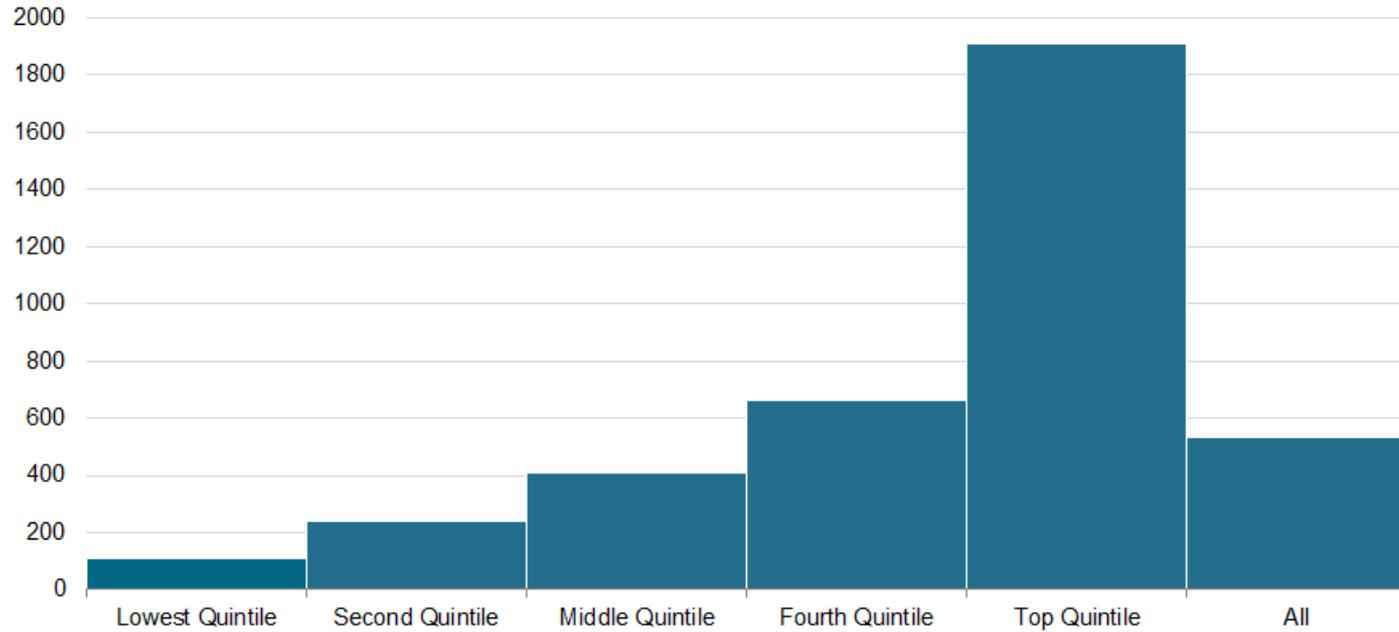
**TABLE 1**

Distribution of Compliance Cost by Expanded Cash Income Percentile, 2017

Expanded Cash Income Percentile	Percent Change in After Tax Income	Share of Total Federal Tax Change	Average Federal Tax Change (\$)	Average Federal Tax Rate <sup>e</sup>	
				Change (Percentage Points)	Under the Proposal
<b>Lowest Quintile</b>	-0.8	5.8	110	0.8	5.0
<b>Middle Quintile</b>	-0.7	14.8	410	0.6	14.6
<b>Top Quintile</b>	-0.8	49.4	1,910	0.6	26.3
<b>Top 1 Percent</b>	-0.6	10.7	8,780	0.4	33.3
<b>All</b>	-0.8	100.0	530	0.6	20.6

Source: Urban Brookings Tax Policy Center Microsimulation Model (version 0217-1)

**FIGURE 3**  
Average Dollar Value of Compliance Burden by ECI Quintile, 2017

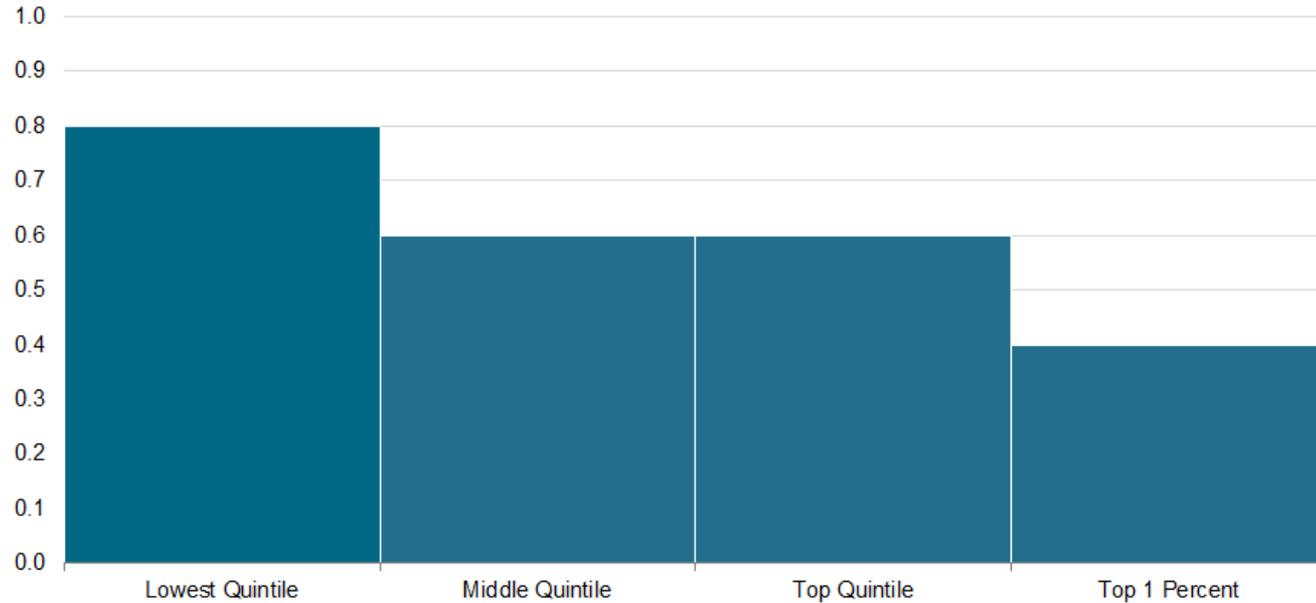


Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0217-1).

# Compliance Cost Share of Pretax Income,



**FIGURE 4**  
Average Share of Pretax Income by ECI Quintile, 2017



Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0217-1).

Marcus et al. 2013

- Ways to limit compliance costs
  - Minimize / Eliminate reporting where information of little use to tax policy or administration
  - Consider whether the policy outweighs the cost of compliance for taxpayers
  - Target Drivers of taxpayer compliance
- TPC's reform options focus on the third mechanism of lowering compliance costs

# Reform Option 1



- Revenue neutral repeal of itemized deductions by proportionally increasing the standard

**TABLE 2**

**Change in tax and compliance cost as a share of pretax income, 2017**

Expanded cash income percentile	Compliance Cost	Tax & Compliance Cost
Lowest Quintile	-0.2	-0.7
Middle Quintile	-0.2	-1.7
Top Quintile	-0.2	1.0
Top 1 Percent	-0.2	2.3
All	-0.2	-0.2

Source: Urban Brookings Tax Policy Center Microsimulation Model (version 0217-1)

## Reform Option 2



- Revenue neutral repeal of itemized deductions except the mortgage interest and charitable giving deductions by proportionally increasing the standard deduction

**TABLE 3**

### Change in tax and compliance cost as a share of pretax income, 2017

Expanded cash income percentile	Compliance Cost	Tax & Compliance Cost
Lowest Quintile	-0.2	-0.7
Middle Quintile	-0.2	-1.3
Top Quintile	-0.2	0.7
Top 1 Percent	-0.2	1.5
All	-0.2	-0.2

Source: Urban Brookings Tax Policy Center Microsimulation Model (version 0217-1)

# Reform Option 3



- Revenue neutral repeal of the Alternative Minimum Tax by pairing down the state and local tax deduction

**TABLE 4**

**Change in tax and compliance cost as a share of pretax income, 2017**

Expanded cash income percentile	Compliance Cost	Tax & Compliance Cost
Lowest Quintile	0.0	0.0
Second Quintile	0.0	0.0
Middle Quintile	0.0	0.1
Fourth Quintile	0.0	0.1
Top Quintile	-0.1	-0.2
All	0.0	-0.1
<b>Addendum</b>		
80-90	0.0	0.2
90-95	0.0	0.1
95-99	-0.2	-0.7
Top 1 Percent	-0.1	-0.2
Top 0.1 Percent	0.0	-0.2

- TPC estimates that individual taxpayer compliance costs for 2017 were \$92 billion or an average of \$530 per tax filer
- While compliance costs increase with Expanded Cash Income (ECI), the lowest ECI quintile's costs are the highest as a share of pre-tax income
- Simplifying the tax can lead to lower burden costs, and mitigate costs for taxpayers that might otherwise see tax increases

- IRS will continue to work with TPC to calibrate and test the PUF model to better align with the IRS full model results
- IRS will provide public documentation of the burden model to accompany the PUF

# THANK YOU

**For more information please contact:**

Daniel Berger  
dberger@urban.org

View other studies at  
[www.taxpolicycenter.org](http://www.taxpolicycenter.org)



# SAFELY EXPANDING RESEARCH ACCESS TO ADMINISTRATIVE TAX DATA

## Creating a Synthetic Public Use File and Validation Server

June 20, 2018

Leonard E. Burman, Alex Engler, Surachai Khitatrakun, James R. Nunns, Sarah Armstrong, John Iselin, Graham MacDonald, and Philip Stallworth



**TAX POLICY CENTER**  
URBAN INSTITUTE & BROOKINGS INSTITUTION

- Produce synthetic data file with the same record layout as IRS Administrative Data that:
  - Protects the confidentiality of tax records
  - May be used for statistically valid analysis for certain research purposes
  - May be used as a “training data set” to develop programs to run on confidential data
- Develop a safe procedure for selected researchers to remotely submit programs to perform statistical analysis on administrative data when the synthetic files are inadequate

- Synthetic dataset may include information suppressed on current PUF, such as state of residence, ages, very high-income returns
- Multiple synthetic datasets could be drawn based on random subsamples of the master file
  - For example, one might contain detailed geographic information, but less detail about sources of income and deductions
  - Could create synthetic panel datasets (like SYNLBD)
- Expanding research access to tax data would advance the state of public economics research
  - Learn about the effects and effectiveness of current tax policies
    - E.g., State is a potentially enormously valuable source of (mostly) exogenous policy variation

- Tax data are useful in many fields (not just public economics)
  - Chetty and Saez on economic mobility, for example
- Bipartisan consensus on the importance of safely using administrative data to improve public policy
  - Bipartisan Evidence-Based Policymaking Commission

- Reduce the cost of producing safe public use data files (and reduce lag time to release)
- Apply machine-learning methods to cleaning tax return data, which could lower the cost of producing INSOLE
- Facilitate collaboration between SOI analysts and outside researchers on issues related to tax administration

- In a nutshell, the goal is to simulate the process that created the administrative data and use that process to replace actual data with synthetic data that has similar statistical properties.
- The data may be synthesized based on parametric methods (e.g., regression) or non-parametric methods (e.g., CART) or a combination
- Two types of synthetic datasets: partial and full

- Partially synthetic datasets use actual data for variables that are not sensitive to disclosure and synthesize the rest
  - Most applications have produced partially synthetic datasets
- Actual (nonsynthesized) data anchor the simulations, improving overall quality, but increases disclosure risk

- In a fully synthetic dataset, all of the data are synthesized
  - If there are  $k$  variables,  $Y_1, \dots, Y_k$ , create synthetic  $\widehat{Y}_1$  drawn from empirical distribution of  $Y_1$ ;  $\widehat{Y}_2$  conditional on  $\widehat{Y}_1$  and empirical distribution of  $\varepsilon_2$ ; and so on until  $\widehat{Y}_k$  is synthesized based on  $\widehat{Y}_1, \dots, \widehat{Y}_{k-1}$
- Minimizes disclosure risk since all data are synthesized
- Concerns about data quality; very little experience with fully synthetic datasets

- Underlying administrative dataset would be a cleaned version of the RTF plus information derived from other sources (wage splits, gender, age, nonfilers, etc.)
- This is unique in that we are starting with the whole population
  - There are few missing variables or errors in variables (especially for electronically filed returns)
  - Would need to clean the data (develop machine learning process to automate SOI data editing)
  - Potentially could create a very good synthetic file, as good as random sample for some purposes
- However, the file is huge (almost 200 million records)

# Challenge: modeling relationships between variables



- A synthetic PUF needs to preserve the relationships among variables, forms, schedules reflected on tax returns
  - Also, the tax law limits the range of certain variables.
- Need to carefully chart which variables are independent and which are dependent (and how)
- May want to target some calculated variables in the synthesis process (e.g., AGI, taxable income, income tax before credits)



- Linear (or log-linear) regression for continuous variables; probit, logit, or ordered probit or logit for categorical variables; or other maximum-likelihood estimators
  - May capture some nonlinear relationships by including polynomial expansion (e.g.,  $X$ ,  $X^2$ ,  $X^3$ )
  - Under certain circumstances, a sequence of conditional distributions (e.g., regressions) may approximate a complex multivariate distribution
  - Advantage: computationally manageable
  - Disadvantage: may be very sensitive to model structure and assumed error distribution
  - If errors correlated, calculated variables (e.g., AGI) may be inaccurate

- Nonparametric methods make no prior assumptions about the underlying distribution or the process that generated the data.
- E.g., CART (Classification and Regression Trees)
  - Divide data into relatively homogeneous partitions conditional on  $X$ . (These partitions are metaphorical branches of the data “tree.”)
  - Randomly draw one of the  $Y_1$  observations from all the observations (leaves) on the branch. That is the synthesized value.
  - Repeat for  $Y_2$  conditioning on the synthesized  $Y_1$  and  $X$ ; etc.
  - To protect against disclosure, draw values from a smoothed empirical density.
  - Prune the tree so that the branches contain enough leaves to preserve confidentiality
- Drawback: computationally intensive

# We propose to use a hybrid process



- Parametric estimation for some variables and nonparametric for others
- For example, CART for discrete variables such as number of children
- Regression for continuous variables (Tobit type estimators for censored variables, such as interest income)

- Because data are fully synthetic, there is no risk of disclosing actual values from tax returns
- However, process may disclose information about the distribution of variables
- Drawing from smoothed version of distribution, sampling, and regression-based method provide substantial protection from disclosure for most of sample
- Special treatment of the tails of distribution (outliers)
- Top coding discrete variables (e.g., number of children)
- Some variables might have to be suppressed

- Researchers would develop their programs using the synthetic dataset and submit the programs electronically to the IRS
  - Output subjected to disclosure review before release to researcher
- Procedure would be similar to access to the confidential version of the SIPP
- Disclosure risk could be reduced by basing estimates on random subsamples of the restricted dataset (i.e., by drawing random samples with replacement from the full dataset)
  - Programs could include generated random seed so that a particular analysis could be replicated, but any new analysis would start with a different seed
- Costs defrayed by fees



# “FUTURE DIRECTIONS IN TAX ADMINISTRATION”

COMMENTS BY LAURA KAWANO  
OFFICE OF TAX POLICY RESEARCH, UNIVERSITY OF MICHIGAN

IRS-TPC RESEARCH CONFERENCE, JUNE 2018

# FUTURE DIRECTIONS IN TAX ADMINISTRATION

- These are three very nice papers that address important issues in tax administration.
- The first two papers ask whether information campaigns could better tax administration:
  - Complying with a complicated tax policy rule
  - Utilizing free resources to get assistance in filing taxes
- The last paper proposes a method for providing administrative tax data that could be used by the broader research community outside of government.



# “WHO IS MINDING THE NANNY TAX?”

COMMENTS BY LAURA KAWANO  
OFFICE OF TAX POLICY RESEARCH, UNIVERSITY OF MICHIGAN

# OVERVIEW

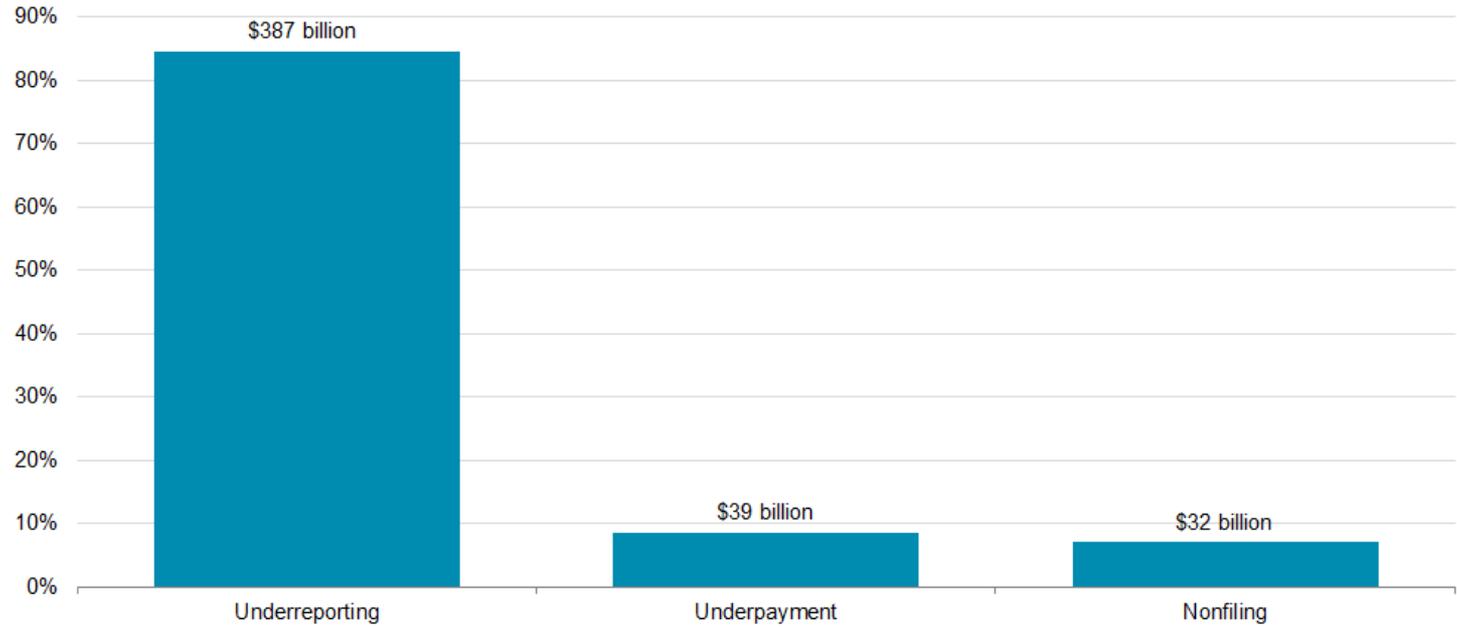
- Survey data implies a low compliance rate for the nanny tax, resulting in a large tax gap.
- ... much of this is driven by a knowledge gap.



FIGURE 1

## Components of the \$458 Billion Gross Tax Gap 2008-10

Share of gap



Source: Internal Revenue Service.

# RESPONSES TO SIMPLIFICATION

- Fascinating detail that measures to simplify the reporting by household employers led to *fewer* taxpayers reporting such relationships.
  - What explains this behavioral response?
  - If the compliance problem is really about informational barriers, this is not the response that we would expect since these are taxpayers that were already aware of their tax obligations.
  - Can you extend the figure on Schedule H Filings back using Form 940 filings?
- Were there any notable filings responses to “Nannygate” cases?

# NRP

- The differences between the estimated compliance numbers using survey data and using the IRS National Research Program are quite large.

## Why?

- What questions are being asked by the auditors that fails to detect noncompliance?
- If the IRS were to dedicate more resources to Nanny Tax enforcement, would it be effective?
- What assumptions would you need to make about the behavior captured in survey data to obtain a 98%?
- One potential consequence of the paper's recommendations is shifting towards services like Molly Maid, e.g. What are the implications of this?



# “CAN IRS MOVE PAPER FILERS TO ASSISTED TAX PREPARATION?”

COMMENTS BY LAURA KAWANO  
OFFICE OF TAX POLICY RESEARCH, UNIVERSITY OF MICHIGAN

# SUGGESTIONS FOR NEXT STEPS

- These results are preliminary, so I will offer suggestions on what be interesting next steps.
- Are there differential changes in reported incomes, tax credit take-up, deductions, refund amounts, etc. by treatment status?
  - In addition to the cost savings to the IRS, are there benefits from increasing compliance or knowledge about tax policies that take place by switching to using assisted preparers?
- Some individuals go to a VITA site and learn they actually more than they predicted given their past returns – likely because they were making some kind of mistake. Some of these individuals walk away and don't file with the VITA site. Can you test for this behavior using those who are in the treatment group but are observed as not taking up the treatment?

# SUGGESTIONS FOR NEXT STEPS

- Does utilization of VITA cause some taxpayers to amend past tax returns?
  - Some people may learn that they qualify for the EITC, and were eligible in previous years.
- When looking at persistence, does this vary by the type of changes in tax liabilities/refunds?
- Overall, I look forward to seeing future versions of this paper.



# “SAFELY EXPANDING RESEARCH ACCESS TO ADMINISTRATIVE TAX DATA: CREATING A SYNTHETIC PUBLIC USE FILE AND A VALIDATION SERVER”

COMMENTS BY LAURA KAWANO

OFFICE OF TAX POLICY RESEARCH, UNIVERSITY OF MICHIGAN

# OVERVIEW

- I really like the goals of this paper.
  - Access to administrative tax records is highly restricted, but additional research using such information is immensely valuable.
- The paper provides a nice, detailed overview of what are contained in various IRS data products.
  - In particular, the paper highlights disclosure risks, and the ways that the publicly available data respond to these risks.
- Offers several methods for solving the data access problem.

# THOUGHTS ON RUNNING PROGRAMS ON A SUBSET OF POPULATION FILES

- Consider using the IRS Databank as the basis for running programs
- Advantages:
  - Easy to link taxpayers over time because it's balanced (e.g., adjusting for cases where primary and secondary filers switch over tax years)
  - Contains non-filer population
  - Some information from information returns are already aggregated and linked
  - Easier to link families
  - Might be easier for the user to draw targeted subsamples
- It would require some translation of IMF variables into names that are consistent with the Databank.



## Session 4. Future Directions in Tax Administration

**Moderator:**

**Stephanie Parrish**  
*IRS, Online Services*

**Who is Minding the Nanny Tax?**

**Brian Erard**  
*B. Erard & Associates*

**Can IRS Move Paper Filers to Free Assisted Tax Preparation?**

**Rizwan Javaid**  
*IRS, RAAS*

**Creating a Fully Synthetic Individual Income Tax Return Data File and Validation Server**

**Len Burman**  
*Urban-Brookings Tax Policy Center*

**Discussant:**

**Laura Kawano**  
*University of Michigan*



Research, Applied Analytics,  
and Statistics



TAX POLICY CENTER  
URBAN INSTITUTE & BROOKINGS INSTITUTION

---

## *8<sup>th</sup> Annual IRS/TPC Joint Research Conference on Tax Administration*

# *Wrap-Up*

***Barry Johnson***

*Director, IRS Statistics of Income Division,  
Research, Applied Analytics, and Statistics*