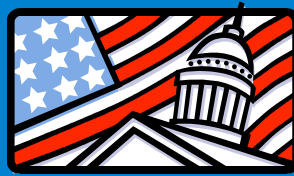


SOI Studies, Data, and Future Plans

*Tom Petska, Director
Statistics of Income (SOI) IRS*



*“Turning Administration Systems
into Information Systems”*

SOI Studies, Data and Future Plans

1. Background on SOI in the IRS
2. SOI studies and data availability, focusing on regional data
3. SOI’s website and ways to find SOI data
4. Overview of future plans



Internal Revenue Code of 1986

Section 6108(a)



“The Secretary (i.e., of Treasury) shall prepare and publish not less than annually....statistics available with respect to the operations of the internal revenue laws”

Who Uses SOI Data?



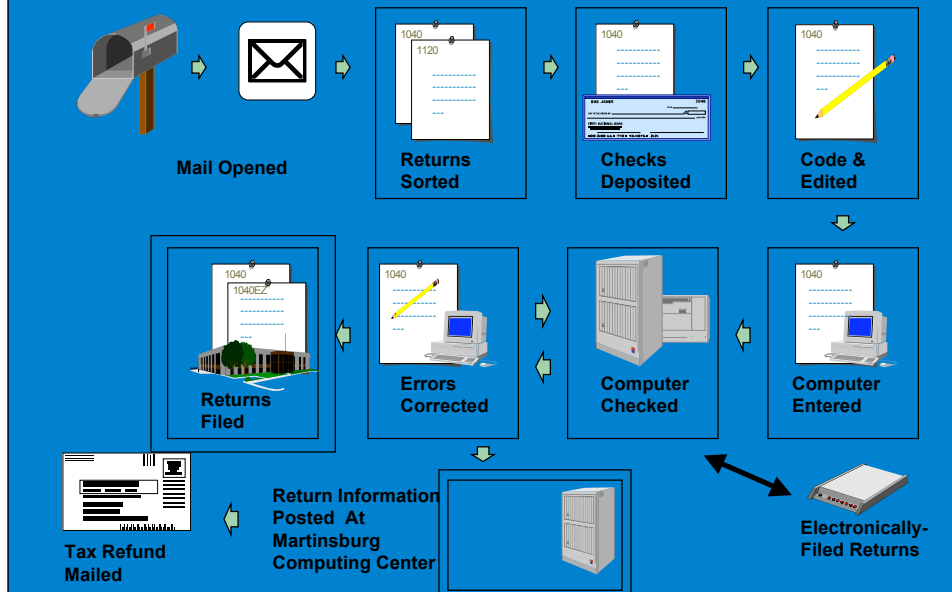
- Treasury’s Office of Tax Analysis (OTA)
- Congressional Joint Committee on Taxation (JCT)
- Commerce’s Bureau of Economic Analysis (BEA) & Census Bureau
- IRS
- Other government agencies
- Academic & private researchers

SOI Budget & Staffing

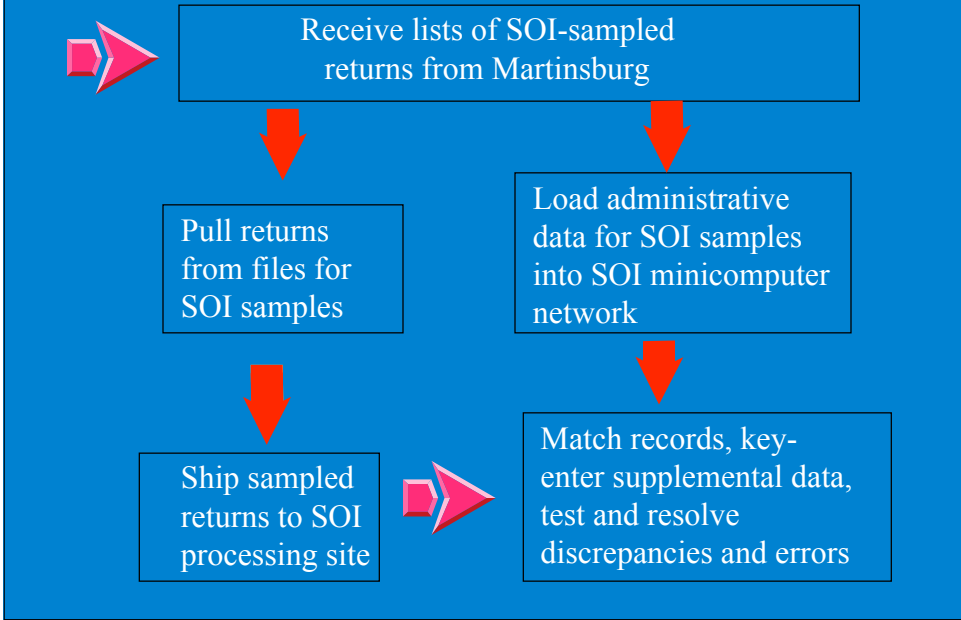
- Overall budget **\$40.5 million**
- Total staffing (staff years) 445
 - National Office 171
 - Economists / Subject Matter (45%)
 - Computer Specialists (30%)
 - Math Statisticians (10%)
 - Other (15%)
 - Field Sites 266
 - Editors (90%)
 - Computer Specialists (10%)



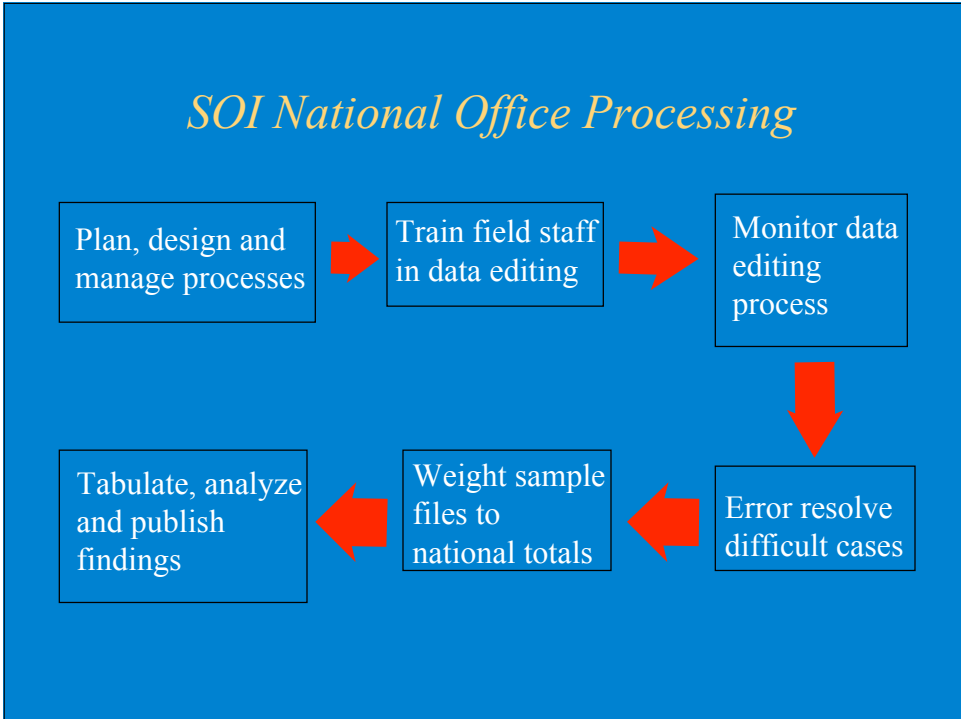
IRS Returns Processing Pipeline



SOI Service Center Processing



SOI National Office Processing



SOI Statistical Processing



SOI "Piggybacks" on Tax Compliance Operations

- IRS mission is to administer the tax laws
- SOI accounts for < 0.4% of IRS \$11 billion budget
- SOI is an administrative records system
 - *Pros* -- data are compiled for other purposes
 - *Cons* -- system is not designed for research purposes and is hard to change



SOI's Relationship to IRS Tax Compliance Processing

- Samples drawn from IRS Master Files
- Administrative data for sampled cases copied into SOI's minicomputer network
- Tax return item content augmented
- Testing and validation are substantially greater
- SOI data are all *pre-audit*



Two Types of "SOI" Data

- SOI edited sample files
 - *Pros* – rich in content and highly accurate
 - *Cons* – expensive, less timely, and not robust for subnational estimates
- IRS administrative (i.e., population) files
 - *Pros* – less expensive and more timely
 - *Cons* – less content and less accurate



SOI Data Compared to IRS Master File Data

- SOI uses samples; Master Files are populations
- SOI samples select large returns at high rates (SOI avg. size & complexity are greater)
- SOI item content is substantially richer
- SOI testing and validation are much greater

Challenges in Producing Tax Statistics

- Statistics are secondary priority in large administrative agency
- Primary customers are external
- Complex organizational structure affects statistical programs
- Tax law changes frequently
- Financial information reported can be complex



Population and SOI Sample Sizes

<u>Study</u>	<u>Population</u>	<u>Sample</u>
Individual (TY 2005)	134,494,440	526,601
Cross-sectional		292,966
BEA Supplementary		2,439
Individual Edit Panel		127,096
Individual Unedited Panel		421,893
Corporation (TY 2004)	5,980,702	146,269
Partnership (TY 2005)	2,949,576	44,241
Tax-exempt Org. (TY 2004)	390,425	21,739
Private Foundation (TY 2004)	84,216	11,451
Unrelated Business (TY 2004)	38,275	7,898
Estate (TY 2006)	56,696	13,488
Gift (TY 2006)	264,013	9,217
Split Interest Trust (TY 2006)	124,632	12,466



A Typical SOI Sample Design

Asset size	Population	Sample	Rate (r)	Weight (1/r)
Totals	1,600,000	25,000	N.A.	N.A.

\$10 mil+	10,000	10,000	100.0%	1
\$10m<1m	90,000	9,000	10.0%	10
\$1m<0.1m	500,000	5,000	1.0%	100
< \$0.1m	1,000,000	1,000	0.1%	1,000

Who Has Access to IRS Microdata?

- OTA & JCT have full access
- BEA primarily for corporations
- CBO primarily for Social Security analysis
- Census has access to some Master File data
- SSA, for earnings histories
- USDA, to conduct the Agriculture Census
- Everyone, to Individual and Tax Exempt public-use files



SOI Branches^(*), Study Areas, and Publications

Individual

Proprietorships
Panel Studies
High Income/
No Tax

Complete Report

Corporation

Corporations
S corporations
Partnerships

Complete Report
Source Book

Special Studies

International
Tax-exempts
Estate and Gift
Activity Bonds
Excise Taxes

SOI Bulletin
Research Report
IRS Data Book
Tax Stats Website



(*) In addition to the Statistical
Computing Branch

SOI Publications



- *SOI Bulletin*
- Individual and Corporation Complete Reports
- Corporation Source Book
- *Special Studies in Federal Tax Statistics* (SOI Research Papers)

Other Publications not based on SOI data

- *IRS Data Book*
- *Research Bulletin* (proceedings of the IRS Research Conference)

Other Data Products and Services

- Tax Stats Web Site -- SOI data on the Internet
- Over 6,600 Files Available - Individual, Corporation, Partnership, Exempt Organizations, and more.
- <http://www.irs.gov/taxstats/index.html>
- Statistical Information Services (SIS) Office
- Reimbursable studies



Microsoft Internet Explorer window showing the Internal Revenue Service Tax Statistics page. The browser title is "Tax Statistics - Produced by the Statistics of Income Division and Other Areas of the Internal Revenue Service". The address bar shows "http://www.irs.gov/taxstats/index.html".

The page header includes the Internal Revenue Service logo and the text "United States Department of the Treasury". A search bar is located in the top right corner.

The main navigation menu includes: INDIVIDUALS, BUSINESSES, CHARITIES & NON-PROFITS, GOVERNMENT ENTITIES, TAX PROFESSIONALS, RETIREMENT PLANS COMMUNITY, and TAX EXEMPT BOND COMMUNITY.

The page content is titled "Tax Statistics - Produced by the Statistics of Income Division and Other Areas of the Internal Revenue Service". It includes a welcome message and a list of links for various categories:

- Statistics:** Business Tax, Charitable, Compliance, Individual Tax, By Form, Products and Publications.
- Tax Stats Topics:** SOI Table Wizard, Estate and Gift, Taxpayer Compliance, More Topics...
- IRS Resources:** Compliance & Enforcement, Contact My Local Office, e-file, Forms and Publications, Frequently Asked Questions, News, Taxpayer Advocacy.
- Business Tax Statistics:** Corporations, International, Partnerships, Nonfarm Sole Proprietorships, S Corporations, All Topics >>
- Individual Tax Statistics:** Filing Season Statistics, Estate & Gift Tax, Individual Income Tax, International, Personal Wealth, All Topics >>
- Charitable & Exempt Org. Statistics:** Charities, Exempt Org. Master File, Private Foundations, Trusts, Tax-Exempt Bonds, All Topics >>
- Products, Publications, & Papers:** SOI Bulletin, IRS Data Books, Conference Papers, All Topics >>
- Statistics by Form:** 706, 709, 990, 990-E, 990-T, 1040, 1041, 1065, 1118, 1120, 5471, 5472, 8038, All Forms >>
- Additional Information:** Tax Statistics @ a Glance, Join SOI Tax Stats E-Mail List, Payments for SOI Data, Questions On Tax Statistics?

At the bottom of the page, it states: "Contains commands for working with the selected items."

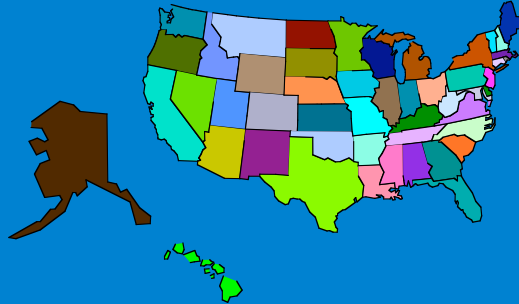
Data by Geographic Area

- Individuals
 - State
 - County
 - Migration Flows
 - ZIP Code
- Tax exempt organizations
- Businesses (forthcoming)



2005 Individual Income Tax Data By State

By Size of AGI including Charitable Contributions,
Earned Income Credit, Alternative Minimum Tax,
and Income Tax



Source: Spring 2007 SOI Bulletin

2005/2006 Migration Data – State-to-State and County-to-County

- Number of Returns
- Number of Exemptions
- Aggregate Adjusted Gross Income (AGI)



– *Source: Contact SIS Office*

2005 County Income Data

- Number of Exemptions
- Aggregate Adjusted Gross Income
- Wages and Salaries
- Dividends
- Interest Received



– *Source:* Contact SIS Office

2005 Zip Code Data

- Income and Tax Items, such as --
 - Charitable Contributions and
 - Alternative Minimum Tax
- Number of returns with Schedules C and F
- All by AGI Size Classes



– *Source:* Contact SIS Office

2006 IRS Data Book



- Data from various offices of IRS
- Available on the Tax Stats Website
- FY 2007 data released by April 2008

Recent Improvements for Customers



- Integrated Business data
- Corporation Source Book multi-year flat files
- Continued improvements to the Tax Stats
- List-serve announcements of new data products and studies
- Payments by Credit Card

Number of Requests by Type of Organization *July, 2006 - June, 2007*

Organization Type	Number of Requests
Total.....	2,300
Consultant/Research.....	483
Academia.....	313
Internal Revenue Service.....	297
Media Relations.....	192
Private Citizen.....	247
State/Local Government.....	189
Media (Direct).....	181
Corporation.....	158
Federal Government.....	109
Nonprofit.....	99
Financial Services.....	96
Congress.....	53
Other.....	75

Frontiers in Tax Statistics

- Increased emphasis on longitudinal analysis
- Greater use of master file population data
- Expanded use of e-filed returns
- Expanded web presence
- Less people-intensive data edits
- Expanded Public Use Files
- Synthetic data?



Planning our Future – 
SOI's 2016 Strategic Planning Initiative

1. Customers -- continued expansion
2. Products and services – more data to the public
3. Information processing – leveraging technologies and e-filing
4. Human capital – refining staffing needs

1. Expanding our Customer Base 

- Recognition of OTA as JCT as primary customers
- Designation of important national needs of BEA , FRB, and Census
- Formal recognition of IRS, researchers, and others (including the public) as important customers
- Continued expansion beyond current group
- How does this impact resources?



2. Improving Products and Services

- New data files (and linkages) for primary customers
- Better, more comprehensive, metadata
- Improvements in public dissemination
 - Enhancements to Website
 - More Public Use microdata
 - Rethinking publications



3. Leveraging Technologies

- Greater use of IRS population files
- Leveraging expanded e-filing
 - Scanning residual paper
- Internalizing sampling operations
- Developing smart data edits



4. *Developing Human Capital*



- Is there a more efficient organizational structure?
- What critical job skills and career paths will we need?
- How to attract, develop, and retain the right people?
- Can we expand the *location of work* and be more people/family friendly?
- How to recruit and retain the best diverse workforce?

Thanks for your attention!

For questions about SOI data or services—

Contact our SIS Office:

Phone: 202-874-0410

E-mail: sis@irs.gov



Or contact —

Tom Petska, SOI Director

PO Box 2608

Washington, DC 20013-2608

thomas.b.petska@irs.gov