

Overview of IRS and SOI Exempt Organization Data Products

July 1, 2014



IRS-SOI EO Data Products

- EO BMF Extract (IRS)
- EO Financial Extracts (IRS-SOI)
- EO Sample Studies (SOI)

April 30, 2014



EO BMF Extract

- Cumulative listing
- Includes organizations with active exemption, regardless of Form 990/990-EZ or 990-PF filing requirement
- All subsection codes
- Most information derived from Entity portion of BMF
- Data also transcribed from administrative filings
- Records: 1,538,313 (as of 5/12/2014)
- Only 28 fields; 3 are financial

April 30, 2014



EO Financial Extracts

- Include selected organization-level data for population of returns from EO Master File.
- All subsection codes
- Minimal data cleaning to correct major transcription and/or arithmetic errors
- Processing year file

April 30, 2014



EO Financial Extracts

- Processing Year 2012 file was first one made widely available to the public.
- For Processing Year 2013, the number of fields was significantly expanded

Form	Records	Fields	% Financial
990	289,603	245	61%
990-EZ	218,981	71	80%
990-PF	100,484	179	71%
Total	609,068	495	68%

April 30, 2014



EO Sample Studies

Form 990
Exempt
Organizations

Subsections
501(c)(3), (c)(4)-(9)

Form 990-EZ
Small Exempt
Organizations

Form 990-PF
Private
Foundations

Subsection
501(c)(3)

Sample of **31,500** filers
from a
Population of **467,500** filers

April 30, 2014



EO Sample Studies

- “Stratified” random sample – sampling rates

Size of assets (\$ M = millions)	Sampling rate	Number of organizations selected for sample
Very Large >\$50m	100%	9,034
Large \$20m - \$50m	39% - 100%	6,033
Medium \$5m - \$20m	7% - 100%	7,334
Small \$1m - \$5m	3% - 14%	4,404
Very Small <\$1m	1% - 5%	4,829
Total		31,634

April 30, 2014



EO Sample Studies

- NTEE codes perfected by tax examiners based on information reported on Form 990, Part III (Statement of Program Service Accomplishments)
- Extensive data cleaning and allocations to correct errors and inconsistencies
- Tax Year 2010

Form	Records	Fields	% Financial
990	19,476	856	70%
990-EZ	1,260	202	65%
990-PF	10,898	308	69%
Total	31,634	1,366	69%

April 30, 2014



Overview of National Center for Charitable Statistics (NCCS)

Center on Nonprofits and Philanthropy, The Urban Institute

The Urban Institute



Nonprofit Data Available through NCCS

- IRS data: BMF (multiple versions / year) and SOI (annually)
- NCCS Core Files (1989-present)
- NCCS Cumulative Master File (1998-present)
- NCCS-GuideStar National Nonprofit Research Database (NNRD or “digitized data”) (1998-2003)
- Parts 3 and 8 Supplement (1992)
- NCCS Trend Analysis Files
- NCCS NTEE Master File

The Urban Institute



NCCS Processing of NTEE Codes

- Original classifications done by NCCS in mid-1980s
- NCCS has led periodic revisions of NTEE codes; also manages the Advisory Committee and Oversight Committee
- Since 1995: Codes assigned by IRS' EO division
- Since 2004: NCCS' EON (Exempt Organization NTEE) program assigns NTEE codes and confidence ratings
- NCCS has also developed the NPC (Nonprofit Program Classification) system to code programmatic activities



Part III: Program Service Accomplishments

Form 990 (2013) Page 2

Part III Statement of Program Service Accomplishments
Check if Schedule O contains a response or note to any line in this Part III

1 Briefly describe the organization's mission:

2 Did the organization undertake any significant program services during the year which were not listed on the prior Form 990 or 990-EZ? Yes No
If "Yes," describe these new services on Schedule O.

3 Did the organization cease conducting, or make significant changes in how it conducts, any program services? Yes No
If "Yes," describe these changes on Schedule O.



Part III, Continued: Details of Program Service Accomplishments

4 Describe the organization's program service accomplishments for each of its three largest program services, as measured by expenses. Section 501(c)(3) and 501(c)(4) organizations are required to report the amount of grants and allocations to others, the total expenses, and revenue, if any, for each program service reported.

4a (Code: _____) (Expenses \$ _____ including grants of \$ _____) (Revenue \$ _____)

4b (Code: _____) (Expenses \$ _____ including grants of \$ _____) (Revenue \$ _____)

The Urban Institute



Proposal to Identify R&D Producing Nonprofits

- Task 1: Analyze narrative data from 990s for NSF's initial sample of nonprofit organizations that are likely producers of research and development
- Task 2: Identify candidate keywords that could be used to locate other NPs that are likely producers of R&D
- Task 3: Use the final list of keywords to locate R&D-producing organizations among the larger NP population (estimated 50,000 to 75,000 organizations)

The Urban Institute



Potential Benefits and Challenges of Proposed Approach

- Benefits:
 - Makes use of newly available database of scanned 990 forms
 - Capitalizes on NCCS' experience with similar projects
 - Results can be used for further analysis
- Challenges:
 - False positives
 - Will keywords identified in Task 2 be appropriate for Task 3?

