

**Marquette University**  
**e-Publications@Marquette**

---

Accounting Faculty Research and Publications

Accounting, Department of

---

10-1-2014

# Using Archival Data Sources to Conduct Nonprofit Accounting Research

Nancy Chun Feng  
*Suffolk University*

Qianhua Ling  
*Marquette University, [qianhua.ling@marquette.edu](mailto:qianhua.ling@marquette.edu)*

Daniel Neely  
*University of Wisconsin - Milwaukee*

Andrea Alston Roberts  
*University of Virginia - Main Campus*

---

Published version. *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 26, No. 3 (Fall 2014): 458-493. [Permalink](#). © 2014 PrAcademics Press. Used with permission.

## USING ARCHIVAL DATA SOURCES TO CONDUCT NONPROFIT ACCOUNTING RESEARCH

Nancy Chun Feng, Qianhua (Q.) Ling,  
Daniel Gordon Neely and Andrea Alston Roberts\*

**ABSTRACT.** Research in nonprofit accounting is steadily increasing as more data is available. In an effort to broaden the awareness of the data sources and ensure the quality of nonprofit research, we discuss archival data sources available to nonprofit researchers, data issues, and potential resolutions to those problems. Overall, our paper should raise awareness of data sources in the nonprofit area, increase production, and enhance the quality of nonprofit research.

### INTRODUCTION

The nonprofit sector is a vital part of the U.S. economy, contributing two percent of national GDP and employing close to ten percent of employees (Wing, Pollak, & Blackwood, 2008). Although it is an important sector of the U.S. economy, it is only in the last decade that nonprofit accounting research has built momentum toward developing a critical mass.<sup>1</sup> This increase in accounting papers is most likely linked with the availability of Form 990 data

---

\* Nancy Chun Feng, DBA, is an Assistant Professor, Accounting Department, Suffolk University. Her research interests include auditing, governance, and financial reporting in nonprofit and for-profit organizations. Qianhua (Q.) Ling, Ph.D., is an Assistant Professor of Accounting at Marquette University. Her research interests are in financial and non-financial reporting, governance, and the role of information intermediaries as these relate to the governmental and nonprofit sectors. Daniel Gordon Neely, Ph.D., CPA, is an Assistant Professor of Accounting at the University of Wisconsin-Milwaukee. His research interests are in governmental and nonprofit accounting. Andrea Alston Roberts, Ph.D., is an Associate Professor at the University of Virginia. Her research interests center around reporting quality and governance in nonprofit organizations.

provided by the National Center for Charitable Statistics (NCCS) described by Gordon et al. in 1999.

The number of published nonprofit accounting papers has increased but this number as well as the number of nonprofit scholars continues to be relatively small.<sup>2</sup> Because of the dearth of nonprofit accounting researchers and because data has been only available for a relatively short time period (especially when compared to data sources such as CompuStat) it can be challenging to become aware of data issues that may impact research designs and the analysis of data.

The overall objective of this paper is to increase the awareness of publicly available archival data sources that can be used to study a wide range of nonprofit types. To accomplish this, we survey recent (i.e., papers published since 1999) nonprofit archival accounting papers (using ABI/Inform, EBSCO, Google Scholar, and Business Source Complete) that study public charities and we focus on the data sources used and data issues described in these studies. We also describe data issues that we have identified in connection with our various nonprofit accounting research projects but are not found in the extant literature.

Gordon et al. (1999) in their seminal work describe NCCS databases. Our paper differs from Gordon et al. (1999) in two major aspects. First, in addition to describing the NCCS data, we discuss other data sources such as those available from the IRS, charity oversight agencies, and state charity officials. We also describe ways to obtain audit data. Second, we discuss data issues related to these data sources and, where possible, provide potential ways to resolve these problems.

This paper has three major contributions. First, by identifying available data sources, we help researchers evaluate the possibility of conducting research projects in the nonprofit sector. Second, by describing issues with the data and providing resolutions, we facilitate sample selection, programming and data analysis efforts, which reduce the possibility of making incorrect inferences. Third, by synthesizing issues related to archival data into one central location, we aid in making the research process more efficient. It is our hope that this paper will provide a basis for the next wave of groundbreaking research in the nonprofit sector, and that increasing

the production of quality nonprofit research will further our understanding of this important sector of the economy.

Most frequently, researchers use data from Form 990, and thus, we use the next section to focus on issues related to the 990. Following the section about Form 990, we describe a variety of ways to obtain 990 data (as well as other non-Form 990 financial data). We then discuss ways to obtain audit data. Within each major section, we discuss issues associated with each of the data sources. The final section concludes.

### **FORM 990 DATA**

Form 990, *Return of Organization Exempt from Income Tax*, is an informational return required to be filed annually with the Internal Revenue Service (IRS) and is the only nonprofit financial report that U.S. law requires nonprofits make available to the public.<sup>3</sup> The 990 includes both financial and governance information.<sup>4</sup> The most frequently used sections (in accounting research) have been sections that: (1) report revenues and expenses (Part I in old 990); (2) report detailed expenses (such as executive compensation, salaries, accounting fees, and supplies) classified as program, management and general, and fundraising expenses (Part II in old 990); and (3) report the organization's financial position (Part IV in old 990).

There has been noticeable growth in the number of accounting papers that use 990 data in the last decade. In 1996, a provision in the Taxpayer Bill of Rights 2 required nonprofits to mail copies of their 990s to anyone upon written request. Prior to this, the law required organizations make 990s available at the organization's headquarters or the IRS headquarters in Washington D.C. only (see Gordon et al. 1999 for a more detailed history on this sequence of events). This surge in nonprofit accounting research after 1999 can be attributed to the fact that by then, the NCCS started to make 990 data easily assessable and available in database form.

#### **The Old and New Form 990 Described**

The content and design of the 990 has changed several times since its inception. The most significant redesign is for tax years beginning 2008.<sup>5</sup> The new 990 includes many reorganizational changes as well as new information that may be of special interest to

scholars (we discuss changes related to compensation disclosures in a separate section). It requires more details about functional expenses and much more narrative information pertaining to the organization's operations and programs. The new form adds: (1) a section dedicated to governance matters (that requires disclosure about management policies and disclosure practices) (Part VI); (2) detailed information for endowment funds (Schedule D); (3) a schedule that reports international activities (Schedule F); (4) hospital charity care reports (Schedule H); (5) supplemental information for tax exempt bonds (Schedule K); and (6) a schedule that details non-cash contributions (Schedule M).

### **Issues with Form 990**

Although research documents that the 990 is generally a reliable source of data, the concern that 990s include preparation errors (both intentional and unintentional) and may not agree to financial statements prepared in accordance with GAAP is well-documented. Most research that studies the reliability of 990 data compares the reports to audited financial statements or other third party reports (Froelich & Knoepfle, 1996; Gantz, 1999; Froelich et al., 2000; Fischer et al., 2002; Grønbjerg, 2002; Wing & Hager, 2004; Gordon et al., 2007; Yetman et al., 2009). Each of these studies reports errors ranging from mathematical and transposition mistakes to omissions of required information. Both Fischer et al. (2002) and Gordon et al. (2007) provide explanations for some of the discrepancies between Form 990 and audited financial statements. They indicate that differences exist because: (1) GAAP requires the consolidation of related entities that may have separate tax-exempt status; (2) there are differences between some 990 reporting rules and nonprofit GAAP; and (3) there is a lack of understanding on the part of some preparers that results in failure to comply with 990 instructions. In fact, research documents that large organizations (Froelich & Knoepfle, 1996; Keating et al., 2008), those that pay executives relatively high salaries (Keating et al., 2008); and organizations that pay an external accountant (Krishnan et al., 2006; Keating et al., 2008) are more likely prepare high quality 990s.

Most often, to reduce the introduction of noise into the data, studies eliminate observations that contain obvious computational errors that would impact the variables under study. For example, Roberts (2005) eliminates observations when the sum of program

expense, administrative expense, and fundraising expense (on Part II) does not equal total expense; and Thornton and Belski (2010) eliminate observations when donation revenue exceeds total revenue or program expenses exceed total expenses.<sup>6</sup>

In addition to mathematical mistakes, a common 990 error is the omission of data (omitted data often is shown as zero values in the databases). For example, Tinkelman (1999) and Jacobs and Marudas (2009) eliminate observations when total revenue, total expenses, total assets, or donations equal zero. Further, scholars have documented a large number of nonprofits that report zero fundraising expenses (Tinkelman, 1999; Froelich et al., 2000; Wing et al., 2004; and Krishnan et al., 2006). While some nonprofits do not engage in fundraising activities, others incorrectly report zero fundraising expense (Wing et al., 2004; Krishnan et al., 2006; Keating et al., 2008).<sup>7</sup>

To date, scholars who explicitly consider this issue when designing their studies, eliminate observations where fundraising expense equals zero (Roberts, 2005; Jacobs & Marudas, 2009; Tinkelman, 1999, 2006).<sup>8</sup> We advise researchers to consider when it is appropriate to exclude all observations where fundraising expense equals zero. Wing et al. (2004), Krishnan et al. (2006), and Yetman and Yetman (2013) suggest plausible explanations for reporting zero fundraising (or administrative expenses).<sup>9</sup> One explanation is that the entity is affiliated with another that incurs and reports the bulk or all of the fundraising and administrative costs. Line H on the 990 (old and new) reports whether the organization is affiliated with another organization. We examine the Digitized Data files maintained by the NCCS (described later) and find that 25 percent of organizations that report zero fundraising (administrative) expenses have an affiliate. Of the cases that report zero program expenses approximately 9.3 percent are affiliated with another organization.

In addition to the 990 issues identified in the extant literature, there are four additional items that we have observed that may impact research activities. First, total expenses from Part I and Part II of the old 990 in many instances do not equal. Most frequently the difference relates to '*payments to affiliates*' on Part I. We find (using the Digitized Data, described later) that 55 percent of the cases where total expense on Parts I and II do not equal are due to payments to affiliates on Part I. This will not be an issue in the new

990, because the breakdown of the expenses by function is only available in Part IX.<sup>10</sup> Second, in rare instances we observe that line item expenses on Part II (old 990) in the administrative or fundraising categories are negative (this can continue to be an issue on Part IX of the new 990). Third, the categorical expense data (old 990 – Part II; new 990 –Part IX) may be incomplete. In fact, many organizations use the ‘other expenses’ line to report a significant portion of their expenses (The Urban Institute, 2003). We find that the average organization reports approximately 33 percent of its expenses as ‘other expenses.’ Based on casual observations, much of the ‘other expenses’ relates to professional service fees (e.g., fees paid to hospital doctors). Since the new 990 requires organizations to provide more ‘other expense’ disclosure and include lines dedicated specifically to professional service fees, lobbying, advertising and promotion, this change will likely reduce the proportion of ‘other expenses’ in total expenses.

Researchers should also be aware that changes in the Form 990 present some challenges for time series analysis for several reasons. First, some information collected on old 990s is no longer directly reported on the new 990s. For example, the direct public support line (line 1b on the old 990) is not available on the new 990. Second, over the years thresholds for organizations required to file Form 990 have changed. Most recently, beginning with tax year 2008, the filing threshold has increased from gross receipts of \$100,000 and total assets of \$250,000 prior to 2008 to gross receipts of \$200,000 and total assets of \$500,000 currently.<sup>11</sup> This means that organizations may drop from samples when they fall below thresholds. Finally, the requirement to file the new 990 has been phased in over a three-year period and is based on the size of the organization (larger organizations were required to file the new 990 first). Organizations not required to file the new 990 in the first two years of the phase-in period were permitted to file Form 990-EZ during these years. Since these organizations had been required to file the old 990, the filing phase-in may affect time-series analyses that include both pre- and post-2008 data.

The final issue relates to an accounting change that impacted 990 reporting (not 990 preparations). The adoption of SFAS 116 in 1996 changed the way nonprofits accounted for donations (nonprofits must account for donations when pledged and not

received), and thus, results can be impacted if the research design includes contribution data in the pre- and post-adoption periods (see Derrick 2005).

### Issues with Form 990 Compensation Information

The use of the compensation disclosure in the 990 has become increasingly popular among researchers (e.g., Baber et al., 2002; Hallock, 2002; Eldenburg & Krishnan, 2003; Carroll et al., 2005; Brickley & Van Horn, 2002; Core et al. 2006; Krishnan et al., 2006). Form 990 discloses compensation information in several places. We first discuss the old 990. Part II, line 25a (new 990 Part IX, line 5) reports total compensation to current executives, directors, trustees, and key employees and line 26 (new 990 Part IX, line 7) reports total wages.<sup>12</sup> Part V '*Officers, Directors, Trustees, and Key Employees*' discloses the name, title, salary, fees, bonus, payments to benefit plans, and payments to expense accounts (or other allowances) for officers, directors, trustees, and key employees.<sup>13</sup> The 990 instructions (prior to 2008) define key employees in terms of responsibility, power, or influence and require entities to report the compensation of executive directors (and those with similar titles), irrespective of pay, on Part V. Schedule A, Part I, '*Compensation of the Five Highest Paid Employees Other than Officers, Directors, and Trustees*' reports the same compensation information as reported in Part V for the top five paid employees that are *not* current officers, directors, trustees, or key employees and whose pay exceeds \$50,000. Schedule A, Part II, '*Compensation of the Five Highest Paid Independent Contractors for Professional Services*,' reports compensation information for those indicated by the title. In the new 990, all compensation detail is reported in one place, Part VII, '*Compensation of Officers, Directors, Trustees, Key Employees, Highest Compensated Employees, and Independent Contractors*'. Section A reports information for officers, directors, key employees, and the five highest paid employees and Section B reports the five highest paid independent contractors.

There are two notable changes regarding compensation disclosures beginning with tax year 2008. First, thresholds for reporting compensation have changed. The threshold related to key employees has changed from zero to \$150,000. We note that in the instructions to the old 990, top management (e.g., the executive director) is a key employee and there is no reporting threshold



requirement for key employees. In the instructions to the new 990, top management is an officer or director and there continues to be no reporting threshold requirement. The threshold related to the highest paid non-key employees and independent contractors has increased from \$50,000 to \$100,000.

Second, detailed compensation amounts reported should largely come directly from W-2 or 1099 forms (IRS, 2008). Because entities are required to use a calendar year to report compensation detail (on Part VII of the new 990) and to use a fiscal year to report total executive compensation (on Part IX line 5 of the new 990), these totals are unlikely to agree unless the fiscal year ends on December 31. Old 990 preparers had the option to use a calendar or fiscal year when reporting compensation detail; hence, total executive compensation from Part V and line 25 may not equal.

There are three overall challenges for researchers to consider when using compensation data from Form 990. The first challenge relates to preparer errors on line 25 and Part V of the old 990. The most frequent preparer error is that this information is left blank or is zero. Based on casual observation, it appears that in some instances total executive compensation is included on line 26 (total wages). These issues may continue with the new 990. Although the old 990 instructions clearly indicate an executive director is a key employee, and thus, irrespective of their total pay, information related to their compensation should be reported on Part V, we observe instances when compensation data for top management is reported on Schedule A and not Part V.<sup>14</sup> Thus, compensation related to top management may not be disclosed on Part V either because the data are missing or because it is reported on Schedule A. This issue is not likely to occur with the new 990.

The second major challenge relates to the identification of top management. Identifying management is a challenge for a number of reasons. First, the highest paid individuals may not necessarily be managers (e.g., athletic coaches at universities many times are paid more than the president of the same university). Second, in many instances titles differ (and some do not report titles) among and within organizations over time. Although the most popular titles are executive director, chief executive officer, president, and founder, the data include a wide range of possible titles. Further, several executives may share the same title and the name of the same officer

may be reported in different ways throughout a time period. For example, one year the name reported is James Smith and another year the name reported is Jim Smith (or names change due to marriage or divorce). Researchers must standardize the titles and names if the methodology requires consistency throughout a time period.

Third, an organization may file an affiliated return and the executive director for each affiliate is listed as an executive director (or whichever title the organization uses to identify the top manager).<sup>15</sup> Additionally, both current and departing executives frequently share the same title and both are reported in the same section of the 990. Beginning with tax year 2005, however, current officers are reported in Part V-A and former officers are reported in Part V-B (this current and former distinction continues with the new 990 but they are reported in the same place in Part VII). The next section discusses ways to obtain 990 data.

#### **AVAILABILITY OF FORM 990 AND OTHER FINANCIAL DATA**

Form 990 data can be obtained from several sources – the Internal Revenue Service (IRS), the National Center for Charitable Statistics (NCCS), charity oversight agencies, and state charity officials' websites. We describe each of these sources below and discuss the advantages and disadvantages associated with each source. Toward the end of this section – specifically within the section that describes state charity official's websites – we discuss the availability of financial data, in addition to Form 990 data.

#### **IRS Database**

In addition to Form 990 data, The IRS provides information on its website that may interest researchers (<http://www.irs.gov/charities/article/0,,id=96136,00.html>). For example, the website includes IRS Publication 78, *Cumulative List of Organizations described in Section 170(c) of the Internal Revenue Code of 1986*. Organizations in this database can be searched by name, location, and deductibility code. In addition to the cumulative list, which is not updated frequently, there are three other lists: (1) *'Additions to Cumulative List'* that separately reports newly approved tax exempt organizations; (2) *'Recent Revocations and Deletions from Cumulative List'* that reports organizations that lost their tax exempt status beginning in or after

January 2005; and (3) 'Suspensions Pursuant to Code Section 501(p)' that reports suspended organizations.

Further, in addition to many statistical tables, the Statistics of Income (SOI) Division of the IRS makes available for free several nonprofit data files that can be used for nonprofit research (<http://www.irs.gov/taxstats/>). The files described below are available in ASCII text format.

**Master Files.** These files (go to the *Exempt Org Masterfile* link on the IRS website indicated above) come from the IRS' Business Master File (BMF) and include information for all active organizations with tax-exempt status (these files are also available in Excel).<sup>16</sup> The Master Files include contact information, descriptive information and three financial variables – total assets, total revenues, and total income (refer to the 'Exempt Organizations Master Listing' for description of variables included in the data files). The files are divided by state, area, or region and are updated monthly.

**Microdata Files.** These files (go to Charities link on the website) contain nonfinancial and substantially all financial variables for a sample of 501(c)(3) organizations and section 501(c)(4)–(9) organizations (a description of the variables included in each file is included when the data are downloaded). Sampling rates for the 501(c)(3) range from 1 percent for small-asset classes to 100 percent for large-asset classes. All 501(c)(3) nonprofits with an asset size greater than or equal to \$50 million in assets (\$10 million before 2000) are included.<sup>17</sup> The Microdata files are available for the years 1992 – 2009. In separate files, data are also available for private foundations required to file Form 990-PF.

The IRS databases are the original sources of many NCCS data. They are publicly available, easily accessible, and free. The IRS data is in ASCII text format, and thus, is not as user friendly as the databases described below.

**NCCS Databases.** The NCCS (<http://nccsdataweb.urban.org>) provides a wealth of financial information (we urge those interested in nonprofit research to explore). For example, in addition to the data files we discuss below, aggregated individual tax information (including information related to charitable deductions that can be classified by zip code) and census data are available. It also includes Form 990 data for 501(c)(3), non-501(c)(3) organizations and Form

990-PF data for private foundations. To date, the most widely used NCCS data are from Form 990 and relate to 501(c)(3) nonprofits.<sup>18</sup>

The IRS has been providing this data to the NCCS since 1999. In turn, the NCCS makes this data available to the public in several different database files accessible through tools on their DataWeb interface. Currently, the NCCS charges academic subscribers an annual flat fee to gain full access to the databases (limited access is available free of charge). Data are user-friendly and can be extracted and downloaded into Excel, SAS, SPSS, Stata, ASCII, dBASE and SQL formats. Because the use of the NCCS data is relatively new, data issues continue to be identified. We describe the major and most frequently used NCCS data files and the issues associated with these data files below.<sup>19</sup>

***Business Master Files (BMF).*** These files are available for 1995 through December, 2012. The NCCS downloads and updates this data quarterly from the Master File maintained by the IRS (described in the prior section). The NCCS posts three separate files: (1) a file that includes all tax-exempt organizations with a limited set of fields, (2) a file that includes 501(c)(3) organizations with the same detail found in the IRS files, and (3) a file that includes non-501(c)(3) organizations with the same detail found in the IRS files. In addition to the file formats, the major differences between the master files maintained by the NCCS and IRS are: (1) the NCCS categorizes the files according to whether the organization is a 501(c)(3) or a non-501(c)(3) and the IRS categorizes the files based on state, area, or region, (2) the NCCS provides individual file data (for each quarter) and the IRS file is cumulative, and (3) the NCCS files contain more classification information, such as FIPS, NAICS, and Metropolitan Statistical Area codes.<sup>20</sup>

***Statistics of Income (SOI) Files.*** These files come from the Microdata files maintained by the SOI Division of the IRS (see prior section) and are available for 1982 - 1983 and 1985 - 2009. The NCCS updates the files as the IRS makes the files available. In addition to the file formats, the major differences between the NCCS SOI files and the IRS Microdata files are: (1) the NCCS files contain more classification information, such as FIPS, NAICS, and Metropolitan Statistical Area codes, (2) the NCCS files include older information (the NCCS files begin in 1982 and the IRS files begin in 1992), and (3) the NCCS

maintains separate files for 501(c)(3), non-501(c)(3), 990-EZ, and 990-PF filers.

The SOI database includes most financial variables on Form 990 and is considered to be the most reliable dataset because of the substantial error checking conducted by the Statistics of Income Division of the IRS. As mentioned previously, the files include data for organizations with assets greater than or equal to \$50 million in assets (\$10 million before 2000), plus a random sample of smaller organizations stratified and weighted by asset level. Studies that use this dataset are biased toward larger organizations. Tinkelman (2004) indicates that this bias becomes more pronounced in longer windows when a time series of data are required since the probability of a “small” organization being selected for several consecutive years is low. We randomly choose 100 small organizations (with asset size less than \$10 million) in 1987 and track these organizations through 2007 to document whether small organizations are in more than one data file year. We find that in 27 cases the organization is in each data file between 1987 and 2007, in 14 cases the organization is in at least half of the data files, and in 44 cases the organization is in a data file more than once (in only a few instances does the asset size exceeds \$10 million). Although we do not downplay the size bias issue, the severity of the bias over long windows may be somewhat tempered.

**Core Files.** These files are available for 1989 through 2010 and are only available through the NCCS. These files are updated annually in the spring or summer. Core files contain nonfinancial and many financial variables for all organizations required to file the 990. In separate files, data are available for organizations required to file the 990-PF. The Core files (for Form 990 and 990-EZ) do not include financial data on program or administrative expenses, two variables that, to date, have been a primary focus of accounting research. Beginning with the 2007 Core files, however, estimates of program and administrative expenses (in addition to other line items) are included in the files.<sup>21</sup> Because the Core files include data for all nonprofits that file the 990, results from studies that use this dataset are considered more generalizable to the population of organizations under study.<sup>22</sup>

**Digitized Data Files.** These files are available for 1998 – 2003 and contain all information from the Form 990 and 990-EZ for the

population of nonprofits that file Form 990 during this time period. The Digitized Data, a joint project between Guidestar (described later) and the NCCS, resolves the size bias inherent in the SOI files. The disadvantage of this dataset is that the time period covered is short (and 1998 does not include all nonprofits that filed Form 990 in that year). Since Guidestar now sells customized datasets and the NCCS lacks funding to obtain additional data for Digitized Data files, it is unlikely that the NCCS will update this dataset.

### **Issues with NCCS Files**

Because Form 990 comprises the NCCS data files, errors and issues with the 990 work their way into these datasets. Each of the data files are checked for errors to varying degrees. Data included in the SOI files are substantially checked for errors by the IRS.<sup>23</sup> The data included in the Core files are reviewed by the NCCS for quality before it is released to the public.<sup>24</sup> Additionally, the NCCS corrected major errors before the Digitized Data was released to the public.<sup>25</sup> Although the NCCS data are generally reliable, Tinkelman and Neely (2011) document that results are sensitive to the inclusion of outliers. And since errors can also result from the process of transcribing data into the files, the NCCS recommends comparing any suspicious observations to the full text versions of the 990 (available at [www.Guidestar.org](http://www.Guidestar.org)).<sup>26</sup> We discuss the next set of issues with NCCS files for researchers to consider in their sample selection, programming or data analysis efforts (these issues may also apply to the IRS files).

### ***Missing or Duplicate Years in the Data Files***

Prior research eliminates observations when identical financial information is reported for different organizations (Keating et al. 2005). There are also instances when the data files contain records where: (1) the years are identical but the 990 data differs; (2) years are skipped or missing; or (3) consecutive year financial data are identical (Tinkelman, 2004). We examine the Digitized Data for each file year and find that the issues described are infrequent and occur in most instances less than 0.01 percent of the time annually.<sup>27</sup>

### ***Variable Names***

The variable names for all SOI file entries change beginning in file year 2000 (this is not the case in the Core files). For example, the variable code for total revenue before 2000 is E047 and the code is R270 beginning with file year 2000. These codes changed again beginning with tax year 2008.

### ***Data File Year***

The SOI and Core data file year is often not an organization's fiscal year (Tinkelman, 2004; NCCS, 2006) and may not be the tax year (the year on the 990 form). The SOI file year represents the year that starts the organization's fiscal year (e.g., an organization with a fiscal beginning July 1, 2007 and ending June 30, 2008 will be included in the 2007 SOI file). The years included in the Core files are limited to fiscal periods starting in the most recent calendar year but may also include the prior two calendar years. For example, the Core 2007 file includes fiscal years starting in 2007 (similar to the SOI file). But, if an organization is missing a 990, then a return for 2006 or 2005 (if 2006 is missing) will be added.<sup>28</sup> To identify new 990 and new 990-EZ data in the Core 2008 files, use the DocCd field to identify the type of return. The new Form 990 records have DocCd = '93' and new 990-EZ records have DocCD = '92'.<sup>29</sup>

Each data file includes the fiscal year (except for the 1982 and 1983 SOI files that include month-end only). The fiscal year formats, however, differ for each of the files and over time. In particular, the fiscal year is in "yyyymm" format (199412=Dec. 1994) for Core file years beginning in 1997 with earlier years using a 2-digit year ("9412"). The fiscal year is in "yymm" (year and month) for SOI file years beginning 2000 with earlier years using two separate fields – one for the year-end and another for the month. Researchers must standardize the years if combining data files or using a long time series of data.

### ***Start Date***

The organization's start date indicated on the NCCS files is the date recorded on the Business Master File data files. This start date is not the date that an organization was founded but it is the date that the organization filed for exempt status. Most researchers use this as a proxy when computing the age of the organization. Since the

IRS began operations in 1913, however, age is necessarily truncated (Tinkelman, 2004). This data field is also sometimes empty.

### ***NTEE Classifications.***

The National Taxonomy of Exempt Entities (NTEE) system, developed by the NCCS, is used to classify nonprofit organizations into 'industry' groups (see <http://nccs.urban.org/classification/NTEE.cfm>). The taxonomy categorizes organizations into twenty-six major groups (alphabetic code) and then further categorizes the organizations into subdivisions (using decile and centile level codes). Organizations that exist across all major groups and engage in a common set of activities – such as fundraising or research – are given one of seven 'common codes'. Most nonprofit accounting studies employ the use of NTEE major group codes in some form.<sup>30</sup> Currently, both the IRS and NCCS assign organizations NTEE codes and in some instances these codes differ. The NCCS' SOI files include both versions, the data field NTEE represents the code assigned by the NCCS and the NTEESOI is the code assigned by the IRS. Further, NTEE codes can change over time due to refinements in NTEE coding policy and practice or changes in the nature of an organization. We recommend the use of NTEE codes from the most recent BMF or data files.<sup>31</sup>

Further, the NTEE system is used for classifying organizations and not programs. The Nonprofit Program Classification (NPC) System, however, provides a means for capturing and classifying the program activities reported on the 990 (see <http://nccs.urban.org/classification/NPC.cfm>). Currently, researchers would need to review the organizations' programs and assign each program a code. The NCCS is currently working on a project that will assign NPC codes to organizations similar to the way it assigns NTEE codes.

### ***Geographic Information***

The NCCS recommends that geographic information be verified for accuracy (or, at minimum, plausibility). On relatively rare occasions, a parent organization files returns for multiple affiliates. These affiliates may use the legal address of the parent organization, thus inflating the number of organizations and the nonprofit activity in a particular city. Additionally, some data files might contain incorrect state abbreviations.



### **Compensation Information**

The NCCS data files include compensation information to varying degrees. The Core files include total executive compensation and total wages (lines 25 and 26 on the old Form 990 and Part VIII, lines 5 and 7 on the new Form 990). The SOI files include the number of compensated officers and directors from Part V (along with other basic information). For Schedule A, the SOI files include the dollar amounts related to salary, payments to benefit plans, and payments to expense accounts for each of the five highest paid employees that are not officers, directors, or key employees. The SOI data, however, does not include names and titles. The Digitized Data includes all information reported on Parts V and Schedule A, including names and titles. We find (using the Digitized Data) that for approximately 30 percent of the observations (124,303 of 420,045), line 25 is zero when total revenue exceeds \$3.6 million (the mean total revenue for the population of nonprofits that file Form 990).<sup>32</sup> In addition, based on our analysis of the Digitized Data, we determine that for as many as 15 percent of the population of nonprofits, the executive director's compensation is reported on Schedule A instead of Part V.<sup>33</sup>

### **Charity Oversight Agencies Databases**

Several charity search engines provide financial information and evaluate select groups of charities. These include: (1) Charity Navigator (<http://www.charitynavigator.org/>); (2) the American Institute of Philanthropy (AIP) (<http://www.charitywatch.org/azlist.html>); (3) the Better Business Bureau's Wise Giving Alliance (BBB) (<http://www.bbb.org/us/charity/>); and (4) Ministry Watch (<http://ministrywatch.org/home.aspx>). Access to the Charity Navigator, BBB, and Ministry Watch (evaluates faith-based charities) reports are free and access to the complete set of AIP reports are available for a fee.<sup>34</sup>

Guidestar (<http://www2.guidestar.org/>), the most comprehensive of the search engines, provides free access to the most recent pdf versions of the 990 for all organizations that file Form 990, 990-PF or 990-EZ. Academics can obtain complimentary access to Form 990s (as far back as 1998) as well as other financial information by applying for the *Guidestar for Education* package. Guidestar, for a fee, will extract data from Form 990 and build customized datasets. This service can be particularly useful when the most current 990

data is needed or when it is not cost beneficial to pay for the access to the NCCS data for a year.<sup>35</sup>

### **State Charity Officials' Websites**

Another valuable source of nonprofit data is State Charity Officials websites. Nonprofits are regulated by the states where they solicit funds (go to the National Association of State Charity Officials website, <http://www.nasconet.org/documents/u-s-charity-offices/>, to be linked to charity official websites or see States' Charity Officials' Online Web Based Resources Table below). Depending on the state's regulatory requirements, nonprofits may be required to submit annual financial information (see Freemont-Smith 2005 for discussion of state regulatory requirements). With some exceptions, much of the financial information available comes from the 990 and can be obtained by using either the NCCS data files or Guidestar's (described before) more comprehensive database. Some states, however, provide access to additional financial information (see table below for a description of the data available for each state) that cannot be gathered using other sources. Much of this information, however, is not in database form.

For instance, Illinois, Massachusetts, New York, North Carolina, and West Virginia provide access to audited financial statements of the nonprofits that solicit in their state. Many more states make available commercial fundraiser reports. Many of these reports (in pdf format) include the name of the commercial fundraiser, the name of the nonprofit, amounts raised by the commercial fundraiser and either the amount or percentage remitted to the charity (the commercial fundraiser reports, CFR, in Table 1 include at a minimum both the names of the commercial fundraiser and the nonprofit). The requirements for commercial fundraiser reports vary across states (e.g., some reports include only information about campaigns conducted within the state while others include information about nationwide fundraising proceeds) and the spelling of the nonprofits' and professional solicitors' names can differ. Researchers must standardize the reports if combining datasets across states. Keating et al. (2008) use this approach to gather data and describe in detail issues associated with using these reports and procedures to undertake so that the reports are useable sources of data.

**TABLE 1**  
**States' Charity Officials' Online Web Based Resources**

<b>State</b>	<b>Search List</b>	<b>CFR</b>	<b>AFS</b>	<b>Link to Nonprofit Information</b>
Alabama	No	No	No	
Alaska	No	No	No	
Arizona	Yes	No	No	<a href="http://www.azsos.gov/scripts/Charity_Search.dll">http://www.azsos.gov/scripts/Charity_Search.dll</a>
Arkansas	Yes	No	No	<a href="http://www.sos.arkansas.gov/corps/search_all.php">http://www.sos.arkansas.gov/corps/search_all.php</a> <a href="http://www.arkansasag.gov/charity/search.php">http://www.arkansasag.gov/charity/search.php</a>
California	Yes	Yes	No	<a href="http://rct.doj.ca.gov/MyLicenseVerification/Search.aspx?facility=Y">http://rct.doj.ca.gov/MyLicenseVerification/Search.aspx?facility=Y</a>
Colorado	Yes	Yes	No	<a href="http://www.sos.state.co.us/ccsa/">http://www.sos.state.co.us/ccsa/</a>
Connecticut	Yes	No	No	<a href="http://www.ct.gov/dcp/cwp/view.asp?a=1654&amp;q=468080">http://www.ct.gov/dcp/cwp/view.asp?a=1654&amp;q=468080</a>
Delaware	No	No	No	
District of Columbia	No	No	No	
<b>State</b>	<b>Search List</b>	<b>CFR</b>	<b>AFS</b>	<b>Link to Nonprofit Information</b>
Florida	Yes	No	No	<a href="http://ccfcorp.dos.state.fl.us/corinam.html">http://ccfcorp.dos.state.fl.us/corinam.html</a>
Georgia	Yes	No	No	<a href="https://secure.sos.state.ga.us/SBR_Weblookup_Prod/Search.aspx?facility=Y">https://secure.sos.state.ga.us/SBR_Weblookup_Prod/Search.aspx?facility=Y</a>
Hawaii	Yes	Yes	No	<a href="http://ag.ehawaii.gov/charity/search,BW,,,...html">http://ag.ehawaii.gov/charity/search,BW,,,...html</a>
Idaho	No	No	No	
Illinois	Yes	No	Yes	<a href="http://www.illinoisattorneygeneral.gov/charities/search/index.jsp">http://www.illinoisattorneygeneral.gov/charities/search/index.jsp</a>
Indiana	Yes	Yes	No	<a href="http://www.stats.indiana.edu/topic/nonprofits.asp">http://www.stats.indiana.edu/topic/nonprofits.asp</a>
Iowa	Yes	No	No	<a href="http://inrc.continuetolearn.uiowa.edu/community/database.asp">http://inrc.continuetolearn.uiowa.edu/community/database.asp</a>
Kansas	Yes	No	No	<a href="http://www.kscharitycheck.org/search.asp">http://www.kscharitycheck.org/search.asp</a>
Kentucky	Yes	No	No	<a href="http://ag.ky.gov/ag/agdownloads/charity.pdf">http://ag.ky.gov/ag/agdownloads/charity.pdf</a>
Louisiana	No	No	No	
Maine	Yes	No	No	<a href="http://pfr.informe.org/ALMSOnline/ALMSQuery/SearchCompany.aspx">http://pfr.informe.org/ALMSOnline/ALMSQuery/SearchCompany.aspx</a>
Maryland	Yes	No	No	<a href="http://www.sos.state.md.us/Charity/SearchCharity.aspx">http://www.sos.state.md.us/Charity/SearchCharity.aspx</a>

TABLE 1 (Continued)

State	Search List	CFR	AFS	Link to Nonprofit Information
Michigan	Yes	No	No	<a href="http://www.michigan.gov/documents/lic_chrt1_40579_7.pdf">http://www.michigan.gov/documents/lic_chrt1_40579_7.pdf</a>
State	Search List	CFR	AFS	Link to Nonprofit Information
Minnesota	Yes	No	No	<a href="http://www.ag.state.mn.us/charities/CharitySearch.asp">http://www.ag.state.mn.us/charities/CharitySearch.asp</a>
Mississippi	Yes	No	No	<a href="http://www.sos.ms.gov/securities_and_charities_charities.aspx">http://www.sos.ms.gov/securities_and_charities_charities.aspx</a>
Missouri	Yes	No	No	<a href="http://ago.mo.gov/checkacharity/">http://ago.mo.gov/checkacharity/</a>
Montana	No	No	No	<a href="http://www.doj.mt.gov/consumer/charity/default.aspx">http://www.doj.mt.gov/consumer/charity/default.aspx</a>
Nebraska	No	No	No	
Nevada	No	No	No	
New Hampshire	Yes	No	No	<a href="http://doj.nh.gov/charitable-trusts/charities.htm">http://doj.nh.gov/charitable-trusts/charities.htm</a>
New Jersey	Yes	No	No	<a href="http://www.njconsumeraffairs.gov/charity/chardir.htm">http://www.njconsumeraffairs.gov/charity/chardir.htm</a>
New Mexico	Yes	No	No	<a href="http://www.nmag.gov/office/Divisions/Civ/charity/CharitySearch/default.aspx">http://www.nmag.gov/office/Divisions/Civ/charity/CharitySearch/default.aspx</a>
New York	Yes	Yes	Yes	<a href="http://bartlett.ag.ny.gov/Char_Forms/search_charities.jsp">http://bartlett.ag.ny.gov/Char_Forms/search_charities.jsp</a>
North Carolina	Yes	Yes	Yes	<a href="http://www.secretary.state.nc.us/csl/Search.aspx">http://www.secretary.state.nc.us/csl/Search.aspx</a>
North Dakota	Yes	No	No	<a href="https://secure.apps.state.nd.us/sc/busnsrch/busnSearch.htm">https://secure.apps.state.nd.us/sc/busnsrch/busnSearch.htm</a>
Ohio	Yes	Yes	No	<a href="http://www.ohioattorneygeneral.gov/files/Publications/Publications-for-Non-Profit">http://www.ohioattorneygeneral.gov/files/Publications/Publications-for-Non-Profit</a>
State	Search List	CFR	AFS	Link to Nonprofit Information
Oklahoma	No	No	No	
Oregon	Yes	No	No	<a href="http://www.doj.state.or.us/cgi-bin/charigroup_db_start.pl">http://www.doj.state.or.us/cgi-bin/charigroup_db_start.pl</a>
Pennsylvania	Yes	No	No	<a href="http://web.dos.state.pa.us/cgi-bin/Charities/char_form.cgi">http://web.dos.state.pa.us/cgi-bin/Charities/char_form.cgi</a>
Rhode Island	No	No	No	
South Carolina	Yes	No	No	<a href="http://www.scsos.com/Public_Charities">http://www.scsos.com/Public_Charities</a>

TABLE 1 (Continued)

State	Search List	CFR	AFS	Link to Nonprofit Information
South Dakota	Yes	No	No	<a href="http://apps.sd.gov/applications/st32cprs/soscorplookup.aspx">http://apps.sd.gov/applications/st32cprs/soscorplookup.aspx</a>
Tennessee	Yes	No	No	<a href="http://www.state.tn.us/sos/charity/index.htm">http://www.state.tn.us/sos/charity/index.htm</a>
Texas	No	No	No	
Utah	Yes	No	No	<a href="http://www.consumerprotection.utah.gov/consumerinfo/lists.html?list=CH">http://www.consumerprotection.utah.gov/consumerinfo/lists.html?list=CH</a>
Vermont	Yes	Yes	No	<a href="http://www.sec.state.vt.us/centennial_nonprofit.html">http://www.sec.state.vt.us/centennial_nonprofit.html</a>
Virginia	No	No	No	
Washington	Yes	Yes	No	<a href="http://www.sos.wa.gov/charities/search.aspx">http://www.sos.wa.gov/charities/search.aspx</a>
West Virginia	Yes	Yes	Yes	<a href="http://apps.sos.wv.gov/business/charities/">http://apps.sos.wv.gov/business/charities/</a>
Wisconsin	No	No	No	
State	Search List	CFR	AFS	Link to Nonprofit Information
Wyoming	No	No	No	

Note: CFR is commercial fundraising report and AFS is audited financial statement.

In summary, researchers can obtain large Form 990 datasets from Guidestar, the NCCS, and the IRS. Guidestar includes the most recent data for nonprofits that file Form 990 and will customize datasets for a fee per dataset. The NCCS offers a variety of non-customized datasets (i.e., Business Master files, Core files, and the Digitized Data file) each with their limitations that researchers can access for an annual fee. Generally, NCCS datasets include fewer data fields than Guidestar can (assuming a customized database is created) but depending on the number of fields and observations required, NCCS may be less expensive to use. The final source for some Form 990 data, in database form, is directly from the IRS. These IRS databases are not as user-friendly as the NCCS data but are free. Next, we discuss audit data.

### AUDIT DATA

Obtaining access to audited financial information is notably more challenging than obtaining 990 data. It is challenging because there is no central location or clearinghouse (like [edgar.gov](http://edgar.gov) for publicly traded companies or Guidestar for 990s) to obtain audited reports. This is most likely because the requirement to obtain an audit is mandated at the state and not the federal level and federal law requiring an audit pertains only to those organizations receiving federal funding over certain thresholds.<sup>36</sup> Further, obtaining data on audit fees is problematic since auditors are not required to report this information on any federal document. In fact, to date, only Grein and Tate (2011) get audit fee data using a database. They obtain this information from the Department of Housing and Urban Development (HUD) for all nonprofit public housing authorities (Vermeer et al. 2009 use audit fee data obtained from surveys). In this section, we discuss sources to obtain audited financial statements, auditor data and auditor opinions.

#### Audited Financial Statements

There are three ways to obtain audited financial statements. One option is to use reports gathered directly from the organization. This is accomplished by either requesting the statements directly from the organization or visiting the organization's website (see Gordon et al. 2010 for a discussion about characteristics of nonprofits that disclose audited financial statements on their website). Since nonprofits generally are not required to provide this information upon request and disclosing audited financial statements is voluntary, the disadvantage of this approach is that there may be an inherent selection bias in the research results.

A second option is to use the audited financial statements available on select state charity official websites (see Table 1 for a list of state charity officials that make audit reports available on their websites for nonprofits that solicit in their state).<sup>37</sup> Although it is likely that many of the nonprofits that file audited financial reports with state agencies solicit nationally, it is the researcher's responsibility to demonstrate that results from studies that use this data can be generalized to organizations beyond the particular state chosen (assuming this is the researcher's goal).<sup>38</sup> One possible way to document that the organizations in the sample solicit nationally is to

examine whether sample organizations file Form 990 in other states (use Part VI Line 90a on old Form 990 or Part VI, Line 17 on new Form 990).

The third option is to use nonprofits that receive A-133 audits. According to Circular A-133, nonprofits with at least \$500,000 (\$300,000 prior to December 31, 2003) of total annual federal expenditures are required to have a financial statement audit. Because the federal government funds these particular organizations, audited financial statements can be obtained under the U.S. Freedom of Information Act (FOIA). The National Security Archive at [http://www.gwu.edu/~nsarchiv/nsa/the\\_archive.html](http://www.gwu.edu/~nsarchiv/nsa/the_archive.html) of The George Washington University provides step-by-step guidance (Adair and Nielsen 2009) and tips on how to use FOIA to effectively collect information about a nonprofit, such as the organization's audited financial statements. To obtain the reports, researchers can contact the FOIA office under the nonprofit's oversight agency.<sup>39</sup> The oversight agency is the federal awarding agency that provides the predominant amount of direct funding to a recipient. For example, the federal oversight agencies for the *American Cancer Society* are the Departments of Health and Human Services and Defense. To identify the nonprofits' oversight agency use either the oversight agency code or directly search for the agency name under Part III, item 8 on the data collection form on the Federal Audit Clearinghouse website (<http://harvester.census.gov/sac/>). Results from studies that use this approach will be biased toward organizations that receive substantial federal funding. Further, the FOIA request approach can be long and arduous because the response is not timely (i.e., some responses can exceed a year). We discuss the Federal Audit Clearinghouse database next.

### **The Federal Audit Clearinghouse (FAC) Database**

The Federal Audit Clearinghouse builds the FAC database using the data on the data collection forms. Most frequently the FAC Database is used to collect auditor name (Keating et al., 2005; Petrovits et al., 2011); auditor address (Vemeer, 2008); and audit opinion (Keating et al., 2005; Tate, 2009; Petrovits, 2011).<sup>40</sup> The FAC database consists of information about Single Audits of Federal awards as required by Office of Management and Budget (OMB) Circular A-133 (revised June 27, 2003) for organizations that receive federal grants that meet the Single Audit threshold (at the time of this

writing it is \$500,000). The database contains information about auditor name, address (street, city, state, zip code), auditor contact person (the contact person's name, title, phone, fax, and email), audit type (Single Audit versus program specific audit), and auditor opinions for years 1997 - 2010. In addition to financial statement opinions (e.g., unqualified opinions, going concern, etc.), the auditor opinion section of the database includes reportable conditions, material weakness, major noncompliance, questioned costs, findings, and judgments on a nonprofit's low-risk status. The auditor opinions for reportable conditions, material weakness, major noncompliance, questioned costs and findings are given both at the organization level and major program level. In addition to the downloadable electronic database, the FAC also includes a search engine that allows researchers to obtain the scanned data collection forms (Form SF-SAC) filed by nonprofits and signed by the auditors.

Researchers should be aware of the following issues when working with this data. First, the FAC includes only those organizations that receive government grants of a certain size, using audit information from this source only may lead to external validity issues as results may not generalize to the population of nonprofit organizations. Some studies combine FAC with NCCS data (Tate, 2007; Vemeeer, 2008; Petrovits, 2011). Not all organizations in the FAC are in NCCS data files, however. This is because the FAC database includes governmental entities that are required to receive A-133 audit but not required to file Form 990.

Second, for fiscal years ending on or before December 31, 2000, only a single Employer Identification Number (EIN) is reported for each entity's report. For fiscal years ending on or after January 1, 2001, multiple records related to a single EIN may be included in the database (e.g., both the original and amended returns). Fiscal year, EIN, and DB Key (a unique database key when paired with fiscal year field for each row in the database) can be used as a reliable, unique key for each entity.

Third, auditor names are not standardized. For example, Ernst & Young can also appear in the database as E.Y. One way to solve this issue is to use the standardized unique auditor keys found in Audit Analytics. To obtain the auditor keys, we suggest merging Audit Analytics with a NCCS database by the nonprofit's EIN.



Nonprofit auditor information is included in the Non-profit Single Audits database within the Other Independent Audits (OIA) Data Module of Audit Analytics. Auditor data included in Audit Analytics is cleansed FAC data. Hence, the generalizability issue remains when using Audit Analytics. Although the use of Audit Analytics dataset helps to enhance efficiency and accuracy of research, it is costly to subscribe to Audit Analytics. The FAC database is free.

Finally, we have observed instances when the audit opinion on the audit report differs from the opinion indicated in the database. This can happen if the organization incorrectly fills out the data collection form or the FAC has a data entry error, thus, leading to incorrect information in the FAC database.

### CONCLUSION

The objective of this paper is to broaden the awareness of the nonprofit archival data sources and better ensure the quality of nonprofit research. In beginning any new research endeavor researchers should assess their available research budget. This is particularly an issue as many universities may not subscribe to nonprofit databases (and may not have it in their budgets to do so). As described earlier, there are several ways to access the Form 990 data including purchasing data extracts from NCCS and/or Guidestar, obtaining free raw data files from the IRS, and/or hand collecting individual Form 990s from Guidestar, State Charity Official's websites, and/or the individual nonprofit organizations' own website. The tradeoff becomes a question of monetary resources versus time. We leave it up to the researcher to decide the use of specific data source(s) given the research question(s) and the resources available.

We note in our study data issues found with the Form 990 in general and with certain data sources in particular. Researchers should pay particular attention to the sections that describe the data source they have selected. Beyond the Form 990 data, there are several interesting pieces of data available to researchers including the audited financial statements, auditor data, and commercial fundraising reports. We describe where researchers may obtain the data and issues to consider when utilizing the data. Researchers interested in the audited financial statements, and/or auditor information should pay particular attention to these sections.

Finally, we hope researchers, current and prospective, are encouraged by the wealth of data currently available on nonprofit organizations. Since Gordon et al.'s (1999) seminal work on the NCCS data, researchers have collectively published works in leading scholarly journals improving our understanding of the nonprofit sector. Overall, our study updates the collective knowledge on nonprofit data and should provide a basis for the next wave of groundbreaking research in the nonprofit sector.

#### ACKNOWLEDGEMENTS

We thank the National Center for Charitable Statistics (NCCS) for their help with this project. In particular we are grateful to: Tom Pollak, Program Director; Katie L. Roeger, Assistant Program Director; and Jon Durnford, Consultant to NCCS. We also thank Dan Tinkelman, Stefanie Tate, and Sara Garven for their helpful comments. This research was funded, in part, by the McIntire Foundation's PricewaterhouseCoopers Faculty Research Fund.

#### NOTES

1. We loosely define accounting research as any study that uses financial data to address questions that might be classified as accounting. Accounting research can be found in traditional accounting journals, as well as various social science outlets.
2. We make this claim by comparing the number of members in each of the following sections of the American Accounting Association: Government and Nonprofit – 290 members; Auditing – 1,381 members; Financial Accounting and Reporting – 1,345; Information Systems – 429; International Accounting – 711; Management Accounting – 939; and Tax – 578 members.
3. Currently (beginning in tax year 2010), nonprofits with greater than \$50,000 in annual gross receipts are required to file either a Form 990, Form 990 EZ, or Form 990-N. See Form 990 Instructions for list of exclusions.
4. In this paper, we refer to 2008 (and post-2008) Form 990 as the “new form” and pre-2008 Form 990 as the “old form.” We urge scholars to study the new forms. See the Background Paper – *Form 990, Moving from the Old to the New* on the IRS website:

<http://www.irs.gov/Charities-&-Non-Profits/Form-990-Redesign--Background-Documents>.

5. See Smith and Shaver (2009) for a description of changes to Form 990 in addition to the changes described in this paper.
6. These may not represent errors and may indicate defunct organizations. One way to identify if an organization is defunct is to look for it on the Business Master Files (made available by the IRS or NCCS) in subsequent years.
7. A few other papers examine whether organizations intentionally misreport 990 data. Jones and Roberts (2006) provide evidence that nonprofits that engage in combined program and fundraising activities use joint costs to avoid reporting changes in the program ratio. Yetman (2001) and Omer and Yetman (2003) document that nonprofits with taxable activities allocate expenses from their tax-exempt activities to reduce their tax liabilities.
8. An exception is Yetman and Yetman (2013). They consider plausible explanations for fundraising expense to equal zero. Unlike the other studies, the central focus of their study, however, is zero fundraising expenses.
9. If the study requires the sample to include organizations that engage in fundraising activities (such as studies that focus on donor responses to accounting information), then it is more appropriate to exclude observations where fundraising expense equals zero since donors are more likely to respond to accounting information when the organization engages in fundraising activities (such as done in Tinkelman, 1999, 2006).
10. Further, the number of times where total expenses in Parts I and II do not equal decreases each year between 1998 and 2003. For example, the number of times where total expenses in Parts I and II do not equal is 25,130 in 1998 and 8,328 in 2003.
11. Organizations with their gross receipts greater than or equal to \$1 million or total assets greater than or equal to \$2.5 million are required to file the new Form 990 for the 2008 tax year. The thresholds are lowered to \$500,000 and \$1.25 million, respectively for the 2009 tax year, and \$200,000 and \$500,000, respectively for the 2010 tax year.

12. Prior to 2006, line 25 is the total compensation to both current and former executives. It becomes lines 25a (current executives), 25b (former executives), and 25c (disqualified persons) in 2006. Total compensation paid to former executives is not reported separately in the new 990.
13. For tax years 2005, 2006 and 2007, Part V becomes Part V-A for current and Part V-B for former officers, directors, trustees, and key employees.
14. For example, the instructions to the 2005 Form 990 state on page 28 “A key employee is any person having responsibility, powers, or influence similar to those of officers, directors, or trustees. The term includes chief management and administrative officials of an organization (such as an executive director or a chancellor).”
15. For example, the 2005 American Cancer Society 990 reports compensation for each executive director for each region on its affiliated return.
16. Although the Business Master File (BMF) includes information for all active nonprofits, it does not necessarily include all organizations that solicit donations. For example, because the IRS does not require organizations to register separately if they are considered part of a municipality and benefits from municipal government tax exemptions, many organizations affiliated with secondary schools (such as parent teacher organizations) or police or fire departments are not included in the BMF. Keating et. al (2008) document this in the context of reviewing the BMF maintained by the NCCS (described in the next section).
17. The IRS’ large-asset class has evolved over time. For example, prior to 2000 it was \$10 million. See <http://www.irs.gov/uac/SOI-Tax-Stats> - Exempt-Organizations-Study-Data-Sources-and-Limitations for sampling information.
18. Exceptions are Petrovits (2006), Sansing and Yetman (2006), and Yoder et al. (2011) who use 990-PF data. Keating et al. (2008) use non-501(c) (3) data files.
19. Refer to the data dictionaries on the NCCS website, the *Guide to Using NCCS Data* (2006), and Gordon et al. (1999) for detailed descriptions of each of the databases. The Digitized Database

discussed below was made available after 1999 and was thus not described in Gordon et al. (1999).

20. The FIPS code is the Federal Information Processing Standards code issued by the National Institute of Standards and Technology. It is a 5-digit code for the organization's county. The first two digits of the county FIPS code identify the state and the last three identify the county. The NAICS is the North American Industry Classification System and is the standard used by federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data.
21. The NCCS gathers this particular data from several different sources; thus, 'estimate' refers to the fact that the NCCS does not guarantee the amount relates to the year indicated in the Core file. Refer to the following link for more details  
<http://nccsdataweb.urban.org/knowledgebase/detail.php?linkID=1770&category=84&xrefID=6566&close=0>.
22. The NCCS Core Supplement (aka: Super Core) files contain Form 990 fields that are not found on the Core files described. The "Super Core" file is an archive of Form 990 return records that have been keypunched for a variety of NCCS projects beginning in 2007. This file is cumulative and is updated as new keypunched records become available. Old tax year returns are not replaced when new ones are added, so the number of records increases with each update. Unlike many NCCS files that contain one record for each unique organization EIN, Super Core may contain more than one tax year return for each unique organization EIN, and may even contain more than one return within the same tax year for legitimate reasons (e.g. amended returns, short fiscal period, by mistake, etc.).
23. For details see <http://www.irs.gov/taxstats/article/Oid=120304,00.html>.
24. The Core files are based on the IRS' Return Transaction Files (RTF), which the IRS uses primarily for regulatory purposes and not research purposes. As a result, the RTF data entry process is geared toward speed and data entry errors occur. While the IRS produces the RTF annually, the NCCS typically only distributes the NCCS Core file, based upon the RTF, once it is reviewed for

quality. See *Guide to Using the NCCS Data* (page 7) for the data cleaning process.

25. The NCCS corrected large-scale errors in financial variables (typically errors over \$1,000,000 and a difference of more than 25 percent from expected value, i.e. value of components versus totals) See *Guide to Using the NCCS Data* (page 8) for details.
26. Petrovits et al. (2011) take this approach and make a small number of corrections. They note that the errors primarily relate to the units reported (i.e., the file lists \$5 instead of \$5 million).
27. Data files may include more than one 990 for a given year if an amended return is filed or a fiscal year change. Guidestar can be used to determine the most recent year by noting the 'date received' stamp on the first page of the 990.
28. The Digitized Data does not have a file year issue because years are combined into one data file. The database is divided by Form 990 section and not by year. Further, it uses a fiscal year field, defined in terms of the ending year.
29. <http://nccsdataweb.urban.org/knowledgebase/detail.php?linkID=953&category=84&xrefID=4809>.
30. An exception is Yetman and Yetman (2013). They use the common codes to identify organizations whose primary focus is fundraising.
31. Tinkelman (2004) notes this issue and indicates that the lack of clarity in the classification process affects the usefulness of research to policy-makers. Neely (2011) documents instances when a nonprofit's NTEE code changes over a three year period (2003 to 2005).
32. There are likely plausible explanations for an entity to report zero executive compensation; but we do not investigate these reasons in this paper.
33. We arrive at 15 percent by first determining that 30 percent of the nonprofits in the digitized database report compensation data related to an executive director (or someone similarly titled) on Schedule A. We then chose a random sample of 100 nonprofits that disclose executive director compensation data on Schedule A and find that: fifty-one do not disclose the compensation data on

Part V; five disclose the name of the executive on Part V but not the related compensation; and thirty-eight disclose the compensation data on both Schedule A and Part V. We do not verify six because the 990 data are incomplete.

34. Refer to Gordon et al. (2009) for a detailed description of Charity Navigator, BBB, and AIP ratings agencies.
35. To obtain names and addresses for a large number of executives, the Guidestar service is preferred to the Business Master File maintained by the NCCS because although the BMF may include contact information it does not necessarily include the name of the executive.
36. About half the states require audited reports from NPOs that have revenues above a minimum threshold. The revenues thresholds vary from state to state. For example, Connecticut requires audits when gross revenues exceed \$200,000 while New Hampshire and Massachusetts demand audits when gross revenues are above \$500,000.
37. Krishnan and Schauer (2000) visit United Ways in Southeastern PA and Southern NJ to obtain audited financial reports.
38. Many of the audit reports for organizations on these state websites likely pertain to national organizations. For example, the audit report for the *American Diabetes Association* is filed in Massachusetts but the organization is headquartered in Virginia.
39. According to Circular A-133, if the amount of annual federal award expenditures is more than \$25 million (\$50 million for fiscal years ending after December 31, 2003) and includes direct awards, a nonprofit has a cognizant agency. If the amount of annual federal award expenditures is \$25 million (\$50 million for fiscal years ending after December 31, 2003) or less, or includes no direct awards, a nonprofit has an oversight agency.
40. An exception is Kitching (2009). She hand-collects auditor names from Form 990.

## REFERENCES

- Adair, K., & Nielsen, C. (2009). *Effective FOIA Requesting for Everyone*. [Online]. Available at [www.gwu.edu/~nsarchiv/nsa/foia/foia\\_guide.html](http://www.gwu.edu/~nsarchiv/nsa/foia/foia_guide.html). [Retrieved June 14, 2013].
- Baber, W. R., Daniel, P. L., & Roberts, A. A. (2002). "Compensation to Managers of Charitable Organizations: An Empirical Study of the Role of Accounting Measures of Program Activities." *The Accounting Review*, 77 (3): 679-693.
- Brickley, J.A., & Van Horn, R. L. (2002). "Managerial Incentives in Nonprofit Organizations." *Journal of Law and Economics*, 45 (1): 227-250.
- Carroll, T., Hughes, P., & Luksetich, W. (2005). "Managers of Nonprofit Organizations Are Rewarded for Performance." *Nonprofit Management and Leadership*, 16(1): 19-41.
- Core, J. E., Guay, W. R., & Verdi, R. S. (2006). "Agency Problems Of Excess Endowment Holdings In Not-For-Profit Firms." *Journal of Accounting and Economics*, 41: 307-333.
- Derrick, P.L. (2005). *Accounting for Promises: The Impact of SFAS No. 116 on Charities*. Ph.D Dissertation. Washington, DC: The George Washington University.
- Eldenbug, L., & Krishnan, R. (2003). "Public versus Private Governance: A Study of Incentives and Operational Performance." *Journal of Accounting & Economics*, 35 (3): 377-404.
- Fischer, M., Gordon, T.P., & Kraut, M. (2002). "An Examination of Differences between Financial Information Provided in IRS Form 990 and Audited Financial Statement Data of U.S. Private Colleges and Universities." *Academy of Accounting and Financial Studies Journal*, 6 (1): 87-104.
- Fremont-Smith, M. (2005). *Governing Nonprofit Organizations: Federal And State Law and Regulation*. Cambridge, MA: Harvard University Press.
- Froelich, K. A., & Knoepfle, T.W. (1996). "Internal Revenue Service 990 Data: Fact or Fiction?" *Nonprofit and Voluntary Sector Quarterly*, 25 (1): 40-52.



- Froelich, K. A., Knoepfle, T.W., & Pollak, T. (2000). "Financial Measures in Nonprofit Organization Research: Comparing IRS 990 Return and Audited Financial Statement Data." *Nonprofit and Voluntary Sector Quarterly*, 29 (2): 232-254.
- Frumkin, P. and Kim, M. T. (2001). "Strategic Positioning and the Financing of Nonprofit Organizations: Is Efficiency Rewarded in The Contributions Marketplace?" *Public Administration Review*, 61 (3): 266-275.
- Gantz, M. (1999). "Who Do You Trust? Comparing Data on Skilled-Nursing Facilities from the Internal Revenue Service and Health Care Financing Administration." *Nonprofit and Voluntary Sector Quarterly*, 28(4): 476-490.
- Gordon, T.P., Greenlee, J., & Nitterhouse, D. (1999). "Tax-Exempt Organization Financial Data: Availability and Limitations." *Accounting Horizons*, 13: 113-128.
- Gordon, T.P., Khumawala, S.B., Kraut, M., & Meade, J. (2007). "The Quality and Reliability of Form 990 Data: Are Users Being Misled?" *Academy of Accounting and Financial Studies Journal*, 11: 27-49.
- Gordon, T.P., Knock, C.L., & Neely, D.G. (2009). "The Role of Rating Agencies in the Market For Charitable Contributions: An Empirical Test." *Journal of Accounting and Public Policy*, 28 (6): 469-484.
- Gordon, T.P., Khumawala, S.B., Kraut, M., & Neely, D.G. (2010). "Five Dimensions of Effectiveness for Nonprofit Annual Reports." *Nonprofit Management and Leadership*, 21 (2): 209-228.
- Greenlee, J., & Gordon, T.P. (1998). "The Impact of Professional Solicitors on Fundraising in Charitable Organizations." *Nonprofit and Voluntary Sector Quarterly*, 27 (3): 277-299.
- Grein, B.M., & Tate, S.L. (2011). "Monitoring By Auditors: The Case of Public Housing Authorities." *The Accounting Review*, 86 (4): 1289-1319.
- Gronbjerg, K.A. (2002). "Evaluating Nonprofit Databases." *The American Behavioral Scientist*, 45 (11): 1741- 1777.
- Hallock, K. F. (2002). "Managerial Pay and Governance in American Nonprofits." *Industrial Relations: A Journal of Economy and Society*, 41 (3): 377-406.

- IRS. (2008). *Background Paper – Form 990, Moving from the Old to the New*. [Online] Available at: [http://www.irs.gov/pub/irs-tege/moving\\_from\\_old\\_to\\_new.pdf](http://www.irs.gov/pub/irs-tege/moving_from_old_to_new.pdf). (Retrieved June 14, 2013).
- Jacobs, F. A., & Marudas, N. P. (2009). "The Combined Effect Of Donation Price and Administrative Inefficiency on Donations to US Nonprofit Organizations." *Financial Accountability & Management*, 25 (1): 33-53.
- Jones, C.L., & Roberts, A.A. (2006). "Management of Financial Information in Charitable Organizations: The Case of Joint-Cost Allocations." *The Accounting Review*, 81 (1): 159-178.
- Keating, L., Fischer, M., Gordon, T.P., & Greenlee J. (2005). "The Single Audit Act: How Compliant Are Nonprofit Organizations?" *Journal of Public Budgeting, Accounting, & Financial Management*, 17(3), 285-309.
- Keating, L., Parsons, L.M., & Roberts, A. A. (2008). "Misreporting Fundraising: How Do Nonprofit Organizations Account for Telemarketing Campaigns?" *The Accounting Review*, 83 (2): 417-446.
- Kitching, K. (2009). "Audit Value and Charitable Organizations." *Journal of Accounting and Public Policy*, 28 (6): 510-524.
- Krishnan, J., & Schauer, P.C. (2000). "The Differentiation of Quality among Auditors: Evidence from the Not-For-Profit Sector." *Auditing: A Journal of Practice & Theory*, 19 (2): 9-26.
- Krishnan, R., Yetman, M.H., & Yetman, R. J. (2006). "Expense Misreporting in Nonprofit Organizations." *The Accounting Review*, 81 (2): 399-420.
- Marudas, N. P. (2004). "Effects of Nonprofit Organization Wealth and Efficiency on Private Donations to Large Nonprofit Organizations." *Research in Government and Nonprofit Accounting*, 11: 71-91.
- National Center Charitable Statistics (NCCS). (2006). *Guide to Using NCCS Data*. [Online]. Available at <http://nccsdataweb.urban.org/kbfiles/742/NCCS-data-guide-2006c.pdf>. [Retrieved June 14, 2013].
- Neely, D.G. (2011). "The Impact of Regulation on the U.S. Nonprofit Sector: Initial Evidence from the Nonprofit Integrity Act of 2004." *Accounting Horizons*, 25 (1): 107-125.

- Okten, C., & Weisbrod, B. A. (2000). "Determinants of Donations in Private Nonprofit Markets." *Journal of Public Economics*, 75 (2):255.
- Omer, T.C., & Yetman, R.J. (2003). "Near Zero Taxable Income Reporting by Nonprofit Organizations." *Journal of the American Taxation Association*, 25 (2): 19-34.
- Petrovits, C. (2006). "Corporate-Sponsored Foundations and Earnings Management." *Journal of Accounting & Economics*, 41: 335-362.
- Petrovits, C., Shakespeare, C., & Shih, A. (2011). "The Causes and Consequences of Internal Control Problems in Nonprofit Organizations." *The Accounting Review*, 86 (1): 325-357.
- Roberts, A. (2005). "Implications of Joint Cost Standards for Charity Reporting." *Accounting Horizons*, 19: 11-24.
- Sansing, R., & Yetman, R.J. (2006). "Governing Private Foundations Using the Tax Law." *Journal of Accounting & Economics*, 41: 363-384.
- Smith, P. & Shaver, D. (2009). "Transparency, Compliance, and Filing Burden - Principles for The Revised form 990." *Journal of Legal Tax Research*, 7: 133-151.
- Tate, S.L. (2007). "Auditor Change and Auditor Choice in Nonprofit Organizations." *Auditing: A Journal of Practice & Theory*, 26 (1): 47-70.
- Tate, S.L. (2009). "Differences in Non-Profit Organization Financial Statement and Compliance Audit Assessments between Big 5 and Non-Big 5 Auditors." *Research in Governmental and Nonprofit Accounting*, 12: 3-33.
- The Urban Institute. (2003). *Form 990 Accounting and Reporting Issues for Discussion*. [Online] Available at [www.qual990.org/990issues-03.pdf](http://www.qual990.org/990issues-03.pdf). (Retrieved June 14, 2013).
- Thornton, J.P., & Belski, W.H. (2010). "Financial Reporting Quality and Price Competition among Nonprofit Firms." *Applied Economics*, 42: 2699-2713.
- Tinkelman, D. (1997). "Differences in Sensitivity of Financial Statement Users to Joint Cost Allocations: The Case of Nonprofit

- Organizations." *Journal of Accounting, Auditing, and Finance*, 13 (4): 377-393.
- Tinkelman, D. (1999). "Factors Affecting the Relation Between Donations to Not-For-Profit Organizations and an Efficiency Ratio." *Research in Government and Nonprofit Accounting*, 10: 135-161.
- Tinkelman, D. (2004). "Using Nonprofit Organization-Level Financial Data to Infer Managers' Fund-Raising Strategies." *Journal of Public Economics*, 88: 2181-2192.
- Tinkelman, D. (2006). "The Decision-Usefulness of Nonprofit Fundraising Ratios: Some Contrary Evidence." *Journal of Accounting, Auditing & Finance*, 21 (4): 441-462
- Tinkelman, D., & Mankaney, K. (2007). "When Is Administrative Efficiency Associated with Charitable Donations?" *Nonprofit and Voluntary Sector Quarterly*, 36 (1): 41-64.
- Tinkelman, D., & Neely, D.G. (2011). "Some Econometric Issues in Studying Nonprofit Revenue Interactions Using NCCS Data." *Nonprofit and Voluntary Sector Quarterly*, 40 (4): 751-761.
- Trussel, J.M. (2003). "Assessing Potential Accounting Manipulation: The Financial Characteristics of Charitable Organizations with Higher Than Expected Program-Spending Ratios." *Nonprofit and Voluntary Sector Quarterly*, 32 (4): 616-634.
- Vemeer, T.E. (2008). "Market for Former Andersen Clients: Evidence from Government and Non-Profit Sectors." *Journal of Accounting and Public Policy*, 27 (5): 394-408.
- Vemeer, T.E., Raghunandan, K., & Forgione, D.A. (2009). "Audit Fees at U.S. Non-Profit Organizations." *Auditing: A Journal of Practice & Theory*, 28 (2): 289-303.
- Wing, K., & Hager, M.A. (2004). *The Quality of Financial Reporting by Nonprofits: Finding and Nonprofit Overhead Cost Project (Brief No. 4)*. [Online]. Available at <http://www.urban.org/publications/311045.html>. [Retrieved June 14, 2013]
- Wing, K., Hager, M.A., Rooney, P., & Pollak, T. (2004). *Lessons for Boards from the Nonprofit Overhead Cost Project. (Guide No. 2)*. [Online] Available at: <http://www.urban.org/url.cfm?ID=411119>. [Retrieved June 14, 2013]

- Wing, K., Pollak, T.H., & Blackwood, A. (2008). *The Nonprofit Almanac 2008*. Baltimore, MD: Urban Institute Press.
- Yetman, R.J. (2001). "Tax-Motivated Expense Allocations By Nonprofit Organizations." *The Accounting Review*, 76 (3): 297-311.
- Yetman, M., Yetman, R.J., & Badertscher, B. (2009). "Calibrating the Reliability of Publicly Available Nonprofit Taxable Activity Disclosures: Comparing IRS 990 and IRS 990-T Data." *Nonprofit and Voluntary Sector Quarterly*, 38 (1): 95-116.
- Yetman, M., & Yetman, R.J. (2013). "Do Donors Discount Low Quality Accounting Information?" *Accounting Review*, 88 (3): 1041-1067.
- Yoder, T.R., Addy, N., & McAllister, B.P. (2011). "Tax-Motivated Increases in Qualifying Distributions by Private Foundations." *Journal of the American Taxation Association*, 33 (1): 79-108.