Distorted measures of employment in charitable organizations: Causes, impact, and remedies

Martin H. David

Associate Scholar, Urban Institute, Washington DC

Emeritus professor, University of Wisconsin – Madison

email: david@ssc.wisc.edu telephone: 608.238.2181

2603 Middleton Beach Rd., Middleton WI 53562

22 August 2007

Abstract

Keywords: Charitable organizations, employment, administrative records

Public sector failures lead to a large understatement of employment in charitable organizations. Multiple forces lead to this understatement. Partitioning private business into charities, other exempt organizations, and for profit business has a low priority in Federal Statistical Agencies. Regulatory failures in IRS oversight of exempt organizations compromise available statistics – the count of active organizations, data on employment, coverage of available employment reports, and consistency in its reporting. The incentive for IRS to regulate exempt entities is negative as the activity does not generate net revenue. Because exempt organizations constitute a small part of private businesses, publication of estimates for their establishments is limited by the imperative not to disclose proprietary information. Finally, regulation of burden in completing government forms leads to peculiar censoring of data within the population of exempt entities.

This analysis demonstrates that existing published estimates of employment in charitable organizations is understated. We link IRS information returns to the BLS/QCEW. A substantial proportion of employers can not be matched. Employment on IRS returns contains substantial nonresponse. Imputation of QCEW employment to matched organizations and augmenting the available census of IRS returns with employment in exempt organizations that are not covered produces aggregates that are substantially larger than the published *Economic Census* for 2002.

Understated employment can be overcome by a combination of more sophisticated imputation of information returns and matching of IRS information to records of payroll tax submissions, IRS/Form 941.

1 Importance of nonprofit employment

Charitable organizations constitute about 5% of the corporate business sector. They are private corporations and associations exempted from taxation by the Federal government under *Internal Revenue Code* (IRC) §501(c)(3). State and local codes often reduce sales tax liability and property taxes for charitable organizations (Brody 2002). Gifts to exempt *charitable organizations* substantially reduce tax burdens of individual donors.[1]

Charitable organizations operate in a broad range of industries -- from the arts, advocacy, and social assistance to education, training, and hospital services to research and international services. Labor is the largest input to the sector. Wages paid are likely to be less than in for-profit companies.[2]

Good measures of the number of employees in charitable organizations, classified by industry subsectors, are a first step to understanding the value of these organizations to the economy. This paper demonstrates that estimates of employment can be more timely and more comprehensive than understated estimates available in Census publications.

2 Sources of employment measures

2.1 Published estimates of exempt sector employment

The *Economic Census* publishes employment information on private business entities that are exempted from corporation taxes every five years. Exempt includes charitable organizations as well as neighborhood associations, clubs, labor unions, credit unions, and cooperatives. Three NAICS sectors are classified by exempt status of the organization (*Educational services*, 61; *Health Care and Social Assistance*, 62; and *Arts, Entertainment, and Recreation*, 71). Sector 61 excludes elementary and secondary schools and colleges and universities. One additional

[1] Tax expenditures due to deductions for education, health, and social services ate estimated at more than $55 billion in the *2008 Budget of the United States* government (Table 19-3).

[2] A plausible hypothesis is that workers in nonprofit organizations derive more satisfaction from their employment than workers in for-profit employment; they are willing to work at wages less than the wage paid in comparable for-profit positions [Helliwell and Huang 2005].

subsector, *Religious, Grantmaking, Civic, Professional, and similar Organizations* (813), is dominated by exempt organizations. Many are charitable organizations; some are not.[3]

Employment reported for those (sub-) sectors totals 10.4 million (Table 1). That statistic is not only incomplete, it is timely only once in five years. Further it mixes employment of *charitable organizations* with employment of other exempt organizations. More extensive tabulation of charitable organizations is probably foreclosed by the disclosure review procedures used by the *Economic Census*.[4]

2.2 Unpublished microdata containing employment reports

Regulations promulgated for IRC 501(c) organizations stipulate that most must file annual information returns.[5] Organizations averaging revenue more than $25,000 must file; religious congregations are excused from filing information returns. However, religious organizations contracting to provide services funded by the Federal government also file information returns. Returns of organizations with more than $100,000 of revenue elicit employment in the week of March 12. Thus, employment of larger organizations can be estimated. (Charitable organizations can be distinguished from other exempted organizations.)

The IRS count of employee is incomplete. What proportion of sector employment is included? How well is employment reported in different industries? Partial answers to these questions are provided in this report.[6]

Microdata from these returns are accessible through the Urban Institute *Dataweb*. A census of all Form 990 and Form 990-EZ filed by charitable organizations is available for 1999-2003 (*GuideStar-NCCS National Nonprofit Research Database*, ver. 1 [1999-2003], cited as NCCS:<year>). Form 990 elicits the number of employees; Form 990-EZ does not.

---

[3] Most religious congregations included under this NAICS sector do not appear in any other official statistics of the US.

[4] That is ironic, as exempt organizations are required to file information returns for public scrutiny.

[5] The returns refer to fiscal years. Fiscal years ending after 1 December 2002 and before 1 December 2003 will be filed on the Form 990 for the year 2002.

[6] Answers to this question may encourage the Statistics of Income Division of IRS to publish employment estimates in its annual statistical report on nonprofit organizations (Arnsberger 2006).

3 Linking public IRS records to other employment data

Matching any two datasets entails five steps: a) Finding a common identifier, b) determining the multiplicity of the matches, c) assessing the extent of matching errors, d) correcting the match to reduce bias from failed matches, and e) removing mismatches from the analysis files. We follow these steps in linking IRS information returns below.

3.1 Gains from linking IRS nonprofit information returns to other sources

Employment is reported on three administrative records: IRS Form 941, the BLS *Quarterly Census of Employment and Wages (QCEW)*, and IRS W-2. The QCEW is derived from employer reports to state Unemployment Compensation (UC) agencies. Form 941 and QCEW both contain reports of employees on payroll in a particular month. The March reference month for those two series is identical to the reference month for employment elicited on Form 990.

The identifier common to Form 941, *QCEW*, Forms 990/990-EZ, and the IRS *Registry* is the Federal Employer Identifying Number, or ein. Multiple establishments in the *QCEW* match some IRS ein's. In most cases only one establishment matches. The IRS files contain one record for each ein.

Errors in ein's emanate from the employer who may enter an incorrect ein or omit an ein, and from the state and Federal administrative agencies that process the ein. An omitted ein leads to a match failure. An incorrect ein leads to a match failure or a mismatched pair of records.

The candidates to link to nonprofit returns vary in coverage and reporting period:

(a) Form 941, filed quarterly, includes most employees whose wages have been withheld for accruing income taxes or Social Security payroll taxes. Employers that owe less than $2500 annually for withholding and payroll taxes do not file Form 941. (They file Form 944 at the end of the year.) Thus employers of few, or part-time, employees may not file Form 941.

(b) QCEW covers many, but not all, Form 941 employees. (It also includes a few of the employers who file Form 944.) However, QCEW excludes nonprofit employers with less than 4 employees in some states. QCEW excludes some employees – part-time workers, students, and interns – in some states. One or both exclusions apply in about 30 states.

QCEW 'breaks-out' organization employment into employment at worksites dispersed over different counties. For larger organizations multiple worksites in different states can be identified. As services and products of exempt organizations are often tied to local sites, this

geographic information is an extremely valuable addition to the Form 990/990-EZ report (Salamon-Sokolowski 2005].

(c) IRS W-2 reports total employment for a 12-month reference period. That is, every employee receiving wages in the prior calendar year will be counted. Multiple job-holders and employees switching jobs during the *year* will be counted twice. Multiple job-holders and some employees switching jobs during a *month* are counted twice on the QCEW and Form 941. While W-2 and Form 941 contain payroll information for different reporting periods, combining the four Forms 941 filed during a calendar year produces payroll comparable to Form W-2. Monthly employment for employers that do not file Form 941 might be estimable from the available annual information.

In conclusion, a match of Forms 990/990-EZ to IRS Form 941 for March will yield a larger count of nonprofit employees than the QCEW. The QCEW match gives insight to the distribution of multi-establishment employment across worksites.[7]

3.2 Losses from linking nonprofit information returns to other sources

Any errors in ein's being matched reduce the quality of matched data. Errors in identifying numbers arise from reporting errors by filers, from changes in the legal organization that entail a new identifier for continuing (perhaps expanded) activities, and from failures to maintain correct identifiers in the two record systems.

Errors in identifiers have two consequences: failed matches and mismatches. An incorrect identifier in the IRS file leads to a failed match when no corresponding identifier exists in the second data source, and vice versa. Mismatches result when the incorrect IRS identifier matches the identifier of a disparate organization in the second data source, and vice versa. The proportion of nonprofit organizations is small relative to the populations in the QCEW and Form 941 universes, so mismatches are most likely to relate information about taxable organizations to the exempt organization. Failed matches reduce the coverage of statistics based on both records. Mismatches badly distort statistics. Both levels of the matched data and correlates of those data are distorted (Scheuren-Winkler 1997). For example, when employment is imputed to Form 990-

[7] As we describe later, non-filing employers can be identified from the IRS *Registry* of exempt organizations, an extract from the IRS *Business Master File*.

EZ from Form 941, any difference in the mean employment of exempt organizations and the mean employment of mismatched employers will bias the imputed mean.

4 Matching QCEW and IRS information on exempt organizations

Our investigation matches IRS public information on nonprofits to the QCEW. Most estimates refer to matches and employment for March, 2003.

4.1 Joining QCEW and IRS information

Matching Forms 990/990-EZ, the *Registry*, and the *QCEW* yields eight outcomes. Outcomes that link *QCEW* to IRS information are indicated by *m* in the text table below. Unmatched records are indicated by "*". Outcomes 1-5 are useful. Outcome 1 gives the most complete information since all three data files match. (Outcome 3 occurs primarily as organizations whose application for exempt status is pending, file Form 990/990-EZ.) Outcomes 2 and 4 do not match *QCEW*, but information on Forms 990/990-EZ is informative.

	Outcomes							
Record system	1	2	3	4	5	6	7	8
Forms 990/990-EZ	m	*	m	*				
Registry	m	*			m	*		
QCEW	m		m		m		*	

Outcomes 6-8 are uninformative. The function of the *Registry* is to identify active exempt organizations. Unfortunately, organizations that are inactive or defunct sometimes linger for a substantial time before they are identified and removed. Thus unmatched *Registry* records (outcome 6) overstate the number of active unmatched organizations. Unmatched *QCEW* records (outcome 7) include nonprofit entities, but they can not be identified reliably.[8] Some nonprofits are not UC liable and do not file information returns (outcome 8). Neither IRS nor *QCEW* records afford insight into this group.

[8] Employers are asked to identify tax-exempt status, but the information is not well-reported.

Outcomes 1-4 include all exempt organizations filing Forms 990/990-EZ and processed into the NCCS:1999-2003.[9] They constitute a census of operating charitable organizations, covering all states. Unmatched Forms 990/990-EZ (Outcome 2 and 4) provide information about organizations that do not fall in the QCEW universe. Understanding the unmatched cases is a lever that allows us to estimate employment outside the UC system and the proportion of organizations that operate without paid employees, i.e. nonemployers.

Outcome 5 matches the *Registry* to three types of organizations: exempt organizations other than charities (exempted under subsections other than 501(c)(3)), private foundations, and operating charities failing to file timely returns.

Each matched Form 990 reveals whether the organization failed to report employment on line 90(b), a false negative (FN). Each matched Form 990-EZ and *Registry* record reveals the number of employees covered by the UC system at each worksite. No matches provide any information on employees excluded from UC coverage.

4.2 Assessing matching errors

4.2.1 Failed and invalid matches

Some failed matches can be detected by *invalid ein's*. Those ein's have less than the required nine characters or they have ciphers that indicate the ein is unknown. We scanned both the *Registry* and the *QCEW* for invalid ein's. The *Registry* contained less than 500 or 3/10,000 invalid ein's. Though the *Registry* contains ein errors, we regard it as a "gold standard". Less than 90 failed matches arise when it matches a population of less than 300,000 organizations.

The *QCEW* contained an average of 1.7% invalid ein's in the years 1999-2003. These invalid ein's were not randomly distributed across the 35 million records that we scanned. The probability of invalid ein's was higher in establishments with few employees than elsewhere. The probability varied substantially over the 49 jurisdictions (46 states, The District of

[9] Operating charities and public foundations exempted under 501(c)(3) are included. Coverage is limited by cut-off of data processing before late returns are filed, long after they are due to the IRS.

Information from Forms 990 for other organizations is available from *NCCS Core File* [year] and *NCCS Private Foundation File* [year] at http://nccsdataweb.urban.org.

Columbia, Puerto Rico, and the Virgin Islands) in our universe. Three states had probabilities over 0.03; six had probabilities less than 0.005.

We created weights for the *QCEW* that increase counts of establishments and the number of employees. Weighting offsets downward bias of unweighted counts (David 2007).[10] Weights on the QCEW were transferred to all matching IRS records. They are applied to Form 990-EZ and *Registry* matches in the estimates below.

Additional match failures occur because valid ein numbers are corrupted in the process of filing and transmitting the tax-related reports that are the basis for the *QCEW*. Digits can be transposed, duplicated, or erroneously entered. We do not know what order of magnitude to assign to this problem.

4.2.2 Removing mismatches -- Forms 990/990-EZ

Before appropriate estimates could be made, matches were scanned for evidence of mismatches. The legal name of the organization and its industry class were critical information for detecting mismatches. Forms 990/990-EZ and the *QCEW* use different industry classifications, the *National Taxonomy of Exempt Entities (NTEE)* and NAICS respectively. These classifications are relatively similar at the sector level (Appendix A). In examining matched records we discovered several pathologies. Some parent-teacher organizations were associated with large numbers of establishments in banks. Some private organizations awarding fellowships and scholarships were associated with employee counts for state university systems. And some private nonprofit entities were associated with school systems or governments. In many of these cases the name on the Form 990 was substantially different than the name on the *QCEW*. The pairing of large numbers of establishments to the nonprofit organization was also suggestive of mismatching. Hundreds of establishments exist for only a few of nonprofit charitable organizations. Because Forms 990/990-EZ contain more information than the *Registry*, the procedure for removing mismatches of Forms 990/990-EZ differs from the procedure applied where the *Registry* matches *QCEW* and no Form 990 exists.

[10] Salamon-Sokoloweski (2005) did not weight their estimates from matching the *Registry* to *QCEW* for the selective character of invalid ein's.

We test the relationship between payroll reported to the *QCEW* and nonprofit expenditures on Form 990/990-EZ. If first-quarter establishment payroll exceeds 20 percent of total expenditures for the organization, the establishment-organization link is suspect. All links between establishments and organization with *any* suspect links to the QCEW were severed. The logic for this procedure incorporates two assumptions: A) Most organizations with a *QCEW* record for the first quarter of the year were in operation for a whole fiscal year. B) Annual compensation for the nonprofit is at least 20 percent less than total expenditures for the fiscal year. Prorating annual compensation levels to the first quarter, we expect that one-quarter of eighty percent of annual expenditures is a reasonable upper bound for first quarter compensation.

The test identified 2,564 organizations as suspect in 2003. The total of matched Forms 990/990-EZ was 85,852 prior to testing for suspect matches. Failing the test caused 3.0 percent of tentatively matched organizations to be recoded as unmatched. The suspect matches were dominated by links between tiny organizations and entities classified as NAICS 522 (credit intermediation) or NAICS 5412 (Accounting, tax preparation, bookkeeping, and payroll services). Three-quarters of the establishments involved with suspect links were tied to entities in those NAICS classifications.

Over the five years, 1999-2003, an average of more than 8%, or 11,000, *establishments* matched to Form 990/990-EZ are suspect. This average is low, as information on total expenses was missing in 4% of the 1999 Form 990's and 7% of the 2003 Form 990's. Mismatching associated with organizations filing Form 990-EZ is astronomically larger than for organizations filing Form 990. Figure 1 presents the rate of mismatched establishments separately for Form 990 and Form 990-EZ. Over 50% of establishments linked to Form 990-EZ proved to be suspect in years 2000-2002 where expense data were almost universally available.

Following the removal of suspect matches to the *Registry* (next section), we re-examined Forms 990/990-EZ matches to determine whether any NAICS 52 (Finance and Insurance) organizations passed the expense test. 15 organizations were identified in the sector and delinked from QCEW. The count of mismatched Forms 990/990-EZ increased from 2,564 to 2,579.

4.2.3 Removing mismatches – Registry matches, no Forms 990/990-EZ

QCEW matches to the *Registry* contain no information on organization expenses.[11] We investigated two classes of matches: Organizations that are charitable and exempted under section 501(c)(3) of the *IRC*, and organizations whose 501(c) subsection was unknown. The first class is more extensive than the operating charities for which we have Forms 990/990-EZ. Private foundations and trusts are included. The 3,500 organizations with subsection unknown are likely to include a majority with 501(c)(3) activity.[12]

Registry-QWEW matches reveal industry and name of the organization on both records. NAICS signals many mismatches. Organizations matched to NAICS 52 (finance and insurance) appear to be suspect, as they included multiple establishment links of banks to parent-teacher organizations, paralleling the most egregious mismatches identified among the Forms 990/990-EZ. We excluded all public sector, business associations, and labor unions (NAICS 92, 81391, 81393) as they are not generally §501(c)(3). We excluded broader classes than with Forms 990 because organizations with unknown subsection could be 501(c)(4) or (9), not relevant to our analysis of 501(c)(3). A total of 1210 organizations with matches to the *Registry* were declared suspect, representing 5.1% of the 23,878 tentatively matched organizations.

Taken together the delinked organizations are 3.5% of tentative matches. Estimates of employment that include identified mismatches would be wildly overstated.

Lastly, in cases where the subsection was unknown, we surmise that some organizations are not charitable organizations. Organizations linked to NAICS subsectors including the public sector, business associations, and labor unions are unlikely to be operating charitable organizations. Organizations in all of these sectors were deemed *not 501(c)(3)*, and excluded from the *501(c)(3)* universe. Table 2 describes the division of subsection "NA" between those tabulated with 501(c)(3) and those excluded. It also shows that less than 500 501(c)(3) were reallocated out of that class. The editing procedures described produced an *enhanced match* that we discuss below.

---

[11] Only matches to the *Registry* where no Form 990 information is in the Forms 990 database are considered here. Such matches include organizations exempted under many subsections other than 501(c)(3).

[12] 501(c)(3) is the dominant exempt group sampled by the IRS. Arnsberger (2006).

4.3 Structure of matched and unmatched data

4.3.1 Available data

Matched data bring together information defined by different measures and collected from different universes. Important conceptual differences are displayed below.

Attribute	Form 990	Form 990-EZ	QCEW
Employment	Elicited from all	Not collected	Employment collected. Excludes: Part-time workers of nonprofits in all but 19 states, and students working for school, student nurses. Interns (some states)
Compensation	Detail on wages, benefits and payroll taxes	All employment related payments	Payroll
Industry classification	NTEE	NTEE	NAICS
Universe	Exempted entities	Exempted entities	Employers liable to pay UC benefits
Minimum threshold for inclusion	3-yr. avg. revenue >$100,000	$25,000 < 3-yr. avg. revenue <$100,000	1+ employee working more than 20 weeks in a calendar year, US standard
Excluded states (this investigation)	None	None	MA, MI, NY, WY
Excluded entities	Most religious	Most religious	Employers with 1-3 workers (29 states), some religious organizations, small agriculture employers, some local governments
Periodicity	Fiscal year	Fiscal year	Calendar quarters

The measure of employment is critical to our investigation. Form 990 elicits March 12 employment; Form 990-EZ does not. All workers on the payroll should be counted. *QCEW* does

not count part-time employees in many states; students at work in their schools, student nurses in hospitals, and interns are excluded in some states.

Industry is classified by the *National Taxonomy of Exempt Entities* (NTEE) on Forms 990/990-EZ and the *Registry*. Industry is classified by the NAICS on the QCEW.

Average revenue over the last three years determines which IRS form is used. A few organizations file returns because they are required to do so by Federal contracting rules, or because they have provisional permission to operate as a charitable organization or association.

4.3.2 Censoring pertinent to the QCEW universe

Exclusion of nonprofit *employers* with 1-3 workers in 29 states, and part-time *employees* of nonprofits in 33 states (largely, but not the same as the previous states) drastically reduces coverage of the *QCEW* for small organizations.

Exclusion of states. Four states (MA, MI, NY, WY) did not release their information for this study. Those states provide the BLS with QCEW information for all periodic estimates in the BLS publication program. BLS could tabulate nonprofit UC employees directly for the entire US universe from its QCEW data. Geographic censoring of QCEW records affects our partition of organizations into matched and unmatched. Clearly, matches are precluded whenever an organization operates exclusively in the four states that declined to participate, indicating that substantial numbers of UC eligible nonprofit employees were not available to us.[13] In addition, some matches are truncated.

Exclusion of establishments. Organizations that operate in both included and excluded states will appear with fewer matching establishments than their full complement of worksites. For example, an organization operating in the New York metropolitan area will match to its NJ and CT establishments; it will not match to any NY establishments. Some insight to the extent of this problem comes from the proportion of matched organizations that have multiple establishments.

[13] Nonprofit employment in NY, MA, and MI are ranked as 1, 8, and 9 by Salamon and Sokolowski (2006) in 2004. WY employment ranks 50 out of 51 jurisdictions. The authors estimated MA and WY employment. The number of nonprofit employees, in millions, in those states is: New York (1.329), Massachusetts (0.474), and Michigan (0.470), and Wyoming (0.037).

Truncated multi-establishment organizations. Nearly 9/10 organizations matching the QCEW and filing Form 990 have only one establishment in one state.[14] Four percent of Form 990 filers operate in several states. These estimates of multi-establishment rates are lower bounds, as we can not count worksites in excluded states. Virtually no organization that files Form 990-EZ has more than one establishment. See Table 3. Registry matches include nearly half as many multi-establishment organizations as Form 990 filers.[15]

This background about multi-establishment organizations colors the meaning of a partition of the data that we use repeatedly.

5 Learning from the matched data

5.1 Classifying matched organizations

The universe is partitioned into three groups: *included states, matched*; *included states, not matched*; and *excluded states*. An organization with one worksite belongs in only one of these classes. An organization with worksites in both included and excluded states could be located in two of the three classes. We resolve the ambiguity by placing all matched cases together, in *included states, matched*, irrespective of the location at which the Forms 990/990-EZ were filed. As a consequence, organizations classified under *excluded states* are a subset of exempt organizations that operate in those states. A few more complex organizations that operate in excluded states will have establishments in *included states, matched* and will be tabulated under that heading. Logically, those organizations include some whose activity is concentrated in included states; others, whose activity is in concentrated in excluded states.

Table 4 describes the link of IRS information to the *QCEW* in our *enhanced match* for organizations. Two features of the table are highlighted by boxes:

- Column 1, *matched to QCEW? Yes*, includes all IRS records that match the QCEW;

[14] Almost without exception, if the organization operates one establishment, it files in the same state where it operates. Organizations may operate in states different than the state in which Form 990 is filed. This may or may not connote multiple worksites. For example the accounting firm for the organization may file the Form 990 in DC, but the organization has a single worksite in MD.

[15] Two circumstances could lead to this finding. *Registry* matches include some large organizations that have not yet filed. Alternatively, some mismatched organizations have not been deleted. No evidence for either of these alternatives is available.

- Rows A – C include all organizations for which we have Forms 990/990-EZ.

The intersection of the two boxes gives the most information. Outside the intersection row D lacks Form 990 data and the column *excluded states* lacks QCEW data. The total number of IRC 501(c)(3) organizations is 277,015 (Row E, last column). More than 90 percent are organizations filing Forms 990/990-EZ (254,347 Row C, last column).

One-third of Forms 990/990-EZ match the *QCEW*. That percentage rises to 38% when we compare matches to the total number of organizations with a presence in included states. This larger percentage is closer to the yield of matches that could be obtained were all the states in the US included.

Row A reveals that only 8% of organizations filing Form 990-EZ match the *QCEW*. That low rate is consistent with the small expenditures of EZ organizations and the exclusion of many nonprofit workers from UC benefits.

Thirteen percent of organizations do not match *QCEW* and file Forms 990/990-EZ from *excluded states*. The column labeled *excluded states* excludes multi-state organizations operating establishments in *included states*. For example, an organization that has an establishment in CA and files Form 990 from its headquarters establishment in NY will be included in the column labeled *Yes* because the CA establishment matches the *QCEW*.

Row D shows that Forms 990/990-EZ matches are augmented by matches from the *Registry* increasing matches by 27%. Row F contains an estimate of Form 990/990-EZ matches that could be obtained from excluded states.[16] When estimates are added to actual matches, Form 990/990-EZ matching in the QCEW universe rises to 43% (Row G).

Arnsberger (2006) estimates 9,000 more §501(c)(3) organizations than we find among Form 990/990-EZ's. Our *Registry* match raises the universe in Table 4 to 14,000 organizations greater than estimated in Arnsberger (2006). Both differences are far beyond sampling errors. A cutoff date for processing Forms 990 into the NCCS database may be responsible for shortfall of §501(c)(3) filers. The excess produced by *Registry* matches has quite different sources. The *Registry* match may capture some large organizations that do not file Form 990 within two years

[16] Matching rates for each type of Form in the included states and the distribution of Form 990 and Form 990-EZ in excluded states are used to estimate the division of organizations in excluded states between the *matched* and *unmatched* columns (Row F).

of the end their fiscal year.[17] Alternatively, some organizations are not required to file and participate in the UC system. Large religious organizations are an example. At the other extreme, many small organizations not required to file Forms 990 have no employees and would not have a UC record that matches the *Registry.* The number of organizations detected by *Registry* matches is a puzzle.[18]

5.2 Employment 501(c)(3) organizations, 2003

5.2.1 Aggregates for the US

Rows A, B, and D of Table 4 include seven cells defined by the IRS data available (Form 990, Form 990-EZ, *Registry*), match status (match [yes, no]) in included states, and the remainder, unmatched organizations filing from excluded states. We use the same logical structure to display employment aggregates in Table 5. The number of organizations (from Table 4) is repeated in the leftmost column as a guide to the reader. The column headed *establishments, raw* shows the number of establishments operated by matched organizations. The column headed e*stablishments, weighted* indicates the extent to which *invalid ein's* conceal presence of matching establishments. The remaining columns pertain to estimates of employment.

The subtotal row sums Form 990/990-EZ filers with worksites in included states.[19] The row *Available states* adds employment from matches to the *Registry* to the subtotal. The total row counts both matched and unmatched employment in all states. As no QCEW data are available for excluded states, totals can not be calculated for QCEW employment or establishments.

The IRS column labeled *employment, raw* contains reports from Form 990. The top row indicates that IRS reports on Form 990 exceed raw *QCEW* employment by 50,000 (0.74%) for

---

[17] SOI sampling does not substitute prior year records for large organizations that are missing from the stratum that is sampled with certainty.

[18] An alternative hypothesis is that some of the organizations treated as 501(c)(3) here should be classified elsewhere. That possibility can only be resolved with detailed study of more recent data, where the 501(c) subsection is more completely classified.

[19] The subtotal for organizations appeared in Table 4, column *subtotal*, row C.

matched records. [20] That difference has two principal sources: (1) workers not covered by UC do not appear in *QCEW*; and (2) workers reported on Form 990 include workers in excluded states. Multi-state organizations that operate in both excluded and included states will show more employees in the first row than their QCEW report which excludes some states.

The column *Imputed* shows more employment than *raw* because we substituted QCEW *raw employment* for zeros on Form 990's that failed to report employment. Attribution of *QCEW* employees increases aggregate employment by 633,000 on Form 990's, an increase of 9.3%.[21] This imputation is too small because many part-time workers and student workers are not counted in QCEW. Also, employees of organizations that have establishments in excluded states are understated to the extent that the organization has employees in the excluded states.

Additional employees should be imputed to unmatched Form 990 records (in 2nd row and *990, exc.*). Employment is presumably unreported and at a greater rate than matched Form 990 filers. Smaller organizations are likely to have more difficulty in correctly completing Form 990 than matched organizations. Ratio estimation of the imputation is inappropriate, as a proportion of unmatched organizations are nonemployers, while nearly all matched records currently employ workers. A statistical model that encompasses both the decision to employ workers and the number employed is needed to impute employees to the unmatched records.

The third row displays matched Form 990-EZ. No employment information can be garnered from those records. Weighting *QCEW* employment makes imputed numbers more representative of the universe, while unweighted employment would understate levels in the exempt sector (David 2007a). Weighted QCEW employment is transferred to the *Imputed* column, in the third row.

Matching *QCEW* to the *Registry* identifies 22,668 organizations that did not file Form 990/990-EZ (Row 6). The *Registry* organizations contribute a weighted count of 32,749 worksites and weighted employment of 1,777,000. Weighted employment from the *Registry* is added to the count from Forms 990/990-EZ in the column labeled *augmented*. A considerable

---

[20] Weighted *QCEW* employment exceeds raw IRS employment. However, weights are not appropriate in this context. The cells compared pertain to records that match, precluding invalid ein's. Employment for unmatched Form 990 in included states is represented in the second row.

[21] Raw employment is used in the imputation as invalid ein's can not occur when matches are successful.

part of the 1.8 million may be employed in large, late-filing organizations. That is suggested by the relatively large numbers of establishments associated with the *Registry* matches. The character of the remainder of the organizations that we identify by *Registry* matches is unknown.

In the total row, covering all states, augmented employment is 11.7 million (50%) higher than the raw, 7.8 million, count of Form 990 IRS employment. The difference lies in two enhancements to the Form 990 data: (A) the 1.8 million employees represented by *Registry* rather than Form 990/990-EZ matches in *included states*; and (B) the 0.6 million imputed employees unreported to IRS on Form 990, again, in *included states*.

Three other aspects of employment aggregates are noteworthy. First, Form 990 counts 1.0 million workers in organizations that file from included states and do not appear in the QCEW. [22] This count understates employment, as explained above. Second, UC coverage for Form 990-EZ filers is small. We estimate that a quarter of Form 990-EZ filers are employers and only 8% match QCEW. (Employment in small organizations must be elicited directly from a revised Form 990-EZ.) Third, extrapolation of imputed employment, or modeling of an employment imputation, would increase employment numbers in excluded states by roughly 8 percent (the ratio of imputed to raw in the subtotal row). It appears prudent to investigate differences among industry classes before attempting that imputation, as reporting of employees is not uniform across industries.

5.3 Employment in major industries

5.3.1 Industry classifications

Forms 990/990-EZ are classified by the *National Taxonomy of Exempt Entities (NTEE)*. This classification was adopted by the nonprofit sector in the 1980's and is used in IRS/SOI statistics on the sector. Although more detail is available, we used a classification that collapsed the code to 15 classes and unclassified. Labels for the NTEE classes appear in the text table below.

---

[22] A match of *QCEW* to the *Registry* could impute employment to the IRS *Business Master File*, where employment does not appear. The resulting employment count could be weighted adding 200,000 (2.4%) to totals available from the *QCEW*. This experiment makes clear that *QCEW* without the Form 990 employment count lacks substantial coverage of 501(c)(3) employment shown in table 5.

Abbreviation	NTEE classes	Label
A	A	Arts, Culture & Humanities'
B not B4	B except B4	Education, excludes post-secondary
B4	B4	Higher education (post-secondary)
C, D	C, D	Environment, Animal-related
E, not E2	E,F,G,H except E2	Health, not hospitals
E2	E2	Hospitals (and support organizations)
I, M	I,M	Crime, public safety
J, K, L	J, K, L	Employment, Food & Agriculture, Housing
N, O	N, O	Recreation sports, Youth development
P	P	Human services, multi-purpose, and other
Q	Q	International, foreign affairs
R – W	R, S, T, U, V, W	Advocacy, Philanthropy, Science, Society benefit
X	X	Religion related
Y	Y	Mutual/ membership benefit
Z	Z	Unclassified

The *NTEE 15* provides detail that is similar to NAICS sectors. NAICS industry classes are available *only* for records matched to the *QCEW*. Those classes apply to *establishments* and can vary over the distinct worksites identified through the BLS disaggregation of multi-establishment organizations to counties. Multiple classifications occur in less than 10% of matched cases (derived from Table 3).

Table 6 shows the distribution of organizations and matched establishments over NTEE classes. The difference in the distribution of matched establishments and the distribution of organizations reflects (a) variation in the match rate by NTEE and (b) differences in the proportion of organizations that have multiple establishments in different NTEE classes. For example, A, Arts organizations, tend to be small with too few employees to be matched. They also are unlikely to have multiple sites. The result is that matched establishments are 6% of all establishments, while arts organizations are 10% of the total. E2, Hospitals, reflect the opposite

situation; match rates are high and many have multiple establishments. Hospitals are 4% of all matched establishments, and 1% of all organizations.

5.3.2 Employment by *NTEE*

Table 7 disaggregates employment into 15 NTEE classes. Weighted counts from the *QCEW* in column 3 are unbiased for each state and size of workforce.[23] Columns 4-6 display NTEE aggregates of *raw*, *imputed* and *augmented* IRS employment (defined as in Table 5). Column 4 is limited to Form 990 information. The imputed and augmented columns enhance Form 990, Form 990-EZ, and *Registry* matches with available *QCEW* data. Augmented employment is larger than the weighted *QCEW* for each of the 15 NTEE classes. More employees are identified by *QCEW* than by Form 990 in three industries (*B not B4*; *X*; and Y).

The outcomes of imputation and augmentation vary widely over the 15 NTEE classes. Both Education, *B not B4*, and Mutual membership benefit, *Y*, acquire more than half of their employment from *Registry* matches. Mutual membership benefit, *Y*, also shows a large increase because employment is unreported on Form 990. At the other extreme, International, foreign affairs, *Q*, report employment well and are seldom identified by *Registry* matches. Together, imputation and augmentation increase reported employment for *Q* by 7.3%.

Table 8 reveals the impact of imputation and augmentation through ratios of employment counts to various bases. Column 1, *imputation rate*, of Table 8 displays imputed employment as a proportion of employment reported by *matched* organizations filing Form 990. This ratio is smaller than the 20% of organizations that fail to report employment on matched Forms 990 because the level of employment is lower in nonreporting organizations. The range of column 1 is from 4% for *E2*, Hospitals and their related support organizations to 90% for *Y*, mutual/ membership benefit organizations.

Column 2, *Augmentation rate*, displays the proportional increase of imputed employment that comes from *Registry* matches (over the included QCEW states). The concentration of unclassified industries in *Registry* matches assures that the augmentation rate is extreme for class *Z*. The reasons for remaining rates over 100% are unclear. Educational institutions other than higher education, class *B not B4*, may administer small trusts established to finance

[23] Weights do not control for industry, and introduce variance into these estimates (David 2007).

scholarships. Employees in this case may be educators in the school, while the trusts are nonemployers. The trusts may be exempt from filing information returns. Further study of NTEE classes with high rates of augmentation is appropriate.

Sorting the NTEE classes according to the proportionate change induced by imputation (column 1) gives the middle panel of Table 8. Sorting the NTEE classes according to the proportionate change induced by augmentation (column 2) gives the right panel of the table. The bold NTEE classes in the sorted panels are defined by the partition of all classes into thirds. *E2* and *Q* lie in the bottom third of both imputation and augmentation. *B not B4*; *N, O*; and *Y* lie in the top third of both imputation and augmentation. The association may be due to a concentration of small and relatively new organizations in the top third.

We conclude:

- Imputation of employment on Forms 990 is essential. Imputed employment counts exceed matched Form 990 in most industries by more than 10%. Higher education, hospitals, and international organizations are the exception.
- A match of *QCEW* to the *Registry* is needed to augment IRS employment estimates for Education (not higher education) and for religious organizations.
- Augmented employment is sufficiently larger than a match which relies only on *QCEW* in most NTEE classes that *QCEW* can not be the sole source for employment information on charitable organizations.

6 Evaluation

6.1 Nonemployers and matching: How much do we know?

Accounting for nonemployers gives another perspective on coverage and employment reported in Tables 4 and 5. Both Form 990 and Form 990-EZ elicit reports of compensation. If that information correctly predicts absence of employees, it can be used in imputing employees. Comparison of Form 990 records matched to *QCEW* records revealed that 20% of organizations fail to report employment. 97% of the non-reporting organizations reveal compensation. The consistency of compensation reporting in the two record systems makes compensation a useful predictor of employment. However 1.1% of matched forms report neither compensation nor

employees on Form 990. More than three-fifths of this group *should* report employment on Form 990. Absence of compensation and employees on Form 990 do not always predict nonemployers.

Unmatched Forms 990 in included states show 51% with no compensation; organizations in excluded states show 31% with no compensation. Using those estimates we calculate that almost 50,000 of the 201,000 Form 990 filers are not employers; and 37,000 of the 53,000 Form 990-EZ filers are not employers. In all, 87,000 of the 254,000 Form 990/990-EZ filers can be tentatively identified as nonemployers. Those estimates are too high, as we can not quantify the proportion of filers who bury compensation expenses among *other expenses*.[24]

Our estimate of nonemployers can be applied to rows C and F in Table 4. 83,300 matches plus the 12,800 matches estimated for excluded states gives 96,100. Dividing those matches by the difference between all filers and nonemployers, 254,000 less 87,000, yields a match rate of 57%. That compares to the 43% rate shown in row G. A clear reality is that some employers are exempt from UC and will never match. However, unknown errors also contribute to low match rates. The level of failed matches is certainly more than the 1.7% we have identified. An important contributor to those errors are birth, death, and merger of organizations. Each of those events can create discrepancies between ein's used in filing Forms 900/990-EZ and *QCEW*.

6.2 Adequacy of employment estimates for charitable organizations

At this point we see a glass half-full, half-empty. The *2002 Economic Census* counts 9.5 million employees in exempt organizations in those industries in which charitable organizations are concentrated (Table 1). Some additional charities and religious organizations are counted in a sector that includes many exempt organizations that are not charities. The failure to divide exempt into two groups – (a) charities (501(c)(3)) and religious congregations and (b) other exempt entities –compromises the policy value of the estimates. Exclusion of K-12 and higher education compromises those estimates.

The match of *QCEW* to IRS counts almost as many 501(c)(3) employees as the *Economic Census*, in a universe that excludes QCEW information for four states. The match shows 1.8 million employees in education that are largely excluded from the Census universe.

---

[24] IRS/SOI has identified this pathology in completing Form 990 as an important source of error.

IRS counts 1 million more employees in charitable organizations than *QCEW*. IRS counts 0.7 million less than Census, when primary and secondary education are excluded. IRS fails to impute Form 990 and elicit employment for Form 990-EZ. Substituting QCEW employment for missing information fills most of the Census-IRS gap, via the 0.6 million employees we have imputed. More imputation is needed to estimate unreported employment for unmatched Form 990 records.

Our largest employment estimate for charitable organizations, the augmented composite of QCEW and Form 990, is 11.7 million, or 9.0 million, when education is excluded. That estimate is for 2003, while the comparable Census number reflects a smaller population of organizations with fewer employees in 2002. Including exempt organizations other than 501(c)(3) brings our augmented estimate to 13.3 million, substantially larger than the Census counts including all of NAICS 813.

The current, quinquenial Census estimate does not adequately track increasing employment in a sector that is growing rapidly (David, Pollak, Arnsberger 2006). Total nonprofit employment is the same order of magnitude as the health sector of the economy (which includes a major group of exempt organizations).

6.3 Timeliness and employment dynamics

Openness and accountability motivate the mandate for exempt entities to file information returns open to the public. Donors, potential donors, and persons valuing an equitable and efficient tax structure need to know that *every* exemption is not a scam.

The gestation period for both Census and IRS/SOI statistics on exempt organizations is more than two years, an interval that does not enable public review at a time when malfeasance can be nipped in the bud. Estimates of employment in exempt organizations could be produced by QCEW nine months after the quarter. The gestation period for estimates from Form 941 would be no longer. Tallies of those administrative records would also reveal the number of exempt organizations (or establishments), the exempt subsection (e.g., 501(c)(3)), and the NTEE classification of those organizations.

We know that births and deaths of enterprises account for a large part of the flux in job creation and job destruction. (Haltiwanger, Davis, Schuh 1996; BED, BLS) The annual rate of births and deaths is staggering. A speculation is that both births and deaths of charitable

organizations exceed levels in the private business sector. We can not know, until charitable organizations are partitioned from other private, nonfarm business employers.

6.4 Summing up

A great deal of value attaches to statistics on employment, and employment growth, in exempt organizations with classification by industry and geography. The potential to produce those statistics exists in current administrative record systems. The substantial tax subsidy to donors to charitable organizations, and a need to know more about how these organizations perform argues eloquently to publish employment estimates for charitable organizations, other exempt organizations, and the residual of private business.

QCEW has been shown to enhance statistics derived from Forms 990/990-EZ and the *Registry*. The Form 941 appears to have a greater capacity to count employees, but that has not yet been proven. Coverage and presentation of charitable organizations in *Economic Census* needs to be improved to allow policy analysts to associate changes in employment with tax expenditures estimated for the sector.

7 A work plan for the future

7.1 Statistical Agencies

7.1.1 First-best activity

Both IRS/SOI and the Census have access to Form 941, with its employment and payroll information. Both agencies have unlimited access to information returns filed on behalf of exempt organizations and the *Register*.[25] The steps taken in this paper can be replicated on Form 941. The tri-partite match – Forms 990/990-EZ, *Registry*, and Form 941 – can be executed in both agencies. Census may have an advantage in editing and linking records (Winkler 2004); IRS/SOI may have earlier and more comprehensive access to records.

Both agencies currently receive updates from IRS *Business Master Files* monthly. Those files contain extracts from Form 941 and Form 990/990-EZ. The *registry* is also updated monthly. Both agencies could devise a quarterly estimate of exempt employment that is

[25] Publication of identifiable information at the organizations level is sanctioned. So no disclosure review is necessary at the organization (ein) level of detail.

subsequently benchmarked to the more thoroughly edited annual samples of Form 990/Form 990-EZ produced by IRS/SOI.

Indicators of exempt status and applicable IRC section can be taken from Form 990/990-EZ and the IRS *Registry*. They can be inserted into the *Census Business Register*. That would enable sampling using exempt status (not exempt, 501(c)(3), other exempt) and, more importantly, analyses of differences between exempt and other organizations.

7.1.2 Second-best activities

A. IRS/SOI can impute unreported employment on Form 990. This should yield far more employment than the 633,000 we substituted from the *QCEW*. We know from this investigation that even incomplete imputation can produce substantially larger employment for Form 990 than employment reported to the *QCEW* (Table 5).

The BLS has its own business register that is derived from the *QCEW*. Under current law BLS can not access the Form 941, but it can receive the continuous stream of Forms 990/990-EZ as they become available in the *IRS Business Master File*, because all of that information is public. Thus it is feasible for BLS to continue the tri-partite match – Forms 990/990-EZ, *Registry*, and *QCEW* – pioneered in this research. The *IRS Business Master File* does not include employment so that employment estimates would only cover the *QCEW* universe. Such estimates could be produced within the current nine-month interval after the quarter that is the timetable for *QCEW* reports. The *QCEW* exempt employment series should be compared to SOI employment estimates that lag the reporting year by about 30 months.

The gain from a continuing BLS activity would be to partition *Business Employment Dynamics* (BLS webpage) into estimates of private for-profit and a private nonprofit job creation and destruction.

7.2 Operating Agencies

7.2.1 IRS/TEGO

IRS/TEGO administers approvals for tax exemption, updates the *Registry* and classifies exempt organizations by NTEE. IRS/TEGO has the power to reject Forms 990/990-EZ that are incomplete. It can identify organizations that fail to file timely information returns. It can reject all Form 990 returns where employment is not reported. IRS/TEGO should also reject Form 990 where no compensation is paid and employees are present.

IRS/TEGO is in the process of redesigning Form 990 to assure compliance of exempt organizations with terms of their approval. They can ask organizations to report employment on Form 941; that would assure that employment is public information for larger organizations.[26] It can ask to have Forms redesigned so that answering the employment question is tied directly to reports of positive compensation. Then Form 990-EZ could have an employment question that is conditional on positive compensation.

Electronic filing of Form 990 is required for large organizations and facilitated by free software for small organizations. With e-filing, Forms that fail to report employment, when requested, can be rejected by edits in the e-filing software. That eventuality will ultimately greatly reduce current unreported employment.

7.2.2 OMB/OIRA

OMB/OIRA administers the *Paperwork Reduction Act*. It contains the Office of the Chief Statistician whose role is to coordinate the US statistical agencies and provide guidance to good statistical practice. The Chief Statistician needs to be convinced that higher priority be given to employment statistics for charitable organizations.

Acknowledgements

I thank the Urban Institute and the Bureau of Labor Statistics for their far-sighted access to microdata pertaining to nonprofit organizations. The National Center for Charitable Statistics and its Dataweb (NCCS 2003) provided the *Registry* and Form 990/990-EZ from 1999-2003 for this research. Linda Lampkin and Tom Pollak offered critical suggestions on what was needed for policy-making on charitable organizations. Jen Auer ably assisted me in identifying mismatches and thinking about multi-establishment organizations. Kendal Golladay prepared extracts of the population files and detected invalid EIN's.

---

[26] They can also ask for number of employees reported in connection with filing of W-2 at the end of the calendar year. That would establish number of employees in a reference year, a larger count than the number of employees in a reference pay period. Both measures are needed to estimate employment among small employers who currently file Form 990-EZ.

Amy Knaup and Merissa Piazza reviewed work-in-progress and provided disclosure review for the tables presented. Rick Clayton and David Talan educated me about the QCEW program and reviewed this paper and David (2007). Access to BLS data through its secure research site is overseen by thoughtful peer review, chaired by James Spletzer, and is a major resource for scientific research on data collected by the BLS.

The views expressed here do not reflect policies of the BLS. I am responsible for the design of this study and errors in its execution. I hope others can advance beyond these limited results.

References

Arnsberger, Paul. Fall, 2006. Charities and other tax-exempt organizations. *SOI Bulletin*. Washington DC: IRS/Statistics of Income. 26:2, 231-245.

Brody, Evelyn (ed.). 2002. *Property-tax Exemptions for Charities*. Washington DC: Urban Institute Press.

BLS webpage. Business Employment Dynamics: www.bls.gov/bdm/home.htm.

David, Martin H. 2007. Employment in Nonprofit Entities: Coverage, Bias, and Measurement Errors in QCEW and public IRS Information, 1999-2003. *2006 ASA Proceedings*. Alexandria VA: American Statistical Association. (CD- ROM publication).

David, Martin H, Thomas Pollak and Paul Arnsberger. 2006 Compliance with information reporting: Exempt Organizations. *IRS Research Bulletin: Recent Research on Tax Administration and Compliance* (Publication 1500) 231-245.

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. Cambridge MA: MIT Press.

Fellegi, I. P., and A. B. Sunter. 1969. A theory for record linkage. *Journal of the American Statistical Association*, 64(328), 1183--1210.

Foster, Lucia, Joel Elvery, C. J. Krizan, David Talan. 2007. Preliminary Micro Data Results from the Business List Comparison Project. *Proceedings of the American Statistical Association*, 2006 (forthcoming).

Gronbjerg, Kirsten A. and Erich T. Eschmannn. May 2005. *Indiana nonprofit employment*. Ctr. on Philanthropy, School of Public and Environmental Affairs, Indiana University and the Johns Hopkins Nonprofit employment data project.

Helliwell, John and H. Huang. 2005. How's the job? Well-being and social capital in the workplace. NBER working paper 11759. Cambridge MA: NBER (on-line).

Independent Sector. 2003. *The Nonprofit Almanac: 2002*.

Michigan Nonprofit research program. 2005. *Economic benefits of Michigan's nonprofit sector: 2004*.

NCCS:<year>. *GuideStar-NCCS National Nonprofit Research Database* [Years 1999-2003] Washington DC: Urban Institute.

Okolie, Cordelia. 2004 July. Why size class methodology matters in analyses of net and gross job flows. *Monthly Labor Review*, 3–12.

Salamon, Lester M. and S. Wojciech Sokolowski. 2005 Sept. Nonprofit organizations: new insights from QCEW data. *Monthly Labor Review*, 19–26.

Salamon, Lester M. and S. Wojciech Sokolowski. 2006 Dec. Employment in America's charities: a profile. *Nonprofit Employment Bulletin* 26. Baltimore MD: Johns Hopkins Center for Civil Society Studies, Table 3.

Scheuren, Fritz and William E. Winkler. Regression analysis of data files that are computer matched – Part II *Survey Methodology*, 1997.

Spletzer, James and Joel Elvery. 21 October 2005. Presentation to CNSTAT workshop "Benefits of Interagency Business Data Sharing".

United States. *2008 Budget of the United States*. Table 19-3.

Vilhuber, Lars, Bryce E. Stephens, John M. Abowd, Fredrik Andersson, Kevin L. McKinney, Marc Roemer and Simon Woodcock (forthcoming, draft 2005 CRIW conference at NBER.org) The LEHD Infrastructure files and the creation of the Quarterly Workforce Indicators.

Winkler, William E. 2004. Methods for evaluating and creating data quality. *Information Systems*, 29: 531-550.

Appendix A. Comparison of NTEE and NAICS classifications

The granularity of industry coding used in this analysis allows for 15 bins of the NTEE (excluding unclassified) and 15 bins for the NAICS. If the information in both coding systems were identical and no errors were made in coding, all NAICS sector codes should map into one NTEE bin, and conversely. Figure 2 shows the allocation of NAICS industries to NTEE classes for NTEE-NAICS combinations that include a total of 90% or more of NTEE employment.[27] Three NAICS sectors are "well-behaved" as the sector maps uniquely to one NTEE class. Seven additional sectors match to two NTEE classes. Four sectors (54, 61, 62, and 81) are substantially partitioned as NAICS is distributed to five or more NTEE bins.

[27] Limiting the combinations of NAICS to NTEE to classes that contain the preponderance of employment reduces combinations that result from a low level of classification error. The share of employment included was chosen to avoid problems with disclosure of employment information within each NTEE-NAICS combination.

Table 1. Nonprofit employment reported in Economic Census 2002

Naics classifi-cation	Employment A Exempt	C All	D Ratio: A/C	Establishments E Exempt	G All	H Ratio: E/G
61	120	431	0.28	12	50	0.24
62	7,980	15,048	0.53	136	703	0.19
71	1,363	1,847	0.74	49	109	0.45
813	936	936	1.00	11	11	1.00
Total exc. 813	9,463	17,326		197	862	
inc. 813	10,400	18,262		208	874	

Source: A1

Table 2. Reallocation of NA and mismatched 501(c) subsection matched organizations, 2003q1*

Reallocated subsection	501(c) subsection	Type of match		Total	
		Forms 990	Registry	Count	Percent
501c3 plus	**NA**	171	2,338	2,509	2.4
	501(c)(3)	83,117	20,330	103,447	97.6
	Total	83,288	22,668	105,956	100.0
		Type of reclassification			
		None	**LbrBusGov**		
Not 501c3	**NA**	5	527	532	0.8
	501(c)(3)	0	409	409	0.6
	Other	64,872	0	64,872	98.6
	Total	64,877	936	65,813	100.0

* Excludes all Naics 52 (Finance and Insurance).

Source: Table 10. 02oct06.
Wss_tables16oct06_d21may22mar07.xls

Table 3. Proportion of matched organizations with multiple worksites, interstate operations

By type of match, 501c3 plus organizations, 2003q1

Match	Organ-izations	Rate: Multi-establishment		Rate: multi-state	
Form 990	79045	9086	0.115	3161	0.040
Form 990-EZ	4243	d	*	d	*
Registry	22668	1197	0.053	482	0.021

d, not disclosable

* Less than 0.005

Source: **12. 03oct06.**

Table 4. Strata defined by QCEW matches to Forms 990, 501(c)(3) organizations, 2003q1*

		Included states			Excluded states	Total
		Matched to QCEW?		Subtotal		
		Yes	No			
A	**Form 990-EZ**	4,243	42,166	46,409	6,733	53,142
	proportion of Forms 990	0.080	0.793		0.127	1.000
B	**Form 990**	79,045	95,294	174,339	26,866	201,205
	proportion of Forms 990	0.393	0.474		0.134	1.000
C	**Subtotal NCCS Census**	83,288	137,460	220,748	33,599	254,347
	proportion of Forms 990	0.327	0.540		0.132	1.000
D	**Master QCEW matches**	22,668		22,668		22,668
	augmentation rate			0.10268723		
E	**TOTAL**	105,956	137,460	243,416	33,599	277,015
	proportion of total	0.382	0.496		0.121	1.000
F	**Estimated matches****	*12,797*	*20,802*			
G	**UNIVERSE**	*118,753*	*158,262*	277,015		277,015
	proportion of universe	*0.429*	*0.571*	*1.000*		*1.000*
H	990dd/all_dd			0.790	0.800	
I	match rate 990			0.377	*0.381*	

*Includes 4,836 cases where 501(c) section is not known. 2338 are in the Master matches; the remainder are relatively equally distributed across the filing population.

**Entry in 'Yes' column is the product of the estimated match rate in excluded states (row I) and the 33,600 excluded filers (row E). 0

No estimate of additional augmentation is included.

Source: Table 1

Table 5. Imputation and augmentation of Form 990 Employment, 2003, Section 501(c)(3), in 1,000's

			QCEW emp.		IRS employment		
Source	**Match?**	**Orgs.**	**Raw**	**Wtd.**	**Raw**	**Imputed**	**Augmented**
990, inc.							
	Yes	79	6,780	6,930	6,831	*7,463*	*7,463*
	No	95			1,005	1,005	1,005
990-EZ, inc.							
	Yes	4	10	*10*	*NA*	*10*	*10*
	No	42					
Subtotal		220	6,790	6,940	7,836	*8,478*	*8,478*
Registry	**Yes**	23	1,724	*1,777*	NA	NA	*1,777*
Available states		243	8,514	*8,717*			*10,255*
990, exc.		27			1,485	1,485	1,485
990-EZ, exc.		7					NA
Total		277	NA	NA	9,321	*9,963*	*11,740*
Excess of IRS employment over included QCEW					1.16	1.22	

Italics reflect model-based weighting of QCEW and substitution of QCEW for missing employment reports to IRS.
Imputation would increase employment in unmatched and excluded states.

Table 6. Organizations and matched establishments by NTEE
2003 q1, 501c3 plus select NA subsection

	IRS Organizations		*QCEW* establishments		
NTEE 15	**Count**	**Percent**	**Count**	**Percent**	**Wtd.**
A	28,582	10.3%	10,468	6.0%	10,628
B not B4	47,890	17.3%	20,487	11.8%	20,749
B4	1,376	0.5%	2,029	1.2%	2,061
C,D	10,273	3.7%	4,977	2.9%	5,047
E not E2	33,480	12.1%	30,167	17.4%	30,496
E2	3,378	1.2%	6,968	4.0%	7,067
I	4,811	1.7%	3,578	2.1%	3,619
J,K,L	21,069	7.6%	13,254	7.7%	13,405
M	4,003	1.4%	731	0.4%	741
N,O	25,772	9.3%	9,124	5.3%	9,249
P	38,627	13.9%	41,943	24.2%	42,374
Q	4,982	1.8%	2,112	1.2%	2,149
R - W	32,980	11.9%	17,028	9.8%	17,256
X	17,484	6.3%	7,681	4.4%	7,820
Y	659	0.2%	489	0.3%	495
Z	1,649	0.6%	2,179	1.3%	2,218
Total	277,015	100.0%	173,215	100.0%	175,373

Source: T5A_14sep06.
Wss_tables16oct06_d21may22mar07.xls

Table 7. Employment by NTEE major sectors, 2003 q1, 501c3 plus select NA subsection

NTEE 15	IRS organ-izations	*QCEW* employment		IRS employment		
		Raw	Wtd.	Raw	Imputed*	Augmented**
A	28,582	184,565	187,908	243,222	260,772	289,694
B not B4	47,890	1,119,588	1,158,436	577,391	644,346	1,381,306
B4	1,376	728,814	757,030	1,107,722	1,184,338	1,270,880
C,D	10,273	65,756	66,928	70,176	78,222	83,430
E not E2	33,480	1,272,976	1,289,056	1,329,935	1,439,382	1,713,259
E2	3,378	2,884,275	2,971,198	3,525,270	3,635,550	3,875,725
I	4,811	47,447	47,882	55,246	62,594	65,325
J,K,L	21,069	270,879	273,897	362,594	400,201	428,882
M	4,003	7,181	7,226	6,940	8,167	9,199
N,O	25,772	140,490	142,051	123,994	141,340	185,005
P	38,627	1,258,663	1,271,529	1,544,368	1,693,076	1,761,028
Q	4,982	20,429	20,687	39,046	40,835	41,882
R – W	32,980	277,378	282,398	265,133	297,781	359,063
X	17,484	125,731	128,391	64,997	69,936	160,529
Y	659	13,057	13,207	3,424	5,411	14,855
Z	1,649	96,660	99,043	1,175	1,261	99,769
Total	277,015	8,513,889	8,716,863	9,320,633	9,963,211	11,739,831

* Substitutes QCEW employment for matched Form 990, employment NA, and matched Form 990-EZ.
** Adds Registry matches to imputed.

Source: **T5B_14sep06.**

Table 8. Imputation and augmentation, 2003q1, 501(c)(3)
Matched Form 990's

Ordered by NTEE 15			Ordered by Imputation			Ordered by Augmentation	
NTEE 15	Imputation rate*	Augmentation rate**	NTEE 15	rate*		NTEE 15	rate**
A	0.107	0.111	**E2**	0.043		**Q**	0.026
B not B4	0.173	1.144	**Q**	0.055			0.040
B4	0.086	0.073		0.086	P		0.044
C,D	0.132	0.067	B4	0.107	I	**E2**	0.066
E not E2	0.117	0.190	A E not E2	0.117		C,D	0.067
E2	0.043	0.066		0.118		J,K,L	0.072
I	0.178	0.044	X	0.127			0.073
J,K,L	0.139	0.072	P C,D	0.132	B4		0.111
M	0.197	0.126	J,K,L	0.139	A		0.126
N,O	0.177	0.309	R - W	0.154	M	E not E2	0.190
P	0.127	0.040	***Z***	*0.164*		R - W	0.206
Q	0.055	0.026	**B not B4**	0.173		**N,O**	0.309
R - W	0.154	0.206	**N,O**	0.177		**B not B4**	1.144
X	0.118	1.295		0.178			1.295
Y	0.906	1.745	I	0.197	X	**Y**	1.745
Z	*0.164*	*78.110*	M **Y**	0.906		***Z***	*78.110*
All	0.093	0.210	All	0.093		All	0.210

*Observed for Form 990 matched to QCEW.

**Increase over imputed total, for all matches, included states.

Bold emphasizes consistency of ranks for industries in the top of bottom third of NTEE classes.

Italics emphasize that no industry class is assigned to *Z.*

Source: F6_T6_Wss_tables16oct06.xls revised

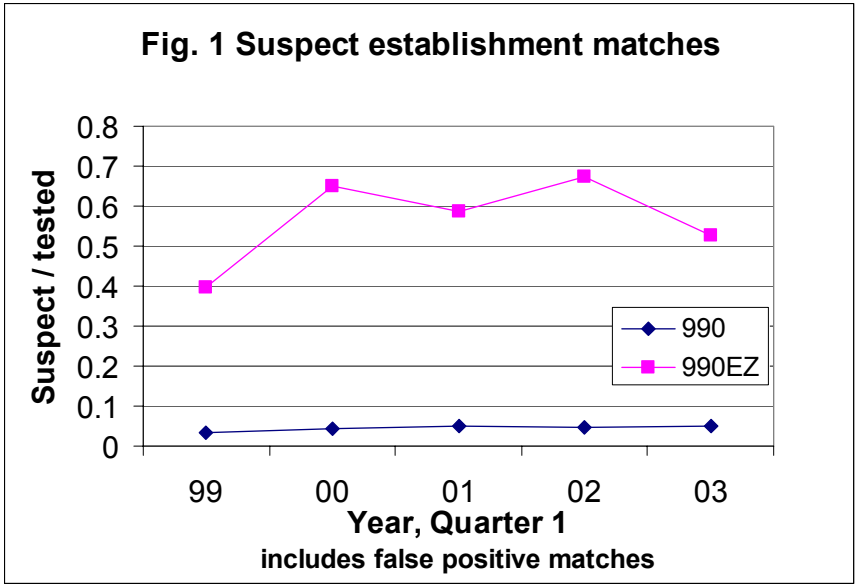
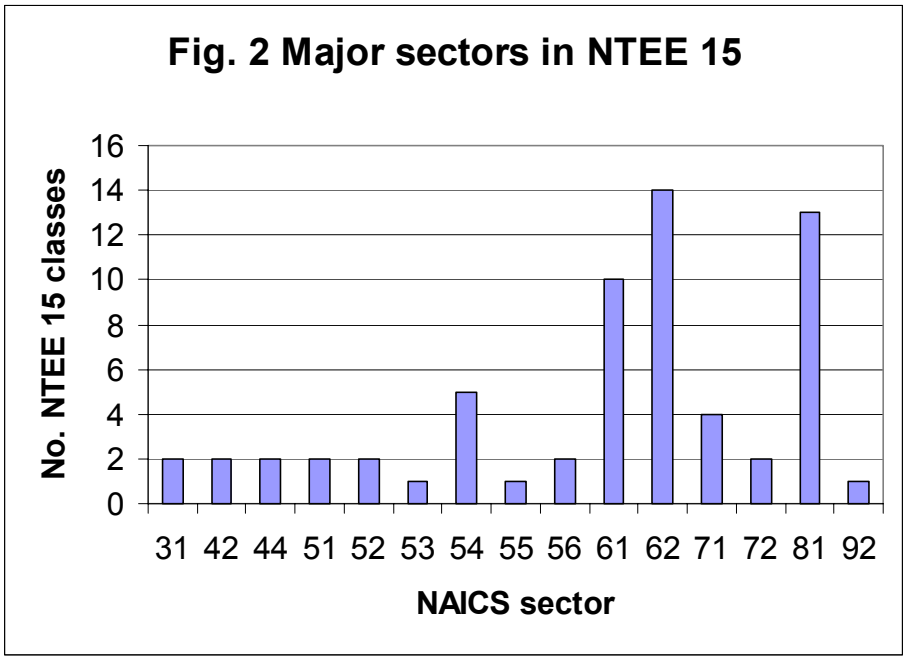


Fig. 1 Suspect establishment matches
Suspect / tested
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
990
990EZ
99
00
01
02
03
Year, Quarter 1
includes false positive matches



Source: ddldb_nt2p_naics2dis28sep06.xls