Analysis of Nonresponse in the Statistics of Income's 1999 Individual Tax Return Panel

Tara R. Wells
Statistics of Income, IRS 77 K Street NE Washington DC, 20002

Abstract In 1999, the Statistics of Income (SOI) Division of the Internal Revenue Service began collecting individual tax returns for SOI's 1999 Individual Tax Return Panel. Longitudinal tax data is essential to study how the tax system affects taxpayers over an extended period of time. However, as in all panels, SOI's 1999 Individual Tax Return Panel is impaired by panel attrition. Previous papers have evaluated the presence and motivation for attrition in prior SOI individual tax return panels, yet there has not been research that has measured the nonresponse error caused by the attrition. In this research, I use an exploratory approach to estimate the nonresponse error in specific tax-related variables collected from SOI's 1999 Individual Tax Return Panel from 2000 to 2003. Then I use a propensity score method of subclassification to investigate if the nonresponse bias can be removed. My results show that the nonresponse bias in the tax-related variables can be reduced with the use of propensity score adjustments.

Key Words: Nonresponse bias, propensity score adjustment, panel attrition

1. Introduction

Tax policy research heavily relies on panel data to investigate how taxpayers react to amendments in tax laws and how the tax system affects taxpayer reporting of income and earnings over time (Bryant 2008; Feldstein 1995). One of the key functions of the Statistics of Income Division (SOI) of the Internal Revenue Service (IRS) is to provide longitudinal data from individual tax returns to the Department of the Treasury. These data allow tax analysts to assess how the income tax system is performing and to project how it might perform under different proposals for changes in tax law (Sailer, Weber, and Wong 1999). However, while SOI's 1999 Individual panel data provides a wealth of information on taxpayers' behaviors, the data are not error-free. Panel attrition affects estimates based on these data and may undermine their validity, leading to false inferences that influence tax policy (Bryant 2008). This study explores the effects of panel attrition in SOI's Edited 1999 Individual Panel.

Panel surveys play an essential role in understanding causal processes in the social world (Lazarsfeld 1948; Kalton and Citro 2000), such as comprehending how particular tax laws affect individual taxpayers. Panel surveys collect measurements on the same sample members at different points in time and are particularly useful for estimating gross and other components of change for individuals over time (Kalton, Kasprzyk, and McMillen 1989). Specifically, the major benefit of a panel survey is that it generates the data needed for longitudinal analysis, thus offering much greater analytic potential than a cross-sectional survey (Kalton 2009). In his analysis of the effect of the 1986 Tax Reform Act, Feldstein (1995) states that one important benefit of a tax return panel is that an individual is observed both before and after the change in tax rates. Unfortunately, it is

inevitable in panel surveys that certain panel members will drop out as time progresses. Because of this, panel surveys are susceptible to a special type of nonresponse that compounds over time: panel attrition. If these 'lost' panel members are similar to the remaining members, then the nonresponse is ignorable (i.e., the data are missing at random). The mathematics of nonresponse confirm that as nonresponse increases, so does the potential for nonresponse bias (see, for example: Bethlehem 2002). Additionally, if the topic of interest within the survey is correlated with the reasons for attrition, this too can bias estimates (Lessler and Kalsbeek 1992).

In panel surveys, nonresponse may occur not only at the initial wave, but also at each subsequent wave of the panel. Kalton and Brick (2000) classify four response patterns in panel surveys: (1) total respondents, those who provide data on every wave; (2) temporary drop-outs, who return to the panel after missing one or more waves; (3) attrition nonrespondents, who drop out of the panel at some point after the first wave and never return; and (4) total nonrespondents, who provide data for none of the waves. In the SOI 1999 Individual Panel, members behave in three of the possible patterns: some file continuously from 1999 through 2003 (total respondents), some file intermittently (temporary dropouts), and some drop out (attrition nonrespondents). Bryant (2008) used Figures 1A through 1C to provide a graphic depiction of the three taxpayer filing behaviors present in the 1999 Individual Panel. A 'balanced' panel (as termed by Arellano 2003), where each base year taxpayer is present in all years as shown in Figure A, is the most straightforward type of panel, requiring no data manipulation when analyzing taxpayer behavior (Christian and Frischmann 1989). Figure B illustrates an unbalanced panel where taxpayers are present in the base year, are missing for at least one year after, and then later return for at least one additional year. Lastly, attrition within an unbalanced panel is displayed in Figure C, where the filers are present in the base year and every subsequent year until dropping out before 2003 and not returning.

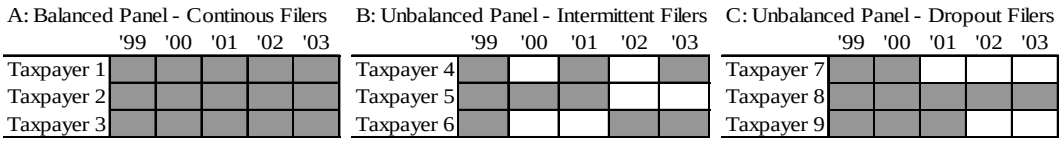


Figure 1: Patterns of Nonresponse in the SOI 1999 Individual Panel

Nonresponse bias in survey estimates reflects two components—the amount of nonresponse and the difference in the estimate of interest between the respondents and nonrespondents. For survey means, nonresponse bias is given as:

$$\text{Nonresponse Bias } (\bar{y}_r) = (1 - w_r)(\bar{y}_r - \bar{y}_{nr}), \qquad (1)$$

where w_r is the proportion of respondents, $\bar{y}_r$ is the estimated mean for the respondents only, and $\bar{y}_{nr}$ is the estimated mean of nonrespondents. Any estimate from a panel study can be subject to bias if nonresponders differ significantly from responders with respect to characteristics of interest. The four major sources of attrition in the SOI Individual Panel are: taxpayer death, taxpayer income below the required filing threshold, a taxpayer SSN that cannot be matched properly, and taxpayer noncompliance (Bryant 2008, see also Clotfelter 1980). These sources of panel member loss trigger the elimination of specific types of panel members, thus increasing the potential for biased estimates. Christian and Frishmann (1989) studied the effects of attrition in the SOI 1979 Individual Panel and found that certain estimates (age, marital status, and adjusted gross income

(AGI)) were biased by nonresponse. Similarly, the research presented in this paper examines whether specific 1999 Individual panel estimates are biased due to the nonresponse caused by panel attrition.

Commonly, when estimates are identified as biased, a remedy to eliminate or adjust for the bias is necessary. The best approach is typically to avoid attrition from the onset by improving response rates through the use of pre-notification letters, incentives, advanced locating techniques, etc. (Groves 2004; Lessler and Kalsbeek 1992). Unfortunately, while administrative record panels such as SOI's 1999 Panel experience nonresponse, they cannot take advantage of classic methods aimed at improving respondents' response propensities because participation is entirely passive. Individuals, who submit tax returns in the panel's base year, are unknowingly selected for SOI's panel through a stratified probability selection mechanism. Thus, it is especially important to measure the amount of bias in administrative record panels because there are limited methods for solving the problem of nonresponse using conventional procedures.

There are a variety of methods to adjust for nonresponse bias in panel data and the best choice is not always obvious (Kalton 1986). This research will focus on the use of subclassification by propensity scores. The propensity score can be estimated using a logistic regression model because it makes no assumptions about the distributions of the covariates on the dichotomous outcome (D'Agostino 1998). A single propensity score is estimated for every individual in the study and used to adjust for the differences between the two groups on the observed covariates in the study. In this case, the two groups are the complete respondents and the dropouts. The propensity score is often thought of as a 'balancing score' allowing researchers to control for a large number of covariates simultaneously (Rosenbaum and Rubin 1983). The three most common propensity score methods are: (1) matching, (2) subclassification, and (3) regression adjustment. Subclassification has the advantage of being easy to implement, requires only the assignment of units into a small number of subclasses, and provides a framework that extends naturally to incorporate survey weights from complex survey designs (Zanutto, Lu, and Hornick 2005).

Originally, the propensity score subclassification technique was developed to allow for the calculation of unbiased estimates of treatment effects derived from observation studies with nonrandom assignment. However, Rubin (1985) proposed the use of propensity scores to adjust for unit nonresponse. Use of the propensity score method is limited to situations where information is available for all nonrespondents, for example, in administrative panel surveys (Little and David 1983). Czaijka and Radbill (1992) applied propensity score weighting on early versus late tax submission data to estimate values for final returns in the SOI 1979 Individual Panel. They achieved improvements in accuracy of estimates on a range of variables over the existing post-stratification method that was used at the time to adjust for unit nonresponse. Because information is available from all nonrespondents in the base year of SOI's 1999 Panel, propensity score subclassification is an ideal method to investigate whether the nonreponse bias can be removed.

1.1 Research Hypothesis

This paper estimates the nonresponse bias caused by panel attrition in SOI's 1999 Individual Panel. In particular, this research will attempt to answer the following: (1) *How do the intermittent filers and dropouts differ in their demographic characteristics?* (2) *How do the two groups of nonrespondents differ from the continuous filers?* (3) *What*

is the estimated nonresponse error of the 1999 Individual SOI Panel data? (4) *Does this error increase as the panel ages?* (5) *Can this bias be removed by the use of propensity scores?* Findings will provide evidence on the validity of the 1999 Individual Panel data, knowledge about the existence of nonresponse bias in the panel data, and functional guidance on modifications that can be made for future SOI Individual Panels.

2. Method

2.1 Background of the Data

Since 1916, SOI's primary product has been an annual cross-section of individual income tax returns, also known as the Individual Complete Report File (ICRF). The ICRF has been the basis for most of Federal tax policy analyses. While these cross-sectional data were enormously instructive, they did not offer the benefits of longitudinal data. Longitudinal panels have a much shorter history of use in SOI due to their statistical and operational complexity and limitations in SOI's budget (Weber and Bryant 2005). In 1979, SOI formed its first panel study, but it lacked the ability to provide reliable estimates for high income taxpayers. In 1987, the SOI Individual Panel addressed this by employing a stratified sample design to adequately represent high income taxpayers. Although the 1979 and 1987 panels provided useful information on taxpayers from a longitudinal perspective, the panel continued to experience problems such as under-representing high income taxpayers. The new 1999 panel attempted to address these limitations. SOI employed a new set of methods to increase the statistical reliability of the panel and to collect a finer level of financial information on taxpayers' sales of capital assets than previous panels and cross-sections had done.

The target population for the SOI 1999 Individual panel study is unaudited Individual Income Tax Returns, which includes Forms 1040, 1040A, and 1040EZ filed by U.S. citizens and residents during Calendar Year 2000. The target population includes approximately 177 million taxpayers. The study follows all individuals who file tax returns except for dependents listed on the returns. The sample design is a stratified probability sample, where the strata are defined by five variables. The five variables are based on AGI, source of income, inclusion of special tax forms and schedules, positive or negative income, and potential usefulness as defined by Office of Taxation and Analysis (see Walker and Puckett 2000). A total of 26 strata were formed. Individual tax returns processed during calendar year 2000 were used to assign each taxpayer's record to the appropriate stratum and to determine whether the record should be included in the sample. Tax returns are selected for the sample based on the last four or five digits of the social security number (SSN) (see Walker and Puckett 2000 for details). After the data were captured and cleaned, weights were obtained by dividing the population count of returns in a stratum by the number of sample returns for that stratum. The weights were adjusted to correct for misclassified returns, but were not adjusted for nonresponse. The 1999 SOI Individual Panel began with 124,657 taxpayers; by 2005, the panel had reduced to 108,405 taxpayers.

2.2 Subset of the 1999 SOI Individual Panel

A subset of the 124,657 taxpayers selected for the 1999 SOI Individual panel is used in these analyses. Because the deceased cannot ever be included in the panel again, I eliminated 2,923 taxpayers who died between 1999 and 2003. Additionally, I removed all cases where gender is missing or unknown or if age is equal to 0 or greater than 120 (n=228). The final sample of the panel that is considered in this analysis had 121,506

cases for the 1999 year. Table 1 shows the size of the panel subset used in this analysis through from 1999 through 2003. A useful feature of the dataset is that information is collected from all participants in the base-year (1999); the base-year data can be used to identify possible causes and correlates of nonresponse (Lepkowski and Couper 2001). Note that taxpayers do not always file on time; as a result, the panel sample for a respective year is kept open for an additional two years in order to include these late returns. For instance, sampling for the Tax Year 1999 panel data would not be complete until as late as December 31, 2002. The SOI 1999 Individual data is complete through 2003, but late returns are still in the process of being edited and checked for years 2004 and 2005 and these years were not considered in this analysis.

Table 1. SOI's 1999 Individual Panel (1999 - 2003)

Year	Sample	Sample Missing	Percent of Base
1999	121,506	---	100.00%
2000	117,983	3,523	97.10%
2001	116,586	4,920	95.95%
2002	115,087	6,419	94.72%
2003	113,933	7,573	93.77%

2.2 Variables of Interest

Data collected from individuals' IRS 1040 Tax Return Form provide all of the variables in these analyses. I use two correlated dependent variables in separate sections of these analyses. The first dependent variable is taxpayer response behavior: continuous filing, intermittent filing, and dropping out. This variable permits the examination of the demographic characteristics of the different response groups. The second dependent variable is a dichotomous variable indicating whether a taxpayer filed a return continuously through the lifetime of the complete panel (1999 through 2003) or dropped out of the panel before or in 2003. Christian and Frishmann (1989) established that taxpayer age, marital status, gender, and AGI were related to attrition in the 1979 SOI Individual Panel data. In addition to these variables, I incorporate specific tax-related variables that are used for tax policy research. These variables are described in Table 2.

2.3 Techniques for Analyses

I followed Christian and Frishmann's (1989) exploratory approach to quantify the nonresponse bias caused by attrition in the SOI 1999 Panel. Those who dropped out at any time during 2000 are removed from the 1999 base year of the panel and the key variables are re-estimated. I calculated the absolute relative nonresponse bias:

$$\text{Absolute Relative Bias } (\bar{y}_r) = \left| \frac{Bias\ (\bar{y}_r)}{\bar{y}_r} \right|, \qquad (2)$$

between the base year estimates for 1999 and the adjusted 1999 estimates. The adjusted estimates were based on the 2000 respondents. This approach was completed for years 2000 through 2003. If attrition in the panel is a random event, any difference between the descriptive statistics from the full 1999 panel and the adjusted 1999 panel should be within sampling error.

Once I quantified the nonresponse relative bias, I experimented with a technique to compensate for this type of nonresponse error using response propensity scores. This is accomplished by the propensity score method of subclassification, where taxpayers who

file annually in the panel are compared to those who drop out of the panel. First, I estimated the propensity scores by a logistic regression. The dependent variable is whether the taxpayer filed continuously from 1999 through 2003 (continuous filer) or dropped out at some time before 2003 (dropout). Cochran (1968) showed that five subclasses are sufficient to remove over 90 percent of the bias due to the covariates, thus I used quintiles of the propensity scores to determine the cut-offs for the different strata or subclasses. Next, I verified that the quintiles function appropriately and balanced the differences between continuous filers and dropouts by using a two-way analysis of variance (ANOVA) model, which includes the main effects and interactions for the propensity score quintile and taxpayer filing behavior. After confirming covariate balance, I created a single class adjustment for each of the quintiles and used this to reweight the data. To determine whether propensity score weighting is a beneficial form of nonresponse adjustment, I estimated the differences between continuous filers and dropouts on the different covariates, before and after adjusting for their propensity quintile weight. If the propensity score method worked correctly, the difference between the estimates for the respondents and nonrespondents should be reduced.

Table 2. 1999 SOI Individual Panel Data Variables Used in the Analyses

Variable	Description
AGI ($)	Adjusted Gross Income, a taxpayer's gross income (from all sources) after subtracting deductions.
Dependent Status	A binary variable that denotes whether a taxpayer is filing as a dependent (someone who is supported materially and/or financially).
Exemptions	The number of exemptions claimed on the tax return.
Late Filing	A binary variable indicating whether a taxpayer filed by April 15th of the respective tax year.
Medical Exemptions	A binary variable indicating whether a taxpayer claimed medical exemptions.
Refund (-) / Owe (+) ($)	The amount a taxpayer has overpaid (-) / underpaid (+) on their federal tax liability.
Temporary Citizen	A binary variable indicating whether a taxpayer is a temporary American citizen (indicated as a '9' in the first digit of their SSN).
Total Payments ($)	The amount a taxpayer has paid during the tax year towards his/her federal tax liability, including refundable credits received to offset tax liability.
Wages, Salaries, and Tips ($)	The amount of income received from an employer and reported on a W-2.

Note: All dollar values presented in these analyses are normalized to 1999 dollars using the August 1999 CPI.

3. Results

In analyzing the effects of nonresponse in the SOI 1999 Individual Panel, I examined a variety of demographic and tax-related estimates and compare the three tax response behavior types. Then, I estimated the nonresponse bias in the panel for specific tax-

related variables and analyzed how the bias changes from 2000 through 2003. Finally, I used propensity scores in an attempt to remove the nonresponse bias.

3.1 Panel Sample Loss in SOI'S 1999 Individual Panel

As shown in Table 2, only 93 percent of the sample remained by the fifth wave of the panel in 2003. Most of the taxpayers (88 percent) filed continuously, six percent filed intermittently, and the remaining seven percent dropped out of the panel at some point on or before 2003. Dropping out of the panel occurred consistently over time with approximately two percent of taxpayers dropping out per year. Table 3 shows that approximately 2,500 intermittent filers dropped out of the panel each year, but returned to file taxes in another year. However, approximately 1,200 taxpayers dropped out of the sample each year following 1999 and never filed a tax return again in the life of the panel.

Table 3. Taxpayer Filing Behavior in SOI's 1999 Individual Panel

Taxpayer Filing Behavior	Number of Taxpayers
Intermittent filers	6,953
Filed only for 1999	1,275
Filed continuously 1999-2000	1,126
Filed continuously 1999-2001	1,198
Filed continuously 1999-2002	1,281
Filed continuously 1999-2003	109,673
N	121,506

Table 4. Reasons for Sample Loss in SOI's 1999 Individual Panel

Year	Intermittent Filers	Filers who Dropped Out for Unknown Reason
1999	---	---
2000	2,248	1,275
2001	2,519	1,126
2002	2,820	1,198
2003	2,693	1,281
N	---	4,880

Note: Intermittent filers can leave the panel and return to the panel multiple times. Thus, the total number cannot be represented in this table. The total number of Intermittent filers is 6,953.

3.2 Estimates from the SOI 1999 Individual Panel

As expected, the average age of the SOI 1999 Individual Panel member increases as the panel ages (see Table 5). Taxpayers have consistently claimed more exemptions over panel waves and more taxpayers declare medical exemptions. Wages, Salaries, and Tips (WST), AGI, and Total Payments peak in 2000 and slowly decrease. Taxpayers' refunds from 1999 to 2003 nearly tripled in size, on average. As the panel matured, the distributions of certain demographic variables changed in the panel. By 2003, the panel had lost approximately 4,100 males and 3,400 females. In 1999, approximately 42 percent of taxpayers reported themselves as married, but in 2003, over 62 percent of

taxpayers reported that they were married. The proportion of temporary citizens and taxpayers who file late tax returns remained constant over the life of the panel.

3.3 Differences between Taxpayers in the SOI 1999 Individual Panel

Table 6 shows estimates by respondent groups for 1999. All of the respondent groups' estimates are significantly different from one another ($p < 0.01$ or less), so Table 6 omits p-values. Continuous filers claim the most exemptions (2.4), have the highest percentage claiming medical exemptions (5.2 percent), have significantly higher WST ($42,848), AGI ($61,030), and pay more federal tax ($10,388) than the other groups. As expected, the continuous filers receive the lowest refund amount ($367), on average. The intermittent filers and dropouts' report much lower income – their average WST and AGI are approximately $40,000 less than the averages for the continuous filers – but these groups only pay an average of approximately $6,000 less in federal tax and receive only an average of approximately $300 more than the continuous filers in refunds. The average age of a continuous filer is 43 years old, while intermittent filers are the youngest at 33 years old, and dropouts are in between at 38 years old. On average, the older continuous filers have the lowest dependent status at 4.3 percent and the younger intermittent filers have the highest dependent status as 21.8 percent. There are a higher percentage of women among the continuous filers (52.0 percent) and a higher percentage who report they are married (61.4 percent). In the intermittent filer and dropout subgroups, there are a higher percentage of men, 52.0 and 53.4 percent, respectively, and at least 20 percent fewer married taxpayers. Not surprisingly, the continuous filing taxpayers have the lowest percentages of temporary citizens (1.4 percent) and late filing (1.6 percent). However, late filing is not an overall concern in the panel; more than 93 percent of all taxpayer behavior response groups file taxes on time, within the respective tax year.

3.4 Nonresponse Error in the SOI 1999 Individual Panel

The estimated nonreponse bias increases as the respective nonrespondents for each year (2000 through 2003) are removed from the 1999 full panel estimates. Table 7 displays the relative nonresponse bias (see equation [2]) in the tax-related variables for the SOI 1999 Individual Panel. The estimated relative nonresponse bias shows the impact of excluding the nonrespondents from the estimates based on the 1999 data for each of the panel years. The average absolute relative nonresponse bias in the five variables incrementally increases each year, starting at 2.2 percent in 2000 and reaching 5.0 percent in 2003. There are differences across the different estimates in the amount of absolute relative bias increase. From 2000 to 2003, the absolute relative nonresponse bias of refund amount increases by 261 percent. Also, the absolute relative nonresponse bias of WST, AGI, and total payments increases by at least 110 percent from 2000 to 2003. Lastly, the absolute relative nonresponse bias of the average number of taxpayer exemptions increases by 57 percent. Figure 2 shows the estimated absolute relative nonresponse bias trends for the tax-related variables.

3.5 Using Propensity Scores to Remove Nonresponse Bias

I used a logistic regression model to estimate a taxpayer's propensity to file continuously in the SOI 1999 Individual Panel using the dichotomous dependent variable. The logistic regression included all of the independent variables discussed in these analyses, since they all are thought to have a relationship with response propensities. A key to causal inference in observation studies using propensity score methods is to examine the overlap in propensity scores between the two groups of interest (Oakes 2007). In Figure 3, I

examined the overlap in their propensity scores to determine whether the continuous filers can be adequately compared to the dropouts. Although the continuous filers have a higher average propensity to file continuously (91.8 percent) than the dropouts (79.3 percent), overlap exists for all propensity scores which provides the region of 'common support' in the data; this means that bias could possibly be reduced via a propensity score adjustment (Rubin 1977).

Cases are sorted into quintiles based on their propensity scores. The quintiles are of equal sizes – there are 24,301 taxpayers in quintiles one, two, four, and five and 24,302 taxpayers in quintile three. The majority of continuous filers are in quintile five, while most dropouts are in quintile one. Using two-way ANOVA models, I investigated the balance in the covariates of interest: exemptions, WST, AGI, total payment, and refund. These analyses establish that the covariate means do not differ significantly for continuous filers and dropouts within quintile classes, which confirms the covariate balance. The first two rows of Table 8 illustrate the extreme differences between the estimated means of the continuous and dropouts in the covariates of interest in the overall sample. However, after carrying out the propensity score subclassification, the estimated means are much closer for the two response groups within the quintiles. The estimated means of exemptions for continuous filers and dropouts are very close in the five quintiles. WST, AGI, total payments, and refunds are similar in the first four quintiles, whereas in the fifth quintile the groups are somewhat separated. Nevertheless, as can be seen in the table, the two groups of filers are more similar within each propensity score quintile than they were before stratification.

In the last part of my analyses, I estimated a single class adjustment within each quintile, using a weighted average estimated propensity:

$$\widehat{\emptyset}_c = \sum_{i \in s_c} w_i \widehat{\emptyset}(x_i) / \sum_{i \in s_c} w_i, \qquad (3)$$

where $\widehat{\emptyset}_c$ is the average weighted propensity in quintile c, w_i is the design weight, and $\widehat{\emptyset}(x_i)$ is the estimated propensity score for unit x_i. New weights were calculated for the SOI 1999 Individual Panel data by multiplying the original design weight by the inverse of the response propensity adjustment for each individual within the panel. Table 9 displays the differences between respondent (continuous filers) and nonrespondent (dropouts) estimated means in the panel before and after the propensity weighting adjustment. The difference between the means is reduced after the propensity weighting. Reduction of the difference between continuous filer and dropout averages is most apparent in the estimate of refunds, which has over a 20 percent reduction.

The other covariates reveal a lessening of the difference between the estimated means of the two response groups, but the difference is not nearly as drastic as with the estimates for refunds. The reduction of absolute relative bias is trivial in the estimates for 2000 and 2001; nevertheless, the advantage of the propensity weight adjustment is obvious in the estimates for 2002 and 2003. In 2002, approximately 24 percent of the average absolute relative nonresponse bias is eliminated on average – and in 2003, that amount increases to approximately 41 percent.

Table 5. SOI 1999 Individual Panel Estimates, By Year

	1999		2000		2001		2002		2003	
	mean	s.e.	mean	s.e.	mean	s.e.	mean	s.e.	mean	s.e.
Age	42.381	0.050	43.578	0.050	44.616	0.050	45.562	0.050	46.440	0.050
Exemptions	2.283	0.004	2.331	0.004	2.365	0.004	2.395	0.004	2.413	0.005
Medical Exemptions (%)	4.88	0.07	5.67	0.08	6.63	0.09	7.85	0.09	8.18	0.10
Wages, Salaries, and Tips ($)	39,391	89	42,022	127	42,257	181	41,494	133	41,513	160
AGI ($)	55,790	82	60,127	166	57,393	206	55,595	164	56,249	197
Total Payments ($)	9,429	22	10,368	39	10,082	55	9,286	42	8,926	49
Refund (-) / Owe (+) ($)	-404	15	-462	24	-1,125	24	-1,266	21	-1,481	20
Gender (%)										
Female	51.39	0.17	51.58	0.17	51.68	0.17	51.75	0.17	51.76	0.18
Male	48.61	0.17	48.42	0.17	48.32	0.17	48.25	0.17	48.24	0.18
Dependent Status (%)										
No	93.35	0.08	95.13	0.07	96.38	0.06	97.27	0.06	97.94	0.05
Yes	6.65	0.08	4.87	0.07	3.62	0.06	2.73	0.06	2.06	0.05
Marital Status (%)										
Other	42.18	0.14	40.40	0.14	39.13	0.15	38.12	0.15	37.40	0.15
Married	57.82	0.14	59.60	0.14	60.87	0.15	61.88	0.15	62.60	0.15
Temporary Citizen (%)										
No	98.41	0.04	98.45	0.04	98.48	0.04	98.53	0.04	98.53	0.04
Yes	1.59	0.04	1.55	0.04	1.52	0.04	1.47	0.04	1.47	0.04
Late filing (%)										
No	97.87	0.05	98.12	0.05	98.17	0.05	98.13	0.05	98.09	0.05
Yes	2.13	0.05	1.88	0.05	1.83	0.05	1.87	0.05	1.91	0.05
N	121,506		117,983		116,586		115,087		113,933	

Note: All estimates are calculated using the trimmed panel weight for the respective year.

Table 6. Selected Demographic Characteristics, By Response Groups -- SOI 1999 Individual Returns

	Continous Filers		Intermittent Filers		Dropout Filers		Intermittent + Dropout	
	mean	s.e.	mean	s.e.	mean	s.e.	mean	s.e.
Age	43.0	0.053	33.088	0.200	45.887	0.288	38.368	0.176
Exemptions	2.4	0.005	1.626	0.017	1.907	0.019	1.742	0.013
Medical Exemptions (%)	5.2	0.08	2.32	0.19	3.76	0.28	2.92	0.16
Wages, Salaries, and Tips ($)	42848	106	15,403	264	16,004	369	15,651	216
AGI ($)	61030	109	18,977	302	20,994	413	19,809	243
Total Payments ($)	10388	28	2,677	56	3,083	91	2,845	50
Refund (-) / Owe (+) ($)	-367.4	17	-661	32	-643	38	-654	24
Gender (%)								
Female	52.0	0.18	48.01	0.62	46.61	0.74	47.43	0.48
Male	48.0	0.18	51.99	0.62	53.39	0.74	52.57	0.48
Dependent Status (%)								
No	94.7	0.08	78.15	0.51	93.12	0.38	84.33	0.34
Yes	5.3	0.08	21.85	0.51	6.88	0.38	15.67	0.34
Marital Status (%)								
Other	38.6	0.15	72.27	0.55	58.78	0.73	66.71	0.44
Married	61.4	0.15	27.73	0.55	41.22	0.73	33.29	0.44
Temporary Citizen (%)								
No	98.6	0.04	98.15	0.17	96.30	0.28	97.39	0.15
Yes	1.4	0.04	1.85	0.17	3.70	0.28	2.61	0.15
Late filing (%)								
No	98.4	0.05	94.73	0.28	93.13	0.37	94.07	0.22
Yes	1.6	0.05	5.27	0.28	6.87	0.37	5.93	0.22
N	109,673		6,953		4,880		11,883	

Note: All estimates are calculated using the trimmed panel weight for the 1999. All Response groups estimates are significantly different ($p < .001$).

4. Discussion

In the SOI 1999 Individual Panel, the continuous filers are more likely to have higher average AGI and WST, while those who earn less are more likely to drop out of the panel. The dropouts differ from the continuous filers in many of the demographic and tax-related variables. The loss of these nonresponders from the panel will affect panel estimates in the later waves. The continuous filers and dropouts differed in age, gender, and in the tax-related variables. The nonresponders are most likely to be younger, unmarried, male, and have a lower income. This finding is intuitive because these variables have all been associated with survey nonresponse (e.g., Groves 2004).

Nonresponse bias does exist in the SOI 1999 Panel and it increases as the panel ages. The bias is most extreme for estimating refund amount, but does exist for all of the tax-related variables. In 2000, the average absolute relative bias is 2.2 percent and by 2003, it had increased to 5 percent. Fortunately, the nonresponse bias can be reduced via the propensity score method. In Table 8, the estimated means of the continuous filers and dropouts within the quintiles are much closer than they originally were in the overall sample. Even as the differences in the estimated means are reduced by subclassifying the data based on propensity scores, the differences between the continuous filers and dropouts remain significant across all variables. The differences between the estimates within the quintiles for the continuous filers and dropouts are reduced by at least 50 percent for exemptions, AGI, and WST. The differences between continuous filers and dropouts' total payments and refunds were extremely reduced in first four quintiles, 93 percent and 42 percent, respectively. However, in quintile five the differences increase. This increase is most likely due to the large dollar amounts (in the millions), thus the respective differences are larger (in the hundred thousands). Quintile five presents some concern with the variation between the two response groups, suggesting that future work may require the use of more subclasses. The results from this research confirm that nonresponse bias exists in the SOI 1999 Individual Panel and that the current design weights would benefit from a nonresponse propensity adjustment.

There are several limitations in the data in these analyses. The IRS taxpayer filing requirement thresholds do not mandate extremely low income taxpayers (under approximately 9,000 dollars in 2010) to file. For example, a taxpayer could earn 20,000 dollars and file in 1999, but in 2000 through 2003, quit his/her job and earn less than the amount that requires an individual to file taxes. Such taxpayers do not need to be in the sample, and these types of situations lead us to overstate the level of nonresponse error in the panel. Currently, SOI does not have access to any additional information on why taxpayers drop out of the panel. If taxpayers are no longer filing for a specific reason (such as legally missing or incarcerated), SOI would not require inclusion of their tax return in the panel. Unfortunately, as in other panel surveys, it is not customary for the IRS to send a follow-up questionnaire to find out the reason(s) that a taxpayer did not file. However, a follow-up is done for certain taxpayers who have a tax liability. Finally, the method that I used to remove the nonresponse bias has one noteworthy shortcoming. In this research, propensity score subclassification could only measure the propensity to file taxes continuously in the context of the observed variables, not in the unobserved ones. Thus, these analyses may have ignored specific variables that are more predictive of a taxpayer's decision to file taxes continuously versus file intermittently or simply drop out of the panel. Regardless of these limitations, this study suggests that SOI would benefit from further research that focuses on the best method for removing the nonresponse bias caused by panel attrition in the SOI 1999 Individual Panel data.

Table 7. Estimated Absolute Relative Bias in SOI's 1999 Individual Panel Estimates - Removed Dropouts for Respective Years

	1999		2000			2001			2002			2003		
Variables	mean	s.e.	mean	s.e.	|Bias| %	mean	s.e.	|Bias| %	mean	s.e.	|Bias| %	mean	s.e.	|Bias| %
Exemptions	2.283	0.004	2.309	0.004	1.14	2.319	0.004	1.57	2.324	0.004	1.80	2.324	0.004	1.79
WST ($)	39,391	89	40,412	93	2.59	40,891	95	3.81	41,274	97	4.78	41,536	99	5.45
AGI ($)	55,790	82	57,312	89	2.73	58,023	92	4.00	58,610	95	5.05	59,003	97	5.76
Total Payments ($)	9,429	22	9,699	24	2.86	9,835	24	4.30	9,940	25	5.41	10,020	26	6.26
Refund/Owe ($)	-404	15	-398	15	1.50	-395	16	2.17	-387	16	4.20	-382	16	5.43
N			117,983			116,586			115,087			113,933		

Note: All estimates are calculated using the trimmed panel weight for 1999.

Table 8. Comparison of Quintile Means for Tax Related Variables in SOI's 1999 Individual Panel

		N	Exemptions	WST ($)	AGI ($)	Total Payments ($)	Refund / Owe ($)
Overall	Continous	109,673	2.362	42,848	61,030	10,388	-367
	Dropouts	11,833	1.742	15,651	19,809	2,845	-654
After Stratification into quintiles based on propensity scores							
Quintile 1	Continous	17,571	1.281	8,271	10,217	1,547	-587
	Dropouts	6,730	1.254	6,409	7,929	1,092	-565
Quintile 2	Continous	21,273	2.140	17,271	27,856	3,330	-760
	Dropouts	3,028	2.284	15,475	20,755	2,427	-913
Quintile 3	Continous	23,250	2.791	40,091	53,766	6,905	-829
	Dropouts	1,052	2.811	40,383	46,083	5,567	-939
Quintile 4	Continous	23,770	3.124	100,107	129,819	23,800	140
	Dropouts	531	3.198	108,079	130,969	24,264	-278
Quintile 5	Continous	23,809	3.115	1,070,095	2,526,352	669,315	59,559
	Dropouts	492	3.278	1,134,034	2,483,491	622,792	79,012

Note: All estimates are calculated using the trimmed panel weight for 1999.

Table 9. Difference in Continuous and Dropout Filers Means in SOI's 1999 Indivudal Panel - Before and After Propensity Weighting

	2000		2001		2002		2003	
Variables	Before	After	Before	After	Before	After	Before	After
Exemptions	0.69	0.67	0.67	0.64	0.59	0.57	0.50	0.48
Wages, Salaries and Tips ($)	27,228	25,343	27,905	25,889	27,046	25,142	26,129	24,285
AGI ($)	40,570	37,722	41,532	38,461	40,493	37,541	39,129	36,230
Total Payments ($)	7,189	6,609	7,541	6,898	7,330	6,719	7,192	6,585
Refund (-) / Owe (+) ($)	161	126	163	133	244	206	267	224
N	117,983		116,586		115,087		113,933	

Note: All estimates are calculated using the trimmed panel weight for 1999. All differences are significant (p <.001)

Figure 2. Estimated Absolute Relative Bias in Selected Characteristics of the SOI 1999 Individual Panel, By Year

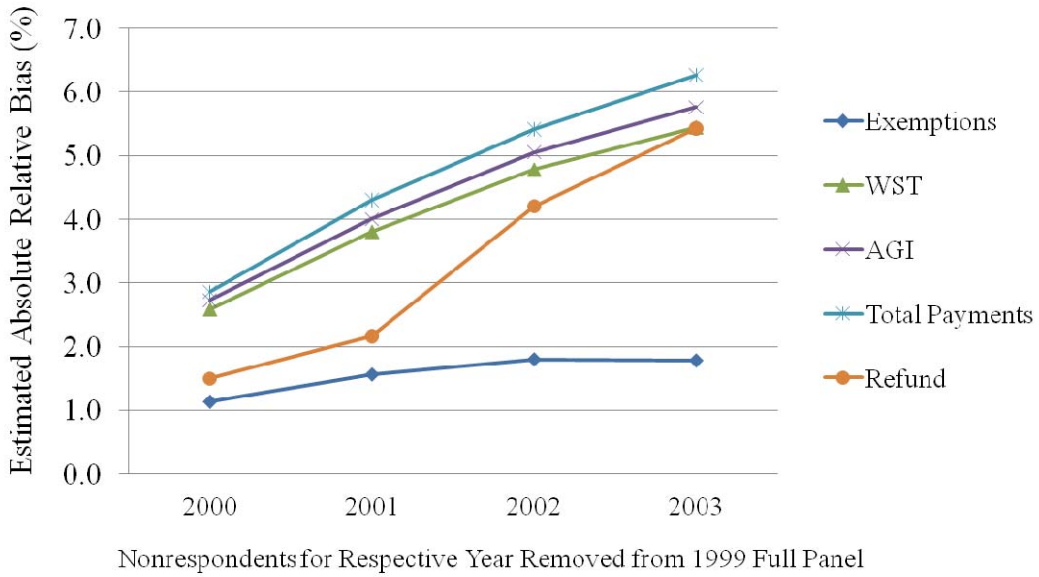
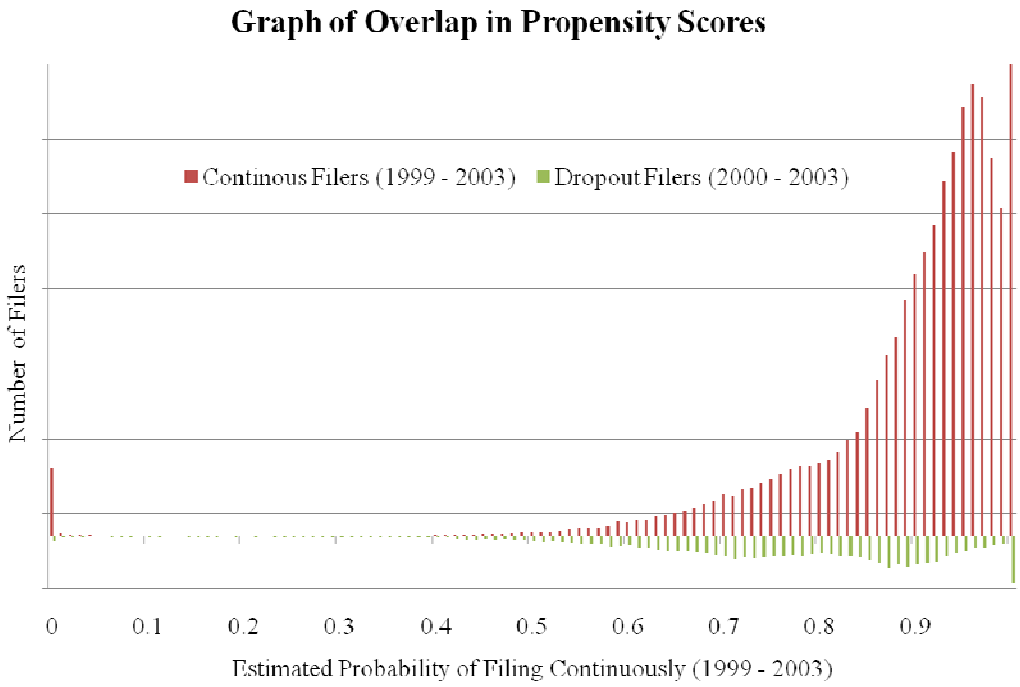


Figure 3. Graph of the Overlap in Propensity Scores – Predicting Propensity to Continuously File in the SOI 1999 Individual Panel



References

Arellano, M. (2003). Panel Data Econometrics, Oxford University Press 2003.

Bethlehem, J. (2002). "Weighting Nonresponse Adjustments Based on Auxiliary Information." In Groves, R., Dillman, D., Eltinge, J., and Little, R. (eds.) *Survey Nonresponse* (pp. 275-288). NewYork: Wiley, 2002.

Bryant, V. (2008). "Attrition in the Individual Income Tax Return Panel Tax Years 1999-2005." Proceedings of the American Statistical Association. American Statistical Association, Aug. 2008.

Christian, C. and Frischmann, P. (1989). "Attrition in the Statistics of Income Panel of Individual Returns." National Tax Journal. 24 (4), 495-501.

Clotfelter, C. (1980). "Tax Incentives and Charitable Giving: Evidence from a Panel of Taxpayers," Empirical Economics. vol. 13(3), pp. 319-40.

Cochran, W. (1968). "Removing Bias in Observational Studies." Biometrics, 24, 295 – 313.

Czajka, J., Hirabayashi, S., Little, R., and Rubin, D. (1987). "Evaluation of a New Procedure for Estimating Income and Tax Aggregates from Advance Data." *Statistics of Income and Related Administrative Record Research: 1986-1987,* Internal Revenue Service, November 1987, pp. 109-135.

Czajka, J. and Radbill, L. (1992). "Weighting Panel Data for Longitudinal Analysis," Proceedings of the Section on Survey Research Methods, American Statistical Association, Washington, DC.

D'Agostino, R.H. (1998). Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-randomized Control Group, Statistics in Medicine, 17, 2265 – 2281.

Ekholm, A. and Laaksonen, S. (1991). "Weighting via Response Modeling in the Finnish Household Budget Survey." Journal of Official Statistics, 7, 3, 1991, pp. 325 - 337.

Feldstein, M. (1995). "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," Journal of Political Economy, 103: 551-72.

Kalton, G. and Brick, M. (2000). Weighting in household panel surveys. In D. Rose (Ed.), Researching social and economic change. London: Routledge.

Kalton, G. and Citro, C. (1993). "Panel Surveys: Adding the Fourth Dimension." Survey Methodology, 19 (2): 205-215.

Kalton, G., Kasprzyk, D., and McMillen, D. (1989). "Nonsampling Error in Panel Surveys." In D. Kasprzyk et al. (eds.), Panel Surveys, New York: Wiley.

Kalton. G. and Miller. M. (1986). "Effects of Adjustments for Wave Nonresponse on Panel Survey Estimates," Proceedings of the Survey Research Methods Section, American Statistical Association

Lazarsfeld, P. (1948). “The Use of Panels in Social Research.” Proceedings of the American Philosophical Society 92 (5): 405:10.

Lepkowski, J. and Couper, M. (2002). “Nonresponse in the Second Wave of Longitudinal Household Surveys.”Survey Nonresponse, R.M. Groves et al. (eds). New York: John Wiley.

Lepkowski J., Kalton G., and Kasprzyk D. (1989). “Weighting Adjustments for Partial Nonresponse in the 1984 SlPP Panel” *Proceedings of the Section on Survey Research Methods,* American Statistical Association, 296- 301.

Lessler, J. and Kalsbeek, W. (1992). Nonsampling Errors in Surveys. New York: John Wiley.

Little, R. and David, M. (1983). “Weighting Adjustments for Non-response in Panel Surveys.” Washington D.C.: U.S. Bureau of the Census.

Oakes, J. and. Johnson, P. (2006). “Propensity score matching methods for social epidemiology.” In Oakes J & Kaufman J, eds. *Methods in Social Epidemiology*. San Francisco: Jossey-Bass/Wiley.

Rosenbaum, P. and Rubin, D. (1984). “Reducing Bias in Observational Studies Using Subclassifcation on Propensity Score,” Journal of American Statistical Association, 79, 516 – 524.

Rubin, D. (1985). The Use of Propensity Scores in Applied Bayesian Inference. *Bayesian Statistics, 2.* Bernardo, Degroot, Lindley, and Smith (eds.). North Holland, pp. 463-472.

Sailer, P., Weber, M., and Wong, W. (2000). “Attrition in a Panel of Individual Income Tax Returns, 1992-1997,” 2000 Proceedings of the Section on Government Statistics. American Statistical Association, Aug. 2000.

Walker, B. and Testa, V. (2003), Statistics of Income – 2001 Individual Income Tax Returns, Internal Revenue Service, Publication 1304, pp. 23-27.

Weber, M. and Bryant, V. (2005). “The 1999 Individual Income Tax Return Edited Panel,” 2005 Proceedings of the American Statistical Association. American Statistical Association, Aug. 2005.

Weber, M. (2006). “The Tax year 1999-2003 Individual Income Tax Return Panel: A First Look at the Data,” 2006 Proceedings of the American Statistical Association. American Statistical Association, Aug. 2006.

Zanutto, E., Lu, B., and Hornick, R. (2005). Using Propensity Score Subclassification for Multiple Treatment Doses to Evaluate a National Antidrug Media Campaign. Journal of Educational and Behavioral Statistics, Vol. 30, No. 1 (Spring 2005), pp. 59- 73.