

Special Studies in Federal Tax Statistics

2005



Selected Papers Given in 2005
at the Annual Meetings of the
American Statistical Association
and the National Tax Association

Compiled and Edited by
James Dalton and Beth Kilss*
Statistics of Income Division
Internal Revenue Service

*Prepared under the direction of Janet McCubbin, Chief, Special Studies Branch

PREFACE

This edition in the IRS Methodology Report series, *Special Studies in Federal Tax Statistics*, includes papers presented at the 2005 Annual Meetings of the American Statistical Association (ASA) held August 7-11, 2005, in Minneapolis, Minnesota, and at the National Tax Association (NTA) Conference held November 17-19, 2005, in Miami, Florida.

This year's compilation has been divided into seven areas of interest:

- ❑ The volume begins with three papers -- one on analyzing business organizational structure from tax data, one on current research in the nonprofit sector, and one on geographic variation in filing rates for Schedule H, the IRS form used to report social security and medicare wages paid to household employees;
- ❑ The second section presents a paper on Schedule M-1 corporate book-tax difference data, 1990-2003;
- ❑ The third section presents a paper on the effects of taxation on corporate financial policy;
- ❑ The fourth section contains three papers on measuring nonsampling error in the SOI Individual Tax Return Study; how imputed returns on the Corporate File compare to actual returns; and the impact of followup on Tax Year 2002 Foreign Tax Credit Data;
- ❑ The fifth section includes four papers on cluster analysis in describing tax return data; comparing income concepts at IRS, Census, and BLS; the 1999-2003 Statistics

of Income Individual Income Tax Return Edited Panel; and trends in 401(k) and IRA contribution activity, 1999-2002;

- ❑ The sixth section presents a paper on the Estate and Personal Wealth Sample design; and
- ❑ The final section presents a paper on IRS area-to-area migration data.

Nine of the articles in this volume were prepared for publication in the *2005 Proceedings of the American Statistical Association*. Therefore, the format conforms basically to that required by the ASA, with the exception that we have not imposed a strict page limitation. Hence, in some cases, additional explanatory material may be included that is not available in the Proceedings.

The contents of the papers included here are the responsibility of the authors. Views expressed in these papers are those of the authors and do not necessarily represent the views of the Treasury Department or the Internal Revenue Service.

◆ Acknowledgments

The editors of this collection, James Dalton and Beth Kilss, would like to thank Lisa Smith and Dorothy Wallace for their invaluable contribution in laying out the papers in this volume, and Bobbie Vaira for her assistance in the publishing process.

Thomas B. Petska
Director
Statistics of Income Division
Internal Revenue Service

October 2006

Special Studies in Federal Tax Statistics 2005

CONTENTS

	PAGE
Preface	iii
1 ▼ NEW RESEARCH FROM THE IRS	
An Analysis of Business Organizational Structure and Activity from Tax Data, <i>by Tom Petska, Michael Parisi, Kelly Luttrell, Lucy Davitian, and Matt Scoffic</i>	9
Current Research in the Nonprofit Sector, <i>by Paul Arnsberger, Melissa Ludlum, and Margaret Riley</i>	45
Geographic Variation in Schedule H Filing Rates: Why Should Location Influence the Decision To Report "Nanny" Taxes?, <i>by Kim M. Bloomquist and Zhiyong An</i>	57
2 ▼ CORPORATE TAX ISSUES: BOOK-TAX DIFFERENCES AND MEASURING TAX AVOIDANCE	
Prelude to Schedule M-3: Schedule M-1 Corporate Book-Tax Difference Data, 1990-2003, <i>by Charles Boynton, Portia DeFilippes, and Ellen Legel</i>	71
3 ▼ BEHAVIORAL RESPONSES TO CORPORATE TAXATION	
An Essay on the Effects of Taxation on the Corporate Financial Policy, <i>by George Contos</i>	95
4 ▼ ISSUES IN QUANTIFYING MEASUREMENT ERROR	
Measuring Nonsampling Error in the Statistics of Income Individual Tax Return Study, <i>by Jana Scali, Valerie Testa, Maureen Kahr, and Michael Strudler</i>	115
Corporation Supercritical Cases: How Do Imputed Returns on the Corporate File Compare to the Actual Returns?, <i>by Lucy Davitian</i>	123
The Impact of the Followup Process on the 2002 Foreign Tax Credit Study Data, <i>by Rob Singmaster and Lissa Redmiles</i>	131

5 ▼ INTERESTING METHODOLOGICAL TOPICS RELATED TO INTERNAL REVENUE SERVICE TAX STATISTICS

A Cluster Analysis Approach To Describing Tax Data, *by Brian G. Raub and William W. Chen*..... 141

A Comparison of Income Concepts: IRS Statistics of Income, Census Current Population Survey, and BLS Consumer Expenditure Survey, *by Eric L. Henry and Charles D. Day* 149

The 1999 Individual Income Tax Return Edited Panel, *by Michael E. Weber and Victoria L. Bryant*..... 159

Trends in 401(k) and IRA Contribution Activity, 1999-2002--Results from a Panel of Matched Tax Returns and Information Documents, *by Peter Sailer, Victoria L. Bryant, and Sara Holden* 167

6 ▼ ESTATE AND PERSONAL WEALTH SAMPLE DESIGN

Origins of the Estate and Personal Wealth Sample Design, *by Paul B. McMahon, Jr.* ... 175

7 ▼ IRS AREA-TO-AREA MIGRATION DATA

Internal Revenue Service Area-To-Area Migration Data: Strengths, Limitations, and Current Trends, *by Emily Gross* 185

INDEX OF IRS METHODOLOGY REPORTS ON STATISTICAL USES OF ADMINISTRATIVE RECORDS.....191

Special Studies in Federal Tax Statistics: 2005 Online

Special Studies in Federal Tax Statistics: 2005 is available online on the IRS Internet site at: <http://www.irs.gov/taxstats/productsandpubs/article/0,,id=141315,00.html>. The papers included in this volume may also be found on the IRS web site according to the conference at which they were presented, i.e., ASA and NTA: <http://www.irs.gov/taxstats/article/0,,id=106270,00.html>.

1



New Research from the IRS

Petska ♦ Parisi ♦ Luttrell ♦ Davitian ♦ Scoffic
Arnsberger ♦ Ludlum ♦ Riley
Bloomquist ♦ An

An Analysis of Business Organizational Structure and Activity from Tax Data

*Tom Petska, Michael Parisi, Kelly Luttrell, Lucy Davitian, and Matt Scoffie
Internal Revenue Service*

► Introduction

Studies of businesses based on tax and information returns filed with the Internal Revenue Service (IRS) have generally focused on the financial activities or behaviors of one or more business legal or organizational types. The motives for these studies have generally been: (1) to examine and analyze data on one form of business over time, or (2) to examine the dynamics of shifting from one organizational form to another based on various factors, including incentives (or disincentives) in the Internal Revenue Code (IRC). Studies in IRS's Statistics of Income (SOI) Division have most often been the first type. This approach has contributed to the understanding of the effect of taxation on the business sector, but has not taken into consideration the dynamic and "zero sum" dimensions of business activity--that businesses conduct profit-seeking activities in a variety of legal modes, and that they examine various alternative forms of organizational structure to optimize growth and after-tax profits. The SOI Integrated Business Database (IBD) is being developed to provide evidence that businesses do, in fact, pursue optimal organizational structures. This initiative is an extension of earlier work in SOI, expanded to include Tax Years 1980-2002, incorporating the latest years for which complete SOI data are available.¹⁻⁸

This paper is divided into four sections. The first section briefly provides background information on the tax treatment of business income. The second section briefly summarizes major tax law changes that affected the taxation of business income in the period 1980-2002. The third section presents and analyzes data from annual SOI cross-sectional business studies, and the final section notes some conclusions and plans for future research.

► Taxation of Business Income

The tax treatment of the many organizational forms is complicated and varies considerably; so, only brief summaries of Federal taxation of business income are provided. The major legal forms of economic organiza-

tion are: corporations, partnerships, and nonfarm sole proprietorships.

Corporations--Corporations, in this analysis, are subdivided into those taxed at corporate rates (taxable or C corporations), and those electing to be taxed through their shareholders at individual income tax rates. The latter group includes Subchapter S corporations (or simply S corporations), Regulated Investment Companies (RICs), and Real Estate Investment Trusts (REITs), all of which are not taxed at the enterprise level but whose income similarly flows through to their owners, where it is subject to tax. C or taxable corporate income is generally taxed directly at the business level, then again at the shareholder level, at the applicable rates on dividend income. However, certain provisions in the Federal tax code lessen this effect. First, the corporate income potentially taxable at the shareholder level excludes the taxes paid by the corporation; so, income distributed to corporate shareholders is only taxable on the after-tax profits earned by the corporation. Second, the after-tax income of the corporation is not taxable at the shareholder level until it is paid out in dividends or until the shareholder realizes capital gains by selling shares that appreciated in value.

Subchapter S corporations are usually small, closely held corporations that are not taxed directly. With some exceptions, their incomes are subject to tax only at the owner level, much like the flowthrough treatment of partnerships. Owners of S corporations report their pro rata shares of income or loss on their own tax returns. Although S corporations have attractive features, they do face restrictions, including limitations on the number and type of shareholders and on the classes of stock permitted, and prohibition of foreign or corporate ownership. Similar to S corporations, the profits of RICs and REITs are not taxed at the enterprise level but flow through to their owners, where they are subject to tax.

Partnerships--Like an S corporation, a partnership serves as a conduit between a business and its owners,

in this case, its partners. The partnership entity is thus not taxed directly. Each partnership files an annual information return, which includes an income statement, balance sheet (in most cases), and a schedule of allocations or distributions made to each partner. Partners are predominately, though not exclusively, individuals who report their allocated shares of income and expenses on their own tax returns. Partnerships may be general partnerships, limited partnerships, or limited liability companies (LLCs). General partnerships, and general partners as well, face personal liability limited only by their personal resources and the applicable bankruptcy laws. Limited partners are more like corporate shareholders, with liability limited to the amount invested and with no active participation in management of the business.

A relative newcomer among for-profit businesses is the limited liability company, or LLC. These entities have the limited liability of corporations, but are taxed in the partnership model--income and expenses flow through the LLC to the owners, who are taxed on their pro rata shares. Unlike S corporations, however, LLCs do not have the extensive restrictions on the number and composition of owners. LLCs report their financial activities on their applicable business tax forms, most commonly the partnership information return (Form 1065), and indicate that they are filing as an LLC. The SOI partnership program began identifying these entities for Tax Year 1993. To provide some perspective on their prevalence and the scope of their financial activities, summary data on partnership LLCs are included in the next section.

Sole proprietorships--The profits of nonfarm sole proprietorships are taxed only at the personal (i.e., owner) level. The income statement of sole proprietorships, which summarizes the income and expenses of the business, is completed on Schedule C (or C-EZ) of the owner's individual income tax return. The net income or loss from the business is added to personal income from all other sources and taxed at the applicable individual income tax rates. In effect, the proprietorship also acts as a conduit through which the income of the business is passed through to the business owner where it is subject to tax.

Summary--While it is generally presumed that all corporate income is subject to double taxation, at both the entity and shareholder levels, the profits of S corporations, RICs, and REITs are all untaxed at the entity level and flow through to the owners or shareholders, similar to the treatment for partnerships. As a result, in the third section of the paper, we examine profits for each organizational type and subsequently aggregate data from all entities with flowthrough characteristics (including proprietorships) and compare them to C corporations that are taxed directly and whose incomes are potentially subject to double taxation.

► Tax Law Changes

The Tax Reform Act of 1986 (TRA86), the most comprehensive revision of the Internal Revenue Code since 1954, had a major impact on business decisions in the period after 1986 through broadening of the tax base of both individuals and corporations, tightening the corporation "alternative minimum tax," limiting losses from passive activities, and repealing the long-term capital gain exclusion. The most marked effect has been on the changes made to the individual and corporate marginal tax rates. In pre-TRA86, the highest individual rate (50 percent) exceeded the highest corporation rate (46 percent) by 4 percentage points. TRA86 reversed this trend, starting in 1987 and continuing with the phase-in of lowered rates in 1988-1990 of 34 percent for corporations and 28 percent for individuals. However, for 1991 and 1992, this difference between the corporate and individual marginal rates was cut in half when the top rate for the latter was increased to 31 percent.

Beginning for Tax Year 1993, the top individual rate increased to 39.6 percent, surpassing the rate of 35 percent for the highest corporation incomes, and restoring the pre-TRA relationship where the highest individual rate exceeded the top corporate rate. In fact, the difference of 4.6 percentage points between the individual rate and the corporation rate is similar to the pre-TRA86 difference of 4 percentage points, providing a reversal of the post-TRA incentive to switch to business types taxed solely at the individual level. However, this incentive declined with the lowering of top individual rates beginning for 2001.

The Small Business Job Protection Act of 1996 (SBJPA) made several noteworthy changes that affected S corporation filings. First, the Act increased the maximum number of shareholders from 35 to 75. Second, it enabled financial institutions that did not use the reserve method of accounting for bad debts to make an S election. Third, small business trusts electing to be S corporations were permitted to be shareholders in an S corporation. Finally, restrictions on the percentage of another corporation's stock that an S corporation might hold were eliminated, enabling S corporations to make an election to treat the assets, liabilities, income, deductions, and credits of wholly owned subsidiaries as those of the parent S corporation.

Even though the SBJPA eased restrictions on S corporations, the number of S corporation entities has not grown as rapidly as partnership limited liability companies (LLCs). The IRS ruled in late 1988 (Revenue Ruling 88-76, 1988-2 C.B.360) that any Wyoming LLC would be treated as a partnership, and the door was opened for other States to consider LLC legislation. By 1993, 36 States allowed LLCs as a legal entity, and that number grew to 46 States plus the District of Columbia a year later. By 1997, all 50 States and the District of Columbia had enacted LLC legislation. The "check-the-box" regulations, implemented by IRS in January 1997, relaxed the requirements for LLCs to obtain a favorable partnership tax classification, leading to a wider acceptance of LLCs.

► Analysis of Business Data

The SOI Integrated Business Dataset (IBD) has been compiled at the table level from the annual SOI cross-sectional studies of corporations (C and S corporations), partnerships, and nonfarm sole proprietorships for 1980-2002.⁹ Data from these annual statistical studies are generally publicly available and are published in a variety of SOI reports. (See the References section.) They represent weighted estimates of U.S. totals by year for each legal form or organizational type. The database combines data from these types of organizations for a 22-year period to enable examination of changes in business composition. The IBD is composed of 3 subsets; (1) selected financial data on businesses for all industries for 1980-2002 (Table 1); (2) selected financial data by

size of business receipts for 1998-2002 (Tables 2A-2E); and selected financial data on businesses for 21 North American Industrial Classification System (NAICS) sectors for 1998-2002 (Tables 3A-3E). Although some of the data in the IBD have already been published, this is the first time that they have been compiled for this duration, and work on analysis of significant trends and findings is just beginning.¹⁰

This section is divided into three parts. First, summary data by organizational type for 1980-2002 are presented and analyzed. In the next two subsections, trends in the data between 1998 and 2002 by receipt size and industrial sector are examined. The period for the industry data has been restricted since, beginning with 1998, all SOI business studies adopted the new NAICS industrial classification system. Previously, SOI business studies, and most economic statistics produced by Federal agencies, used an industry coding system based on the Standard Industrial Classification (SIC) System. Although NAICS has substantially improved coverage on newer, emerging industries, there is a major discontinuity between 1997 and 1998, and, for some industries, it is difficult or even impossible to derive a consistent time series.

► Data for All Industries, 1980-2002

The all-industry data compiled and discussed in this section include: the number of entities, total and business receipts, net income (less deficit), net income, and deficit. Although this is limited financial detail, these data comprise a consistent time series for the 22-year period for all types of businesses. Table 1 presents these data in its most detailed format, while Figures A-G highlight some of the most significant trends.¹¹

Number of Business Entities--The number of businesses doubled between 1980 and 2002, from 13 million in 1980 to over 26 million in 2002. Overall, the growth was relatively steady, with increases in all years, including even those with declines in real GDP (1980-1982, 1990-1991, and 2000-2001). However, unlike the steady overall growth in the number of entities, the composition of businesses by organizational type varied considerably. Figure A shows the percent-

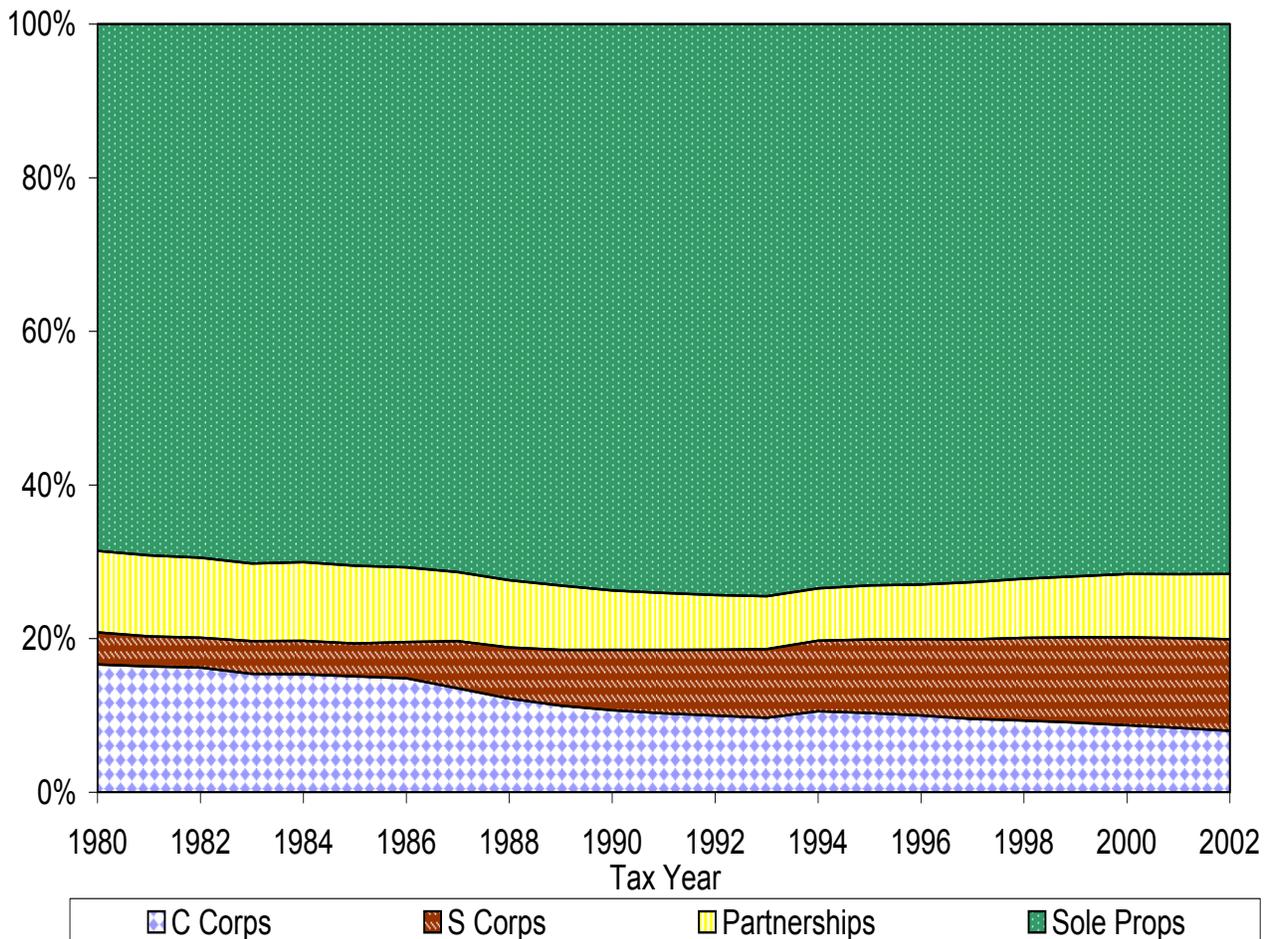
age composition in the number of business entities for C corporations, S corporations, partnerships, and sole proprietorships.

Sole proprietorships were the largest and most stable component of business entities, accounting for between 68.6 percent and 74.5 percent of overall business entities in all years and growing by 3 percentage points in the 22-year period, from 68.6 percent in 1980 to 71.6 percent in 2002. C corporations, on the other hand, accounted for 16.6 percent of business entities in 1980, but their percentage fell steadily to 8.0 percent in 2002. S corporations accounted for only 4.2 percent

of business entities in 1980, but their share increased substantially, particularly in the period following the 1986 Tax Reform, to 11.9 percent in 2002. Partnerships were also a relatively stable portion of the business entity types, declining modestly from 10.6 percent in 1980 to 8.5 percent in 2002. While the number of partnerships increased between 1980 and 1988, their proportion of the overall number of business entities declined, mainly due to the higher growth rates of S corporations and proprietorships.

Figure B presents annualized growth rates in the number of business entities with some additional detail

Figure A--Composition of the Number of Businesses, Tax Years 1980-2002



by business organizational type.¹² Overall, the number of businesses increased at a 3.2-percent annual rate for the 22-year period, but this percentage varied by business type. Although the total number of corporations showed an annual 3.0-percent increase, this was composed of a -0.1-percent annual decline for C corporations and a robust 8.0-percent annual increase by S corporations. C corporations had 2-percent annual increases in 1980-1987 and 1993-1997 but declines in both 1987-1993 and 1997-2002. S corporations increased in all periods, though the annual rate of increase declined steadily from 10.4 percent in the 1980-1987 period, to 6.4 percent for 1993-1997, and 5.0 percent for 1997-2002. Partnerships had an overall 2.2-percent growth rate for the 22-year period but declined in number between 1987-1993 before restoring growth between 4 percent to 5 percent for the later periods. Complete data for all types of partnerships are unavailable for years prior to 1993 but indicate a clear pattern between 1993 and 2002. In these years, general partnerships declined in number at an increasing rate, while limited partnerships grew at increasing rates. However, these data are dominated by the 75.1-increase for LLC's in the 1993-1997 period, which slowed

considerably but still grew at a robust 19.9 percent for 1997-2002. As noted, sole proprietorships were the most stable entity type with an overall rate of growth of 3.4 percent, which was comprised of an annual growth rate of 5.5 percent for 1980-1987 that steadily declined to 1.9 percent for 1997-2002.

Since most types of business income are essentially taxed at the individual level, a total for all business types other than C corporations was computed and is also shown in Figure B. This aggregation includes the data for 1120-RICs, 1120-REITs, S corporations, all types of partnerships, and sole proprietorships--essentially, all business organizational forms except for C corporations. Since proprietorships dominate the statistics on the number of business entities and were also a relatively stable component, it is not surprising that the growth pattern for the aggregation of businesses less C corporations mirrored that of proprietorships. These entities grew at an annual rate of 3.7 percent for the entire period, and the rate of growth steadily declined from 5.4 percent for the earliest period (1980-1987) to a low of 2.6 percent for 1997-2002. However, they avoided the reductions

Figure B--Annual Growth Rates for the Number of Businesses, Tax Years 1980-2002

Form of business	Annual Growth Rates (Percent)				
	Total interval, 1980 to 2002	Tax Years			
		1980 to 1987	1987 to 1993	1993 to 1997	1997 to 2002
	(1)	(2)	(3)	(4)	(5)
All business types	3.2	4.9	2.5	2.6	2.2
Corporations	3.0	4.1	1.6	4.3	2.2
C corporations	-0.1	2.0	-3.2	2.2	-1.4
1120-RIC and 1120-REIT	9.0	11.5	10.6	7.8	4.4
S corporations	8.0	10.4	8.7	6.4	5.0
Partnerships	2.2	2.5	-1.9	4.5	4.9
General	(1)	(1)	(1)	-2.1	-5.0
Limited	(1)	(1)	(1)	4.3	6.5
LLC	(1)	(1)	(1)	75.1	19.9
Sole proprietorships	3.4	5.5	3.2	2.0	1.9
Total less C corporations	3.7	5.4	3.2	2.7	2.6

¹ Data not available for all years.

in numbers that C corporations had in both 1987-1993 and 1997-2002.

Business Receipts--Unlike data on the numbers of business entities, the business receipts data include double counting, since intercompany sales and purchases are included. However, they are still an important metric of business activity by organizational type. Data on the composition and growth of business receipts by type of entity are presented in Figures C and D, respectively. C corporations dominated business receipts for the 22-year period, although their share has declined throughout the period from a high of 87.5 percent for 1981 to 64.9 percent for 2002.

So, where did this share of C corporation business receipts go? First, S corporations increased their share of receipts from about 3 percent for the 1980-1982 period to 18.5 percent for 2002. Although the rate of growth

was steady for most years, between 1986 and 1987, the S corporation share jumped from 5.5 percent to 10.1 percent in this one year, with enactment of the 1986 Tax Reform Act, which lowered the top marginal rate on business income taxed at the individual rate in comparison to the top marginal tax rate on corporate profits. Although the share of business receipts accruing to proprietorships declined from 6.4 percent to 5.0 percent in the period, the share of partnerships grew from 3 percent - 4 percent in the earliest years to 11.6 percent for 2002.

As shown in Figure D, overall business receipts grew at an annual rate of 5.3 percent over the 22-year period, peaking at 7.5 percent for 1993-1997.¹² Similarly, corporation receipts grew at a 5.0-percent annual rate for the entire period and also peaked in the 1993-1997 period at 7.1 percent. Although C corporations held the dominant share of receipts, receipts of S corporations grew at a 13.3-percent rate throughout the period, peaking at 21.9

Figure C--Composition of Business Receipts, Tax Years 1980-2002

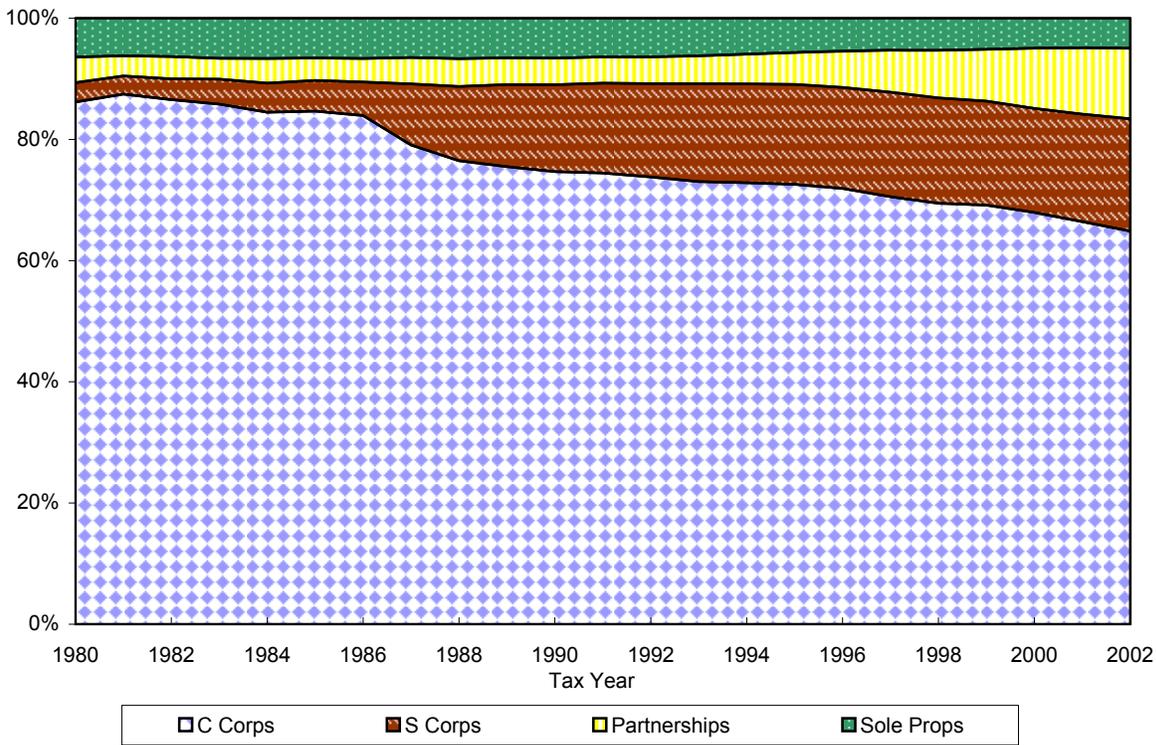


Figure D--Annual Growth Rates for Business Receipts, Tax Years 1980-2002

Form of business	Annual Growth Rates (Percent)				
	Total interval, 1980 to 2002	Tax Years			
		1980 to 1987	1987 to 1993	1993 to 1997	1997 to 2002
	(1)	(2)	(3)	(4)	(5)
All business types	5.3	5.5	4.3	7.5	4.6
Corporations	5.0	5.5	4.3	7.1	3.6
C corporations	4.0	4.3	2.9	6.7	2.9
S corporations	13.3	21.9	12.1	9.2	6.0
Partnerships	9.9	6.0	5.2	17.8	15.0
General	(1)	(1)	(1)	6.4	0.7
Limited	(1)	(1)	(1)	18.1	15.7
LLC	(1)	(1)	(1)	90.7	26.7
Sole proprietorships	4.2	5.7	3.6	3.5	3.4
Total less C corporations	9.6	11.4	8.5	9.7	8.1

¹ Data not available for all years.

percent between 1980-1987 before steadily declining. Partnerships had an overall 9.9-percent rate of growth in business receipts for the 22-year period, which was led by increases of 17.8 percent and 15.0 percent during the 1993-1997 and 1997-2002 periods, respectively. As for the entity data, the growth in partnership data was led by the increases for LLC's, which had 90.7-percent and 26.7-percent annual growth rates for the periods 1993-1997 and 1997-2002, respectively. Proprietorships exhibited the most stable growth, with an overall rate of 4.2 percent, which started at 5.7 percent in the 1980-1987 period and declined steadily to 3.4 percent in the latest years. Unlike for the number of entities, proprietorships do not dominate the receipts data; so, the pattern for the total excluding C corporations was much more like those for S corporations and partnerships, with 9.6-percent growth throughout, ranging from 11.4 percent in the earliest period and staying above 8 percent for all later periods.

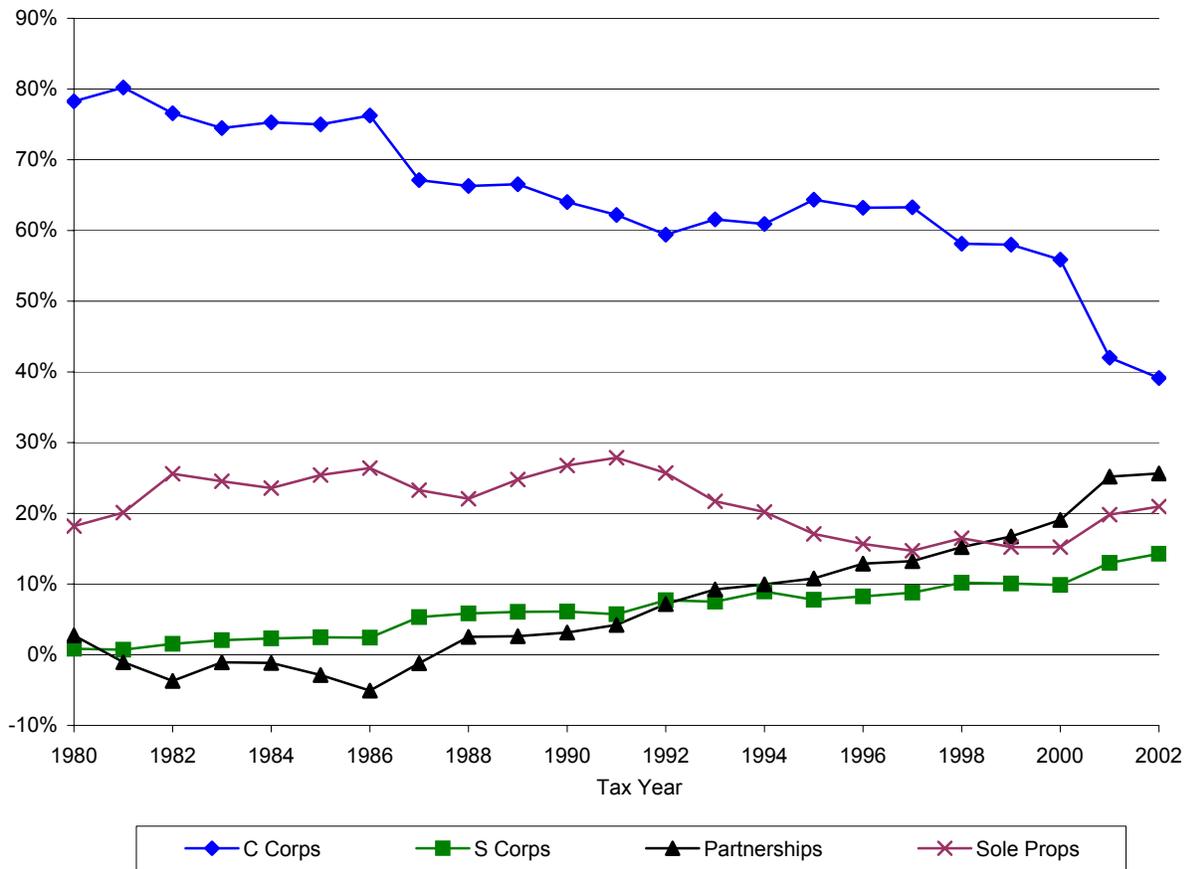
Net Income (Less Deficit)--Figures E and F show data on the composition and growth of net income (less deficit), respectively.¹³ Overall, as for business receipts, data for net income (less deficit) show the dominance of C corporations, although their share of the total declined precipitously, plummeting from 80 percent for 1980-

1981 to 39.1 percent for 2002. This is a very significant turn of events since revenue from the corporation income tax has been a significant component of overall tax collections.¹⁴ This phenomenon is even more noteworthy considering the relative stability of corporate statutory tax rates in the post-TRA period.

Once again, profits of proprietorships were the most stable of any entity type, increasing from 18.2 percent for 1980 to 20.9 percent for 2002; however, the proprietorship share had increased to 25.6 percent for 1982 and stayed above 20 percent through 1994 before bottoming out in 1997. The flowthrough entities, S corporations and partnerships, together accounted for less than 2 percent of net income (less deficit) for 1981-1986, partly because partnerships had losses in all of these years. However, beginning with 1987, their combined net income (less deficit) grew rapidly from about 4 percent for 1987 to nearly 40 percent for 2002, a tenfold increase in just 15 years.

Concerning the growth rates for net income (less deficit), overall business had profits increasing at increasing rates in all of the pre-1997 periods before falling at a 3.7-percent annual rate in the 1997-2002 period, largely due to corporate profit declines in the 2001-2002 eco-

Figure E--Composition of Business Net Income (Less Deficit), Tax Years 1980-2002



conomic downturn.¹² C corporation profits had a similar, though more prominent trend, with steady increases peaking at 12.5 percent for the 1993-1997 period before falling at an annual 17.1-percent rate for 1997-2002. The flowthrough entities, S corporations and partnerships, both had substantial growth in profitability, with overall 19.5-percent and 15.9-percent annual rates of growth throughout the 22-year period, respectively. S corporation profits increased at over 32 percent for the 1980-1987 period and stayed in the double-digit range, until dropping to a modest 3.6-percent rate of increase for 1997-2002. Partnership had overall losses from 1981 through 1987, became profitable in 1988, and then had increases of over a 20-percent level for 1993-1997, before dropping to 9.5 percent for 1997-2002.

Once again, proprietorships were the most stable component experiencing overall growth in profits of

6.3 percent for the entire period, with growth of 9.3 percent for 1980-1987 that steadily declined to 3.4 percent for the 1997-2002 period. For entities excluding C corporations, profitability growth patterns mirrored a combination of the rapid profit growth in the earlier periods of the flowthrough entities with the greater stability of proprietorships. Overall, profit growth was 11.5 percent for the entire 22-year period, with double-digit growth through 1997 before declining to 3.3 percent for 1997-2002.¹⁵

Deficits--Information on business losses or deficits is shown in Figures G and H for all entity types. C corporation losses ranged from about 48 percent to just under 63 percent for the entire period, substantially lower than the percentages for receipts and profits. The only years that C corporation losses exceeded 60 percent of the total were for the last 3 years, 2000-2002, a period

Figure F--Annual Growth Rates for Business Net Income (Less Deficit), Tax Years 1980-2002

Form of business	Annual Growth Rates (Percent)				
	Total interval, 1980 to 2002	Tax Years			
		1980 to 1987	1987 to 1993	1993 to 1997	1997 to 2002
	(1)	(2)	(3)	(4)	(5)
All business types.....	5.8	5.8	8.0	14.5	-3.7
Corporations.....	4.2	4.5	7.4	15.7	-9.4
C corporations.....	0.4	0.8	6.4	12.5	-17.1
1120-RIC and 1120-REIT.....	10.7	18.4	5.7	24.0	-4.8
S corporations.....	19.5	32.3	16.8	20.9	3.6
Partnerships.....	15.9	(²)	(²)	23.1	9.5
General.....	(¹)	(¹)	(¹)	11.8	2.7
Limited.....	(¹)	(¹)	(¹)	42.8	13.1
LLC.....	(¹)	(¹)	(¹)	104.3	20.9
Sole proprietorships.....	6.3	9.3	6.6	4.4	3.4
Total less C corporations.....	11.5	11.7	14.9	16.5	3.3

¹ Data not available for all years.

² Value not computed due to negative values.

Figure G--Composition of Business Losses, Tax Years 1980-2002

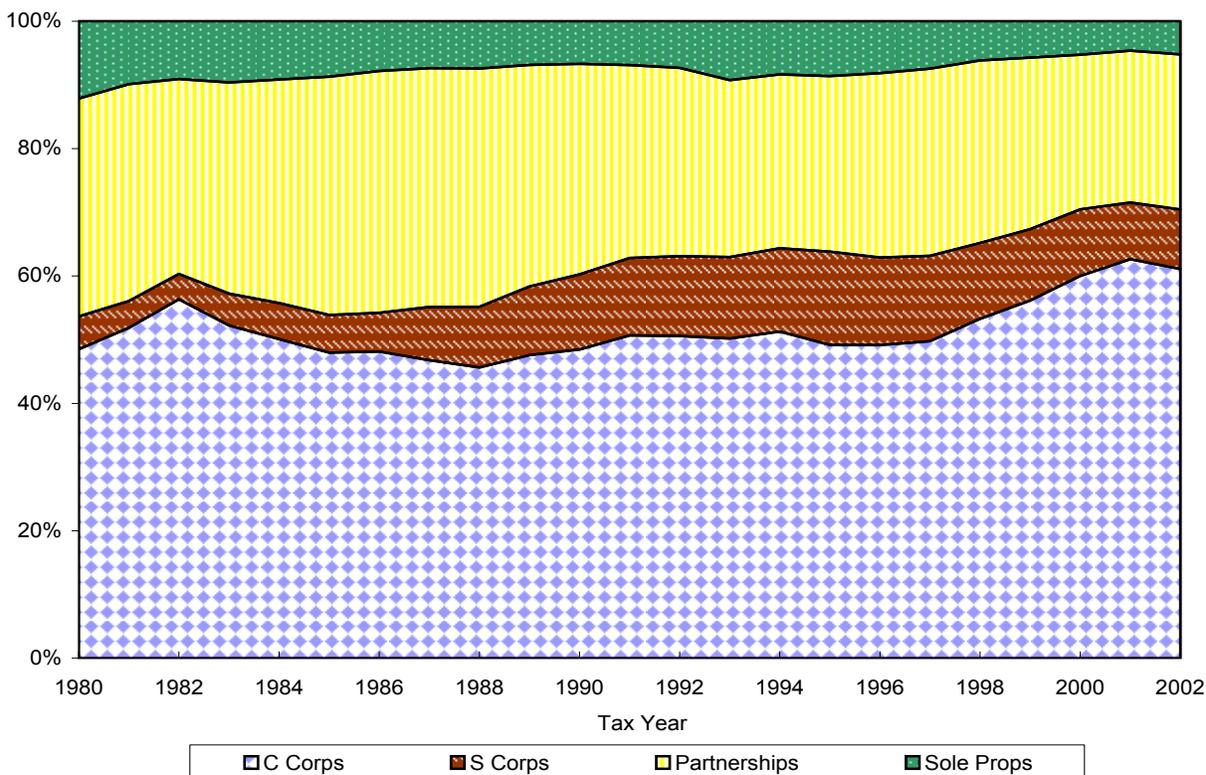


Figure H--Annual Growth Rates for Business Losses, Tax Years 1980-2002

Form of business	Annual Growth Rates (Percent)				
	Total interval, 1980 to 2002	Tax Years			
		1980 to 1987	1987 to 1993	1993 to 1997	1997 to 2002
	(1)	(2)	(3)	(4)	(5)
All business types.....	8.5	8.7	4.2	5.4	15.7
Corporations.....	9.7	12.3	2.6	5.4	18.0
C corporations.....	9.5	11.4	1.5	5.4	19.4
1120-RIC and 1120-REIT.....	23.6	15.0	29.7	15.8	34.5
S corporations.....	11.0	18.9	7.3	4.7	9.6
Partnerships.....	6.9	13.3	-4.6	7.1	11.7
General.....	(1)	(1)	(1)	-4.6	-0.2
Limited.....	(1)	(1)	(1)	3.6	4.3
LLC.....	(1)	(1)	(1)	83.7	26.7
Sole proprietorships.....	4.6	4.8	4.2	0.3	8.4
Total less C corporations.....	7.3	5.6	7.4	5.4	11.1

¹ Data not available for all years.

that included three quarters of decline in real GDP. Other recessionary periods seemed to have had less effect on the C corporation share of losses. S corporation losses grew starting after 1980, peaking in 1995 at 14.6 percent, before beginning a steady decline to around 9 percent for 2001 and 2002.

Interestingly, partnerships have had a substantial share of deficits throughout the 22-year period, growing from the mid-30 percents in the pre-TRA period, peaking at 47 percent for 1987 and 1988, before beginning a gradual decline to the low 20-percent range in the 2000-2002 period. Clearly, the TRA passive loss limitations had an effect. Proprietorships once again held a stable but small share of losses, which peaked for 1980, and gradually declined throughout the period to about 5 percent for the 2000-2002 period.

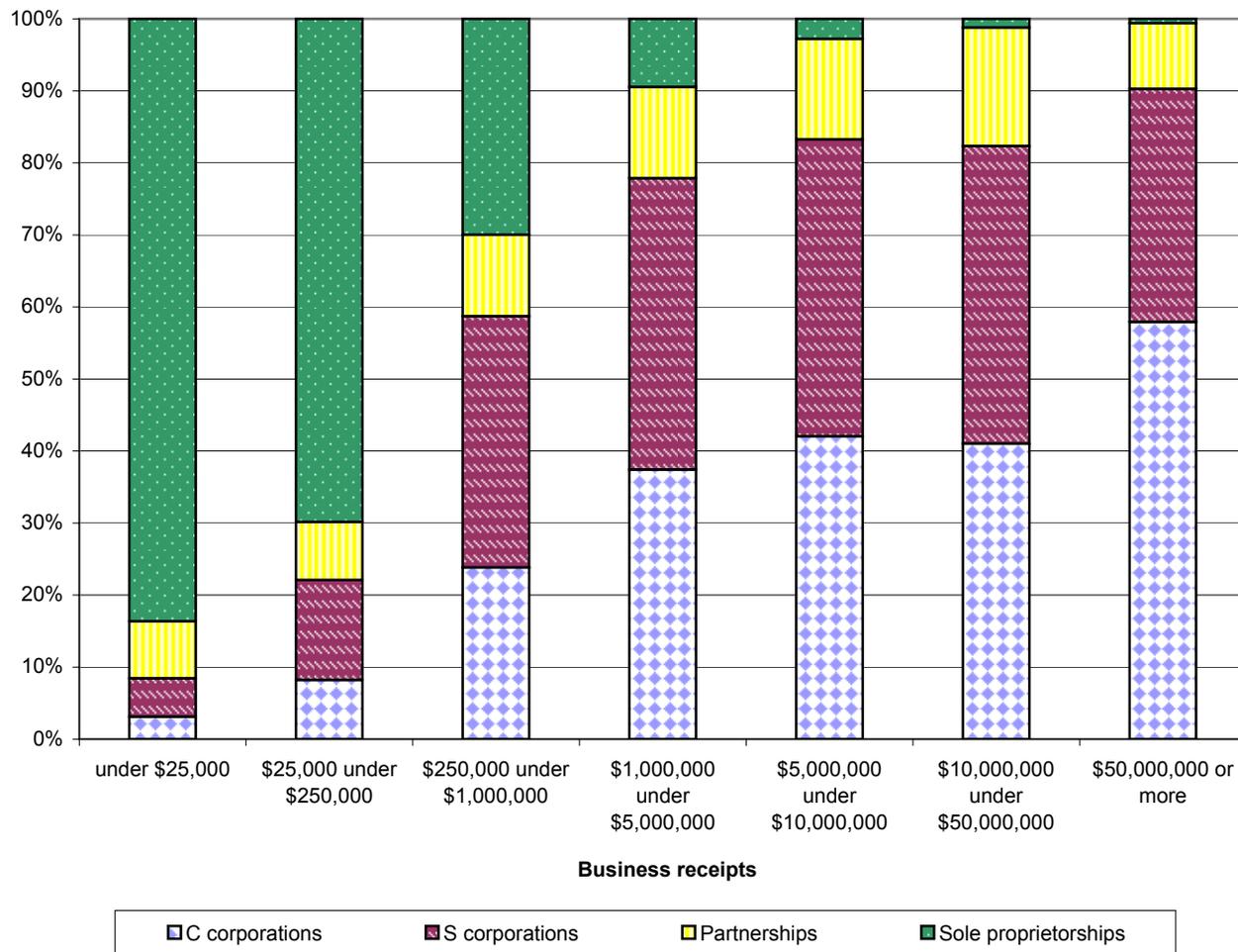
From a growth perspective, overall losses, which increased at nearly 9 percent in the 1980-1987 period, declined to around 5 percent from 1987-1997, then jumped to over 15 percent in the 1997-2002 period.¹² C corporations had a similar pattern, though growth in deficits was larger in periods of large deficit growth and smaller in periods when deficits grew at slower rates, im-

plying more stability for the other types of entities. For businesses other than C corporations, losses averaged 7.3 percent over the entire period, ranging between 5 percent and 7 percent during 1980-1997 before increasing to 11.1 percent for the 1997-2002 period. S corporations had an 18.9-percent increase for 1980-1987, but the growth in losses dropped for 1987-1993 and again for 1997-1997 before increasing to nearly a 10-percent rate for 1997-2002. For partnerships, losses increased in all periods, with the exception of the 1987-1993 period, where the post-TRA passive loss limitations disallowed an increasing share of partnership losses to offset other (positive) income.

► **Data by Size of Business Receipts, 1998-2002**

In this section, we focus on business activity during the period of 1998 through 2002 by size of business receipts. As noted, selected financial data by size of business receipts for 1998-2002 are included in Tables 2A-2E.¹¹ When the data are segmented by size of business receipts, some notable characteristics of business composition are apparent. Composition percentages on the number of businesses by size of business receipts

Figure I--Composition of Number of Businesses by Size of Business Receipts, Tax Year 2002



are shown for Tax Year 2002 in Figure I, while business receipts and net income (less deficit) by size of business receipts are shown in Figure J.

Overall, the numbers of business entities are dominated by small proprietorships, particularly those with receipts under \$1 million. C corporations, on the other hand, comprise less than 25 percent of business entities for each size-class under \$1 million, but their share grows from 37 percent to nearly 58 percent with increasingly larger receipt size-classes. The flowthrough entities, S corporations and partnerships, show their largest composition shares in the middle receipt size-classes. S corporations account for between 35 percent-41 percent

of entities for all classes between \$250,000 and \$50 million, and partnerships also have their largest composition percentages in these mid-sized receipt classes.

From Figure J, and as previously discussed, C corporations dominate activity in business receipts, accounting for nearly 65 percent of receipts for 2002. However, their share of receipts is strongly associated with size of receipts. The smallest C corporations account for only 2 percent of receipts, but this share grows rapidly to nearly 81 percent for businesses with \$50 million or more in business receipts. As with data on the numbers of entities, the flowthrough businesses show their largest composition shares in the middle size-classes, with their

Figure J--Business Receipts and Net Income (Less Deficit) by Size of Business Receipts, Tax Year 2002

[Money amounts are in billions of dollars]

Income item and type of business	Business Receipts							
	Total	under \$25,000	\$25,000 under \$250,000	\$250,000 under \$1,000,000	\$1,000,000 under \$5,000,000	\$5,000,000 under \$10,000,000	\$10,000,000 under \$50,000,000	\$50,000,000 or more
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Business receipts:								
All businesses.....	20,741.0	94.6	641.8	1,070.8	1,876.7	908.3	2,472.2	13,676.6
C corporations.....	13,455.8	2.3	72.3	275.5	732.0	378.9	930.3	11,064.5
S corporations.....	3,841.3	3.5	123.1	402.3	775.7	389.1	1,028.6	1,119.0
Partnerships.....	2,414.2	1.7	34.2	97.1	216.4	114.6	485.6	1,464.5
Sole proprietorships.....	1,029.7	87.2	412.2	295.8	152.5	25.7	27.7	28.5
Net income (less deficit):								
All businesses.....	1,055.4	-46.2	142.9	97.0	73.2	36.9	135.3	616.5
C corporations.....	413.0	-19.1	-8.4	-11.2	-10.4	-0.2	21.9	440.6
S corporations.....	150.6	-8.4	9.4	24.1	33.3	16.3	37.6	38.4
Partnerships.....	270.7	-34.9	13.6	25.1	35.7	19.5	74.9	136.9
Sole proprietorships.....	221.1	16.4	128.3	59.0	14.6	1.3	1.0	0.5

largest composition percentages in receipt size-classes between \$250,000 and \$50 million. Proprietorships, as would be expected, comprise the majority of small organizations, accounting for 92 percent of businesses with receipts under \$25,000 but with a rapidly diminishing share with increases in receipt size. For the largest size receipt size-class (\$50 million or more), proprietorships comprise only 0.2 percent of the total.

The composition of net income (less deficit) or profits among receipt sizes also shows some interesting and well-defined patterns. First, for the under \$25,000 receipt size-class, there was an overall \$46-billion loss for all types of businesses, and only proprietorships had positive net income. Although C corporations accounted for 39 percent of business profits for 2002, they show losses in all receipt size-classes below \$10 million. However, C corporations become profitable for size-classes over \$10 million, and those with receipts above \$50 million earned over \$440 billion in profits, nearly 42 percent of the total. S corporations once again show their largest composition shares in the middle receipt size-classes, with composition shares ranging from nearly 25

percent to almost 46 percent for businesses with receipts between \$250,000 and \$50 million. Partnerships had nearly \$35 billion in losses for the smallest size-class, but were profitable for all larger receipt size-classes. For receipt sizes above \$25,000, partnerships had profits of at least \$13 billion and accounted for 22 percent to 55 percent of total profits. Proprietorships, which include nearly 21 percent of overall profits, are the only business type with profitability in the under \$25,000 receipt size-class. Above \$25,000, proprietorships show a rapidly decreasing share of profits, with nearly 90 percent in the \$25,000-\$250,000 receipt size-class but only \$0.5 billion and 0.1 percent for the largest class.

► Data by Industrial Sector, 1998-2002

In this section, we focus on specific sectors that showed significant activity during the period 1998 through 2002. During this timeframe, a number of national and international events impacted economic activity, including the end of the uninterrupted GDP growth of the 1990's; the technology boom and bust; the September 11, 2001, attacks; real estate volatility;

accounting scandals; and enactment of the Small Business Job Protection Act of 1996. All of these potentially impacted business activity in specific sectors. As noted, selected financial data for 21 NAICS sectors for 1998-2002 are included in Tables 3A-3E and summary data for eight key sectors are presented in Figures K, L, and M and discussed below.

Utilities--As shown in Tables 3A-3E, the number of business entities in the Utilities sector decreased by 2.1 percent from 17,662 for 1998 to 17,283 for 2002. The most notable aspect of the decline was the 19.3-percent decrease in S corporation returns, from 2,124 to 1,715. The number of C corporations and partnerships classified as Utilities increased slightly, with only the large decline in S corporations and a slight decline in proprietorship Utilities, reducing the total for all businesses. The large decline in S corporation Utilities was mostly attributable to the smallest business receipt class, those returns with less than \$25,000 in business receipts.

The Utilities sector experienced a large decline in net income (less deficit) over the period, most of which was attributable to the largest receipt size-class for C corporations. C corporations reporting \$50 million or more in business receipts saw their net income (less deficit) decline from \$30.7 billion for 1998 to a loss of \$95.4 million for 2002. S corporations and partnership net income (less deficit) increased slightly both overall and in the largest receipt size-class.

Construction--The Construction industry accounted for roughly 12 percent of the total number of business entities. The number of businesses in this sector increased 4.8 percent over the 5-year period, from 2.9 million to 3.1 million. However, over the 1998-2002 period, the number of C corporations declined from 246,404 to 229,765 (6.8 percent), while the number of S corporations increased from 305,531 to 418,770 (37.1 percent).

Between 1998 and 2002, businesses showed significant increases in all data items, with the largest increases in S corporations, partnerships, and proprietorships. Business receipts of S corporations increased by 46.0 percent, from \$391.9 billion to \$572.1 billion; those of partnerships increased by 59.5 percent from \$106.3

Figure K--Number of Entities as Percent of Total by Selected Sector, Tax Year 2002

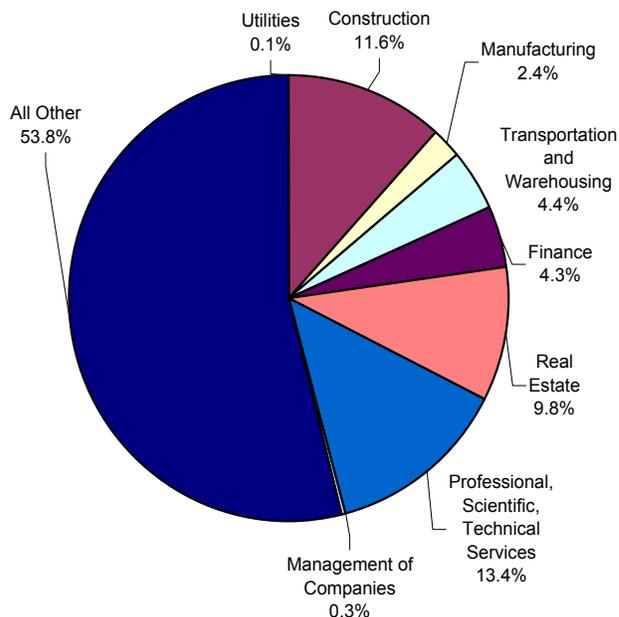


Figure L--Number of Entities as Percent of Total, by Selected Sector, Tax Year 2002

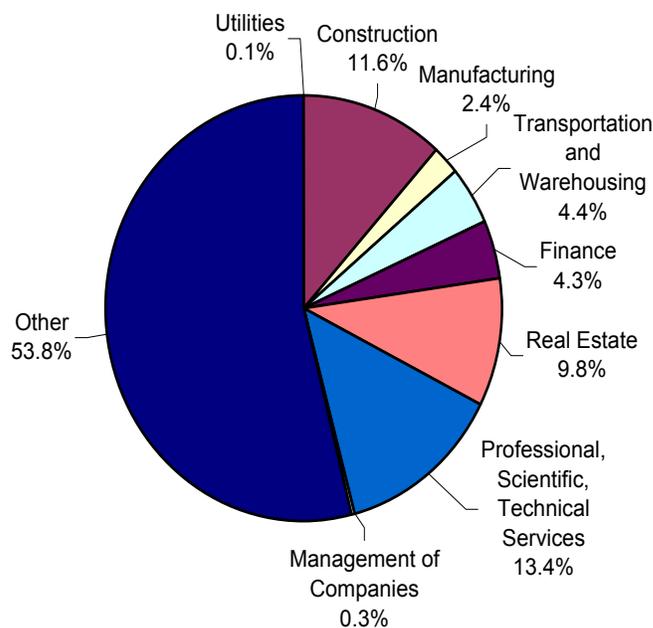
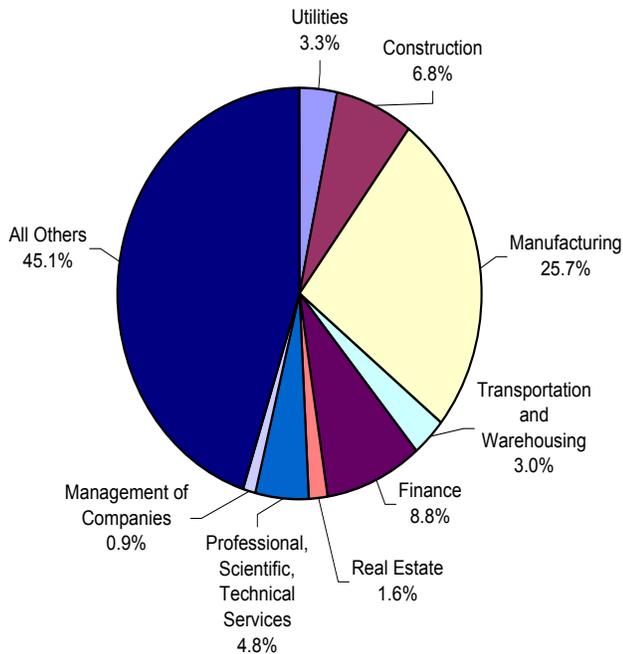


Figure M--Business Receipts as Percent of Total by Selected Sector, Tax Year 2002



billion to \$169.6 billion; and those of proprietorships increased by 17.0 percent, from \$143.9 billion to \$168.5 billion. Significant increases were also seen in salaries and wages of these entities, as well as in depreciation.

Manufacturing--For 1998, 706,002 businesses classified themselves in the Manufacturing sector. By 2002, the number had dropped to 628,868, a 10.9-percent decrease in business return filers for this sector. Of the four entity types, all declined in number with the exception of partnerships, which showed a 10.1-percent increase to 38,364. The increase in number of partnerships did little, however, to alter the distribution of partnerships among receipt size-classes. For 1998, 47.1 percent of partnerships classified in Manufacturing reported business receipts under \$100,000. For 2002, 45.2 percent of manufacturers still fell under this threshold.

C corporations and sole proprietorships accounted for most of the decline in the number of manufacturers. C corporations dropped by 27,141 (16.6 percent), and proprietorships dropped by 50,935 (14.1 percent). The

distribution of C corporation manufacturers across business receipt classes changed little from 1998 to 2002, with all classes but one (\$100,000 under \$250,000) showing decreases. Despite a decreasing number of sole proprietorships engaged in manufacturing, the period 1998–2002 saw growth in the number of large manufacturing proprietorships, with those reporting between \$5 million and \$50 million in business receipts increasing by 52.6 percent from 116 for 1998 to 177 for 2002. These changes in the manufacturing sector did little to change the composition of the sector, with each entity type making up roughly the same share of all Manufacturing for 1998 as for 2002.

Growth in business receipts for partnerships in Manufacturing exceeded that of partnerships in all sectors. Partnership business receipts in Manufacturing grew by 96 percent to \$485.0 million between 1998 and 2002. This growth could be traced to partnerships with \$50 million or more in business receipts. For 1998, 73.6 percent, or \$182.2 million, of business receipts of manufacturing partnerships were in the \$50 million or more business receipt size-class, while, for 2002, 81.4 percent, or \$394.9 million, were in this class.

Transportation and Warehousing--Growth in the overall number of business filers in this sector outpaced the growth of all sectors. The number of business entities classified in Transportation and Warehousing increased from 969,104 to 1,153,198, an increase of 19.0 percent. The number of each separate entity type increased over the period 1998–2002, but the largest percentage increases were seen in partnerships, S corporations, and proprietorships. Partnerships increased by 35.5 percent, or 6,814 returns; S corporations by 21.3 percent, or 17,290 returns; and proprietorships by 20.1 percent, or 159,181 returns. Although C corporations did show positive growth, their numbers increased by only 1 percent, from 78,342 for 1998 to 79,150 for 2002.

Well over half of all growth in Transportation and Warehousing partnerships can be traced to the smallest two receipt size-classes. The number of partnerships reporting \$100,000 or less in business receipts accounted for 59.5 percent, or 4,051, of new partnership returns in this sector. Sole proprietorships showed increases in all receipt size-classes, but growth was concentrated on the

lower end, with 99.3 percent, or 157,999, of new returns reporting less than \$250,000 in business receipts. S corporation growth was more evenly distributed among the various receipt size-classes. As with Manufacturing, the composition of the Transportation and Warehousing sector changed little. Of the 5 years studied, each entity's share of this sector remained relatively constant. Business receipts increased 13.6 percent to \$617.9 billion across all entities, while net income (less deficit) decreased 91.9 percent to \$2.5 billion over this period. Both C corporations and S corporations were responsible for the decrease in net income (less deficit).

Finance and Insurance--C corporations represent the majority of business income for the Finance and Insurance sector, while all other business entities combined represent 88.8 percent of all businesses in the sector. The number of C corporations declined over the period 1998-2002 by 12.0 percent, from 115,309 to 101,495. This decline was particularly noticeable in the smallest receipt size-classes. C corporations reporting less than \$25,000 in business receipts declined from 30,440 to 22,464. Partnerships reported the largest increase in number of businesses from 209,150 for 1998 to 263,024 for 2002, or 25.8 percent. Growth in the number of partnerships was also concentrated in smaller receipt size-classes, with the number of returns reporting less than \$25,000 in business receipts, increasing from 152,559 to 176,425.

Although net income (less deficit) for the Finance and Insurance sector declined from 1998 to 2002, partnerships were an exception. Net income (less deficit) for partnerships in this sector increased by 41.1 percent, from \$63.3 billion to \$89.3 billion. However, partnership net income (less deficit) represented only 25.2 percent of the \$354.8 billion in net income (less deficit) for all entity types for 2002.

Real Estate--The overall number of business entities in Real Estate increased 17.2 percent to 2,585,914 between 1998 and 2002. With this increase in the number of entities, there was also an increase of business receipts, which increased by 25.3 percent to \$326.4 billion. For all businesses, interest paid increased until 2002, when the overall interest paid declined by 26.0 percent from 2001 to \$19.6 billion.

The number of partnerships in Real Estate grew by 23.1 percent to 999,786 entities during the period 1998 through 2002, faster than any other entity type. Partnerships also displayed the largest amount of net income (less deficit) (\$55 billion) for the same time period, representing 68.0 percent of net income (less deficit) for all business entities. This growth could be traced to the \$5 million to under \$10 million class of business receipts, where net income (less deficit) increased from \$4.1 billion to \$8.0 billion, a 95.5-percent increase. C corporations were the only entity type in Real Estate to experience a decline in numbers. C corporation net income (less deficit) declined for the period 1998-2002, decreasing from \$4.9 billion in 1998 to almost -\$0.9 billion in 2002. Nearly all this decline was found in the C corporations reporting business receipts with \$50.0 million or more.

Professional, Scientific, and Technical Services--Overall, the number of businesses in the Professional, Scientific, and Technical Services sector showed a 12-percent increase, from 3.2 million for 1998 to 3.6 million for 2002. The increase was due to a 29.4-percent increase in S corporations, from 371,152 to 480,120, and a 9.9-percent increase for proprietorships, from 2.4 million to 2.7 million. Most of the growth for both S corporations and proprietorships could be traced to smaller receipt size-classes rather than to a single class.

For 2001, partnerships surpassed proprietorships as the leader in net income (less deficit), accounting for \$49.9 billion of the nearly \$93.2 billion reported for all business entities. Beginning for 1999, total net income (less deficit) for C corporations decreased to a \$4.5-billion loss and has remained negative for each year through 2002 when C corporations reported -\$19.7 billion. Despite this decline, C corporations continued to show the largest total receipts, business receipts, and total business deductions for this sector.

Since 1999, all entities excluding C corporations have displayed positive amounts for the total net income (less deficit), while C corporations displayed negative amounts for total net income (less deficit) during the same time period. Entities other than C corporations represented over 50 percent of all total receipts and business receipts for all business entities.

Management of Companies--The number of business entities in the Management of Companies (holding companies) sector increased 55.7 percent over the period 1998-2002, from 42,918 to 66,826 entities. However, one entity type, proprietorships, is not represented in this sector. S corporations displayed the largest percentage increase in number of businesses for this industry, 89.9 percent, an increase from 11,471 for 1998 to 21,779 for 2002. The largest increases were in smaller receipt size-classes, i.e., entities with business receipts under \$25,000 grew from 9,460 entities to 17,729 entities. This growth of S corporations can be attributed partly to the Small Business Job Protection Act of 1996, which permitted financial institutions that use the specific chargeoff method of Section 166 to account for the writeoff of bad debts to elect Subchapter S status. This provision has also led to a significant increase in the number of bank holding companies, which are also included in this sector.

Cost of goods sold for all Management of Companies more than tripled over the 5-year period of 1998-2002. C corporations nearly tripled their cost of goods sold for this period with an increase of \$7.2 billion, from nearly \$3.8 billion for 1998 to \$11.1 billion for 2002. Almost all of this growth was concentrated in C corporations with \$50 million or more in business receipts. C corporations in this class alone saw cost of goods sold rise from \$3.5 billion for 1998 to \$10.9 billion for 2002. Partnerships accounted for the largest percentage increase for cost of goods sold during this 5-year span, increasing 576.9 percent, to \$6.5 billion.

► **Conclusions and Plans for Future Research**

The most significant findings for the 22-year period are the shift in overall business activity away from C corporations to those organizations whose profits are taxed at the individual level. Overall, the data for net income (less deficit) show the dominance of C corporations, although their share of the total declined precipitously, plummeting from 80 percent for 1980-1981 to 39 percent for 2002. This is a very significant development since revenue from the corporation income tax has been a significant source of overall tax collections. This phenomenon is even more noteworthy considering the relatively stable corporation statutory tax rates, especially

in the post-TRA period. C corporations accounted for nearly 17 percent of business entities in 1980, but their percentage fell steadily to 8 percent in 2002. Although C corporations dominated business receipts, their share likewise declined throughout the period from a high of 87 percent in 1981 to 65 percent in 2002. Sole proprietorships were the largest and most stable component of business entities for this period, accounting for between 69 percent and 74 percent of overall business entities in all years. When the data are classified by size of business receipts, the largest number of entities fell into the smallest receipt size-class, but the vast majority of business receipts for most entity types generally accrued to those in the largest receipt class. C corporations dominated the receipts data in the largest class, accounting for approximately 80 percent of business receipts and nearly 72 percent of profits.

Although economic events affected different industrial sectors in very different ways, the data showed a particularly substantial trend in the 1998-2002 period. The data by industrial sector illustrated that the trend of shifting overall business activity away from C corporations to those organizations whose profits are taxed at the individual level was prevalent throughout all sectors of the economy. The most notable trend by industrial sector was the rapid growth in the number of businesses organized as flowthrough entities. In many industrial sectors, the number of C corporations grew very slightly or even declined. Across industrial sectors, almost without exception, S corporations and partnerships showed rapid growth in number of entities. S corporations showed large nominal increases, while partnerships typically grew at the fastest rates. In almost all sectors, the most notable growth in net income (less deficit) was also isolated in businesses organized as flowthrough entities.

Finally, opinions expressed in this paper are those of the authors and should not be attributed to the Internal Revenue Service or the U.S. Department of the Treasury although comments are welcome.

► **Endnotes**

¹ Legel, Ellen; Bennett, Kelly; and Parisi, Michael (2004), *The Effects of Tax Reform on the*

- Structure of U.S. Business, *2003 Proceedings of the American Statistical Association, Section on Government Statistics*.
- ² Petska, Tom (1998), Taxes and Business Organizational Choice: Deja Vu All Over Again? *1997 Proceedings of the American Statistical Association, Section on Business and Economic Statistics*.
- ³ Wittman, Susan M. and Gill, Amy, S Corporation Elections After the Tax Reform Act of 1986, presented at the 1996 Allied Social Science Meetings and published in *Statistics of Income Bulletin*, Spring 1998, Volume 17, Number 4.
- ⁴ Nutter, Sarah E.; Young, Jim; and Wilkie, Patrick, Tax Legislation and Business Form Choice: C Corporation Behavior Before and After TRA86, presented at the 1996 Allied Social Science Meetings and published in *Statistics of Income Bulletin*, Winter 1995-96, Volume 15, Number 3.
- ⁵ Petska, Tom, Taxes and Organizational Choice: An Analysis of Trends, 1985-1992, *Statistics of Income Bulletin*, Spring 1996, Volume 15, Number 4.
- ⁶ Petska, Tom, Do Taxes Affect Business Legal Structure? An Analysis of IRS Data, presented at the 1996 Allied Social Science Meetings.
- ⁷ Petska, Tom and Wilson, Robert, Trends in Business Structure and Activity, 1980-1990, *Statistics of Income Bulletin*, Spring 1994, Volume 13, Number 4.
- ⁸ Petska, Tom (1994), The Effects of Tax Reform on the Structure of U.S. Business, *1993 Proceedings of the American Statistical Association, Section on Business and Economic Statistics*.
- ⁹ As noted, the Integrated Business Dataset is a compilation of table level data from SOI cross-sectional business studies. Future plans are to construct a true Integrated Business Database consisting of microdata from SOI C and S corporations, partnerships, and nonfarm sole proprietorships.
- ¹⁰ Data on financial activity by size of business receipts by NAICS sectors are included in an extended version of Tables 2A-2E for this paper on the SOI Tax Stats Web site at <http://www.irs.gov/taxstats/bustaxstats/article/0,,id=152029,00.html>.
- ¹¹ In Table 1, Regulated Investment Companies (RIC's) and Real Estate Investment Trusts (REIT's), which are not taxed at the enterprise level but whose income similarly flows through to their owners, are excluded from C corporations and shown separately. However, in all other tables and figures, they are included with C corporations.
- ¹² Annual growth rates were computed as follows:
- $$G_t = (\ln X_t - \ln X_{t-n}) 100 / n$$
- where G_t = the annual growth rate in the value of X between periods t and n,
- $\ln X_t$ = the natural logarithm of the value of X for period t,
- $\ln X_{t-n}$ = the natural logarithm of the value of X for period t-n, and
- n = the number of years on which the computation is based.
- ¹³ Unlike data in the SOI *Corporation Income Tax Returns* and *Source Book of Corporation Income Tax Returns*, net income (less deficit) used in this paper includes the more comprehensive "total net income" for S corporations. This item includes trade or business income plus portfolio income, as well as real estate and rental activity incomes distributed directly to shareholders.
- ¹⁴ From Table 7 in the IRS 2004 *Data Book*, for 1980, the corporation income tax accounted for nearly 14 percent of total Internal Revenue collections. For 2002, this share had declined to about 10.5 percent.
- ¹⁵ In this paper, we assume that all partnership profits and losses accrue to individuals. However, from the Partnership Schedule K, data are available on distributions by type of partner. For 2002, \$156.1 billion, or 54.2 percent, of allocated income was

distributed to nonindividual partners (which include corporate, partnership, tax-exempt, and nominees). The *SOI Bulletin* article, Partnership Returns, 2002, referenced below, has additional information.

► **References**

- Internal Revenue Service, *Statistics of Income--Corporation Income Tax Returns*, Publication 16, 1980-2002.
- Internal Revenue Service, *Statistics of Income--Individual Income Tax Returns*, Publication 1304, 1980-2002.
- Internal Revenue Service, *Statistics of Income--Source Book of Corporation Income Tax Returns*, Publication 1053, 1980-2002.
- Internal Revenue Service, *Statistics of Income--Special Studies in Federal Tax Statistics*, various years.
- Luttrell, Kelly, S Corporation Returns, 2002, *Statistics of Income Bulletin*, Spring 2005, Volume 24, Number 4. (See <http://www.irs.gov/taxstats/productsandpubs/article/0,,id=130681,00.html> for *SOI Bulletin* articles and data on additional years.)
- Nelson, Susan C., S Corporations: The Record of Growth After Tax Reform, *Journal of S Corporation Taxation*, Fall 1993, Volume 5, Number 2.
- Nelson, Susan, and Petska, Tom (1990), Partnerships, Passive Losses, and Tax Reform, *1989 Proceedings of the American Statistical Association, Section on Survey Research Methods*.
- Pierce, Kevin, Sole Proprietorship Returns, 2003, *Statistics of Income Bulletin*, Summer 2005, Volume 25, Number 1. (See <http://www.irs.gov/taxstats/productsandpubs/article/0,,id=130681,00.html> for *SOI Bulletin* articles and data on additional years.)
- Petska, Tom, Partnerships, Partners, and Tax Shelters After Tax Reform, 1987-1989, *Statistics of Income Bulletin*, Summer 1992, Volume 12, Number 1.
- Petska, Tom and Nelson, Susan (1991), Partnerships and Tax Shelters: An Analysis of the Impact of the 1986 Tax Reform, *1990 Proceedings of the American Statistical Association, Section on Survey Research Methods*.
- Plesko, George A., Corporation Taxation and the Financial Characteristics of Firms, *Public Finance Quarterly*, July 1994.
- Scholes, Myron S., and Wolfson, Mark A. (1992), *Taxes and Business Strategy: A Planning Approach*.
- Wheeler, Timothy and Parsons, Maureen, Partnership Returns, 2002, *Statistics of Income Bulletin*, Fall 2004, Volume 24, Number 2. (See <http://www.irs.gov/taxstats/productsandpubs/article/0,,id=130681,00.html> for *Bulletin* articles and data on additional years.)

Table 1.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business, Tax Years 1980-2002

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Tax Year							
	1980	1981	1982	1983	1984	1985	1986	1987
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All Businesses								
Number of businesses.....	13,021,904	13,857,712	14,545,660	15,244,531	16,076,714	16,919,395	17,525,167	18,351,297
Total receipts.....	7,064,487,840	7,725,544,701	7,754,452,966	7,891,981,399	8,751,940,681	9,305,441,171	9,626,065,304	10,634,345,667
Business receipts.....	6,413,930,882	6,901,768,455	6,842,267,893	7,043,019,718	7,782,861,217	8,212,317,757	8,422,295,127	9,436,817,505
Net income (less deficit).....	316,874,165	263,985,693	197,592,719	246,063,040	300,167,182	310,007,924	342,583,143	434,130,755
Net income.....	424,569,277	420,560,759	396,557,182	435,858,670	508,725,907	539,687,640	599,572,585	680,068,330
Deficit.....	107,695,112	156,575,064	198,964,461	189,795,629	208,558,725	229,679,718	256,989,442	245,937,575
Corporations								
Number of businesses.....	2,710,538	2,812,420	2,925,933	2,999,071	3,170,743	3,277,219	3,428,515	3,612,133
Total receipts.....	6,361,284,012	7,026,351,839	7,024,097,766	7,135,494,059	7,860,711,226	8,398,278,426	8,669,378,501	9,580,720,701
Business receipts.....	5,731,616,337	6,244,678,064	6,156,994,009	6,334,602,711	6,948,481,893	7,369,538,953	7,535,482,221	8,414,537,647
Net income (less deficit)(*).....	253,678,291	213,648,962	154,334,143	188,313,928	232,900,596	240,119,020	269,530,240	334,089,233
Net income.....	311,497,470	301,440,778	274,352,942	296,932,146	349,179,415	363,867,384	408,860,760	468,631,779
Deficit.....	57,819,180	87,791,816	120,018,799	108,618,218	116,278,819	123,748,365	139,330,520	134,542,546
C Corporations								
Number of businesses.....	2,163,458	2,268,966	2,359,272	2,348,162	2,465,843	2,549,091	2,598,271	2,480,440
Total receipts.....	6,133,036,929	6,782,602,310	6,746,286,554	6,801,022,254	7,440,141,155	7,920,235,884	8,115,394,384	8,538,869,502
Business receipts.....	5,526,725,253	6,038,269,090	5,921,937,283	6,043,788,300	6,575,574,080	6,953,447,173	7,068,730,197	7,463,209,264
Net income (less deficit).....	236,487,630	185,868,913	120,180,204	154,156,433	196,435,483	192,991,940	203,018,630	250,706,247
Net income.....	288,701,762	266,981,510	232,171,007	253,219,429	300,847,319	303,127,497	326,576,008	366,764,203
Deficit.....	52,214,132	81,112,597	111,990,802	99,062,994	104,411,836	110,135,558	123,557,378	116,057,956
1120-RIC and 1120-REIT								
Number of businesses.....	1,691	1,965	2,442	2,642	3,561	3,379	4,030	3,788
Total receipts.....	17,924,659	31,235,499	34,754,643	34,223,383	35,543,228	47,400,761	69,997,816	69,604,933
Business receipts.....	3,716	51,060	45,971	49,473	175,374	50,592	39,187	22,551
Net income (less deficit).....	14,671,749	25,909,303	31,105,996	29,082,144	29,558,446	39,524,630	58,218,369	53,365,950
Net income.....	14,710,269	26,005,246	31,189,913	29,137,568	29,625,752	39,580,022	58,342,246	53,476,411
Deficit.....	38,521	95,943	83,918	55,426	67,306	55,392	123,877	110,461
S Corporations								
Number of businesses.....	545,389	541,489	564,219	648,267	701,339	724,749	826,214	1,127,905
Total receipts.....	210,322,424	212,514,030	243,056,569	300,248,422	385,026,843	430,641,781	483,986,301	972,246,266
Business receipts.....	204,887,368	206,357,914	235,010,755	290,764,938	372,732,439	416,041,188	466,712,837	951,305,832
Total net income (less deficit) (*).....	2,518,912	1,870,746	3,047,943	5,075,351	6,906,667	7,602,450	8,293,241	30,017,036
Net income.....	8,085,439	8,454,022	10,992,022	14,575,149	18,706,344	21,159,865	23,942,506	48,391,165
Deficit.....	5,566,527	6,583,276	7,944,079	9,499,798	11,799,677	13,557,415	15,649,265	18,374,129
Partnerships								
Number of businesses.....	1,379,654	1,460,502	1,514,212	1,541,539	1,643,581	1,713,603	1,702,952	1,648,032
Total receipts (*).....	291,998,115	272,129,807	296,690,303	291,318,703	375,192,511	367,117,315	397,302,544	442,802,234
Business receipts.....	271,108,832	230,027,336	251,608,987	243,248,370	318,342,380	302,733,374	327,428,647	411,457,126
Net income (less deficit).....	8,248,655	-2,734,897	-7,314,587	-2,610,041	-3,500,024	-8,883,674	-17,370,860	-5,419,105
Net income.....	45,061,756	50,567,190	53,556,856	60,308,114	69,696,922	77,044,693	80,214,873	87,654,011
Deficit.....	36,813,100	53,302,086	60,871,442	62,918,155	73,196,946	85,928,367	97,585,733	93,073,116
General (*)								
Number of businesses.....	1,209,318	1,252,298	1,288,328	n.a.	n.a.	n.a.	1,429,876	1,385,824
Total receipts (*).....	n.a.							
Business receipts.....	n.a.							
Net income (less deficit).....	n.a.							
Net income.....	n.a.							
Deficit.....	n.a.							
Limited (*)								
Number of businesses.....	170,336	208,204	225,886	n.a.	n.a.	n.a.	273,076	262,210
Total receipts (*).....	n.a.							
Business receipts.....	n.a.							
Net income (less deficit).....	n.a.							
Net income.....	n.a.							
Deficit.....	n.a.							
LLC								
Number of businesses.....	n.a.							
Total receipts (*).....	n.a.							
Business receipts.....	n.a.							
Net income (less deficit).....	n.a.							
Net income.....	n.a.							
Deficit.....	n.a.							
Nonfarm Sole Proprietorships								
Number of businesses.....	8,931,712	9,584,790	10,105,515	10,703,921	11,262,390	11,928,573	12,393,700	13,091,132
Total receipts.....	411,205,713	427,063,055	433,664,897	465,168,637	516,036,944	540,045,430	559,384,259	610,822,732
Business receipts.....	411,205,713	427,063,055	433,664,897	465,168,637	516,036,944	540,045,430	559,384,259	610,822,732
Net income (less deficit).....	54,947,219	53,071,628	50,573,163	60,359,153	70,766,610	78,772,578	90,423,763	105,460,627
Net income.....	68,010,051	68,552,791	68,647,384	78,618,410	89,849,570	98,775,563	110,496,952	123,782,540
Deficit.....	13,062,832	15,481,162	18,074,220	18,259,256	19,082,960	20,002,986	20,073,189	18,321,913

Footnotes at end of table.

Table 1.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business, Tax Years 1980-2002--Continued

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Tax Year							
	1988	1989	1990	1991	1992	1993	1994	1995
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
All Businesses								
Number of businesses.....	18,896,336	19,560,585	20,052,917	20,498,855	20,849,195	21,280,315	21,990,203	22,478,939
Total receipts.....	11,435,215,490	12,133,006,886	12,659,120,980	12,664,503,877	13,030,765,631	13,633,127,677	14,854,464,587	16,161,117,843
Business receipts.....	10,085,772,195	10,585,040,288	11,074,465,157	11,161,361,183	11,612,337,830	12,183,757,092	13,330,403,562	14,353,779,041
Net income (less deficit).....	563,932,180	548,157,101	541,253,496	523,452,364	611,007,348	733,369,871	843,984,176	1,012,514,546
Net income.....	818,548,839	829,704,453	n.a.	818,176,732	877,227,604	987,904,144	1,095,275,051	1,270,904,560
Deficit.....	254,616,660	281,547,353	n.a.	294,724,370	266,220,258	254,534,273	251,290,875	258,390,016
Corporations								
Number of businesses.....	3,562,789	3,627,863	3,716,650	3,802,788	3,869,024	3,964,629	4,342,369	4,474,167
Total receipts.....	10,264,867,461	10,934,973,405	11,409,520,074	11,436,474,767	11,742,134,728	12,269,721,709	13,360,007,157	14,539,050,115
Business receipts.....	8,949,846,244	9,427,277,533	9,860,441,633	9,965,628,799	10,360,428,795	10,865,542,520	11,883,614,940	12,785,797,708
Net income (less deficit)(¹).....	423,115,815	401,320,146	383,213,763	360,529,974	414,130,453	510,258,780	595,002,432	736,423,014
Net income.....	561,646,539	563,402,110	n.a.	542,341,802	581,920,697	670,480,179	756,502,169	900,524,657
Deficit.....	138,530,724	162,081,965	n.a.	181,811,828	167,790,244	160,221,400	161,499,736	164,101,644
C Corporations								
Number of businesses.....	2,299,896	2,199,081	2,136,032	2,098,641	2,077,518	2,055,982	2,310,703	2,312,382
Total receipts.....	8,929,061,395	9,381,129,704	9,689,007,338	9,656,969,832	9,821,791,797	10,154,952,821	11,020,933,534	11,955,289,941
Business receipts.....	7,712,940,028	7,992,750,467	8,272,370,751	8,310,147,728	8,569,591,965	8,897,605,783	9,710,160,635	10,419,343,855
Net income (less deficit).....	327,131,666	289,721,555	270,925,138	248,113,316	291,866,888	368,912,105	426,082,290	514,751,182
Net income.....	445,141,000	425,910,498	416,617,439	401,582,120	426,078,044	496,151,930	554,083,672	641,753,805
Deficit.....	118,009,334	136,188,943	145,692,301	153,468,803	134,211,156	127,239,826	128,001,382	127,002,623
1120-RIC and 1120-REIT								
Number of businesses.....	5,702	5,815	5,526	5,876	6,135	7,142	7,912	8,666
Total receipts.....	71,817,689	89,877,386	99,810,072	96,520,359	98,459,970	117,172,085	128,128,279	178,686,713
Business receipts.....	--	--	--	--	--	--	--	--
Net income (less deficit).....	52,447,631	66,819,244	67,457,384	67,671,565	63,933,826	75,113,178	77,243,699	122,543,160
Net income.....	52,596,709	67,087,163	67,983,981	68,188,117	64,704,531	75,770,157	78,447,581	123,812,233
Deficit.....	149,078	267,920	526,597	516,553	770,705	656,979	1,203,881	1,269,074
S Corporations								
Number of businesses.....	1,257,191	1,422,967	1,575,092	1,698,271	1,785,371	1,901,505	2,023,754	2,153,119
Total receipts.....	1,263,988,377	1,463,966,315	1,620,702,664	1,682,984,576	1,821,882,961	1,997,596,803	2,210,945,344	2,405,073,461
Business receipts.....	1,236,906,216	1,434,527,066	1,588,070,882	1,655,481,071	1,790,836,830	1,967,936,737	2,173,454,305	2,366,453,853
Total net income (less deficit) (²).....	43,536,518	44,779,347	44,831,241	44,745,903	58,329,739	66,233,497	91,676,443	99,128,672
Net income.....	63,908,830	70,404,449	n.a.	72,571,565	91,138,122	98,558,092	123,970,916	134,958,619
Deficit.....	20,372,312	25,625,102	n.a.	27,826,472	32,808,383	32,324,595	32,294,473	35,829,947
Partnerships								
Number of businesses.....	1,654,245	1,635,164	1,553,529	1,515,345	1,484,752	1,467,567	1,493,963	1,580,900
Total receipts (³).....	498,378,098	505,222,543	518,994,886	515,461,121	551,548,871	606,190,516	703,827,410	814,704,090
Business receipts.....	463,956,020	464,951,817	483,417,504	483,164,395	514,827,003	560,999,120	656,158,602	760,617,695
Net income (less deficit).....	14,493,114	14,099,275	16,609,540	21,406,607	42,916,649	66,652,288	82,183,076	106,829,196
Net income.....	111,384,545	113,885,966	116,317,801	113,408,221	121,834,358	137,440,684	150,927,743	178,650,950
Deficit.....	96,891,431	99,786,691	99,708,261	92,001,615	78,917,710	70,788,396	68,744,668	71,821,755
General (⁴)								
Number of businesses.....	1,369,093	1,341,527	1,267,760	1,244,665	1,214,004	1,174,395	1,161,800	1,163,376
Total receipts (³).....	n.a.	n.a.	349,839,034	349,793,551	354,750,145	369,030,331	394,825,973	417,535,888
Business receipts.....	n.a.	n.a.	334,184,309	333,189,600	336,912,510	348,350,203	375,032,602	395,396,396
Net income (less deficit).....	38,503,534	35,660,018	37,770,771	38,108,885	46,194,340	55,028,590	58,721,349	63,625,642
Net income.....	n.a.	n.a.	81,903,253	78,330,522	81,313,616	85,128,982	87,680,812	92,586,762
Deficit.....	n.a.	n.a.	44,132,482	40,221,637	35,119,276	30,100,391	28,959,463	28,961,119
Limited (⁵)								
Number of businesses.....	285,152	293,637	285,769	270,681	270,748	275,837	284,346	298,965
Total receipts (³).....	n.a.	n.a.	169,155,852	165,667,570	196,799,726	229,703,974	284,624,411	330,681,486
Business receipts.....	n.a.	n.a.	149,233,195	149,974,795	177,914,493	205,554,303	257,887,113	302,336,684
Net income (less deficit).....	-24,010,711	-21,560,743	-21,161,231	-16,702,278	-3,277,692	11,360,424	21,410,503	38,319,799
Net income.....	n.a.	n.a.	34,414,548	35,077,700	40,520,742	51,238,208	59,544,970	76,029,542
Deficit.....	n.a.	n.a.	55,575,779	51,779,978	43,798,434	39,877,784	38,134,467	37,709,743
LLC								
Number of businesses.....	n.a.	n.a.	n.a.	n.a.	n.a.	17,335	47,816	118,559
Total receipts (³).....	n.a.	n.a.	n.a.	n.a.	n.a.	7,456,210	24,377,026	66,486,715
Business receipts.....	n.a.	n.a.	n.a.	n.a.	n.a.	7,094,614	23,238,886	62,884,616
Net income (less deficit).....	n.a.	n.a.	n.a.	n.a.	n.a.	263,274	2,051,224	4,883,755
Net income.....	n.a.	n.a.	n.a.	n.a.	n.a.	1,073,995	3,701,961	10,034,647
Deficit.....	n.a.	n.a.	n.a.	n.a.	n.a.	810,221	1,650,737	5,150,892
Nonfarm Sole Proprietorships								
Number of businesses.....	13,679,302	14,297,558	14,782,738	15,180,722	15,495,419	15,848,119	16,153,871	16,423,872
Total receipts.....	671,969,931	692,810,938	730,606,020	712,567,989	737,082,032	757,215,452	790,630,020	807,363,638
Business receipts.....	671,969,931	692,810,938	730,606,020	712,567,989	737,082,032	757,215,452	790,630,020	807,363,638
Net income (less deficit).....	126,323,251	132,737,680	141,430,193	141,515,783	153,960,246	156,458,803	166,798,668	169,262,336
Net income.....	145,517,755	152,416,377	161,657,252	162,426,709	173,472,549	179,983,281	187,845,139	191,728,953
Deficit.....	19,194,505	19,678,697	20,227,059	20,910,927	19,512,304	23,524,477	21,046,471	22,466,617

Footnotes at end of table.

Table 1.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business, Tax Years 1980-2002--Continued

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Tax Year						
	1996	1997	1998	1999	2000	2001	2002
	(17)	(18)	(19)	(20)	(21)	(22)	(23)
All Businesses							
Number of businesses.....	23,240,648	23,645,197	24,113,044	24,448,466	25,007,504	25,605,898	26,434,293
Total receipts.....	17,371,531,836	18,729,888,900	19,717,102,456	21,616,705,144	23,845,405,224	23,752,254,090	23,361,178,481
Business receipts.....	15,418,548,555	16,473,284,387	17,285,188,902	18,899,080,667	20,719,272,866	20,799,323,834	20,741,003,999
Net income (less deficit).....	1,160,565,585	1,311,621,607	1,284,131,816	1,421,748,416	1,470,658,335	1,142,478,029	1,088,304,478
Net income.....	1,444,416,590	1,628,080,417	1,668,090,251	1,864,354,418	2,046,212,168	1,851,745,212	1,781,234,413
Deficit.....	283,851,005	316,458,810	383,959,436	442,606,001	575,553,831	709,267,183	692,929,934
Corporations							
Number of businesses.....	4,631,369	4,710,083	4,848,887	4,935,904	5,045,273	5,135,591	5,266,607
Total receipts.....	15,525,718,006	16,609,707,302	17,323,955,004	18,892,385,693	20,605,808,071	20,272,957,625	19,749,426,052
Business receipts.....	13,659,470,309	14,460,928,696	15,010,264,802	16,313,971,384	17,636,551,348	17,504,288,630	17,297,125,146
Net income (less deficit) (¹).....	838,591,644	956,736,971	895,152,469	985,363,334	986,952,279	648,758,089	596,524,023
Net income.....	1,016,135,059	1,155,242,666	1,144,026,382	1,282,481,469	1,391,008,755	1,155,497,714	1,084,179,817
Deficit.....	177,543,415	198,505,695	248,873,914	297,118,135	404,056,474	506,739,630	487,655,794
C Corporations							
Number of businesses.....	2,317,886	2,248,065	2,249,970	2,198,740	2,172,705	2,136,756	2,100,074
Total receipts.....	12,709,004,468	13,445,458,022	13,996,499,545	15,238,422,201	16,607,287,993	16,214,520,589	15,582,601,688
Business receipts.....	11,087,481,313	11,620,304,753	12,006,145,868	13,071,173,955	14,078,901,182	13,813,168,479	13,455,844,040
Net income (less deficit).....	574,553,924	607,541,446	532,246,228	535,289,061	517,937,235	270,774,336	258,673,938
Net income.....	714,272,006	765,753,475	736,810,215	783,499,456	859,530,894	709,003,929	676,337,238
Deficit.....	139,718,081	158,212,028	204,563,988	248,210,395	341,593,657	438,229,593	417,663,300
1120-RIC and 1120-REIT							
Number of businesses.....	9,067	9,764	10,829	11,389	12,090	12,349	12,156
Total receipts.....	198,619,366	269,011,761	266,322,290	353,094,730	381,042,973	296,924,686	255,897,663
Business receipts.....	--	--	--	--	--	--	--
Net income (less deficit).....	138,792,224	196,132,514	181,117,938	256,317,862	270,479,156	190,296,836	154,371,152
Net income.....	139,966,673	197,367,117	183,243,257	258,420,380	277,261,656	197,629,943	161,308,952
Deficit.....	1,174,450	1,234,604	2,125,319	2,102,518	6,782,500	7,333,108	6,937,800
S Corporations							
Number of businesses.....	2,304,416	2,452,254	2,588,088	2,725,775	2,860,478	2,986,486	3,154,377
Total receipts.....	2,618,094,172	2,895,237,519	3,061,133,169	3,300,868,762	3,617,477,105	3,761,512,350	3,910,926,701
Business receipts.....	2,571,988,996	2,840,623,943	3,004,118,934	3,242,797,429	3,557,650,166	3,691,120,151	3,841,281,106
Total net income (less deficit) (²).....	125,245,496	153,063,011	181,788,303	193,756,411	198,535,888	187,686,917	183,478,933
Net income.....	161,896,380	192,122,074	223,972,910	240,561,633	254,216,205	248,863,846	246,533,627
Deficit.....	36,650,884	39,059,063	42,184,607	46,805,222	55,680,317	61,176,929	63,054,694
Partnerships							
Number of businesses.....	1,654,256	1,758,627	1,855,348	1,936,919	2,057,500	2,132,117	2,242,169
Total receipts (³).....	1,002,579,987	1,249,789,312	1,474,879,256	1,754,972,413	2,218,639,870	2,462,461,787	2,582,060,669
Business receipts.....	915,844,403	1,141,963,405	1,356,655,904	1,615,762,245	2,061,764,235	2,278,200,526	2,414,187,093
Net income (less deficit).....	145,218,248	168,240,726	186,704,627	228,438,105	268,990,758	276,334,824	270,667,169
Net income.....	228,157,635	262,373,206	297,874,299	348,467,958	409,972,787	446,069,172	439,761,741
Deficit.....	82,939,388	94,132,480	111,170,672	120,029,853	140,982,029	169,734,347	169,094,572
General (⁴)							
Number of businesses.....	1,121,195	1,081,363	1,015,678	950,608	936,564	885,457	841,299
Total receipts (³).....	458,690,125	482,362,036	428,936,952	414,879,711	460,800,631	508,569,485	506,554,952
Business receipts.....	430,892,523	451,004,863	399,306,152	382,760,263	425,752,004	464,251,886	467,422,866
Net income (less deficit).....	77,446,760	88,235,026	82,766,449	85,767,233	101,786,779	101,830,079	100,914,057
Net income.....	106,074,272	113,264,997	107,709,809	108,487,666	127,059,152	128,591,551	125,748,798
Deficit.....	28,627,513	25,029,971	24,943,359	22,720,432	25,272,374	26,761,472	24,834,741
Limited (⁵)							
Number of businesses.....	311,563	328,210	369,012	396,907	402,232	437,968	454,741
Total receipts (³).....	386,373,126	474,480,710	585,636,689	701,845,221	884,397,372	935,891,900	987,064,490
Business receipts.....	338,916,079	423,968,766	534,248,684	644,246,861	830,429,874	876,234,279	931,055,315
Net income (less deficit).....	55,458,035	62,946,099	79,328,818	107,937,194	119,512,213	127,448,902	121,126,936
Net income.....	97,721,530	109,035,802	131,493,455	157,244,765	170,929,457	187,146,566	178,135,683
Deficit.....	42,263,496	46,089,703	52,164,637	49,307,571	51,417,244	59,697,664	57,008,747
LLC							
Number of businesses.....	221,498	349,054	470,657	589,403	718,704	808,692	946,130
Total receipts (³).....	157,516,736	292,946,566	460,305,616	638,247,481	873,441,868	1,018,000,402	1,088,441,226
Business receipts.....	146,035,802	266,989,776	423,101,069	588,755,121	805,582,357	937,714,361	1,015,708,912
Net income (less deficit).....	12,313,453	17,059,601	24,609,360	34,733,678	47,691,767	47,055,843	48,626,175
Net income.....	24,361,833	40,072,407	58,672,036	82,735,527	111,984,178	130,331,055	135,877,260
Deficit.....	12,048,379	23,012,806	34,062,676	48,001,849	64,292,411	83,275,212	87,251,084
Nonfarm Sole Proprietorships							
Number of businesses.....	16,955,023	17,176,487	17,408,809	17,575,643	17,904,731	18,338,190	18,925,517
Total receipts.....	843,233,843	870,392,286	918,268,196	969,347,038	1,020,957,283	1,016,834,678	1,029,691,760
Business receipts.....	843,233,843	870,392,286	918,268,196	969,347,038	1,020,957,283	1,016,834,678	1,029,691,760
Net income (less deficit).....	176,755,693	186,643,910	202,274,720	207,946,977	214,715,298	217,385,116	221,113,286
Net income.....	200,123,896	210,464,545	226,189,570	233,404,991	245,230,626	250,178,322	257,292,855
Deficit.....	23,368,202	23,820,635	23,914,850	25,458,013	30,515,328	32,793,206	36,179,568

n.a. - not available.

¹ For Tax Years beginning in 1987, Total Corporation "Net income (less deficit)" includes "Total net income (less deficit)" from S Corporations and is more comprehensive than what SOI generally publishes.

² Prior to Tax Year 1987, "Total net income (less deficit)" from S Corporations only includes "Net income (less deficit)" from S Corporations and is not as comprehensive as data in future years.

³ For consistency purposes of this publication, what SOI normally publishes as Partnership "Total income" is labeled as "Total receipts."

⁴ For Tax Years 1980-1995 General Partnerships include Partnerships listed on the tax return as General and not reported. For Tax Years 1996-1999 General Partnerships include Partnerships listed on the tax return as General, Other and not reported. For Tax Years 2000-2002 General Partnerships include Partnerships listed on the tax return as General, Foreign, Other and not reported.

⁵ For Tax Years 1980-1992 Limited Partnerships include Partnerships listed on the tax return as Limited Partnerships. For Tax Years 1993-1995 Limited Partnerships include Partnerships listed on the tax return as Limited Partnerships, General Limited Liability Partnerships, and Limited Liability Partnerships. For Tax Years 1996-1997 Limited Partnerships include Partnerships listed on the tax return as Limited Partnerships. For Tax Years 1998-1999 Limited Partnerships include Partnerships listed on the tax return as Limited Partnerships and Limited Liability Partnerships. For Tax Years 2000-2002 Limited Partnerships include Partnerships listed on the tax return as Domestic Limited Partnerships and Domestic Limited Liability Partnerships.

NOTE: Detail may not add to totals due to rounding.

Table 2A.--Number of Businesses, Business Receipts, Net Income, Deficit, and Other Selected Items, by Form of Business, Industry, and Business Receipt Size, Tax Year 1998
 [All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries										
	Total	Under \$25,000	\$25,000 under \$100,000	\$100,000 under \$250,000	\$250,000 under \$500,000	\$500,000 under \$1,000,000	\$1,000,000 under \$2,500,000	\$2,500,000 under \$5,000,000	\$5,000,000 under \$10,000,000	\$10,000,000 under \$50,000,000	\$50,000,000 or more
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
All Businesses											
Number of businesses.....	24,113,044	13,974,466	4,764,739	2,281,237	1,209,764	804,946	587,772	227,203	125,237	110,594	27,086
Total receipts.....	19,717,102,456	145,339,982	246,322,213	358,333,168	422,586,745	563,180,360	912,928,460	800,598,575	873,218,257	2,266,633,395	13,127,961,301
Business receipts.....	17,285,188,902	85,639,587	240,382,492	349,657,141	412,604,563	549,320,085	884,928,347	768,825,657	826,613,784	2,096,136,446	11,071,080,801
Total business deductions.....	18,591,694,169	155,884,135	183,938,754	299,717,900	380,718,555	528,827,085	878,482,933	774,074,954	840,548,353	2,157,081,858	12,392,419,644
Costs of goods sold.....	10,440,760,907	11,029,952	45,206,264	99,084,131	149,492,786	234,496,702	453,722,764	445,908,980	512,174,772	1,429,001,070	7,060,643,487
Salaries and wages.....	1,842,782,787	12,727,121	13,789,306	39,623,419	59,375,350	81,681,370	120,138,290	92,978,628	95,136,919	209,767,332	1,118,627,067
Taxes paid.....	392,122,646	3,693,573	4,508,918	9,533,042	12,204,453	16,303,360	24,973,128	19,578,789	18,747,658	39,635,590	242,943,776
Interest paid.....	1,051,224,941	8,451,747	4,647,506	6,673,745	7,441,285	8,696,842	15,346,138	14,420,702	20,333,600	64,271,555	902,741,822
Depreciation.....	614,850,618	10,060,618	11,850,625	12,590,142	12,049,870	13,433,911	19,989,400	16,210,411	16,648,705	42,525,385	459,491,746
Net income (less deficit).....	1,284,131,816	-3,653,627	68,583,858	69,311,628	54,164,474	44,141,002	46,483,727	47,507,870	44,552,749	129,836,572	793,203,564
Net income.....	1,668,091,252	80,919,485	86,299,688	85,829,167	68,956,178	60,519,836	69,185,361	54,528,990	63,068,635	173,540,405	925,243,507
Deficit.....	383,959,436	84,573,112	17,715,830	16,517,539	14,791,704	16,378,834	22,701,634	17,021,120	18,515,886	43,703,835	132,039,943
Corporations											
Number of businesses.....	4,848,888	1,169,591	748,636	840,006	644,396	553,217	467,642	193,800	110,302	97,757	23,542
Total receipts.....	17,323,955,004	39,623,284	48,941,501	145,033,825	236,229,500	398,875,148	740,658,271	691,966,841	776,786,140	2,018,310,740	12,227,529,752
Business receipts.....	15,010,264,802	5,300,644	43,751,193	137,548,602	228,043,316	386,956,560	717,709,500	664,080,283	735,051,410	1,864,328,275	10,227,495,018
Total business deductions.....	16,489,425,015	54,731,670	48,616,050	140,187,075	229,891,774	388,472,729	721,740,000	673,104,667	751,274,323	1,925,953,257	11,555,453,469
Costs of goods sold.....	9,362,392,237	1,885,708	9,502,361	39,236,176	78,685,074	162,112,523	365,234,999	389,020,858	463,572,404	1,300,437,747	6,552,701,386
Salaries and wages.....	1,613,559,231	7,809,413	4,724,016	17,648,108	33,476,727	56,928,888	97,605,030	79,260,362	84,321,262	184,327,501	1,047,957,825
Taxes paid.....	354,578,692	2,370,314	2,140,059	5,600,640	8,449,897	13,135,995	21,782,476	17,687,538	17,200,808	36,236,540	229,974,425
Interest paid.....	966,659,473	4,681,406	1,761,072	3,260,359	4,388,991	6,022,030	11,828,246	11,719,663	17,321,956	52,631,514	853,044,236
Depreciation.....	542,490,397	2,298,498	2,672,944	4,845,891	6,718,133	9,472,406	16,057,327	13,809,655	14,213,582	35,166,394	427,235,565
Net income (less deficit) (1).....	895,152,469	-910,825	1,702,940	7,161,929	10,775,991	12,972,958	22,109,880	22,492,339	28,983,818	90,719,519	699,144,220
Net income.....	1,144,026,383	26,222,152	8,629,334	15,901,949	20,293,561	24,228,279	38,628,136	34,678,078	42,972,894	122,108,029	810,363,971
Deficit.....	248,873,914	27,132,977	6,926,394	8,740,200	9,517,870	11,255,320	16,518,256	12,185,739	13,989,876	31,388,511	111,219,751
C Corporations (2)											
Number of businesses.....	2,260,799	470,111	329,244	362,513	313,723	280,738	260,136	109,405	63,741	54,310	16,878
Total receipts.....	14,262,821,835	29,249,069	22,437,425	64,782,917	117,893,376	205,782,983	412,764,768	393,832,755	452,086,050	1,127,211,724	11,436,870,768
Business receipts.....	12,006,145,868	2,275,358	18,743,409	58,884,625	111,843,093	197,067,452	395,248,776	371,438,609	415,277,551	985,996,990	9,449,370,004
Total business deductions.....	13,554,140,784	39,762,936	23,869,740	65,662,274	118,618,761	205,965,971	410,601,999	388,293,413	441,210,912	1,067,676,214	10,793,018,563
Costs of goods sold.....	7,428,465,189	861,989	4,140,648	16,719,545	37,439,053	79,976,046	196,040,981	216,389,192	262,788,993	663,093,941	5,951,014,842
Salaries and wages.....	1,308,886,018	6,479,696	2,260,393	7,897,407	16,224,492	30,136,317	54,216,388	44,947,695	47,199,331	96,786,525	992,647,716
Taxes paid.....	291,957,071	1,705,661	1,183,724	2,840,967	4,496,637	7,233,702	12,897,239	10,541,705	10,275,254	21,732,303	219,049,880
Interest paid.....	929,505,767	3,715,329	1,037,833	1,801,328	2,442,550	3,339,876	7,668,306	8,093,580	13,275,510	43,168,291	844,962,899
Depreciation.....	491,004,497	1,440,500	1,344,518	2,455,012	3,702,106	5,408,595	9,725,130	8,350,120	8,769,602	22,909,846	428,899,607
Net income (less deficit) (1).....	713,364,166	-10,319,604	-1,450,902	-911,784	-762,586	-284,015	2,247,725	4,385,537	8,616,383	48,709,017	663,134,394
Net income.....	920,053,473	6,426,278	2,346,378	4,086,802	4,835,441	6,815,357	13,679,620	13,701,808	20,028,156	75,865,733	772,266,900
Deficit.....	206,689,307	16,745,882	3,797,280	4,998,586	5,598,027	7,099,371	11,413,895	9,316,271	11,412,773	27,156,716	109,132,506
S Corporations											
Number of businesses.....	2,588,088	699,480	419,392	477,493	330,673	272,479	207,505	84,395	46,561	43,447	6,664
Total receipts.....	3,061,133,169	10,374,216	26,504,076	80,250,908	118,336,124	193,092,165	327,893,502	298,134,086	324,700,090	891,189,016	790,658,985
Business receipts.....	3,004,118,934	3,025,287	25,007,785	78,663,977	116,200,223	189,889,107	322,460,724	292,641,674	317,773,859	878,331,285	778,125,014
Total business deductions.....	2,935,284,231	14,968,734	24,746,310	74,524,801	111,273,014	182,506,758	311,678,001	284,811,254	310,063,412	858,277,043	762,434,906
Costs of goods sold.....	1,933,927,048	1,023,719	5,361,714	22,519,631	41,246,021	62,136,476	169,194,018	172,631,666	200,783,451	637,943,806	601,686,544
Salaries and wages.....	304,673,212	1,329,717	2,463,623	9,750,701	17,252,235	26,792,571	43,388,642	34,312,668	36,621,871	77,450,976	55,310,209
Taxes paid.....	62,621,621	664,652	956,335	2,759,674	3,953,260	5,902,294	8,885,236	7,145,833	6,925,554	14,504,238	10,924,546
Interest paid.....	37,153,706	966,077	723,239	1,459,030	1,946,442	2,682,154	4,159,939	3,626,083	4,046,183	9,463,224	8,081,336
Depreciation.....	51,485,899	857,999	1,328,426	2,390,879	3,016,027	4,063,811	6,332,197	5,459,535	5,443,980	12,256,549	10,336,948
Net income (less deficit).....	181,788,303	9,408,779	3,153,842	8,073,713	11,538,277	13,256,973	19,862,155	18,106,802	20,367,335	42,010,502	36,009,826
Net income.....	223,972,910	19,795,874	6,282,956	11,815,147	15,458,120	17,412,922	24,948,516	20,976,270	22,943,738	46,242,296	38,097,071
Deficit.....	111,170,627	10,387,095	3,129,114	3,741,434	3,919,843	4,155,949	5,086,361	2,869,468	2,576,303	4,231,795	2,087,245
Partnerships											
Number of businesses.....	1,855,348	1,037,571	314,120	206,432	113,816	76,622	56,792	22,638	12,217	11,696	3,443
Total receipts.....	1,474,879,256	26,723,342	9,234,457	22,587,046	32,041,217	45,937,124	80,480,503	72,533,863	78,474,189	227,575,321	879,292,195
Business receipts.....	1,356,655,904	1,345,587	8,485,044	21,396,242	30,245,219	43,995,438	75,429,160	68,647,503	73,604,446	211,060,836	822,446,430
Total business deductions.....	1,386,111,725	42,155,084	10,514,909	21,657,153	29,327,138	41,783,909	73,947,653	67,078,913	72,509,511	211,233,618	815,903,837
Costs of goods sold.....	737,235,839	599,393	2,330,873	7,047,658	11,076,713	16,722,295	31,427,483	31,850,267	35,733,664	111,683,919	488,763,575
Salaries and wages.....	142,910,961	3,844,711	723,649	2,141,603	3,957,195	6,658,976	11,262,534	9,978,110	9,877,000	24,329,514	70,139,667
Taxes paid.....	23,813,223	594,961	234,092	480,318	729,413	1,010,086	1,730,346	1,437,053	1,366,728	3,278,105	12,952,121
Interest paid.....	73,406,067	2,642,013	503,273	729,857	1,004,563	1,412,735	2,569,823	2,396,681	2,808,954	9,656,507	49,681,662
Depreciation.....	42,579,701	2,883,828	541,827	880,804	916,074	1,429,159	2,334,977	1,922,477	2,223,844	7,225,713	22,220,998
Net income (less deficit).....	186,704,627	-22,949,829	3,538,893	9,308,886	10,584,311	11,400,164	15,379,479	12,804,592	14,391,407	38,264,385	93,982,338
Net income.....	297,875,299	20,084,633	9,807,102	15,056,844	14,679,860	15,852,414	21,078,876	17,419,377	18,755,679	50,371,059	114,769,455
Deficit.....	111,170,627	43,034,462	6,288,209	5,747,958	4,095,549	4,452,520	5,699,397	4,614,785	4,364,272	12,066,714	20,787,117
Nonfarm Sole Proprietorships											
Number of businesses.....	17,408,809	11,767,304	3,701,983	1,234,799	451,552	175,107	63,338	10,765	2,718	1,141	

AN ANALYSIS OF BUSINESS ORGANIZATIONAL STRUCTURE AND ACTIVITY FROM TAX DATA

Table 2B.--Number of Businesses, Business Receipts, Net Income, Deficit, and Other Selected Items, by Form of Business, Industry, and Business Receipt Size, Tax Year 1999
 [All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries										
	Total	Under \$25,000	\$25,000 under \$100,000	\$100,000 under \$250,000	\$250,000 under \$500,000	\$500,000 under \$1,000,000	\$1,000,000 under \$2,500,000	\$2,500,000 under \$5,000,000	\$5,000,000 under \$10,000,000	\$10,000,000 under \$50,000,000	\$50,000,000 or more
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
All Businesses											
Number of businesses.....	24,448,466	14,045,632	4,925,838	2,279,078	1,260,910	812,137	611,773	240,694	129,061	114,155	29,189
Total receipts.....	21,616,705,144	153,000,907	251,787,730	357,382,981	442,443,531	570,557,969	949,707,439	844,241,774	899,797,960	2,355,160,627	14,792,624,228
Business receipts.....	18,899,080,668	86,911,643	245,278,507	348,987,172	432,862,629	556,233,812	924,781,255	813,294,007	853,876,284	2,174,985,391	12,461,869,970
Total business deductions.....	20,375,890,789	163,641,493	190,720,264	299,558,909	400,262,479	535,158,804	915,231,457	816,280,898	866,494,625	2,241,316,628	13,947,225,232
Costs of goods sold.....	11,556,334,280	12,144,724	45,596,195	93,002,502	156,771,916	231,109,602	468,547,249	468,567,517	519,234,963	1,468,827,054	8,092,532,557
Salaries and wages.....	2,042,858,325	14,487,290	14,045,130	39,670,637	63,039,835	84,119,734	127,909,401	100,809,447	102,175,772	224,847,037	1,271,754,043
Taxes paid.....	412,079,823	3,947,901	4,543,953	9,221,743	12,867,658	16,696,108	25,569,657	19,779,595	19,085,301	40,637,428	259,730,481
Interest paid.....	1,104,625,540	8,415,095	4,533,570	6,701,759	7,539,017	9,142,672	14,628,875	14,573,912	19,011,899	10,970,647	958,108,093
Depreciation.....	666,721,794	11,566,406	12,229,006	13,044,327	12,658,547	14,114,750	20,846,207	17,197,829	17,735,800	47,019,748	500,309,173
Net income (less deficit).....	1,421,748,416	-262,352	70,508,986	68,968,646	52,863,279	45,701,921	48,581,874	38,327,547	43,318,053	138,731,743	915,008,716
Net income.....	1,864,354,417	91,122,770	90,552,915	87,776,675	69,838,682	63,077,928	75,083,080	57,534,265	64,421,776	191,281,206	1,073,663,121
Deficit.....	442,606,001	91,385,122	20,043,930	18,810,028	16,975,404	17,376,005	26,501,205	19,206,718	21,103,724	52,549,462	158,654,404
Corporations											
Number of businesses.....	4,935,904	1,188,676	783,455	823,942	676,133	546,171	478,601	202,646	111,873	99,380	25,026
Total receipts.....	18,892,385,693	41,690,487	51,608,391	142,551,136	249,595,329	396,870,550	760,349,543	721,490,198	789,324,352	2,069,063,871	13,669,841,835
Business receipts.....	16,313,971,385	5,320,054	46,336,764	135,409,300	241,723,679	385,772,814	739,541,842	695,125,280	749,373,499	1,909,148,657	11,406,219,496
Total business deductions.....	17,966,972,060	58,065,223	52,863,919	139,219,322	243,501,542	386,818,661	742,977,358	702,199,592	762,043,705	1,971,734,403	12,904,548,335
Costs of goods sold.....	10,284,098,039	2,526,763	10,536,416	36,001,808	83,293,892	157,029,884	372,638,094	403,704,309	464,077,146	1,320,832,938	7,433,456,697
Salaries and wages.....	1,783,025,584	8,436,138	4,997,985	17,960,931	36,019,144	58,272,613	102,652,390	85,166,178	89,189,063	194,446,549	1,185,884,593
Taxes paid.....	371,183,229	2,570,791	2,151,175	5,405,983	9,052,334	13,160,400	22,074,901	17,728,081	17,335,095	36,818,322	244,886,146
Interest paid.....	1,018,972,484	4,284,267	1,748,137	3,076,147	4,651,184	6,159,057	10,837,164	11,729,534	15,743,969	51,026,577	909,716,449
Depreciation.....	583,799,586	2,684,792	2,696,629	5,022,993	7,086,316	9,850,040	16,538,625	14,414,944	14,896,667	36,684,895	473,923,662
Net income (less deficit) (1).....	985,363,333	1,800,919	1,295,110	5,870,499	7,870,261	11,611,525	21,284,660	21,298,855	26,063,707	96,993,707	793,633,962
Net income.....	1,282,481,469	33,088,241	10,191,580	16,237,884	18,522,833	23,788,385	40,701,750	35,373,742	40,443,551	132,416,595	931,716,911
Deficit.....	297,118,135	31,287,322	8,896,470	10,367,385	10,652,572	12,176,859	19,417,089	14,074,887	15,839,716	36,322,888	138,082,949
C Corporations (2)											
Number of businesses.....	2,210,129	473,987	322,385	343,211	304,663	264,643	257,151	110,294	62,635	53,605	17,555
Total receipts.....	15,591,516,931	31,119,510	22,350,928	61,313,454	115,034,275	195,517,231	415,023,993	396,729,766	444,172,775	1,129,134,158	12,781,120,842
Business receipts.....	13,071,173,955	2,308,441	18,691,583	55,835,226	109,397,822	187,234,409	398,979,937	374,940,207	409,382,461	983,094,514	10,531,309,356
Total business deductions.....	14,804,802,646	42,283,764	25,351,909	63,896,939	117,099,156	197,336,398	415,817,327	391,938,482	435,190,488	1,069,025,467	12,046,862,715
Costs of goods sold.....	8,224,778,365	1,700,022	4,117,351	14,198,913	37,954,085	74,122,799	197,282,132	217,225,887	253,935,981	660,354,022	6,764,387,173
Salaries and wages.....	1,447,235,089	6,796,789	2,462,164	8,094,401	15,907,701	29,157,499	56,935,202	46,462,681	50,149,397	112,193,850	1,121,193,406
Taxes paid.....	304,321,709	1,802,407	1,161,487	2,608,037	4,460,985	6,918,740	12,775,553	10,286,950	10,143,624	21,305,174	232,828,171
Interest paid.....	978,621,092	3,288,315	1,023,522	1,636,183	2,319,301	3,364,961	6,109,868	7,822,220	11,391,036	40,986,976	900,109,121
Depreciation.....	526,925,540	1,610,292	1,425,281	2,471,795	3,591,522	5,399,831	9,678,254	8,503,951	8,976,255	23,400,267	461,868,092
Net income (less deficit).....	791,606,922	-10,740,380	-2,979,824	-2,332,113	-2,132,368	-1,918,054	-1,262,008	6,450,674	6,484,654	49,448,455	753,693,512
Net income.....	1,041,919,836	8,179,895	2,422,338	3,620,136	4,457,317	6,440,613	12,915,929	14,525,130	19,361,933	81,561,530	888,435,017
Deficit.....	250,312,913	18,920,275	5,402,162	6,250,249	6,589,685	8,358,667	14,177,937	10,848,100	12,911,260	32,113,075	134,741,505
S Corporations											
Number of businesses.....	2,725,775	714,689	461,070	480,730	371,471	281,528	221,450	92,352	49,238	45,775	7,471
Total receipts.....	3,300,868,762	10,570,977	29,257,463	81,237,683	134,561,054	201,353,320	345,325,550	324,760,432	345,151,577	939,929,713	888,720,993
Business receipts.....	3,242,797,429	3,011,613	27,645,182	79,574,074	132,325,857	198,538,404	340,561,905	320,185,073	339,991,038	982,054,143	874,910,141
Total business deductions.....	3,162,169,414	15,781,458	27,512,010	75,322,383	126,402,386	189,482,263	327,160,031	310,261,111	329,853,217	902,708,936	857,686,619
Costs of goods sold.....	2,059,319,673	826,740	6,419,065	21,802,895	45,339,897	82,907,086	175,355,962	186,478,422	210,641,166	660,476,916	669,069,524
Salaries and wages.....	1,447,235,089	6,796,789	2,462,164	8,094,401	15,907,701	29,157,499	56,935,202	46,462,681	50,149,397	112,193,850	1,121,193,406
Taxes paid.....	304,321,709	1,802,407	1,161,487	2,608,037	4,460,985	6,918,740	12,775,553	10,286,950	10,143,624	21,305,174	232,828,171
Interest paid.....	978,621,092	3,288,315	1,023,522	1,636,183	2,319,301	3,364,961	6,109,868	7,822,220	11,391,036	40,986,976	900,109,121
Depreciation.....	526,925,540	1,610,292	1,425,281	2,471,795	3,591,522	5,399,831	9,678,254	8,503,951	8,976,255	23,400,267	461,868,092
Net income (less deficit).....	193,756,411	12,541,299	4,274,934	8,500,612	10,002,629	13,529,579	22,546,668	17,621,825	18,153,162	46,645,252	39,940,450
Net income.....	240,561,633	24,908,346	7,769,242	12,617,748	14,065,516	17,347,772	27,785,821	20,848,612	21,081,618	50,855,065	43,281,894
Deficit.....	46,805,222	12,367,047	3,494,308	4,117,136	4,062,887	3,818,192	5,239,152	3,226,787	2,928,456	4,209,813	3,341,444
Partnerships											
Number of businesses.....	1,936,919	1,036,339	356,913	212,438	125,787	83,799	64,757	25,094	14,375	13,437	3,981
Total receipts.....	1,754,972,413	31,042,309	10,391,869	23,471,292	34,875,981	50,555,407	91,042,927	79,786,794	91,715,629	261,383,977	1,080,706,229
Business receipts.....	1,615,762,245	1,323,477	9,154,272	22,217,320	33,166,729	47,328,987	86,924,445	75,203,945	85,744,805	241,123,955	1,013,574,310
Total business deductions.....	1,647,491,152	44,613,112	12,262,491	23,283,924	32,366,753	45,779,823	83,689,916	74,113,049	84,175,045	245,974,396	1,001,232,643
Costs of goods sold.....	902,157,018	980,414	2,623,624	7,320,780	12,087,758	17,778,260	36,388,551	34,809,446	42,362,235	127,850,796	619,955,153
Salaries and wages.....	169,905,010	5,002,752	887,756	2,425,227	4,427,127	7,125,707	13,000,939	11,175,584	11,368,440	29,092,877	85,398,582
Taxes paid.....	26,896,235	646,924	236,070	550,037	817,543	1,109,282	1,955,513	1,524,080	1,556,255	3,693,307	14,807,224
Interest paid.....	74,428,567	2,937,392	411,785	997,477	1,469,964	2,331,883	4,127,306	2,851,744	2,465,984	10,786,371	48,375,422
Depreciation.....	51,730,335	3,557,058	655,798	900,880	1,162,308	1,472,709	2,656,720	2,180,415	2,616,705	10,162,888	26,365,144
Net income (less deficit).....	228,438,105	-21,404,559	5,070,087	8,782,761	11,406,100	13,512,509	17,538,266	14,023,177	17,241,255	41,529,671	120,738,837
Net income.....	348,467,958	23,041,871	11,826,153	15,120,849	16,411,103	17,982,694	24,066,610	18,874,845	22,282,030	57,600,793	141,261,009
Deficit.....	120,029,853	44,446,430</									

Table 2C.-- Number of Businesses, Business Receipts, Net Income, Deficit, and Other Selected Items, by Form of Business, Industry, and Business Receipt Size, Tax Year 2000

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries										
	Total	Under \$25,000	\$25,000 under \$100,000	\$100,000 under \$250,000	\$250,000 under \$500,000	\$500,000 under \$1,000,000	\$1,000,000 under \$2,500,000	\$2,500,000 under \$5,000,000	\$5,000,000 under \$10,000,000	\$10,000,000 under \$50,000,000	\$50,000,000 or more
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
All Businesses											
Number of businesses.....	25,007,504	14,323,761	5,044,551	2,339,518	1,261,215	863,856	631,317	256,111	134,946	121,005	31,226
Total receipts.....	23,845,405,223	157,049,812	258,732,605	367,481,725	443,025,048	601,296,628	976,329,999	898,103,068	941,224,209	2,485,556,859	16,716,605,269
Business receipts.....	20,719,272,866	89,207,037	252,373,284	358,161,770	432,215,830	587,639,402	949,118,433	865,986,458	894,354,516	2,295,267,190	13,994,948,943
Total business deductions.....	22,597,449,332	181,945,234	196,815,975	311,124,577	402,389,404	565,292,207	945,806,680	874,005,483	915,105,217	2,390,301,738	15,814,662,818
Costs of goods sold.....	12,748,297,892	11,274,422	43,898,992	96,361,476	150,568,425	234,887,831	474,583,770	487,221,982	536,289,847	1,544,581,653	9,168,629,496
Salaries and wages.....	2,251,927,584	17,949,521	14,359,720	40,975,312	64,401,650	92,135,775	135,667,511	112,347,486	110,038,621	243,242,680	1,420,809,312
Taxes paid.....	435,168,334	4,133,603	4,418,811	9,380,045	12,743,848	17,618,060	26,167,996	20,790,089	19,909,885	42,711,239	277,294,761
Interest paid.....	1,376,663,337	9,050,941	4,872,806	7,272,703	7,737,944	10,055,255	16,229,234	16,173,654	20,435,450	17,629,925	1,213,205,467
Depreciation.....	706,107,104	11,907,931	12,662,894	13,878,648	12,758,995	15,493,569	21,853,826	18,382,838	18,627,208	47,744,527	532,796,667
Net income (less deficit).....	1,470,568,334	-15,047,369	69,578,410	67,673,841	51,430,798	47,117,312	45,585,699	37,423,472	38,259,307	123,941,181	1,004,695,686
Net income.....	2,046,212,168	94,696,194	93,436,476	90,469,148	72,872,893	68,945,708	78,614,695	63,620,215	67,867,370	196,641,163	1,219,048,308
Deficit.....	575,553,833	109,743,562	23,858,065	22,795,306	21,442,095	21,828,397	33,028,997	26,196,744	29,608,064	72,699,981	214,352,622
Corporations											
Number of businesses.....	5,045,274	1,220,003	782,747	837,072	677,480	581,940	487,533	212,496	115,106	104,524	26,372
Total receipts.....	20,605,808,070	44,380,488	51,077,677	146,174,039	250,539,810	418,959,740	770,734,628	754,821,357	814,034,212	2,164,472,050	15,190,608,071
Business receipts.....	17,636,551,348	5,491,907	45,779,274	138,446,952	241,515,388	407,815,578	748,446,965	727,755,456	773,334,342	1,996,366,609	12,551,598,878
Total business deductions.....	19,691,591,726	70,783,003	53,932,571	143,619,156	246,273,323	410,164,819	757,604,529	738,909,949	795,091,119	2,086,976,141	14,388,237,571
Costs of goods sold.....	11,135,287,909	1,799,913	9,711,853	36,293,813	83,003,692	160,966,315	373,506,201	411,062,657	470,727,263	1,374,616,895	8,213,599,305
Salaries and wages.....	1,957,812,570	11,705,836	5,547,031	19,064,459	37,236,221	64,933,511	107,329,549	94,499,429	94,299,994	209,222,591	1,313,975,949
Taxes paid.....	390,067,115	2,763,938	2,166,809	5,450,593	8,862,929	14,038,073	22,333,620	18,498,477	17,779,674	38,739,106	259,793,897
Interest paid.....	1,271,678,744	4,618,482	1,849,778	3,374,102	4,601,284	6,816,168	11,654,376	12,906,464	16,780,821	60,085,625	1,148,991,644
Depreciation.....	614,372,700	3,216,011	2,792,121	5,183,637	7,099,495	10,723,382	16,744,022	14,703,544	15,157,154	38,661,334	500,091,999
Net income (less deficit).....	986,952,279	-9,843,613	-1,130,701	3,563,967	5,989,908	10,827,328	16,598,640	19,667,017	20,005,711	77,528,687	843,745,335
Net income.....	1,391,008,755	32,784,125	9,499,059	16,613,183	19,777,917	26,185,419	40,208,246	38,736,118	41,677,076	131,662,989	1,033,864,623
Deficit.....	404,056,476	42,627,738	10,629,760	13,049,216	13,788,008	15,358,092	23,609,607	19,069,101	21,671,366	54,134,301	190,119,288
C Corporations (1)											
Number of businesses.....	2,184,795	473,111	312,248	343,804	290,666	262,547	255,443	111,573	61,995	55,334	18,073
Total receipts.....	16,988,330,966	34,802,542	21,564,795	62,482,405	109,589,610	193,014,491	412,749,259	400,454,455	442,057,984	1,156,040,424	14,155,575,002
Business receipts.....	14,078,901,182	2,197,494	17,986,624	56,691,627	103,297,434	184,552,959	395,726,244	378,468,011	406,823,175	1,002,716,239	11,530,441,375
Total business deductions.....	16,214,559,976	53,510,760	26,472,510	66,914,698	114,415,388	197,305,969	418,621,397	400,447,623	438,678,408	1,115,030,793	13,383,162,430
Costs of goods sold.....	8,870,607,003	947,771	4,080,571	14,747,835	35,015,985	72,510,542	193,038,476	213,245,821	244,905,415	666,289,369	7,425,825,217
Salaries and wages.....	1,586,268,656	9,342,167	3,103,803	8,714,694	16,396,073	30,422,948	58,991,593	50,702,628	52,333,629	117,502,107	1,238,759,014
Taxes paid.....	318,150,036	1,964,909	1,166,774	2,689,927	4,302,065	6,926,143	12,869,057	10,632,853	10,323,869	21,522,265	245,752,274
Interest paid.....	1,224,269,431	3,494,625	1,002,169	1,731,712	2,417,073	3,381,109	6,895,199	8,474,815	11,710,569	48,010,138	1,137,152,020
Depreciation.....	552,820,948	1,986,320	1,446,096	2,650,532	3,446,940	5,438,385	9,751,529	8,590,291	8,780,380	24,454,335	286,276,141
Net income (less deficit).....	788,416,391	-18,618,635	-4,948,454	-4,463,713	-4,870,710	-4,374,969	-6,336,624	-980,895	964,384	30,140,190	801,905,987
Net income.....	1,136,792,550	8,982,412	2,073,206	4,236,412	4,316,944	6,343,904	11,907,151	14,002,234	19,164,422	78,191,379	987,574,486
Deficit.....	348,376,159	27,601,047	7,021,660	8,700,125	9,187,653	10,718,873	18,243,775	14,983,129	18,200,038	48,051,270	185,668,589
S Corporations											
Number of businesses.....	2,860,478	746,892	470,499	493,268	386,814	319,392	232,090	100,923	53,112	49,190	8,298
Total receipts.....	3,617,477,105	9,577,946	29,512,882	83,691,635	140,950,200	225,945,249	357,985,369	354,366,903	371,982,227	1,008,431,626	1,035,033,069
Business receipts.....	3,557,650,166	3,294,413	27,792,650	81,755,325	138,217,954	223,262,619	352,720,721	349,287,445	366,511,167	993,650,369	1,021,157,503
Total business deductions.....	3,477,031,750	17,272,243	27,460,061	76,704,458	131,857,936	212,858,849	338,983,132	338,461,871	356,412,711	971,945,348	1,005,075,141
Costs of goods sold.....	2,264,680,905	852,142	5,631,282	21,545,978	47,987,707	88,455,773	180,467,725	197,816,835	225,821,848	708,327,526	787,774,088
Salaries and wages.....	371,543,914	2,363,670	2,443,228	10,349,765	20,840,148	34,510,564	48,337,956	43,796,801	41,964,365	91,720,484	75,216,934
Taxes paid.....	71,917,080	799,030	1,000,035	2,760,765	4,560,864	7,111,930	9,464,563	7,865,624	7,455,805	16,856,841	14,041,623
Interest paid.....	47,409,313	1,123,857	847,608	1,642,389	2,184,211	3,435,059	4,759,177	4,431,649	5,070,252	12,075,487	11,839,624
Depreciation.....	61,551,752	1,229,691	1,346,026	2,533,105	3,652,555	5,284,997	6,992,492	6,113,254	6,376,774	14,206,999	13,815,858
Total net income (less deficit).....	198,535,888	8,775,022	3,817,753	8,027,680	10,860,618	15,202,297	22,935,264	20,647,912	19,041,327	47,388,578	41,838,438
Net income.....	254,216,205	23,801,713	7,425,853	12,376,771	15,460,973	19,841,515	28,301,095	24,733,884	22,512,654	53,471,610	46,290,137
Deficit.....	55,680,317	15,026,691	3,608,100	4,349,091	4,600,355	4,639,219	5,365,832	4,085,972	3,471,328	6,083,031	4,450,699
Partnerships											
Number of businesses.....	2,057,500	1,105,074	370,358	225,771	127,043	92,392	71,489	29,579	16,277	14,907	4,610
Total receipts.....	2,218,639,870	30,495,031	10,392,827	24,480,295	34,710,820	56,037,004	99,920,879	95,923,270	103,605,325	291,579,200	1,471,495,219
Business receipts.....	2,061,764,235	1,540,837	9,331,909	22,887,427	32,926,025	53,523,939	94,996,978	90,872,562	97,441,501	269,394,972	1,388,848,085
Total business deductions.....	2,099,471,504	46,629,590	12,896,766	25,086,596	33,393,537	51,552,439	93,386,157	90,925,380	98,039,798	275,222,986	1,372,338,256
Costs of goods sold.....	1,225,628,897	1,097,008	2,188,370	7,655,273	11,683,062	19,332,549	38,313,145	43,226,319	49,228,313	146,672,108	906,232,751
Salaries and wages.....	201,350,844	5,247,107	1,102,352	2,645,448	4,793,832	8,426,283	14,917,390	12,940,994	13,485,353	32,600,848	105,191,239
Taxes paid.....	31,145,304	626,327	306,887	593,011	845,209	1,212,790	2,252,245	1,757,577	1,913,773	4,180,375	17,457,111
Interest paid.....	92,751,748	3,088,369	652,403	1,079,794	1,071,594	1,741,748	3,469,457	2,814,473	3,387,853	11,271,315	64,174,743
Depreciation.....	58,912,624	3,317,587	816,377	1,094,159	1,076,331	1,655,762	3,216,287	3,037,116	3,166,059	8,863,705	32,669,241
Net income (less deficit).....	268,990,758	-22,936,889	3,454,060	9,692,339	10,380,367	13,569,499	18,119,607	14,565,822	16,609,437	45,006,046	160,530,472
Net income.....	409,972,787	25,584,756	11,862,617	16,704,693	16,644,791	19,150,083	26,570,557	21,352,290	24,272,086	63,183,837	184,647,077
Deficit.....	140,982,029	48,521,645	8,408,557	7,012,354	6,264,424	5,580,584	8,450,950	6,786,469	7,662,649	18,177,791	24,116,605
Nonfarm Sole Proprietorships											
Number of businesses.....	17,904,731	11,998									

AN ANALYSIS OF BUSINESS ORGANIZATIONAL STRUCTURE AND ACTIVITY FROM TAX DATA

Table 2D.--Number of Businesses, Business Receipts, Net Income, Deficit, and Other Selected Items, by Form of Business, Industry, and Business Receipt Size, Tax Year 2001

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries										
	Total	Under \$25,000	\$25,000 under \$100,000	\$100,000 under \$250,000	\$250,000 under \$500,000	\$500,000 under \$1,000,000	\$1,000,000 under \$2,500,000	\$2,500,000 under \$5,000,000	\$5,000,000 under \$10,000,000	\$10,000,000 under \$50,000,000	\$50,000,000 or more
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
All Businesses											
Number of businesses.....	25,605,897	14,723,359	5,098,349	2,430,206	1,251,739	902,086	660,467	253,212	135,936	119,490	31,054
Total receipts.....	23,752,254,089	155,030,082	262,947,981	379,486,354	440,619,886	633,667,554	1,022,436,632	886,318,474	950,524,043	2,458,059,014	16,563,164,070
Business receipts.....	20,799,323,834	92,938,069	256,574,208	369,920,033	427,758,387	617,312,718	995,886,086	855,124,273	904,232,340	2,276,128,145	14,003,449,576
Total business deductions.....	22,830,860,232	190,442,838	198,314,117	319,438,007	402,547,429	598,204,439	995,149,387	864,930,664	927,849,625	2,377,854,832	15,956,130,892
Costs of goods sold.....	12,743,003,300	14,128,405	43,283,117	93,869,907	144,764,697	243,022,579	481,365,303	479,919,667	534,659,566	1,516,622,647	9,191,367,303
Salaries and wages.....	2,291,598,828	17,028,803	13,875,078	40,658,938	61,816,869	99,864,817	147,622,542	112,550,757	116,305,869	246,368,804	1,435,980,151
Taxes paid.....	441,299,097	4,133,318	4,204,103	9,495,149	12,490,157	18,518,100	28,184,526	21,309,786	20,252,822	43,492,401	279,218,736
Interest paid.....	1,312,833,856	9,278,832	4,650,049	7,259,948	10,991,405	10,156,970	16,900,300	15,222,860	19,437,878	65,808,414	1,153,127,400
Depreciation.....	756,298,215	14,083,334	12,667,160	14,530,796	13,143,511	16,192,247	24,259,084	18,819,327	20,012,177	52,241,879	570,348,699
Net income (less deficit).....	1,142,478,028	-32,154,371	72,739,997	73,153,340	48,445,392	48,128,875	43,051,010	34,130,677	36,657,969	109,446,591	708,878,551
Net income.....	1,851,745,212	90,989,291	96,276,474	96,036,319	71,257,796	72,636,715	81,092,988	62,065,111	67,068,196	185,195,531	1,029,126,790
Deficit.....	709,267,183	123,143,662	23,536,477	22,882,980	22,812,404	24,507,840	38,041,979	27,934,434	30,410,227	75,748,942	320,248,239
Corporations											
Number of businesses.....	5,135,591	1,248,183	780,365	877,052	672,623	601,468	507,184	207,320	114,073	101,459	25,862
Total receipts.....	20,272,957,624	40,603,449	51,621,782	152,408,375	252,126,775	438,077,167	803,504,379	739,037,691	811,555,585	2,109,015,526	14,875,060,896
Business receipts.....	17,504,288,630	5,727,667	46,329,927	144,783,542	240,759,596	425,848,994	783,400,650	713,185,544	772,832,661	1,950,715,950	12,420,704,193
Total business deductions.....	19,682,982,949	69,345,468	53,946,305	148,664,968	247,239,076	428,837,176	793,247,679	727,151,404	796,928,855	2,046,619,455	14,371,002,562
Costs of goods sold.....	11,041,533,030	2,134,165	9,715,098	36,508,299	79,464,316	164,487,891	377,389,929	403,360,800	464,892,716	1,333,242,273	8,170,337,542
Salaries and wages.....	1,968,876,180	10,937,615	5,213,590	19,382,112	36,361,955	69,522,956	116,513,857	92,848,299	98,927,784	207,820,456	1,311,347,555
Taxes paid.....	392,458,475	2,695,707	2,034,139	5,688,754	8,641,548	14,546,169	23,912,089	18,763,889	17,916,074	38,359,556	259,900,550
Interest paid.....	1,203,045,923	4,947,715	1,698,230	3,486,000	7,659,276	6,464,068	12,241,981	11,902,286	15,846,420	54,540,349	1,084,259,599
Depreciation.....	649,988,724	3,763,369	2,821,601	5,374,174	7,018,765	10,853,042	16,307,972	15,175,483	16,363,442	41,297,611	529,013,264
Net income (less deficit) (*).....	648,758,089	-15,097,850	-946,419	6,110,402	5,214,678	11,034,908	12,408,237	12,621,740	15,549,982	59,181,921	542,880,851
Net income.....	1,155,497,718	28,574,988	9,333,558	18,836,306	18,976,139	27,328,387	39,599,276	33,587,594	38,092,569	114,991,883	826,177,019
Deficit.....	506,739,630	43,672,838	10,279,977	12,726,264	13,761,461	16,293,480	27,191,040	20,965,856	22,542,586	55,809,962	283,496,168
C Corporations (*)											
Number of businesses.....	2,149,104	477,423	300,465	340,776	279,879	260,923	253,822	108,022	57,992	52,640	17,161
Total receipts.....	16,511,445,274	31,246,155	21,138,778	61,060,755	105,672,345	191,922,913	408,903,899	388,441,383	415,663,220	1,099,814,334	13,787,581,492
Business receipts.....	13,813,168,479	2,414,666	17,357,351	56,247,962	99,594,626	183,697,650	394,141,225	367,473,891	382,559,183	958,255,611	11,351,426,314
Total business deductions.....	16,065,395,745	5,511,983	25,211,836	65,720,846	109,888,682	197,359,217	418,305,708	391,768,298	416,232,725	1,072,411,693	13,316,984,757
Costs of goods sold.....	8,722,914,095	1,292,341	3,817,796	14,884,554	33,021,754	70,425,847	184,491,593	207,667,832	227,827,051	629,382,466	7,350,102,862
Salaries and wages.....	1,576,363,400	8,503,556	2,682,308	8,651,822	15,991,526	30,802,892	61,529,306	48,181,418	51,273,084	115,378,756	1,251,368,733
Taxes paid.....	315,490,007	1,959,517	1,065,093	2,613,043	4,012,741	7,031,135	13,152,493	10,633,888	9,640,249	21,175,044	244,206,804
Interest paid.....	1,153,625,673	3,686,169	962,677	1,788,174	2,227,109	3,240,018	7,056,850	7,373,097	10,943,818	42,881,468	1,073,466,195
Depreciation.....	582,949,925	2,393,952	1,362,498	2,649,728	3,334,638	5,497,908	10,080,018	8,787,945	9,202,533	25,749,766	513,890,939
Net income (less deficit).....	461,071,172	-20,142,319	-4,100,916	-4,689,731	-4,255,188	-5,489,860	-9,743,971	-4,334,505	-2,809,559	16,151,549	500,485,672
Net income.....	906,633,872	7,232,399	2,185,230	5,200,436	4,163,942	5,868,182	10,826,052	11,892,197	15,939,651	65,194,122	780,011,663
Deficit.....	445,562,701	27,374,718	6,286,146	8,210,166	8,419,130	11,158,043	20,570,024	16,226,703	18,749,209	49,042,573	279,525,990
S Corporations											
Number of businesses.....	2,986,486	770,761	479,900	536,276	392,744	340,545	253,362	99,298	56,081	48,819	8,702
Total receipts.....	3,761,512,350	9,357,294	30,483,003	91,347,620	146,454,430	246,154,254	394,600,480	350,596,307	395,892,365	1,009,201,192	1,087,425,404
Business receipts.....	3,691,120,151	3,313,001	28,972,577	88,535,580	141,164,970	242,151,344	389,259,425	345,711,653	390,273,385	992,460,339	1,069,277,878
Total business deductions.....	3,617,587,204	17,833,486	28,734,468	82,944,122	137,350,394	231,477,959	374,941,971	335,383,106	380,896,130	974,207,762	1,054,017,805
Costs of goods sold.....	2,318,618,934	841,824	5,897,302	21,623,745	46,442,562	94,062,045	192,898,336	195,692,968	237,065,665	703,859,807	820,234,860
Salaries and wages.....	392,512,780	2,434,060	1,230,282	10,730,291	20,370,429	38,720,064	54,984,552	44,666,881	45,664,700	99,441,700	979,878,822
Taxes paid.....	76,968,469	736,190	969,046	3,075,711	4,628,807	7,515,034	10,759,596	8,130,001	8,275,825	17,184,512	15,693,746
Interest paid.....	49,420,350	1,261,547	735,553	1,697,826	5,432,167	3,224,050	5,185,131	4,529,189	4,902,602	11,658,881	10,793,404
Depreciation.....	67,038,798	1,369,417	1,459,103	2,724,446	3,684,128	5,355,134	8,227,954	6,387,537	7,160,909	15,547,845	15,122,326
Total net income (less deficit).....	187,686,917	5,044,469	3,154,497	10,799,773	9,469,866	16,524,768	22,152,208	16,956,245	18,359,541	43,030,372	42,195,179
Net income.....	248,863,846	21,342,589	7,148,328	15,315,870	14,812,197	21,660,205	28,773,224	21,695,397	22,152,918	49,797,761	46,165,357
Deficit.....	61,176,929	16,298,120	3,993,831	4,516,098	5,342,331	5,135,437	6,621,016	4,739,153	3,793,377	6,767,389	3,970,178
Partnerships											
Number of businesses.....	2,132,117	1,129,884	374,726	233,896	139,446	102,800	79,883	31,848	18,140	16,487	5,008
Total receipts (*).....	2,462,461,787	28,815,937	10,648,113	25,747,293	37,301,575	62,400,182	112,183,539	100,130,199	114,951,657	320,565,219	1,649,718,073
Business receipts.....	2,278,200,526	1,599,705	9,566,195	23,805,805	35,807,255	58,273,519	105,736,722	94,788,144	107,382,972	296,933,926	1,544,306,283
Total business deductions.....	2,348,244,173	52,810,087	13,509,164	26,447,813	36,860,536	59,208,339	106,335,531	93,654,952	108,247,118	303,999,426	1,547,371,206
Costs of goods sold.....	1,338,114,656	3,740,224	2,607,519	7,446,858	12,475,154	21,292,438	42,084,213	44,492,498	52,953,906	161,626,452	989,395,395
Salaries and wages.....	230,874,139	5,011,512	1,004,782	2,994,508	4,825,441	9,561,577	17,220,066	14,518,171	15,170,777	36,699,909	123,867,396
Taxes paid.....	34,626,540	599,935	264,984	629,768	862,046	1,435,189	2,594,931	1,975,419	2,077,813	4,945,384	19,241,071
Interest paid.....	97,278,387	2,959,271	583,777	932,689	1,293,904	2,045,903	3,583,656	2,865,083	3,321,768	10,875,521	68,816,815
Depreciation.....	72,199,421	4,449,111	913,472	1,152,786	1,504,663	1,946,712	3,998,698	2,954,239	3,333,797	10,664,982	41,281,560
Net income (less deficit).....	276,334,824	-34,468,487	3,925,840	9,981,423	10,685,674	14,010,936	19,437,946	18,480,371	19,729,201	49,036,703	165,515,219
Net income.....	446,069,172	25,099,386	12,176,465	17,595,165	18,198,510	20,957,296	29,371,661	24,929,440	27,279,776	68,494,229	201,967,243

Table 2E.--Number of Businesses, Business Receipts, Net Income, Deficit, and Other Selected Items, by Form of Business, Industry, and Business Receipt Size, Tax Year 2002

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries										
	Total	Under \$25,000	\$25,000 under \$100,000	\$100,000 under \$250,000	\$250,000 under \$500,000	\$500,000 under \$1,000,000	\$1,000,000 under \$2,500,000	\$2,500,000 under \$5,000,000	\$5,000,000 under \$10,000,000	\$10,000,000 under \$50,000,000	\$50,000,000 or more
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
All Businesses											
Number of businesses.....	26,434,293	15,202,645	5,318,640	2,484,778	1,294,887	918,926	668,699	256,345	137,587	120,986	30,803
Total receipts.....	23,361,178,481	158,287,542	273,216,155	384,486,871	451,453,173	643,349,197	1,035,489,736	895,421,534	952,525,155	2,486,021,062	16,080,928,056
Business receipts.....	20,741,003,999	94,623,867	266,112,890	375,700,789	442,258,534	628,537,583	1,011,167,258	865,492,105	908,346,592	2,311,588,821	13,837,175,560
Total business deductions.....	22,463,630,938	189,439,969	208,087,137	324,116,768	411,662,071	608,173,280	1,005,173,751	873,716,273	926,726,567	2,394,360,316	15,522,174,806
Costs of goods sold.....	12,389,402,643	11,137,601	43,445,479	95,473,645	146,955,332	244,083,483	485,274,766	471,758,737	538,458,939	1,514,356,816	8,838,457,844
Salaries and wages.....	2,322,634,367	15,183,200	14,222,424	40,570,935	63,612,940	109,579,295	149,773,658	118,758,217	116,327,416	255,606,688	1,445,999,594
Taxes paid.....	447,889,738	4,176,027	4,660,884	9,435,166	12,954,735	19,035,887	28,592,443	22,233,837	20,739,552	44,355,212	281,705,995
Interest paid.....	992,318,790	8,266,714	4,559,313	6,551,377	8,857,269	14,281,321	12,851,625	16,694,192	15,823,571	861,481,252	646,885,849
Depreciation.....	831,111,969	14,167,232	14,332,408	16,109,773	14,246,100	18,581,791	26,880,034	20,741,549	22,135,501	55,554,641	628,362,940
Net income (less deficit).....	1,088,304,478	-35,207,003	71,225,703	73,497,961	51,123,474	48,206,107	44,135,280	32,541,039	38,735,299	117,160,769	646,885,849
Net income.....	1,781,234,412	89,088,110	98,555,316	95,831,974	74,766,223	72,468,090	78,540,264	58,353,005	67,388,663	185,433,067	960,811,300
Deficit.....	692,929,934	124,295,113	27,329,613	22,334,013	23,642,748	24,261,983	34,404,983	25,811,966	28,651,765	68,272,298	313,925,451
Corporations											
Number of businesses.....	5,266,607	1,282,449	828,658	893,875	688,785	610,715	510,424	209,942	114,539	101,777	25,443
Total receipts.....	19,749,426,052	38,458,278	53,727,669	153,639,962	253,596,745	442,377,560	808,490,681	743,724,722	805,258,552	2,108,934,069	14,341,217,514
Business receipts.....	17,297,125,146	5,730,889	48,777,161	146,591,152	246,224,279	431,617,409	789,209,655	718,536,156	767,997,903	1,958,909,658	12,183,530,885
Total business deductions.....	19,198,882,117	66,039,533	56,592,652	149,698,092	250,581,925	432,390,571	795,994,369	731,819,176	786,788,633	2,038,781,145	13,890,196,023
Costs of goods sold.....	10,607,404,004	1,866,005	10,168,249	38,006,411	80,144,537	163,208,766	378,171,863	396,440,900	462,415,265	1,318,982,261	7,757,999,748
Salaries and wages.....	1,988,294,948	8,693,151	5,094,139	18,427,070	36,679,649	70,784,522	116,934,177	97,510,821	98,428,756	213,033,121	1,322,709,535
Taxes paid.....	396,571,738	2,556,149	2,240,508	5,591,856	8,866,155	14,823,779	23,978,465	19,500,665	18,193,816	38,864,217	261,956,129
Interest paid.....	912,751,562	4,417,226	1,645,291	3,009,756	3,993,239	5,850,328	9,987,556	9,754,194	13,059,364	42,625,382	818,409,225
Depreciation.....	710,881,312	3,470,361	3,589,714	6,151,023	7,653,132	12,144,892	20,047,986	16,569,681	17,684,046	44,082,666	579,487,809
Net income (less deficit) (*).....	596,524,023	-16,618,912	-1,956,564	4,819,272	3,749,595	11,469,724	13,713,879	12,678,754	17,945,554	65,391,437	485,331,483
Net income.....	1,084,179,817	24,970,657	8,758,595	16,866,762	19,031,190	27,054,800	37,818,772	31,910,085	37,999,291	115,118,704	764,650,961
Deficit.....	487,655,794	41,589,569	10,715,159	12,047,490	15,281,595	15,585,076	24,104,893	19,231,331	20,053,936	49,727,266	279,319,478
C Corporations (*)											
Number of businesses.....	2,112,230	472,469	304,702	336,437	276,819	250,744	240,579	105,620	57,831	50,702	16,326
Total receipts.....	15,838,499,350	28,504,846	20,731,910	59,710,978	104,169,660	184,240,237	388,704,542	376,515,006	410,275,152	1,062,957,229	13,202,689,791
Business receipts.....	13,455,844,040	2,257,712	17,418,221	54,866,994	98,925,788	176,587,421	375,382,257	356,654,979	378,923,647	930,333,430	11,064,493,590
Total business deductions.....	15,439,803,663	47,818,120	24,943,962	63,845,753	110,153,988	189,373,466	395,508,821	378,798,010	408,300,820	1,030,741,791	12,790,318,932
Costs of goods sold.....	8,220,579,884	803,602	3,896,827	14,480,167	33,921,090	66,012,464	174,303,365	195,474,791	224,588,177	594,582,839	6,912,516,022
Salaries and wages.....	1,569,301,518	5,933,543	2,429,179	8,162,958	15,284,537	28,793,483	57,740,829	49,186,789	51,594,471	117,410,926	1,232,764,802
Taxes paid.....	315,744,047	1,758,778	1,086,518	2,565,602	4,081,924	6,743,821	12,507,701	10,563,613	9,789,494	21,302,098	245,344,498
Interest paid.....	873,968,319	3,391,045	871,385	1,495,425	1,882,783	2,920,419	5,435,940	5,922,761	8,699,533	33,453,167	810,005,862
Depreciation.....	632,581,809	2,067,413	1,465,393	2,727,110	3,479,247	5,629,286	10,264,171	8,834,280	9,725,392	26,535,600	561,853,917
Net income (less deficit).....	413,045,090	-19,148,033	-4,218,962	-4,172,025	-5,995,410	-5,195,653	-7,110,483	-3,334,615	-2,446,489	21,853,336	440,613,623
Net income.....	837,646,190	7,054,427	1,886,041	3,193,175	3,823,261	5,067,070	9,695,354	10,729,547	15,707,804	64,559,522	715,929,248
Deficit.....	424,601,100	26,202,460	6,105,003	7,365,940	9,818,671	10,262,923	16,805,837	14,064,162	15,954,293	42,706,185	275,315,625
S Corporations											
Number of businesses.....	3,154,377	890,980	523,956	557,438	411,966	359,971	269,845	104,321	56,708	51,075	9,117
Total receipts.....	3,910,926,701	9,953,432	32,995,759	93,928,985	149,427,085	258,137,323	419,786,138	367,209,716	394,983,700	1,045,976,840	1,138,527,723
Business receipts.....	3,841,281,106	3,473,177	31,358,940	91,724,158	147,298,491	255,029,988	413,827,398	361,881,176	389,074,256	1,028,576,228	1,119,037,294
Total business deductions.....	3,759,078,454	18,221,412	31,648,689	85,852,338	140,427,937	243,017,105	400,485,548	353,021,166	378,487,813	1,008,039,354	1,099,877,091
Costs of goods sold.....	2,386,824,120	1,062,403	6,271,423	23,526,244	46,223,446	97,196,302	203,868,497	200,966,100	237,826,549	724,399,421	845,483,726
Salaries and wages.....	949,893,431	2,759,608	2,664,960	10,264,112	21,395,112	41,991,045	59,193,350	48,324,031	46,834,285	95,622,195	89,944,733
Taxes paid.....	80,827,691	797,371	1,153,991	3,026,254	4,784,231	8,079,958	11,470,764	8,937,052	8,404,321	17,562,118	16,611,631
Interest paid.....	38,783,242	1,026,182	773,906	1,514,331	2,110,456	3,029,909	4,551,617	3,831,433	3,499,831	9,172,215	8,403,363
Depreciation.....	78,299,503	1,402,949	2,124,321	3,423,912	4,173,886	6,515,606	9,783,815	7,735,401	7,958,655	17,547,066	17,633,892
Total net income (less deficit).....	183,478,933	2,529,121	2,262,398	8,991,297	9,745,005	16,665,577	20,824,362	16,013,369	18,191,843	43,538,101	44,717,860
Net income.....	246,533,627	17,916,230	6,872,554	13,672,847	15,207,929	21,987,730	28,123,418	21,180,538	22,291,487	50,559,182	48,721,713
Deficit.....	63,054,694	15,387,109	4,610,156	4,681,550	5,462,924	5,322,153	7,299,056	5,167,169	4,099,643	7,021,081	4,003,853
Partnerships											
Number of businesses.....	2,242,169	1,203,722	380,403	248,533	145,261	104,958	83,998	33,201	19,198	17,709	5,187
Total receipts.....	2,582,060,669	32,608,125	11,849,604	26,249,643	40,107,371	62,897,320	119,181,641	106,969,882	121,553,349	349,431,600	1,711,212,135
Business receipts.....	2,414,187,093	1,671,840	9,696,847	24,512,371	38,285,197	58,845,858	114,140,189	102,229,020	114,635,734	325,023,769	1,625,146,268
Total business deductions.....	2,455,848,170	52,328,995	15,327,948	26,632,418	39,142,651	60,916,260	113,086,856	100,022,075	115,461,191	328,925,224	1,604,004,552
Costs of goods sold.....	1,430,213,629	1,044,041	3,038,499	7,297,974	14,069,028	21,850,701	46,860,712	45,415,503	57,893,441	174,520,652	1,058,223,077
Salaries and wages.....	237,882,426	5,511,544	1,204,394	2,882,053	5,336,723	9,524,195	18,391,405	16,538,950	15,568,867	40,555,992	122,368,302
Taxes paid.....	36,416,569	804,394	290,282	596,450	1,028,306	1,498,228	2,833,426	2,203,244	2,237,581	5,277,974	19,646,813
Interest paid.....	68,127,690	2,634,518	534,193	927,697	1,149,274	1,696,230	3,251,528	2,729,310	3,371,115	9,007,515	43,013,810
Depreciation.....	82,897,056	4,281,378	1,053,369	1,511,670	1,646,264	2,471,460	4,487,239	3,447,065	4,091,014	11,140,872	48,766,726
Net income (less deficit).....	270,667,169	-34,946,815	1,702,237	11,861,270	11,562,945	13,504,407	18,671,768	16,999,756	19,531,172	50,559,365	161,030,063
Net income.....	439,761,741	25,759,194	12,511,429	19,073,137	18,106,027	20,842,847	27,979,207	23,157,217	27,789,869	68,981,239	195,561,575
Deficit.....	169,094,572	60,706,009	10,809,192	7,211,867	6,543,082	7,338,440	9,307,438	6,157,461	8,258,697	18,230,874	34,531,512
Nonfarm Sole Proprietorships											
Number of businesses.....	18,925,517										

Table 3A.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 1998

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries	Agriculture, forestry, fishing, and hunting	Mining	Utilities	Construction	Manufacturing	Wholesale and retail trade	Transportation and warehousing	Information	Finance and insurance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
All Businesses										
Number of businesses.....	24,113,045	539,643	179,941	17,662	2,920,802	706,002	3,813,207	969,101	335,332	1,026,302
Business receipts.....	17,285,188,902	131,665,240	147,677,818	499,833,981	1,109,402,772	4,865,936,073	5,041,650,550	543,877,331	771,910,696	1,435,257,053
Net income (less deficit).....	1,284,131,818	3,143,718	6,394,789	33,386,649	65,318,955	268,147,759	117,285,087	24,421,428	31,289,380	367,427,885
Net income.....	1,668,091,253	12,483,647	19,331,265	42,448,390	80,156,917	331,010,906	158,707,639	33,697,693	89,056,619	411,027,844
Deficit.....	383,959,435	9,339,929	12,936,476	9,061,740	14,837,964	62,863,149	41,422,552	9,276,265	57,767,239	43,599,957
Corporations										
Number of businesses.....	4,848,888	135,107	31,467	8,067	551,935	309,912	956,803	159,646	100,977	218,193
Business receipts.....	15,010,264,802	100,398,430	116,905,970	450,830,225	859,139,558	4,591,071,027	4,516,670,915	469,626,605	667,610,273	1,285,017,559
Net income (less deficit)(¹).....	895,152,471	1,266,193	2,339,453	31,920,963	31,506,413	254,033,430	95,614,094	16,357,260	35,353,093	291,193,439
Net income.....	1,144,026,384	6,021,329	10,017,694	38,343,308	40,340,592	307,995,283	127,371,881	22,464,817	70,780,532	322,289,879
Deficit.....	248,873,913	4,755,137	7,678,241	6,422,344	8,834,179	53,961,854	31,757,788	6,107,556	35,427,439	31,096,439
C Corporations (²)										
Number of businesses.....	2,260,801	65,689	15,988	5,943	246,404	163,295	472,031	78,341	44,895	115,309
Business receipts.....	12,006,145,868	56,012,640	102,328,023	448,214,333	467,247,448	4,107,930,264	3,241,722,259	384,935,892	620,177,682	1,226,629,994
Net income (less deficit).....	713,364,168	231,736	-76,819	31,407,088	10,249,297	218,465,519	57,410,132	12,794,920	29,887,900	279,336,463
Net income.....	920,053,474	2,787,619	7,016,179	37,725,338	15,321,252	267,572,313	82,152,207	17,336,221	62,927,184	307,951,720
Deficit.....	206,689,306	2,555,883	7,094,998	6,318,249	5,071,954	49,106,794	24,742,076	4,541,300	33,039,284	28,615,256
S Corporations										
Number of businesses.....	2,588,088	69,418	15,479	2,124	305,531	146,617	484,772	81,305	56,082	102,884
Business receipts.....	3,004,118,934	44,385,790	14,577,947	2,615,892	391,892,110	483,140,763	1,274,948,656	84,690,713	47,432,591	58,387,565
Total net income (less deficit).....	181,788,303	1,034,457	2,416,272	513,875	21,257,116	35,567,911	38,203,962	3,562,340	5,465,193	11,856,976
Net income.....	223,972,910	3,233,710	2,999,515	617,970	25,019,340	40,422,970	45,219,674	5,128,596	7,853,348	14,338,159
Deficit.....	42,184,607	2,199,254	583,243	104,095	3,762,225	4,855,060	7,015,712	1,566,256	2,388,155	2,481,183
Partnerships										
Number of businesses.....	1,855,348	115,614	29,098	2,448	125,823	34,836	130,288	19,193	21,900	209,150
Business receipts.....	1,356,655,904	15,572,293	25,711,768	48,837,758	106,320,658	247,438,628	304,069,914	31,009,687	98,387,504	88,996,302
Net income (less deficit).....	186,704,627	500,178	4,201,775	1,398,864	7,808,640	10,237,101	5,722,617	1,505,717	-5,773,299	63,268,132
Net income.....	297,875,299	4,148,941	8,556,138	4,034,991	11,973,217	18,574,043	9,967,695	3,488,257	16,204,642	74,310,012
Deficit.....	111,170,672	3,648,762	4,354,363	2,636,127	4,164,578	8,336,943	4,245,077	1,982,540	21,977,941	11,041,880
General (³)										
Number of businesses.....	1,015,678	90,796	11,181	340	69,173	18,619	88,078	9,791	12,961	113,083
Business receipts.....	399,306,152	5,592,102	8,271,842	10,833,116	38,642,807	73,335,482	72,443,611	6,442,068	32,426,530	23,391,638
Net income (less deficit).....	82,766,449	1,460,571	575,260	784,292	3,217,570	4,455,912	2,580,004	1,082,522	2,303,426	18,626,318
Net income.....	107,709,809	3,133,629	3,495,446	1,166,756	4,492,334	6,421,416	3,491,361	1,330,410	6,264,782	21,121,278
Deficit.....	24,943,359	1,673,058	2,920,186	382,464	1,274,764	1,965,504	911,357	247,888	3,961,356	2,494,961
Limited (⁴)										
Number of businesses.....	369,013	12,368	11,966	731	17,226	3,488	6,101	1,445	2,460	63,643
Business receipts.....	534,248,684	4,684,558	10,448,278	21,525,717	28,525,870	85,139,650	134,538,787	10,149,777	50,911,219	40,964,184
Net income (less deficit).....	79,328,818	-471,446	3,178,831	522,013	1,796,126	3,668,785	1,618,212	1,052,098	-3,686,482	35,132,990
Net income.....	131,493,455	574,188	3,825,546	2,449,036	3,550,642	6,826,390	2,936,466	1,603,172	8,634,336	39,657,223
Deficit.....	52,164,637	1,045,634	646,715	1,927,023	1,754,516	3,157,605	1,318,254	551,074	12,320,818	4,524,233
LLC										
Number of businesses.....	470,657	12,450	5,951	1,376	39,424	12,729	36,109	7,957	6,479	32,425
Business receipts.....	423,101,069	5,295,633	6,991,649	16,478,925	39,151,981	88,963,496	97,087,516	14,417,841	15,049,755	24,640,479
Net income (less deficit).....	24,609,360	-488,947	447,685	92,559	2,794,944	2,112,403	1,524,401	-628,903	-4,390,243	9,508,825
Net income.....	58,672,036	441,124	1,235,146	419,198	3,930,241	5,326,237	3,539,868	554,675	1,305,525	13,531,511
Deficit.....	34,062,676	930,070	787,461	326,639	1,135,297	3,213,834	2,015,466	1,183,578	5,695,767	4,022,686
Nonfarm Sole Proprietorships										
Number of businesses.....	17,408,809	288,922	119,376	7,147	2,243,044	361,254	2,726,116	790,262	212,455	598,959
Business receipts.....	918,268,196	15,694,517	5,060,080	165,998	143,942,556	27,426,418	220,909,721	43,241,039	5,912,919	61,243,192
Net income (less deficit).....	202,274,720	1,377,347	-146,439	66,822	26,003,902	3,877,228	15,948,376	6,558,451	1,709,586	12,966,314
Net income.....	226,189,570	2,313,377	757,433	70,091	27,843,108	4,441,580	21,368,063	7,744,619	2,071,445	14,427,953
Deficit.....	23,914,850	936,030	903,872	3,269	1,839,207	564,352	5,419,687	1,186,169	361,859	1,461,638

Footnotes at end of table.

Table 3A.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 1998--Continued

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Real estate and rental and leasing	Professional, scientific, and technical services	Management of companies (holding companies)	Administrative and support and waste management services	Educational services	Health care and social assistance	Arts, entertainment, and recreation	Accommodation, food services, and drinking places	Other services	Religious, grantmaking, civic, professional, and similar	Unclassified industries
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
All Businesses											
Number of businesses.....	2,205,935	3,173,498	42,918	1,479,954	334,469	1,851,412	1,110,054	606,023	2,221,313	212,939	366,536
Business receipts.....	260,368,200	796,236,596	92,627,484	320,982,170	27,931,863	497,570,878	102,238,841	407,944,777	220,892,768	2,607,373	8,576,438
Net income (less deficit).....	77,861,007	108,112,007	63,284,091	21,606,621	2,248,139	48,479,669	7,790,652	14,100,566	20,493,675	1,420,425	1,919,314
Net income.....	120,638,642	132,646,277	71,848,728	28,656,644	3,096,368	61,528,291	16,520,105	25,832,332	25,137,968	1,495,125	2,769,852
Deficit.....	42,777,637	24,534,269	8,564,637	7,050,023	848,228	13,048,622	8,729,453	11,731,767	4,644,292	74,700	850,538
Corporations											
Number of businesses.....	521,917	623,784	30,931	200,449	36,959	307,258	92,966	245,334	300,313	N/A	16,870
Business receipts.....	175,701,248	540,924,209	90,497,966	263,656,627	23,196,929	357,156,938	60,387,702	295,686,536	143,395,451	N/A	2,391,635
Net income (less deficit) (¹).....	20,032,614	23,825,752	57,992,086	11,319,256	952,453	4,703,934	2,665,760	8,708,478	5,570,156	N/A	-202,357
Net income.....	31,165,308	42,062,770	62,833,377	16,412,398	1,447,682	14,339,113	5,929,494	15,734,084	8,127,748	N/A	349,095
Deficit.....	11,132,696	18,237,017	4,841,291	5,093,142	495,228	9,635,179	3,263,735	7,025,607	2,557,591	N/A	551,452
C Corporations (²)											
Number of businesses.....	221,716	252,632	19,460	77,983	16,432	172,414	38,084	98,243	149,877	N/A	6,066
Business receipts.....	120,932,399	361,250,574	87,283,096	170,634,628	12,899,509	278,203,449	35,087,417	197,416,935	86,169,085	N/A	1,070,241
Net income (less deficit).....	4,871,421	3,531,659	55,306,953	3,180,597	285,851	-1,284,344	805,624	5,083,593	1,836,226	N/A	40,352
Net income.....	12,419,771	17,851,123	59,607,765	7,040,055	662,607	6,713,766	2,380,855	9,194,448	3,335,123	N/A	55,928
Deficit.....	7,548,352	14,319,463	4,300,812	3,859,458	376,756	7,998,110	1,575,232	4,110,856	1,498,897	N/A	15,576
S Corporations											
Number of businesses.....	300,201	371,152	11,471	122,466	20,527	134,844	54,882	147,091	150,437	N/A	10,804
Business receipts.....	54,768,849	179,673,635	3,214,870	93,020,999	10,297,420	78,953,489	25,300,285	98,269,601	57,226,366	N/A	1,321,394
Total net income (less deficit).....	15,161,193	20,294,093	2,685,133	8,138,659	666,602	5,988,278	1,860,136	3,624,885	3,733,930	N/A	-242,709
Net income.....	18,745,537	24,211,647	3,225,612	9,372,343	785,075	7,625,347	3,548,639	6,539,636	4,792,625	N/A	293,167
Deficit.....	3,584,344	3,917,554	540,479	1,233,684	118,472	1,637,069	1,688,503	2,914,751	1,058,694	N/A	535,876
Partnerships											
Number of businesses.....	812,404	118,340	11,987	28,268	4,697	37,767	30,319	57,912	63,763	N/A	1,541
Business receipts.....	41,348,441	147,764,823	2,129,518	22,840,826	1,073,235	59,773,854	22,156,807	78,969,307	14,128,213	N/A	126,369
Net income (less deficit).....	40,187,832	38,732,610	5,292,005	1,213,360	-14,143	7,913,211	26,492	3,374,509	1,070,062	N/A	38,964
Net income.....	70,435,470	42,440,066	9,015,351	2,111,820	116,846	10,456,934	3,412,404	6,946,094	1,637,607	N/A	44,771
Deficit.....	30,247,638	3,707,456	3,723,346	898,460	130,989	2,543,723	3,385,911	3,571,585	567,545	N/A	5,807
General (³)											
Number of businesses.....	399,000	64,124	3,077	15,597	2,734	20,159	16,801	30,899	48,119	N/A	1,146
Business receipts.....	8,109,819	62,707,752	461,622	4,048,364	252,245	17,966,586	6,222,789	21,371,210	6,763,109	N/A	23,458
Net income (less deficit).....	18,616,998	20,162,890	598,743	522,565	26,699	4,681,131	753,217	1,406,109	881,286	N/A	30,936
Net income.....	23,920,224	20,814,456	2,017,696	594,321	32,590	5,000,999	1,531,706	1,833,514	1,010,766	N/A	36,125
Deficit.....	5,303,226	651,566	1,418,953	71,755	5,892	319,868	778,489	427,404	129,480	N/A	5,189
Limited (⁴)											
Number of businesses.....	212,838	12,630	3,944	1,214	98	4,995	2,889	8,588	2,015	N/A	375
Business receipts.....	17,700,146	51,478,821	195,939	7,956,966	289,017	22,588,714	10,056,807	35,117,416	1,976,313	N/A	505
Net income (less deficit).....	14,931,331	14,074,114	3,069,115	581,525	-10,432	1,973,743	-54,373	1,849,712	97,494	N/A	5,463
Net income.....	33,253,393	15,271,996	3,999,473	828,100	39,928	2,920,711	1,356,212	3,565,697	195,430	N/A	5,514
Deficit.....	18,322,062	1,197,882	930,358	246,576	50,361	946,968	1,410,585	1,715,985	97,935	N/A	51
LLC											
Number of businesses.....	200,566	41,587	4,966	11,457	1,864	12,613	10,629	18,425	13,629	N/A	20
Business receipts.....	15,538,476	33,578,249	1,471,957	10,835,496	531,973	19,218,553	5,877,211	22,480,681	5,388,790	N/A	102,407
Net income (less deficit).....	6,639,502	4,495,606	1,624,147	109,270	-30,410	1,258,336	-672,352	118,688	91,000	N/A	2,565
Net income.....	13,261,852	6,353,614	2,998,182	689,399	44,328	2,535,224	524,486	1,546,883	431,411	N/A	3,132
Deficit.....	6,622,350	1,858,008	1,374,035	580,129	74,737	1,276,887	1,196,837	1,428,195	340,129	N/A	567
Nonfarm Sole Proprietorships											
Number of businesses.....	871,614	2,431,374	N/A	1,251,237	292,813	1,506,387	986,769	302,777	1,857,237	212,939	348,125
Business receipts.....	43,318,511	107,547,564	N/A	34,485,717	3,661,699	80,640,086	19,694,332	33,288,934	63,369,104	2,607,373	6,058,434
Net income (less deficit).....	17,640,561	45,553,645	N/A	9,074,005	1,309,829	35,862,524	5,098,400	2,017,579	13,853,457	1,420,425	2,082,707
Net income.....	19,037,864	48,143,441	N/A	10,132,426	1,531,840	36,732,244	7,178,207	3,152,154	15,372,613	1,495,125	2,375,986
Deficit.....	1,397,303	2,589,796	N/A	1,058,421	222,011	869,720	2,079,807	1,134,575	1,519,156	74,700	293,279

N/A - not applicable.

¹ Total Corporation "Net income (less deficit)" includes "Total net income (less deficit)" from S Corporations and is more comprehensive than what SOI generally publishes.

² For this table, the computations for C Corporations also include 1120-RIC and 1120-REIT returns.

³ For Tax Year 1998 General Partnerships include partnerships listed on the tax return as General, Other and blank.

⁴ For Tax Year 1998 Limited Partnerships include Limited Partnerships and Limited Liability Partnerships.

NOTE: Detail may not add to total because of rounding.

Table 3B.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 1999

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries	Agriculture, forestry, fishing, and hunting	Mining	Utilities	Construction	Manufacturing	Wholesale and retail trade	Transportation and warehousing	Information	Finance and insurance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
All Businesses										
Number of businesses.....	24,448,466	563,589	176,043	18,733	2,991,812	694,345	3,759,529	972,915	364,517	1,016,375
Business receipts.....	18,899,080,668	134,816,195	142,755,807	541,101,464	1,253,264,625	5,138,844,358	5,390,347,492	569,394,999	884,229,947	1,928,698,837
Net income (less deficit).....	1,421,748,414	4,963,514	6,843,334	40,915,599	73,542,134	272,261,935	121,558,071	20,354,248	38,286,294	463,708,734
Net income.....	1,864,354,420	13,937,084	20,261,718	46,130,529	88,518,927	344,567,830	173,366,185	32,202,024	119,994,339	506,510,894
Deficit.....	442,605,999	8,973,569	13,418,384	5,214,930	14,976,794	72,305,896	51,808,114	11,847,777	81,708,045	42,802,160
Corporations										
Number of businesses.....	4,935,904	141,678	30,849	7,044	580,302	297,714	948,371	160,195	107,628	217,780
Business receipts.....	16,313,971,385	104,645,084	109,685,715	478,836,511	973,521,174	4,801,823,220	4,789,438,632	485,223,550	760,824,421	1,740,167,487
Net income (less deficit)(*).....	985,363,332	2,375,446	731,214	39,073,530	35,851,126	255,594,801	98,451,496	11,131,614	43,394,087	365,650,230
Net income.....	1,282,481,471	6,614,998	9,280,430	42,368,292	45,139,310	318,701,505	139,309,819	19,385,091	97,518,412	397,080,911
Deficit.....	297,118,133	4,239,551	8,549,216	3,294,762	9,288,185	63,106,705	40,858,323	8,253,478	54,124,325	31,430,681
C Corporations (†)										
Number of businesses.....	2,210,129	70,306	14,772	5,584	246,775	151,824	454,773	72,675	49,160	114,026
Business receipts.....	13,071,173,955	57,328,751	96,063,482	475,658,599	516,969,690	4,303,643,709	3,431,344,964	397,193,258	709,929,597	1,682,078,285
Net income (less deficit).....	791,606,921	1,010,347	-1,306,291	38,831,103	10,875,231	218,512,766	58,979,787	8,366,054	36,717,057	356,062,254
Net income.....	1,041,919,838	2,961,219	6,668,489	42,072,320	16,688,111	276,562,059	92,082,454	14,566,133	88,049,936	384,558,606
Deficit.....	250,312,911	1,950,871	7,974,780	3,241,217	5,812,880	58,049,293	33,102,667	6,200,079	51,332,879	28,496,352
S Corporations										
Number of businesses.....	2,725,775	71,372	16,077	1,460	333,527	145,890	493,598	87,520	58,468	103,754
Business receipts.....	3,242,797,429	47,316,333	13,622,233	3,177,912	456,551,484	498,179,511	1,358,093,668	88,030,292	50,894,824	58,089,202
Total net income (less deficit).....	193,756,411	1,365,099	2,037,505	242,427	24,975,895	37,082,035	39,471,709	2,765,560	6,677,030	9,587,976
Net income.....	240,561,633	3,653,779	2,611,941	295,972	28,451,199	42,139,446	47,227,365	4,818,958	9,468,476	12,522,305
Deficit.....	46,805,222	2,288,680	574,436	53,545	3,475,305	5,057,412	7,755,656	2,053,399	2,791,446	2,934,329
Partnerships										
Number of businesses.....	1,936,919	115,006	28,095	2,612	127,581	37,072	141,851	22,344	20,343	219,233
Business receipts.....	1,615,762,245	13,518,418	28,635,592	62,156,799	125,518,084	309,693,927	372,693,889	38,182,156	116,417,632	102,140,730
Net income (less deficit).....	228,438,105	1,343,662	6,252,201	1,819,162	9,360,698	13,058,214	6,441,214	2,046,745	-6,930,530	83,643,256
Net income.....	348,467,958	4,938,301	10,237,584	3,728,757	13,191,380	21,586,149	11,910,451	4,545,507	20,130,834	93,379,163
Deficit.....	120,029,853	3,594,639	3,985,383	1,909,595	3,830,682	8,527,935	5,469,238	2,498,762	27,061,363	9,735,907
General (‡)										
Number of businesses.....	950,608	85,161	10,815	562	64,934	18,022	85,523	10,210	10,461	106,696
Business receipts.....	382,760,263	4,195,470	8,533,483	5,623,536	38,250,028	58,245,905	87,510,848	6,316,124	31,613,268	25,144,449
Net income (less deficit).....	85,767,233	1,871,577	1,192,332	782,459	3,405,775	3,922,729	2,458,581	1,208,171	2,483,966	23,882,686
Net income.....	108,487,666	3,113,116	3,783,675	1,145,811	4,498,980	5,575,818	3,577,461	1,532,958	6,304,209	25,489,562
Deficit.....	22,720,433	1,241,539	2,591,343	363,351	1,093,205	1,653,089	1,118,880	324,788	3,820,243	1,606,876
Limited (¶)										
Number of businesses.....	396,908	12,532	9,907	1,113	13,998	2,987	8,444	1,947	2,036	68,007
Business receipts.....	644,246,861	3,824,836	12,663,341	35,833,837	32,406,961	115,079,403	148,171,203	10,840,622	62,306,828	47,683,031
Net income (less deficit).....	107,937,194	-361,913	4,342,538	1,171,164	2,538,434	6,212,157	2,246,290	1,368,209	-2,212,176	42,286,392
Net income.....	157,244,765	609,892	4,872,244	2,119,068	3,728,794	8,545,529	3,355,429	2,137,725	11,344,940	46,538,059
Deficit.....	49,307,571	971,805	529,706	947,904	1,190,360	2,333,372	1,109,139	769,516	13,557,116	4,251,666
LLC										
Number of businesses.....	589,403	17,312	7,372	936	48,650	16,062	47,885	10,188	7,846	44,530
Business receipts.....	588,755,121	5,498,111	7,438,788	20,699,426	54,861,096	136,368,619	137,011,837	21,025,410	22,497,536	29,313,251
Net income (less deficit).....	34,733,678	-166,002	717,331	-134,461	3,416,489	2,923,328	1,736,342	-529,635	-7,202,319	17,474,178
Net income.....	82,735,527	1,215,293	1,581,665	463,879	4,963,606	7,464,802	4,977,561	874,824	2,481,685	21,351,542
Deficit.....	48,001,849	1,381,295	864,334	598,339	1,547,118	4,541,474	3,241,219	1,404,459	9,684,004	3,877,364
Nonfarm Sole Proprietorships										
Number of businesses.....	17,575,643	306,905	117,099	9,077	2,283,929	359,559	2,669,307	790,376	236,546	579,362
Business receipts.....	969,347,038	16,652,693	4,434,500	108,154	154,225,367	27,327,211	228,214,971	45,989,293	6,987,894	86,390,620
Net income (less deficit).....	207,946,977	1,244,406	-140,081	22,907	28,330,310	3,608,920	16,665,361	7,175,889	1,822,737	14,415,248
Net income.....	233,404,991	2,383,785	743,704	33,480	30,188,237	4,280,176	22,145,915	8,271,426	2,345,093	16,050,820
Deficit.....	25,458,013	1,139,379	883,785	10,573	1,857,927	671,256	5,480,553	1,095,537	522,357	1,635,572

Footnotes at end of table.

Table 3B.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 1999--Continued

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Real estate and rental and leasing	Professional, scientific, and technical services	Management of companies (holding companies)	Administrative and support and waste management services	Educational services	Health care and social assistance	Arts, entertainment, and recreation	Accommodation, food services, and drinking places	Other services	Religious, grantmaking, civic, professional, and similar	Unclassified industries
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
All Businesses											
Number of businesses.....	2,230,947	3,223,670	55,907	1,693,387	367,654	1,863,824	1,167,836	630,425	2,154,135	210,843	291,981
Business receipts.....	280,466,415	855,476,153	95,722,366	352,129,454	26,048,213	519,887,619	115,655,479	436,626,093	222,514,017	2,209,867	8,891,250
Net income (less deficit).....	82,461,634	103,489,041	73,997,346	20,955,739	2,383,814	49,983,138	8,222,183	15,954,203	19,343,018	1,208,280	1,316,153
Net income.....	129,509,061	136,420,383	85,042,620	27,631,217	3,396,647	64,051,082	17,723,489	27,562,816	24,280,677	1,320,620	1,926,279
Deficit.....	47,047,426	32,931,342	11,045,273	6,675,477	1,012,832	14,067,944	9,501,305	11,608,614	4,937,658	112,340	610,124
Corporations											
Number of businesses.....	521,447	657,153	43,246	205,011	35,196	303,499	93,922	252,113	305,725	N/A	27,031
Business receipts.....	185,450,183	576,276,292	91,583,476	283,700,509	20,532,679	371,442,071	70,756,712	318,528,271	146,498,454	N/A	5,036,944
Net income (less deficit)(¹).....	14,525,074	17,633,962	67,069,382	8,865,906	666,803	5,883,711	2,450,222	11,065,417	4,828,525	N/A	120,785
Net income.....	26,723,002	43,324,463	74,005,614	13,705,989	1,384,796	15,926,481	6,389,531	17,528,785	7,629,604	N/A	464,438
Deficit.....	12,197,926	25,690,502	6,936,231	4,840,083	717,992	10,042,770	3,939,308	6,463,368	2,801,078	N/A	343,651
C Corporations (²)											
Number of businesses.....	214,262	259,460	23,526	71,327	14,353	165,886	35,576	94,577	140,920	N/A	10,348
Business receipts.....	126,943,155	370,936,482	87,892,147	186,717,710	11,615,065	286,220,509	38,832,310	208,611,516	82,132,863	N/A	1,061,862
Net income (less deficit).....	2,810,303	-4,515,568	56,275,439	2,781,004	21,357	-1,716,764	-356,592	7,226,673	1,234,499	N/A	-201,738
Net income.....	11,028,235	17,205,755	60,573,733	6,205,359	628,573	6,715,148	1,774,765	10,656,829	2,878,130	N/A	43,984
Deficit.....	8,217,931	21,721,323	4,298,294	3,424,355	607,215	8,431,912	2,131,357	3,430,156	1,643,630	N/A	245,720
S Corporations											
Number of businesses.....	307,185	397,693	19,720	133,684	20,843	137,613	58,346	157,536	164,805	N/A	16,683
Business receipts.....	58,507,028	205,339,810	3,691,329	96,982,799	8,917,614	85,221,562	31,924,402	109,916,755	64,365,591	N/A	3,975,082
Total net income (less deficit).....	11,714,771	22,149,530	10,793,943	6,084,902	645,446	7,600,475	2,806,814	3,838,744	3,594,026	N/A	322,523
Net income.....	15,694,767	26,118,708	13,431,881	7,500,630	756,223	9,211,333	4,614,766	6,871,956	4,751,474	N/A	420,454
Deficit.....	3,979,995	3,969,179	2,637,937	1,415,728	110,777	1,610,858	1,807,951	3,033,212	1,157,448	N/A	97,931
Partnerships											
Number of businesses.....	858,066	122,773	12,661	32,508	6,015	39,890	33,705	63,162	51,822	N/A	2,182
Business receipts.....	52,143,490	172,277,572	4,138,910	31,147,073	1,359,899	65,685,097	25,444,429	81,804,555	12,298,764	N/A	505,229
Net income (less deficit).....	49,665,658	40,628,476	6,927,964	1,512,770	123,489	8,486,828	421,718	2,733,972	883,768	N/A	18,840
Net income.....	83,003,855	44,880,009	11,037,006	2,387,425	204,424	11,255,870	3,925,572	6,602,193	1,416,643	N/A	106,835
Deficit.....	33,338,198	4,251,533	4,109,042	874,654	80,935	2,769,042	3,503,854	3,868,222	532,875	N/A	87,995
General (³)											
Number of businesses.....	377,717	54,360	2,709	17,423	3,448	17,602	16,184	30,563	37,457	N/A	762
Business receipts.....	9,209,131	52,980,673	294,875	5,339,017	234,885	16,510,480	6,072,807	21,365,619	5,245,444	N/A	74,221
Net income (less deficit).....	19,373,161	15,887,529	1,386,583	510,427	17,022	4,442,354	866,692	1,438,950	655,267	N/A	-19,029
Net income.....	24,778,501	16,525,330	2,557,509	636,020	44,426	4,655,789	1,586,373	1,926,047	747,882	N/A	8,199
Deficit.....	5,405,340	637,801	1,170,926	125,592	27,404	213,435	719,681	487,097	92,615	N/A	27,228
Limited (⁴)											
Number of businesses.....	229,572	16,945	4,745	2,701	180	6,245	4,132	9,016	2,101	N/A	300
Business receipts.....	20,470,814	73,994,646	2,110,770	7,017,025	258,098	25,320,021	11,610,864	32,484,727	2,140,787	N/A	29,047
Net income (less deficit).....	22,566,267	19,626,628	3,396,412	546,176	31,606	2,188,901	296,921	1,527,297	166,385	N/A	-493
Net income.....	40,399,430	20,424,639	4,426,640	740,350	39,237	3,138,158	1,529,429	3,066,015	228,512	N/A	676
Deficit.....	17,833,163	798,011	1,030,229	194,174	7,631	949,257	1,232,508	1,538,718	62,127	N/A	1,169
LLC											
Number of businesses.....	250,777	51,468	5,207	12,384	2,387	16,042	13,389	23,583	12,264	N/A	1,120
Business receipts.....	22,463,545	45,302,253	1,733,265	18,791,031	866,917	23,854,596	7,760,757	27,954,209	4,912,533	N/A	401,961
Net income (less deficit).....	7,726,230	5,114,319	2,144,969	456,167	74,861	1,855,573	-741,895	-232,276	62,116	N/A	38,361
Net income.....	17,825,925	7,930,041	4,052,857	1,011,055	120,761	3,461,923	809,770	1,610,131	440,249	N/A	97,960
Deficit.....	10,099,695	2,815,721	1,907,888	554,888	45,900	1,606,349	1,551,665	1,842,407	378,133	N/A	59,599
Nonfarm Sole Proprietorships											
Number of businesses.....	851,434	2,443,744	N/A	1,455,868	326,443	1,520,435	1,040,209	315,150	1,796,588	210,843	262,768
Business receipts.....	42,872,742	106,922,289	N/A	37,281,872	4,155,635	82,760,451	19,454,338	36,293,267	63,716,799	2,209,867	3,349,077
Net income (less deficit).....	18,270,902	45,226,603	N/A	10,577,063	1,593,522	35,612,599	5,350,243	2,154,814	13,630,725	1,208,280	1,176,528
Net income.....	19,782,204	48,215,911	N/A	11,537,803	1,807,427	36,868,731	7,408,386	3,431,838	15,234,430	1,320,620	1,355,006
Deficit.....	1,511,302	2,989,307	N/A	960,740	213,905	1,256,132	2,058,143	1,277,024	1,603,705	112,340	178,478

N/A - not applicable.

¹ Total Corporation "Net income (less deficit)" includes "Total net income (less deficit)" from S Corporations and is more comprehensive than what SOI generally publishes.

² For this table, the computations for C Corporations also include 1120-RIC and 1120-REIT returns.

³ For Tax Year 1999 General Partnerships include partnerships listed on the tax return as General, Other and blank.

⁴ For Tax Year 1999 Limited Partnerships include Limited Partnerships and Limited Liability Partnerships.

NOTE: Detail may not add to total because of rounding.

Table 3C.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 2000

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries	Agriculture, forestry, fishing, and hunting	Mining	Utilities	Construction	Manufacturing	Wholesale and retail trade	Transportation and warehousing	Information	Finance and insurance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
All Businesses										
Number of businesses.....	25,007,505	532,328	165,304	24,441	2,958,179	678,953	3,797,576	1,076,305	427,654	1,043,242
Business receipts.....	18,659,570,396	122,612,734	146,867,803	708,180,639	1,194,678,304	5,287,885,546	5,490,535,822	558,173,928	824,439,100	1,628,868,725
Net income (less deficit).....	1,201,936,567	4,305,215	12,129,647	29,318,476	62,456,542	283,928,550	107,547,979	16,466,208	12,061,210	401,601,964
Net income.....	1,636,649,354	10,079,049	18,010,790	35,430,405	76,315,388	367,392,144	167,064,668	28,876,412	98,830,177	446,018,129
Deficit.....	434,712,784	5,773,836	5,881,144	6,111,928	13,858,846	83,463,593	59,516,688	12,410,204	86,768,967	44,416,166
Corporations										
Number of businesses.....	5,045,274	140,851	32,578	7,968	597,902	288,506	959,575	160,437	118,073	221,394
Business receipts.....	17,636,551,349	106,085,760	140,917,053	707,815,083	1,034,087,166	5,259,173,394	5,267,581,835	505,713,781	817,186,647	1,525,629,096
Net income (less deficit)(¹).....	986,952,278	2,771,799	11,568,288	29,268,805	35,757,665	279,610,134	92,637,276	8,959,964	10,171,572	387,653,903
Net income.....	1,391,008,755	7,549,336	16,664,668	35,355,913	46,969,598	362,321,332	145,734,841	19,984,584	96,384,845	429,289,049
Deficit.....	404,056,474	4,777,538	5,096,381	6,087,107	11,211,933	82,711,197	53,097,565	11,024,620	86,213,273	41,635,148
C Corporations (²)										
Number of businesses.....	2,184,795	68,555	14,892	5,413	232,294	141,687	453,838	71,417	55,995	104,563
Business receipts.....	14,078,901,184	57,708,101	122,891,531	703,863,380	522,979,306	4,737,156,398	3,767,376,961	414,456,985	764,211,744	1,452,461,321
Net income (less deficit).....	788,416,390	1,099,041	7,610,738	29,085,238	9,873,890	246,352,850	54,099,727	6,716,444	4,031,594	373,773,331
Net income.....	1,136,792,550	3,070,493	12,155,823	35,048,390	16,460,765	323,064,519	96,649,397	14,990,511	86,311,839	411,646,454
Deficit.....	348,376,157	1,971,453	4,545,086	5,963,151	6,586,875	76,711,668	42,549,670	8,274,067	82,280,245	37,873,124
S Corporations										
Number of businesses.....	2,860,478	72,296	17,686	2,555	365,608	146,819	505,737	89,020	62,078	116,831
Business receipts.....	3,557,650,166	48,377,659	18,025,522	3,951,703	511,107,860	522,016,996	1,500,204,874	91,256,796	52,974,903	73,167,775
Total net income (less deficit).....	198,535,888	1,672,758	3,957,550	183,567	25,883,775	33,257,284	38,537,549	2,243,520	6,139,978	13,880,572
Net income.....	254,216,205	4,478,843	4,508,845	307,523	30,508,833	39,256,813	49,085,444	4,994,073	10,073,006	17,642,595
Deficit.....	55,680,317	2,806,085	551,295	123,956	4,625,058	5,999,529	10,547,895	2,750,553	3,933,028	3,762,024
Partnerships										
Number of businesses.....	2,057,500	113,931	26,084	2,453	115,509	37,950	148,305	26,941	26,945	251,657
Business receipts.....	2,061,764	16,320	57,347	107,719	140,387	411,568	493,306	43,745	139,237	131,752
Net income (less deficit).....	268,991	214	15,898	3,608	10,320	17,284	7,045	2,676	-3,497	99,656
Net income.....	409,973	4,668	20,474	5,896	14,034	26,947	14,372	5,491	20,517	115,087
Deficit.....	140,982	4,454	4,576	2,288	3,714	9,663	7,327	2,815	24,014	15,431
General (³)										
Number of businesses.....	936,564	80,041	10,442	261	54,608	17,908	85,311	13,753	13,772	115,364
Business receipts.....	425,752	5,258	13,740	8,015	37,885	67,696	99,816	6,574	39,208	26,317
Net income (less deficit).....	101,787	1,252	5,067	1,253	3,595	4,621	2,435	1,177	2,915	32,836
Net income.....	127,059	2,810	7,770	1,558	4,471	6,088	3,392	1,816	6,312	36,385
Deficit.....	25,272	1,558	2,704	305	876	1,467	957	639	3,397	3,548
Limited (⁴)										
Number of businesses.....	402,232	12,469	7,482	682	10,352	1,933	8,242	1,487	1,503	78,455
Business receipts.....	830,430	3,705	19,978	54,237	36,292	155,576	212,811	12,241	63,814	73,544
Net income (less deficit).....	119,512	-401	7,867	1,553	2,877	8,189	3,959	2,397	580	40,192
Net income.....	170,929	654	8,530	2,725	4,089	10,673	5,238	2,872	10,558	46,406
Deficit.....	51,417	1,055	663	1,172	1,212	2,484	596	475	9,977	6,214
LLC										
Number of businesses.....	718,704	21,421	8,160	1,510	50,548	18,109	54,752	11,702	11,669	57,838
Business receipts.....	805,582	7,357	23,629	45,467	66,210	188,295	180,679	24,930	36,215	31,891
Net income (less deficit).....	47,692	-636	2,964	802	3,848	4,475	651	-898	-6,992	26,628
Net income.....	111,984	1,204	4,174	1,613	5,474	10,187	5,741	802	3,647	32,297
Deficit.....	64,292	1,840	1,210	811	1,626	5,712	5,090	1,701	10,639	5,669
Nonfarm Sole Proprietorships										
Number of businesses.....	17,904,731	277,546	106,642	14,020	2,244,768	352,497	2,689,696	888,927	282,636	570,191
Business receipts.....	1,020,957,283	16,510,654	5,893,403	257,837	160,450,751	28,300,584	222,460,681	52,416,402	7,113,216	103,107,877
Net income (less deficit).....	214,715,298	1,533,202	545,461	46,063	26,688,557	4,301,132	14,903,658	7,503,568	1,893,135	13,848,405
Net income.....	245,230,626	2,525,045	1,325,648	68,596	29,331,756	5,043,865	21,315,455	8,886,337	2,424,815	16,613,993
Deficit.....	30,515,328	991,844	780,187	22,533	2,643,199	742,733	6,411,796	1,382,769	531,680	2,765,587

Footnotes at end of table.

Table 3C.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 2000--Continued

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Real estate and rental and leasing	Professional, scientific, and technical services	Management of companies (holding companies)	Administrative and support and waste management services	Educational services	Health care and social assistance	Arts, entertainment, and recreation	Accommodation, food services, and drinking places	Other services	Religious, grantmaking, civic, professional, and similar	Unclassified industries
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
All Businesses											
Number of businesses.....	2,373,298	3,270,162	64,278	1,798,842	394,803	1,945,785	1,208,571	642,061	2,103,502	226,867	275,355
Business receipts.....	315,146,764	931,622,176	133,216,727	394,519,709	28,858,868	563,805,144	123,169,925	473,426,560	244,741,572	2,933,690	5,748,721
Net income (less deficit).....	87,077,469	87,155,800	92,033,462	19,334,538	2,637,723	54,806,309	6,304,888	15,221,225	20,845,350	1,498,270	1,277,086
Net income.....	142,901,801	147,126,920	102,671,023	29,520,694	4,010,504	68,237,725	17,552,136	28,772,909	26,614,063	1,679,507	1,849,846
Deficit.....	55,824,333	59,971,120	11,037,559	10,186,155	1,372,780	13,431,416	11,247,248	13,551,683	5,768,714	181,237	572,759
Corporations											
Number of businesses.....	532,426	689,412	47,542	211,993	36,756	306,352	97,866	257,525	316,138	N/A	21,980
Business receipts.....	204,519,672	623,368,137	127,242,280	313,932,798	22,021,416	403,580,914	64,157,666	346,989,626	164,175,721	N/A	2,373,305
Net income (less deficit)(¹).....	14,935,833	-3,906,788	84,733,578	5,987,563	816,534	8,171,921	1,232,220	11,155,487	5,828,456	N/A	-1,933
Net income.....	29,322,559	45,913,867	90,006,013	13,622,169	1,811,683	17,428,292	5,630,381	17,775,976	9,022,211	N/A	221,437
Deficit.....	14,386,726	49,820,656	5,672,434	7,634,605	995,148	9,256,371	4,398,161	6,620,488	3,193,756	N/A	223,369
C Corporations (²)											
Number of businesses.....	212,680	263,494	26,357	72,978	15,125	163,465	35,395	93,618	141,282	N/A	11,747
Business receipts.....	138,723,611	400,696,546	122,928,517	204,978,744	13,743,225	304,962,586	34,866,246	227,687,772	85,900,693	N/A	1,307,517
Net income (less deficit).....	1,450,889	-26,918,719	75,886,309	-487,166	96,848	-1,287,652	-405,031	6,497,107	1,021,006	N/A	-80,054
Net income.....	10,906,984	17,834,015	80,199,791	5,527,204	964,208	6,549,491	1,901,142	10,435,175	2,987,290	N/A	89,059
Deficit.....	9,456,094	44,752,734	4,313,482	6,014,369	867,359	7,837,142	2,306,173	3,938,068	1,966,285	N/A	169,112
S Corporations											
Number of businesses.....	319,746	425,918	21,185	139,015	21,631	142,887	62,471	163,907	174,856	N/A	10,233
Business receipts.....	65,796,061	222,671,591	4,313,763	108,954,054	8,278,191	98,618,328	29,291,420	119,301,854	78,275,028	N/A	1,065,788
Total net income (less deficit).....	13,484,944	23,011,931	8,847,269	6,474,729	719,686	9,459,573	1,637,251	4,658,380	4,807,450	N/A	78,121
Net income.....	18,415,575	28,079,852	9,806,222	8,094,965	847,475	10,878,801	3,729,239	7,340,801	6,034,921	N/A	132,378
Deficit.....	4,930,632	5,067,922	1,358,952	1,620,236	127,789	1,419,229	2,091,988	2,682,420	1,227,471	N/A	54,257
Partnerships											
Number of businesses.....	905,796	135,905	16,736	37,696	5,752	44,038	35,091	62,076	61,643	N/A	2,991
Business receipts.....	61,899,580	193,998,910	5,974,447	40,370,566	2,033,451	73,247,847	38,443,515	89,091,640	15,045,733	N/A	276,756
Net income (less deficit).....	51,598,841	42,945,726	7,299,884	1,771,173	75,145	9,758,764	-302,694	1,820,136	824,433	N/A	-4,276
Net income.....	91,406,835	49,516,987	12,665,010	2,978,182	241,673	12,575,743	4,028,424	7,499,357	1,556,234	N/A	18,979
Deficit.....	39,807,995	6,571,261	5,365,125	1,207,009	166,528	2,816,979	4,331,118	5,679,221	731,801	N/A	23,255
General (³)											
Number of businesses.....	366,696	56,581	2,500	18,099	3,308	16,237	16,958	25,712	37,859	N/A	1,154
Business receipts.....	9,456,095	53,075,905	225,176	5,748,549	261,393	16,786,348	13,161,141	16,727,609	5,744,432	N/A	54,501
Net income (less deficit).....	19,264,923	17,847,674	1,415,642	446,421	-11,057	4,648,902	1,140,120	1,252,200	632,191	N/A	20
Net income.....	24,838,280	18,442,731	3,472,805	548,912	35,265	4,820,173	1,789,919	1,752,693	753,859	N/A	2,402
Deficit.....	5,573,357	595,057	2,057,163	102,492	46,322	171,271	649,799	500,493	121,667	N/A	2,382
Limited (⁴)											
Number of businesses.....	227,085	18,488	6,165	3,110	71	8,073	3,667	8,499	3,339	N/A	1,130
Business receipts.....	21,684,835	83,972,072	1,707,402	11,778,902	315,263	28,416,653	12,912,004	35,097,027	2,346,855	N/A	--
Net income (less deficit).....	23,225,545	21,002,777	3,325,341	363,944	101,644	2,698,524	-102,030	1,594,389	81,363	N/A	7,744
Net income.....	43,154,732	22,260,640	4,382,362	589,135	101,644	3,628,344	1,408,839	3,481,317	169,662	N/A	8,046
Deficit.....	19,929,187	1,257,863	1,057,022	225,191	--	929,821	1,510,869	1,886,928	88,298	N/A	302
LLC											
Number of businesses.....	312,016	60,836	8,071	16,487	2,373	19,728	14,466	27,866	20,445	N/A	707
Business receipts.....	30,758,650	56,950,933	4,041,868	22,843,115	1,456,796	28,044,845	12,370,370	37,267,004	6,954,445	N/A	222,255
Net income (less deficit).....	9,108,373	4,095,275	2,558,902	960,809	-15,442	2,411,338	-1,340,784	-1,026,453	110,878	N/A	-12,040
Net income.....	23,413,824	8,813,617	4,809,842	1,840,135	104,764	4,127,225	829,666	2,265,348	632,713	N/A	8,532
Deficit.....	14,305,450	4,718,341	2,250,940	879,327	120,206	1,715,888	2,170,450	3,291,801	521,835	N/A	20,572
Nonfarm Sole Proprietorships											
Number of businesses.....	935,076	2,444,845	N/A	1,549,153	352,295	1,595,395	1,075,614	322,460	1,725,721	226,867	250,384
Business receipts.....	48,727,512	114,255,129	N/A	40,216,345	4,804,001	86,976,383	20,568,744	37,345,294	65,520,118	2,933,690	3,098,660
Net income (less deficit).....	20,542,795	48,116,862	N/A	11,575,802	1,746,044	36,875,624	5,375,362	2,245,602	14,192,461	1,498,270	1,283,295
Net income.....	22,172,407	51,696,066	N/A	12,920,343	1,957,148	38,233,690	7,893,331	3,497,576	16,035,618	1,679,507	1,609,430
Deficit.....	1,629,612	3,579,203	N/A	1,344,541	211,104	1,358,066	2,517,969	1,251,974	1,843,157	181,237	326,135

N/A - not applicable.

¹ Total Corporation "Net income (less deficit)" includes "Total net income (less deficit)" from S Corporations and is more comprehensive than what SOI generally publishes.

² For this table, the computations for C Corporations also include 1120-RIC and 1120-REIT returns.

³ For Tax Year 2000 General Partnerships include partnerships listed on the tax return as General, Foreign, Other and blank.

⁴ For Tax Year 2000 Limited Partnerships include Domestic Limited Partnerships and Domestic Limited Liability Partnerships.

NOTE: Detail may not add to total because of rounding.

Table 3D.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 2001

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries	Agriculture, forestry, fishing, and hunting	Mining	Utilities	Construction	Manufacturing	Wholesale and retail trade	Transportation and warehousing	Information	Finance and insurance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
All Businesses										
Number of businesses.....	25,605,898	528,224	173,580	19,566	3,124,732	662,521	3,674,362	1,129,498	426,500	1,059,181
Business receipts.....	20,799,323,834	137,726,166	218,469,712	1,143,709,184	1,405,785,332	5,348,916,414	5,897,663,230	593,002,438	982,177,427	1,684,991,171
Net income (less deficit).....	1,142,478,028	3,674,270	24,844,876	21,425,917	74,462,287	160,456,099	103,595,361	3,563,871	-44,851,759	386,021,771
Net income.....	1,851,745,213	14,046,103	39,158,379	36,088,138	95,238,192	313,684,796	171,974,002	29,819,784	83,596,193	468,340,204
Deficit.....	709,267,183	10,371,833	14,313,506	14,662,222	20,775,905	153,228,696	68,378,639	26,255,912	128,447,950	82,318,433
Corporations										
Number of businesses.....	5,135,591	140,806	31,776	7,802	624,478	278,995	963,403	164,492	115,435	220,895
Business receipts.....	17,504,288,630	102,909,416	151,151,906	1,004,358,112	1,084,579,920	4,862,174,424	5,183,197,415	493,765,699	815,772,817	1,430,898,834
Net income (less deficit) (*).....	648,758,088	1,672,678	10,279,423	16,965,071	35,393,571	134,837,427	84,526,491	-6,441,292	-35,586,988	272,519,760
Net income.....	1,155,497,719	6,429,255	17,023,541	28,877,598	48,628,978	274,142,214	136,822,456	14,678,456	61,329,395	332,135,408
Deficit.....	506,739,630	4,756,577	6,744,121	11,912,527	13,235,392	139,304,786	52,295,963	21,119,748	96,916,381	59,615,647
C Corporations (†)										
Number of businesses.....	2,149,105	66,284	13,908	5,941	238,116	139,508	440,523	73,304	52,769	99,141
Business receipts.....	13,813,168,479	56,153,283	130,106,865	999,589,343	535,734,095	4,359,364,517	3,647,616,000	399,221,076	765,512,006	1,363,009,858
Net income (less deficit).....	461,071,171	1,221,679	5,628,672	16,585,894	9,018,523	110,021,373	43,425,834	-8,021,537	-36,958,491	260,174,240
Net income.....	906,633,873	3,080,882	11,766,438	28,371,482	16,948,122	241,206,592	85,924,640	10,290,662	54,766,601	316,581,583
Deficit.....	445,562,701	1,859,203	6,137,768	11,785,588	7,929,598	131,185,218	42,498,806	18,312,199	91,725,091	56,407,343
S Corporations										
Number of businesses.....	2,986,486	74,522	17,868	1,861	386,362	139,487	522,880	91,188	62,666	121,754
Business receipts.....	3,691,120,151	46,756,133	21,045,041	4,768,769	548,845,825	502,809,907	1,535,581,415	94,544,623	50,260,811	67,888,976
Total net income (less deficit).....	187,686,917	450,999	4,650,751	379,177	26,375,055	24,816,054	41,100,657	1,580,245	1,371,503	12,345,520
Net income.....	248,863,846	3,348,373	5,257,103	506,116	31,680,849	32,935,822	50,897,816	4,387,794	6,562,794	15,553,825
Deficit.....	61,176,929	2,897,374	606,353	126,939	5,305,794	8,119,568	9,797,157	2,807,549	5,191,290	3,208,304
Partnerships										
Number of businesses.....	2,132,117	117,343	27,269	2,757	127,374	36,514	146,402	25,483	26,091	261,682
Business receipts.....	2,278,200,526	18,573,227	60,502,000	139,090,586	156,967,238	462,062,912	490,913,434	46,548,552	158,779,118	171,469,593
Net income (less deficit).....	276,334,824	678,466	13,958,241	4,390,151	10,538,118	22,184,926	5,478,305	1,914,673	-10,946,478	99,627,703
Net income.....	446,069,172	5,276,110	20,573,102	7,123,443	15,132,697	35,451,133	14,795,537	5,487,560	19,994,802	119,943,530
Deficit.....	169,734,347	4,597,644	6,614,861	2,733,293	4,594,579	13,266,207	9,317,232	3,572,887	30,941,280	20,315,827
General (‡)										
Number of businesses.....	885,457	77,990	10,603	540	55,127	15,935	77,574	10,506	11,563	104,824
Business receipts.....	464,251,886	4,268,379	13,138,627	9,480,774	40,243,629	118,149,292	91,105,525	6,962,623	44,097,606	17,133,339
Net income (less deficit).....	101,830,079	1,761,759	2,912,285	1,276,453	3,618,801	8,855,695	2,287,250	1,223,053	2,102,636	30,644,767
Net income.....	128,591,551	3,197,829	6,431,979	1,748,849	4,592,540	10,282,879	3,358,011	1,873,241	6,528,094	34,437,101
Deficit.....	26,761,472	1,436,070	3,519,694	472,396	973,739	1,427,184	1,070,760	650,189	4,425,457	3,792,334
Limited (§)										
Number of businesses.....	437,968	17,394	7,810	931	11,129	2,903	9,291	2,938	2,167	87,192
Business receipts.....	876,234,279	3,827,239	18,267,977	72,523,323	39,803,876	145,959,928	187,696,593	14,272,618	66,649,516	113,439,079
Net income (less deficit).....	127,448,902	-547,612	7,943,390	2,457,025	3,218,412	7,091,113	3,395,725	1,938,867	-5,262,980	44,697,072
Net income.....	187,146,566	674,613	9,236,149	3,930,377	4,374,005	11,892,494	4,867,844	2,590,253	8,199,391	49,805,651
Deficit.....	59,697,664	1,222,225	1,292,759	1,473,352	1,155,593	4,801,380	1,472,119	651,386	13,462,370	5,108,579
LLC										
Number of businesses.....	808,692	21,959	8,856	1,287	61,117	17,677	59,537	12,038	12,361	69,665
Business receipts.....	937,714,361	10,477,609	29,095,395	57,086,489	76,919,733	197,953,692	212,111,316	25,313,311	48,031,996	40,897,175
Net income (less deficit).....	47,055,843	-535,682	3,102,566	656,672	3,700,905	6,238,117	-204,671	-1,247,247	-7,786,135	24,285,864
Net income.....	130,331,055	1,403,668	4,904,974	1,444,217	6,166,151	13,275,760	6,569,682	1,024,066	5,267,317	35,700,778
Deficit.....	83,275,212	1,939,350	1,802,408	787,545	2,465,247	7,037,643	6,774,353	2,271,313	13,053,452	11,414,914
Nonfarm Sole Proprietorships										
Number of businesses.....	18,338,190	270,075	114,535	9,007	2,372,880	347,012	2,564,557	939,523	284,974	576,604
Business receipts.....	1,016,834,678	16,243,523	6,815,806	260,486	164,238,174	24,679,078	223,552,381	52,688,187	7,625,492	82,622,744
Net income (less deficit).....	217,385,116	1,323,126	607,212	70,695	28,530,591	3,433,746	13,590,565	8,090,490	1,681,707	13,874,308
Net income.....	250,178,322	2,340,738	1,561,736	87,097	31,476,524	4,091,449	20,356,009	9,653,768	2,271,996	16,261,266
Deficit.....	32,793,206	1,017,612	954,524	16,402	2,945,934	657,703	6,765,444	1,563,277	590,289	2,386,959

Footnotes at end of table.

Table 3D.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 2001--Continued

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Real estate and rental and leasing	Professional, scientific, and technical services	Management of companies (holding companies)	Administrative and support and waste management services	Educational services	Health care and social assistance	Arts, entertainment, and recreation	Accommodation, food services, and drinking places	Other services	Religious, grantmaking, civic, professional, and similar	Unclassified industries
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
All Businesses											
Number of businesses.....	2,456,254	3,445,157	63,211	1,829,793	422,180	2,051,024	1,174,566	691,094	2,237,355	231,591	205,507
Business receipts.....	325,077,096	965,106,321	182,587,302	421,976,061	31,760,887	608,972,873	133,977,372	456,267,719	255,091,915	2,837,353	3,227,862
Net income (less deficit).....	93,243,970	97,673,057	91,333,150	22,302,092	2,325,169	63,600,568	4,380,651	11,682,493	20,182,517	1,473,603	1,088,067
Net income.....	153,082,125	155,617,014	105,838,162	32,032,024	4,023,737	75,985,681	17,359,662	26,257,997	26,597,037	1,718,877	1,287,104
Deficit.....	59,838,158	57,943,958	14,505,011	9,729,931	1,698,571	12,385,113	12,979,008	14,575,506	6,414,520	245,275	199,039
Corporations											
Number of businesses.....	539,965	709,837	47,866	223,999	38,480	327,338	102,631	259,465	325,602	N/A	12,325
Business receipts.....	207,454,856	631,691,343	175,450,783	339,002,912	25,148,309	429,190,484	69,089,923	328,552,525	168,989,458	N/A	909,495
Net income (less deficit)(¹).....	13,816,572	-1,095,827	85,179,993	8,299,302	472,261	12,584,750	938,959	9,954,901	4,484,029	N/A	-43,000
Net income.....	28,291,489	45,485,912	93,187,021	15,431,400	1,676,488	20,580,467	5,686,721	17,006,538	8,088,334	N/A	26,055
Deficit.....	14,474,918	46,581,740	8,007,027	7,132,097	1,204,229	7,995,717	4,717,760	7,051,638	3,604,305	N/A	69,057
C Corporations (²)											
Number of businesses.....	208,012	260,025	26,419	72,341	14,407	157,124	35,406	92,568	144,389	N/A	8,916
Business receipts.....	138,430,430	394,400,768	170,384,509	209,587,067	14,424,654	308,545,859	35,905,198	203,384,005	81,561,619	N/A	237,329
Net income (less deficit).....	1,139,392	-26,513,768	79,034,349	1,582,025	-203,819	452,528	-857,275	4,829,631	564,473	N/A	-52,552
Net income.....	10,248,856	15,170,503	84,389,567	7,018,266	751,968	6,990,796	1,676,821	8,789,384	2,645,990	N/A	14,720
Deficit.....	9,109,465	41,684,271	5,355,217	5,436,240	955,789	6,538,268	2,534,094	3,959,754	2,081,517	N/A	67,272
S Corporations											
Number of businesses.....	331,953	449,812	21,447	151,658	24,073	170,214	67,225	166,897	181,213	N/A	3,409
Business receipts.....	69,024,426	237,290,575	5,066,274	129,415,845	10,723,655	120,644,625	33,184,725	125,168,520	87,427,839	N/A	*672,166
Total net income (less deficit).....	12,677,180	25,417,941	6,145,644	6,717,277	676,080	12,132,222	1,796,234	5,125,270	3,919,556	N/A	9,552
Net income.....	18,042,633	30,315,409	8,797,454	8,413,134	924,520	13,589,671	3,979,900	8,217,154	5,442,344	N/A	*11,335
Deficit.....	5,365,453	4,897,469	2,651,810	1,695,857	248,440	1,457,449	2,183,666	3,091,884	1,522,788	N/A	*1,785
Partnerships											
Number of businesses.....	948,200	143,045	15,345	38,516	5,240	44,689	34,594	70,171	58,454	N/A	2,948
Business receipts.....	68,470,179	214,642,623	7,136,519	43,650,320	1,763,853	86,253,831	43,679,315	90,282,581	17,267,790	N/A	146,854
Net income (less deficit).....	59,019,298	49,938,292	6,153,157	2,687,888	113,020	11,321,467	-1,906,125	258,538	901,189	N/A	23,996
Net income.....	102,358,616	57,199,172	12,651,141	3,842,198	286,122	14,439,819	3,763,575	5,977,669	1,727,430	N/A	*45,516
Deficit.....	43,339,319	7,260,881	6,497,984	1,154,310	173,102	3,118,352	5,669,699	5,719,131	826,241	N/A	21,520
General (³)											
Number of businesses.....	349,791	55,333	1,873	14,507	3,093	15,180	15,136	28,867	35,960	N/A	1,057
Business receipts.....	10,515,703	50,109,862	366,440	5,434,223	118,631	16,563,029	15,558,638	15,588,953	5,389,907	N/A	26,706
Net income (less deficit).....	21,108,782	18,677,683	243,766	452,376	36,807	4,284,728	1,037,447	672,090	652,315	N/A	-17,615
Net income.....	25,674,465	19,675,697	1,876,643	619,709	47,723	4,443,501	1,712,566	1,295,820	794,902	N/A	--
Deficit.....	4,565,683	998,014	1,632,878	167,333	10,916	158,774	675,119	623,730	143,587	N/A	17,615
Limited (⁴)											
Number of businesses.....	242,641	16,313	6,059	3,815	265	7,595	4,380	9,710	4,552	N/A	880
Business receipts.....	22,428,847	97,702,096	1,870,339	10,857,367	437,989	32,767,467	12,506,439	32,746,417	2,456,353	N/A	20,815
Net income (less deficit).....	26,599,055	26,578,068	4,033,049	1,044,300	113,397	3,814,195	-709,654	832,826	169,473	N/A	43,180
Net income.....	48,261,080	27,599,589	5,607,019	1,316,813	131,470	4,786,165	1,068,293	2,519,938	239,915	N/A	*45,508
Deficit.....	21,662,025	1,021,521	1,573,970	272,513	18,073	971,970	1,777,947	1,687,113	70,442	N/A	2,327
LLC											
Number of businesses.....	355,768	71,399	7,413	20,195	1,882	21,914	15,078	31,594	17,942	N/A	1,011
Business receipts.....	35,525,630	66,830,666	4,899,740	27,358,730	1,207,233	36,923,335	15,614,238	41,947,211	9,421,530	N/A	99,333
Net income (less deficit).....	11,311,461	4,682,540	1,876,343	1,191,212	-37,184	3,222,544	-2,233,917	-1,246,377	80,401	N/A	-1,569
Net income.....	28,423,072	9,923,886	5,167,479	1,905,676	106,929	5,210,152	982,716	2,161,910	692,613	N/A	*8
Deficit.....	17,111,610	5,241,345	3,291,136	714,464	144,113	1,987,608	3,216,633	3,408,288	612,213	N/A	1,577
Nonfarm Sole Proprietorships											
Number of businesses.....	968,089	2,592,275	N/A	1,567,278	378,460	1,678,997	1,037,341	361,458	1,853,299	231,591	190,234
Business receipts.....	49,152,061	118,772,355	N/A	39,322,829	4,848,725	93,528,558	21,208,134	37,432,613	68,834,667	2,837,353	2,171,513
Net income (less deficit).....	20,408,100	48,830,592	N/A	11,314,902	1,739,888	39,694,351	5,347,817	1,469,054	14,797,299	1,473,603	1,107,071
Net income.....	22,432,020	52,931,930	N/A	12,758,426	2,061,127	40,965,395	7,939,366	3,273,790	16,781,273	1,718,877	1,215,533
Deficit.....	2,023,921	4,101,337	N/A	1,443,524	321,240	1,271,044	2,591,549	1,804,737	1,983,974	245,275	108,462

N/A - not applicable.

¹ Estimate should be used with caution because of the small number of sample returns on which it is based.

² Total Corporation "Net income (less deficit)" includes "Total net income (less deficit)" from S Corporations and is more comprehensive than what SOI generally publishes.

³ For this table, the computations for C Corporations also include 1120-RIC and 1120-REIT returns.

⁴ For Tax Year 2001 General Partnerships include partnerships listed on the tax return as General, Foreign, Other and blank.

⁵ For Tax Year 2001 Limited Partnerships include Domestic Limited Partnerships and Domestic Limited Liability Partnerships.

NOTE: Detail may not add to total because of rounding.

Table 3E.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 2002

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	All industries	Agriculture, forestry, fishing, and hunting	Mining	Utilities	Construction	Manufacturing	Wholesale and retail trade	Transportation and warehousing	Information	Finance and insurance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(10)	(11)	(12)
All Businesses										
Number of businesses.....	26,434,293	562,647	169,687	17,283	3,060,857	628,868	3,753,503	1,153,198	372,249	1,134,714
Business receipts.....	20,741,003,999	141,220,484	203,416,985	684,621,006	1,418,625,997	5,331,158,546	6,031,582,090	617,883,492	973,137,236	1,825,601,822
Net income (less deficit).....	1,088,304,476	-203,658	14,324,289	48,277	69,152,166	148,924,229	112,870,246	2,520,763	-37,650,355	354,829,875
Net income.....	1,781,234,414	11,936,961	29,153,524	28,232,018	94,333,280	289,209,459	175,744,485	30,743,808	70,695,627	433,584,763
Deficit.....	692,929,935	12,140,619	14,829,237	28,183,742	25,181,113	140,285,228	62,874,239	28,223,045	108,345,980	78,754,887
Corporations										
Number of businesses.....	5,266,607	140,223	30,287	7,863	648,535	280,185	964,523	177,745	120,271	224,352
Business receipts.....	17,297,125,146	107,931,393	142,247,484	537,883,736	1,080,555,117	4,822,650,951	5,278,843,887	510,816,998	799,441,224	1,573,271,535
Net income (less deficit)(*).....	596,524,021	181,253	1,828,515	-996,254	30,333,662	122,875,109	92,047,142	-8,071,329	-32,346,204	249,912,504
Net income.....	1,084,179,818	5,375,689	10,246,727	22,610,162	47,104,662	248,294,674	139,521,185	14,939,554	49,906,622	306,820,086
Deficit.....	487,655,795	5,194,437	8,418,213	23,606,417	16,770,999	125,419,563	47,474,044	23,010,883	82,252,824	56,907,581
C Corporations (*)										
Number of businesses.....	2,112,229	62,926	13,689	6,148	229,765	136,154	421,528	79,150	53,442	101,495
Business receipts.....	13,455,844,038	55,913,447	123,353,269	534,775,345	508,439,348	4,310,253,648	3,683,137,171	404,314,605	747,803,342	1,499,651,364
Net income (less deficit).....	413,045,088	-49,355	-694,500	-1,191,723	5,274,233	97,594,117	53,553,028	-10,159,325	-33,801,955	235,885,468
Net income.....	837,646,191	2,174,754	7,032,252	22,301,428	15,510,859	215,419,073	89,774,067	10,257,243	43,415,794	290,625,026
Deficit.....	424,601,101	2,224,109	7,726,753	23,493,152	10,236,625	117,824,954	36,221,039	20,416,569	77,217,748	54,739,558
S Corporations										
Number of businesses.....	3,154,377	77,297	16,598	1,715	418,770	144,031	542,150	98,595	66,829	122,857
Business receipts.....	3,841,281,106	52,017,946	18,894,215	3,108,391	572,115,769	512,397,303	1,595,706,716	106,502,393	51,637,882	73,620,171
Total net income (less deficit).....	183,478,933	230,608	2,523,015	195,469	25,059,429	25,280,992	38,494,114	2,087,996	1,455,751	14,027,036
Net income.....	246,533,627	3,200,935	3,214,475	308,734	31,593,803	32,875,601	49,747,118	4,682,311	6,490,828	16,195,060
Deficit.....	63,054,694	2,970,328	691,460	113,265	6,534,374	7,594,609	11,253,005	2,594,314	5,035,076	2,168,023
Partnerships										
Number of businesses.....	2,242,169	117,667	29,549	2,507	134,114	38,364	159,813	26,007	28,580	263,024
Business receipts.....	2,414,187,093	18,493,176	54,836,750	146,591,432	169,589,554	485,032,481	537,823,272	52,184,396	167,226,832	175,974,554
Net income (less deficit).....	270,667,169	-1,120,675	11,994,183	1,059,594	10,726,523	23,367,624	8,680,372	2,936,996	-6,541,677	89,250,979
Net income.....	439,761,741	4,541,707	17,592,960	5,596,380	15,771,154	37,340,960	16,237,421	6,209,734	19,058,239	108,763,922
Deficit.....	169,094,572	5,662,382	5,598,778	4,536,786	5,044,631	13,973,337	7,557,049	3,272,738	25,599,916	19,512,943
General (*)										
Number of businesses.....	841,299	74,586	10,152	304	49,924	13,524	74,751	7,786	9,363	100,760
Business receipts.....	467,422,866	4,111,608	15,806,315	7,866,688	40,873,429	121,586,703	78,246,760	6,872,176	44,541,936	19,476,261
Net income (less deficit).....	100,914,057	326,094	2,363,373	799,754	3,375,292	7,399,312	2,441,551	1,348,451	2,776,913	30,381,653
Net income.....	125,748,798	2,481,044	5,511,800	1,341,755	4,238,104	8,947,265	3,420,744	1,933,698	5,949,747	33,270,302
Deficit.....	24,834,741	2,154,950	3,148,427	542,001	862,812	1,547,954	979,193	585,247	3,172,835	2,888,649
Limited (*)										
Number of businesses.....	454,741	17,512	8,518	967	13,317	4,313	12,452	2,855	2,883	87,169
Business receipts.....	931,055,315	3,426,772	16,373,002	68,858,403	40,037,930	152,191,353	232,630,290	19,499,553	71,639,619	106,282,223
Net income (less deficit).....	121,126,936	-629,960	6,717,840	-220,262	2,605,478	9,847,500	3,814,619	2,406,473	-112,165	35,320,086
Net income.....	178,135,683	549,170	7,643,989	2,487,910	3,939,865	14,210,050	5,449,308	2,998,487	9,272,006	41,900,615
Deficit.....	57,008,747	1,179,129	926,149	2,708,172	1,334,388	4,362,550	1,634,689	592,014	9,384,171	6,580,529
LLC										
Number of businesses.....	946,130	25,569	10,879	1,236	70,873	20,528	72,610	15,366	16,335	75,095
Business receipts.....	1,015,708,912	10,954,796	22,657,433	69,866,341	88,678,195	211,254,425	226,946,222	25,812,666	51,045,277	50,216,070
Net income (less deficit).....	48,626,175	-816,809	2,912,970	480,102	4,745,754	6,120,812	2,424,202	-817,928	-9,206,425	23,549,240
Net income.....	135,877,260	1,511,493	4,437,171	1,766,715	7,593,185	14,183,645	7,367,370	1,277,548	3,836,485	33,593,005
Deficit.....	87,251,084	2,328,303	1,524,201	1,286,613	2,847,431	8,062,833	4,943,168	2,095,477	13,042,910	10,043,765
Nonfarm Sole Proprietorships										
Number of businesses.....	18,925,517	304,757	109,851	6,913	2,278,208	310,319	2,629,167	949,446	223,398	647,338
Business receipts.....	1,029,691,760	14,795,915	6,332,751	145,838	168,481,326	23,475,114	214,914,931	54,882,098	6,469,180	76,355,733
Net income (less deficit).....	221,113,286	735,764	501,591	-15,063	28,091,981	2,681,496	12,142,732	7,655,096	1,237,526	15,666,392
Net income.....	257,292,855	2,019,565	1,313,837	25,476	31,457,464	3,573,825	19,985,879	9,594,520	1,730,766	18,000,755
Deficit.....	36,179,568	1,283,800	812,246	40,539	3,365,483	892,328	7,843,146	1,939,424	493,240	2,334,363

Footnotes at end of table.

Table 3E.--Number of Businesses, Business Receipts, Net Income, and Deficit, by Form of Business and Industry, Tax Year 2002--Continued

[All figures are estimates based on samples--money amounts are in thousands of dollars]

Form of business, item	Real estate and rental and leasing	Professional, scientific, and technical services	Management of companies (holding companies)	Administrative and support and waste management services	Educational services	Health care and social assistance	Arts, entertainment, and recreation	Accommodation, food services, and drinking places	Other services	Religious, grantmaking, civic, professional, and similar	Unclassified industries
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
All Businesses											
Number of businesses.....	2,585,913	3,553,985	66,826	2,030,303	443,425	2,104,237	1,259,014	711,374	2,347,198	256,606	222,407
Business receipts.....	326,365,476	994,707,323	181,076,985	434,450,537	31,962,231	647,296,654	142,366,794	502,106,590	246,454,926	3,008,913	3,959,910
Net income (less deficit).....	88,486,346	108,603,239	93,713,668	21,559,264	2,497,018	71,897,581	5,026,449	8,914,092	19,504,855	1,767,093	1,519,038
Net income.....	154,030,315	159,757,591	109,455,165	33,137,274	4,453,305	83,381,275	18,266,440	24,841,983	26,595,564	1,964,550	1,717,027
Deficit.....	65,543,969	51,154,352	15,741,496	11,578,010	1,956,286	11,483,693	13,239,991	15,927,891	7,090,710	197,457	197,991
Corporations											
Number of businesses.....	570,639	736,005	48,053	231,412	41,317	334,305	110,609	271,527	321,134	N/A	7,620
Business receipts.....	205,206,751	651,992,903	170,514,329	338,209,323	24,509,009	448,427,967	72,674,159	372,418,853	159,401,281	N/A	128,244
Net income (less deficit) (¹).....	10,916,823	5,529,606	86,974,150	5,569,376	1,074,846	17,201,986	1,287,165	8,690,367	3,530,796	N/A	-15,492
Net income.....	27,306,140	45,588,917	97,401,257	13,923,247	1,920,404	24,207,526	5,595,830	16,130,502	7,260,820	N/A	25,814
Deficit.....	16,389,317	40,059,312	10,427,107	8,353,871	845,557	7,005,539	4,308,665	7,440,134	3,730,024	N/A	41,308
C Corporations (²)											
Number of businesses.....	210,506	255,885	26,274	74,456	16,010	155,300	36,195	93,686	134,581	N/A	5,039
Business receipts.....	129,234,183	393,523,705	165,001,246	210,732,359	14,327,839	319,820,278	38,335,364	240,354,090	76,835,603	N/A	37,832
Net income (less deficit).....	-894,004	-19,657,410	80,499,994	-1,021,791	402,377	2,985,478	-441,061	4,628,666	140,666	N/A	-7,815
Net income.....	9,450,869	14,936,926	89,169,833	5,647,487	969,788	8,441,367	1,603,174	8,696,227	2,209,287	N/A	10,737
Deficit.....	10,344,872	34,594,337	8,669,839	6,669,278	567,411	5,455,888	2,044,235	4,067,560	2,068,822	N/A	18,552
S Corporations											
Number of businesses.....	360,133	480,120	21,779	156,956	25,307	179,005	74,414	177,841	186,553	N/A	2,581
Business receipts.....	75,972,568	258,469,198	5,513,083	127,476,964	10,181,170	128,607,689	34,338,795	132,064,763	82,565,678	N/A	90,412
Total net income (less deficit).....	11,810,827	25,187,016	6,474,156	6,591,167	672,469	14,216,508	1,728,226	4,061,701	3,390,130	N/A	-7,677
Net income.....	17,855,271	30,651,991	8,231,424	8,275,760	950,616	15,766,159	3,992,656	7,434,275	5,051,533	N/A	15,077
Deficit.....	6,044,445	5,464,975	1,757,268	1,684,593	278,146	1,549,651	2,264,430	3,372,574	1,661,402	N/A	22,756
Partnerships											
Number of businesses.....	999,786	145,612	18,773	44,405	6,269	47,468	42,691	77,698	57,121	N/A	2,724
Business receipts.....	67,802,229	217,768,361	10,562,656	51,362,821	2,430,063	101,791,775	46,693,674	92,954,528	14,793,210	N/A	275,329
Net income (less deficit).....	54,988,398	54,436,614	6,739,518	3,671,249	-398,521	13,429,774	-1,828,953	-1,385,726	533,605	N/A	127,291
Net income.....	102,101,478	61,011,977	12,053,908	5,008,766	369,900	16,601,502	4,209,000	5,532,794	1,598,305	N/A	161,634
Deficit.....	47,113,080	6,575,362	5,314,389	1,337,517	768,421	3,171,728	6,037,953	6,918,520	1,064,700	N/A	34,343
General (³)											
Number of businesses.....	330,998	51,653	3,166	18,402	1,706	14,200	17,740	27,750	32,421	N/A	2,114
Business receipts.....	8,961,887	58,420,546	1,215,411	5,515,365	245,495	18,304,199	15,373,595	14,984,086	4,799,322	N/A	221,085
Net income (less deficit).....	18,639,017	21,822,755	1,989,804	595,616	34,903	4,718,857	829,393	513,055	538,678	N/A	19,587
Net income.....	23,063,746	23,018,322	3,150,819	731,826	41,553	4,900,516	1,799,920	1,178,681	727,927	N/A	40,968
Deficit.....	4,424,728	1,195,567	1,161,016	136,210	6,650	181,659	970,527	665,686	189,250	N/A	21,381
Limited (³)											
Number of businesses.....	246,080	20,392	5,780	4,795	451	8,405	4,238	11,400	3,125	N/A	90
Business receipts.....	21,445,241	100,612,413	1,895,174	11,695,703	348,590	37,776,105	12,460,189	31,890,243	1,992,512	N/A	--
Net income (less deficit).....	25,647,581	27,214,119	2,600,821	1,148,316	-354,503	4,718,795	-281,642	503,639	60,922	N/A	119,281
Net income.....	46,905,081	28,159,530	3,848,931	1,279,723	107,709	5,582,047	1,201,222	2,238,646	241,963	N/A	119,430
Deficit.....	21,257,501	945,411	1,248,110	131,407	462,213	863,252	1,482,864	1,735,007	181,041	N/A	149
LLC											
Number of businesses.....	422,708	73,567	9,826	21,208	4,112	24,863	20,713	38,548	21,574	N/A	520
Business receipts.....	37,395,101	58,735,402	7,452,071	34,151,754	1,835,978	45,711,471	18,859,890	46,080,199	8,001,376	N/A	54,244
Net income (less deficit).....	10,701,800	5,399,740	2,148,894	1,927,317	-78,921	3,992,121	-2,376,704	-2,402,420	-65,994	N/A	-11,577
Net income.....	32,132,652	9,834,125	5,054,157	2,997,217	220,637	6,118,939	1,207,858	2,115,407	628,415	N/A	1,235
Deficit.....	21,430,851	4,434,385	2,905,263	1,069,900	299,558	2,126,817	3,584,562	4,517,827	694,409	N/A	12,812
Nonfarm Sole Proprietorships											
Number of businesses.....	1,015,488	2,672,368	N/A	1,754,486	395,839	1,722,464	1,105,714	362,149	1,968,943	256,606	212,063
Business receipts.....	53,356,496	124,946,059	N/A	44,878,393	5,023,159	97,076,912	22,998,961	36,733,209	72,260,435	3,008,913	3,556,337
Net income (less deficit).....	22,581,125	48,637,019	N/A	12,318,639	1,820,693	41,265,821	5,568,237	1,609,451	15,440,454	1,767,093	1,407,239
Net income.....	24,622,697	53,156,697	N/A	14,205,261	2,163,001	42,572,247	8,461,610	3,178,687	17,736,439	1,964,550	1,529,579
Deficit.....	2,041,572	4,519,678	N/A	1,886,622	342,308	1,306,426	2,893,373	1,569,237	2,295,986	197,457	122,340

N/A - not applicable.

¹ Total Corporation "Net income (less deficit)" includes "Total net income (less deficit)" from S Corporations and is more comprehensive than what SOI generally publishes.

² For this table, the computations for C Corporations also include 1120-RIC and 1120-REIT returns.

³ For Tax Year 2002 General Partnerships include partnerships listed on the tax return as General, Foreign, Other and blank.

³ For Tax Year 2002 Limited Partnerships include Domestic Limited Partnerships and Domestic Limited Liability Partnerships.

NOTE: Detail may not add to total because of rounding.

Current Research in the Nonprofit Sector

Paul Arnsberger, Melissa Ludlum, and Margaret Riley, Internal Revenue Service

► The Nonprofit Sector

The nonprofit sector supports and advances a variety of religious, social, and economic endeavors. Tax-exempt nonprofit organizations dedicate billions of dollars annually to operating or supporting various initiatives in education, environmental protection and preservation, the arts and humanities, social welfare, health, and other critical areas. Programs offered by the nonprofit sector may supplement those provided by government agencies or offered by the corporate sector. Nonprofit organizations, which include hospitals, schools, churches, and other public charities as well as private foundations, receive an exemption from income taxes under Internal Revenue Code section 501(c)(3). As of October 2005, there were 909,224 such organizations recorded as active by the Internal Revenue Service (IRS).¹

Nonprofit organizations that receive tax-exempt status are expected to use this status to assist in carrying out their charitable activities, which in turn benefit individuals, households, and communities. Each nonprofit organization is responsible for ensuring that its tax-exemption is not used to benefit individuals having personal or private interest in the organization, such as shareholders or organization founders or their families. Also, nonprofit organizations are limited in their ability to influence political campaigns and lobby. Because private foundations are generally more narrowly controlled and supported than public charities, they are required to meet stricter guidelines than other nonprofit organizations. Nonoperating private foundations, which generally make grants to other charitable organizations, rather than operating charitable programs of their own, are required to pay out a minimum amount for charitable purposes, annually. Additionally, all private foundations are required to pay an excise tax on any net income that they earn from investments. All types of tax-exempt organizations, including nonprofit organizations, are subject to Federal taxation of income produced from activities that are unrelated to their charitable purposes. Nonprofit organizations are required to file annual

information returns with the IRS and to make these documents widely available to the public. They must also file a tax return for any year in which they receive “unrelated business” income or engage in activities that are prohibited under regulation. Information obtained from these documents can provide valuable insight into the composition and financial activities of the nonprofit sector.

The Statistics of Income division (SOI) of the Internal Revenue Service conducts a variety of ongoing research projects using data from information and tax returns filed by nonprofit organizations. This paper will focus on the manner in which this research is being used in analyses that address three key issues in the nonprofit area: the quality of reporting by tax-exempt organizations on their annual information and tax returns, the magnitude of compensation of executives and board members, and the extent to which tax-exempt organizations are known to violate the rules that govern their permissible activities.

► Recent Growth in the Nonprofit Sector

The nonprofit sector is a substantial and growing portion of the overall economy. The aggregate book value of assets, as reported by nonprofit organizations that filed IRS information returns for Tax Year 2002, was \$2.1 trillion. In real terms, this amount was 66 percent larger than the aggregate book value of assets held by nonprofit organizations for Tax Year 1993.² These organizations earned 41 percent more in revenue for Tax Year 2002 than they had earned for Tax Year 1993. Nonprofit organizations directed much of the income from their considerable asset growth and other sources into additional expenditures to promote their charitable programs. Total charitable expenditures reported by nonprofit organizations for Tax Year 2002 were 50 percent larger than those reported for Tax Year 1993 and experienced a real annual rate of growth of nearly 5 percent.³ In contrast, the Gross Domestic Product grew at a real annual rate of 3 percent over the period.⁴

In addition to experiencing significant growth in recent years, the nonprofit sector has also seen increased public interest in its financial dealings and charitable activities. With the development of GuideStar and other Internet sites that provide easy access to nonprofit organizations' IRS returns, public scrutiny of nonprofit organizations has increased, and, in some instances, high-profile cases of potential abuse have been documented. In response to these developments, various government officials and independent organizations have proposed a variety of additional legislative options aimed at curbing abuses of tax-exempt status.

In evaluating proposed tax legislation and initiatives directed toward improving oversight, it is crucial that policymakers and researchers have access to high-quality statistics and microdata for nonprofit organizations. Such information can be useful in determining characteristics of various types of nonprofit organizations, as well as in establishing standards for the administration of charitable programs. In many cases, data collected from tax return records and disseminated by the IRS provide the most comprehensive information available on the financial composition and charitable activities of nonprofit organizations. These data can reveal emerging trends and developments in the nonprofit sector and can be used to evaluate the effectiveness of tax regulation and IRS oversight. Analyses conducted using such data provide a framework for the development of tax policy related to nonprofit organizations and assist practitioners and nonprofit staffs in the establishment of key self-governance principles. Data for nonprofit organizations can be obtained from a number of Web sites and independent organizations. They are also available from IRS sources, such as the Statistics of Income division (SOI).

► **Overview of the Statistics of Income Exempt Organization Program**

SOI provides statistics and microdata derived from a number of administrative records filed with IRS. Sample and population data from information and tax returns are transcribed and corrected using a variety of error resolution and data perfection procedures. Since the 1970's, data for organizations exempt under section 501(c)(3) have been included in the SOI program. Currently, SOI

collects information from Forms 990, 990-PF, 990T, and 4720. Forms 990 and 990-PF are used by tax-exempt organizations to report standard income statement and balance sheet items, as well as additional information on tax-exempt activities and charitable distributions, compliance with the regulations that govern tax-exemption, involvement in various types of nonexempt activities, and certain information regarding employees.

Tax-exempt organizations, other than private foundations, file Form 990; private foundations file Form 990-PF. Form 990-T is filed by nonprofit and other types of tax-exempt organizations to report any unrelated business income (UBI) and taxes. Tax-exempt organizations use Form 4720 to calculate and pay taxes on prohibited activities, such as engaging in excessive lobbying, making political expenditures, or providing private benefit to "disqualified persons," which include organization founders, board members and executives, substantial contributors, and certain other individuals. SOI produces a variety of statistical tables and articles annually for all of the exempt organization programs. Also annually, microdata files that include all information collected for the Form 990 and Form 990-PF samples are made available for purchase. (Microdata derived from Forms 990-T and 4720 cannot be disclosed to the public.)

SOI samples approximately 10 percent of all Forms 990 and 990-PF, and about 20 percent of all Forms 990-T filed for a given tax year.⁵ The Form 990-T study incorporates a special Forms 990/990-T "integrated" sampling routine which ensures the inclusion of any Forms 990-T (with gross UBI of \$1,000 or more, the filing threshold) filed by organizations whose Form 990 or Form 990-EZ information returns were selected for the separate sample of section 501(c)(3) charitable organizations. For any designated tax year, tax-exempt organizations have various fiscal periods that collectively span 2 calendar years; to ensure complete coverage of a single tax year, SOI draws samples of Form 990-series returns over a 2-year timeframe. For example, the Tax Year 2002 studies include returns filed for Tax Year 2002 in Calendar Years 2003 and 2004. The SOI study of Forms 4720 was recently added to the exempt organizations program and includes data collected for the population of Forms 4720 filed over a calendar year. The SOI files

contain most financial items from each return, as well as a number of additional fields dedicated to codes or nonfinancial information. The SOI staff enter data into an online system, which identifies taxpayer and other errors. These are corrected during the data entry process. Often, supplemental information is included with tax returns on schedules and other attachments. Where appropriate, information from these attachments is used to adjust data reported by the filer.

The sample designs and data collection methods that are applied to the SOI files allow clear statistical patterns to emerge. Consistency or variation in such patterns can provide insight into changes in reporting patterns, which may be attributable to tax law modifications or changes in the degree or quality of IRS oversight. Additionally, the largest organizations that appear in each SOI file are sampled with certainty, which creates, in effect, a panel of large tax-exempt organizations. The longitudinal nature of the SOI sample and population files can assist researchers in establishing typical statistical patterns for tax-exempt organizations and identifying cases that deviate from the expected norm. Analyses derived from these data can provide insight into a variety of current issues in the nonprofit sector.

► **Current Research Issues**

Reporting Quality

With the advent of electronic filing and imaging of IRS nonprofit-organization information returns and their widespread availability to the public, the quantity of data available for regulation and research has increased dramatically. Technological improvements that make more data more accessible are certainly desirable, but ensuring that preparers fill out the forms completely and accurately is equally important. Is “more” really better without quality reporting of return information? Ensuring reporting quality is a shared responsibility of both IRS and return preparers. IRS needs to ensure that information and tax forms require essential information for effective regulation, oversight, and public transparency; and it needs to develop form instructions that are complete, explicit, and clear enough for preparers to follow. Preparers need to be meticulous in providing complete responses to the requested information on the

forms, especially itemized financial components. During the past year, SOI has conducted special analyses, using data from its Forms 990 and 990-T statistical files, to assess the quality of information reported by return preparers.

Comparing and Reconciling Unrelated Business Income Data Reported on Forms 990 and 990-T

An analysis of Tax Year 2002 data from 2,894 linked records in the Forms 990 and 990-T integrated sample of section 501(c)(3) public charities concludes that taxable unrelated business income (UBI) reported on Form 990-T oftentimes cannot be reconciled with that reported on Form 990.⁶ Anecdotal information from reviewed cases indicates that the data entered on Form 990-T are much more accurate, perhaps because the purpose of Form 990-T is to calculate tax liability, which carries a greater potential for the assessment of monetary penalties for misreporting than Form 990, whose purpose is to supply information only. Applying Form 990 weights to the sample records produced an estimated population of 8,992 public charities that were required to file both a Form 990 and a Form 990-T. The main sources of data for this analysis were Form 990, Part VII, Analysis of Income-Producing Activities, and Form 990-T, Part I, Unrelated Trade or Business Income.

Form 990, Part VII, provides a three-tiered breakout of an organization’s total revenue (excluding any contributions, gifts, and grants received from Government or public sources): potentially taxable UBI reportable on Form 990-T, UBI excluded from taxation under the Internal Revenue Code, and mission-related (exempt function) income. For each taxable UBI item entered, the filer is instructed to provide an associated business activity code from a list of North American Industrial Classification System (NAICS) codes. Form 990-T, Part I, contains a statement of gross UBI, direct expenses, and net UBI.

As illustrated in Table 1, the Form 990 returns in the integrated sample were separated into three groups based on potentially taxable UBI reported in Part VII: those with positive total UBI (80 percent of all returns), those with zero UBI (13 percent of all returns), and those with negative total UBI (7 percent of all returns). Within

Table 1. Reconciliation of Unrelated Business Income (UBI) Data From Form 990, Part VII, and Form 990-T, Part I, Tax Year 2002

[All figures are estimates based on samples. -- Money amounts are in thousands of dollars]

Item	Number of returns	Percentage of all returns	Form 990 UBI	Form 990-T Gross UBI ¹	Form 990-T Net UBI	Form 990-T Adjusted UBI ²
	(1)	(2)	(3)	(4)	(5)	(6)
Matched returns, total.....	8,992	100.0	3,807,095	4,089,889	3,343,626	3,771,948
Number with Form 990 UBI greater than zero.....	7,194	80.0	3,869,524	3,574,474	3,009,050	3,411,944
Number with UBI that could not be reconciled ³	2,447	27.2	1,870,317	1,521,271	1,253,569	1,433,963
Number with Form 990 UBI equal to zero.....	1,183	13.2	--	270,348	225,634	236,913
Number with UBI that could not be reconciled ³	853	9.5	--	251,173	229,754	234,908
Number with Form 990 UBI less than zero.....	614	6.8	(62,429)	245,067	108,942	123,091
Number with UBI that could not be reconciled ³	124	1.4	(29,903)	181,211	131,100	132,128

¹All returns in the Form 990-T sample had gross unrelated business income of \$1,000 (the filing threshold) or more.

²Adjusted UBI is derived from a combination of Form 990-T gross and net itemized UBI amounts, based on their correlation to the combination of gross and net UBI amounts required to be reported on Form 990.

³The amount of total UBI reported on Form 990, Part VII, does not equal gross UBI, net UBI, or adjusted UBI (within \$100 tolerance) reported on Form 990-T, Part I.

these groups, Form 990 total UBI was matched against both total gross UBI and total net UBI reported in Part I of Form 990-T, and also against a computed amount of total “adjusted UBI.” Adjusted UBI is derived from a combination of Form 990-T gross and net itemized UBI amounts, based on their correlation to the combination of gross and net UBI amounts required to be reported in Part VII, Form 990. If organizations had reported income consistently on both forms, it was expected that the Form 990 total UBI amount would be the same as the Form 990-T adjusted UBI amount, a value that was no more than gross UBI and no less than net UBI, depending on what types of income were reported in each individual case.

UBI reported on nearly 4 out of every 10 Forms 990 could not be reconciled with UBI reported on Form 990-T, meaning that total UBI on Form 990 did not match gross UBI, net UBI, or adjusted UBI on Form 990-T (within a \$100 tolerance). The reasons for the inconsistency are twofold: some filers reported a combination of gross and net taxable income that differed from that specified in the Form 990 instructions; other filers did not report taxable UBI on Form 990 at all. Of the 7,194 returns where the Form 990 UBI amount was positive, 34 percent could not be reconciled. In some observed cases, the Form 990 amounts simply did not

correspond to any Form 990-T amounts. In many other cases, filers of Form 990 erroneously reported gross receipts from sales and services in Part VII, rather than gross profit from sales and services, which is the net of gross receipts minus cost of goods sold. Gross profit, not gross receipts, should be included in total UBI on both Forms 990 and 990-T.

Twenty-eight percent of the 1,183 organizations that reported no taxable UBI amounts on Form 990 filed Forms 990-T with net UBI that was negative. The organization may have presumed that negative net UBI amounts need not be reported on Form 990. These cases were not deemed irreconcilable for this analysis. However, 72 percent of the organizations reporting no taxable UBI on Form 990 filed Form 990-T with positive amounts of gross, net, and adjusted UBI. There is no known reason for this, with the exception of some degree of nonreporting on Form 990.

About one-fifth of the 614 organizations reporting negative UBI on Form 990, Part VII, filed a Form 990-T with positive amounts of gross, net, and adjusted UBI. In some cases, negative amounts entered on Form 990, Part VII, for gain or loss from sales of investment assets were not reported on Form 990-T. Generally, income from investments is not considered unrelated business

income for public charities that file Forms 990 and 990-T. In other cases, negative entries on Form 990 could not be correlated with any amount reported on Form 990-T.

In 36 percent of the linked Forms 990 and 990-T cases, the primary unrelated business activity indicated on the organization's Form 990-T did not match any activity code reported in Part VII of Form 990 for each itemized taxable UBI amount. This, along with UBI reporting inconsistencies, seems indicative of preparers who fill out Form 990 and 990-T exclusive of any attempted reconciliation of reported information on the two forms.

Researchers, both in and outside of IRS, use Form 990 to make assessments of nonprofits' financial activities, operations, and programs. Form 990, Part VII, for example, provides data that should be useful for gauging how much of an organization's income is from taxable unrelated business activities and what types of activities are producing the income. Currently, an IRS team is designing a revised Form 990 that will be geared toward obtaining data that will be useful for better regulation and oversight of nonprofit and other tax-exempt organizations. Taxpayer education, comprehensive IRS form instructions, and complete and accurate reporting by return preparers are vital for making Form 990 a consistent and reliable tool for research and public accountability.

Form 990-T Deductions Allocation Study

The deductions allocation study measures the extent to which high-income organizations (those with gross UBI of \$500,000 or more) misreported specifically defined, itemized deduction components as "Other deductions" on Tax Year 2002 Forms 990-T. During the data entry process, SOI staff check the required Other deductions statement for inaccurately reported items and move (allocate) amounts, when appropriate, to one or more of the specifically defined deduction components, such as Salaries and wages. The study examined the difference between deduction amounts as initially reported by filers and as corrected, through allocation, by SOI staff.⁷

During normal IRS processing of paper and e-file returns, data are captured as reported by the return filer. Misreported amounts are not allocated from residual

"other" categories to the proper, specifically defined return line items. Researchers and IRS staff that use Returns Transaction File (RTF) data for examination or administrative purposes may find this study useful for gauging the extent to which deductions data may be understated, and extrapolating its results to draw conclusions about the possible understatement of itemized income, deductions, assets, and liabilities reported on other types of IRS exempt-organization returns.

Of the 2,381 high-income returns filed, 20 percent required at least one allocation from Other deductions during SOI data entry. Paid preparers completed 79 percent of these 485 returns with taxpayer reporting errors.⁸ Sixty-eight percent of the returns that required SOI allocations of misreported amounts were filed by section 501(c)(3) nonprofit organizations; the remainder were filed by organizations exempt under other sections of the tax code. Section 501(c)(6) business leagues, chambers of commerce, and real estate boards and section 501(c)(7) social and recreational clubs accounted for 11 percent and 7 percent, respectively, of all returns that required allocations from Other deductions to specifically defined components.

After allocation, the increase in the total amount of each specifically defined deduction category reported by high-income filers ranged from 3 percent to 45 percent. Salaries and wages, the largest aggregate itemized deduction reported on Form 990-T, rose by only 3 percent; Contributions to deferred compensation plans rose by 14 percent; and Repairs and maintenance rose by 45 percent. Allocations made to other types of itemized deductions resulted in increases ranging between 4 percent and 9 percent. It is worth noting that no allocations were made to Compensation of officers, directors, and trustees, Excess exempt expenses, or Excess readership costs. Form 990-T filers must provide detailed information on related schedules for these items and then enter schedule totals in the itemized deductions statement. The schedule preparation requirement apparently deters preparers from including these items in Other deductions.

As shown in Table 2, the three deduction items with the largest *aggregate* dollar amount allocated from Other deductions were Salaries and wages (\$32.0 mil-

lion allocated), Repairs and maintenance (\$21.7 million allocated), and Employee benefit programs (\$7.8 million allocated). Allocated amounts accounted for close to half of the SOI-edited amount of Salaries and wages, and three-quarters or more of the other two cited deduction items. The largest *average* dollar amounts allocated from Other deductions were made to Salaries and wages (\$381,269), Repairs and maintenance (\$92,593), Net depreciation (\$92,503), and Employee benefit programs (\$69,921).

The deduction items with the highest frequency of allocation of misreported taxpayer amounts were Repairs and maintenance (243 returns), Taxes and licenses (180 returns), Salaries and wages (93 returns), and Employee benefit programs (92 returns). The top three primary unrelated business activities reported by organizations, based on self-reported NAICS codes and percentage of returns with allocations, were medical and diagnostic laboratories (14 percent), gambling industries (9 percent), and advertising and related services (6 percent). Overall, close to 10 percent of the reported Other deductions amount should have been included in

the more specifically defined deduction items, and the percentage change in itemized deduction amounts, after SOI allocations, ranged from 12.5 (Salaries and wages) to 106.7 (Repairs and maintenance).

The deductions allocation study makes it clear that Form 990-T preparers could do a much better job of accurately reporting all-inclusive amounts within the specifically defined deduction components listed on the form. If IRS plans to use tax processing data to make intelligent decisions regarding regulation, compliance, or potential abuses of tax-exempt status, it is imperative that a high priority be placed on educating nonprofit organizations and their tax practitioners to report detailed items completely and accurately. Also, because organizations are not allowed to file supplementary electronic financial statements with e-filed returns (they must provide financial data in the IRS format), it is feared that if the data provided are incorrect or incomplete, there will be no additional information available with the e-filed returns, as there is with paper returns, that can be used to correct these reporting errors.

Table 2. Form 990-T Returns with Gross Unrelated Business Income of \$500,000 or More and At Least One Allocation Made from Other Deductions, Tax Year 2002

[Money amounts are in thousands of dollars]

Deduction item	Number of returns with allocations	Percentage of all returns ¹ with allocations	SOI edited amount	Taxpayer reported amount	Allocated amount	Percentage of SOI edited amount allocated from Other deductions
	(1)	(2)	(3)	(4)	(5)	(6)
Other deductions.....	485	100.0	753,388	832,164	(78,776)	N/A ²
Compensation of officers, directors, and trustees.....	--	--	--	--	--	--
Salaries and wages.....	93	19.2	68,069	36,043	32,027	47.1
Repairs and maintenance.....	243	50.1	28,840	7,174	21,667	75.1
Bad debts.....	32	6.6	1,618	10	1,608	99.4
Interest.....	39	8.0	2,094	4	2,090	99.8
Taxes and licenses.....	180	37.1	16,213	10,296	5,917	36.5
Charitable contributions.....	22	4.5	1,524	37	1,487	97.6
Net depreciation.....	54	11.1	6,004	1,009	4,995	83.2
Depletion.....	--	--	--	--	--	--
Contributions to deferred compensation plans.....	26	5.4	1,242	34	1,207	97.2
Employee benefit programs.....	92	19.0	9,897	2,119	7,778	78.6
Excess exempt expenses.....	--	--	--	--	--	--
Excess readership costs.....	--	--	--	--	--	--

¹Detail does not add to 100 percent because some returns had allocations made to more than one deduction item.

²N/A - not applicable. However, 9.5 percent of the total amount of aggregate Other deductions reported by taxpayers was allocated to one or more specifically defined deduction items.

Form 990 Asset Allocation Study

An asset allocation study, similar to 990-T deductions allocation study but on a smaller scale, was conducted for public charities that filed Form 990. The goal was to measure the degree to which assets were misreported by filers as “Other assets” on Form 990, rather than in the appropriate specifically-defined asset categories. For this study, SOI data were compared to a file made available by GuideStar containing data transcribed from the same information returns. The GuideStar data were chosen because, like the IRS Returns Transaction File, reporting errors were not resolved based on research on attached financial statements during the transcription process. For this reason, the GuideStar data provided a useful record of what each filer reported on the form.

Over 6,600 Form 990 returns from Tax Year 2002, representing virtually all of the certainty strata of the SOI sample, were matched with the same filings from the GuideStar dataset. Eleven returns, for which the balance sheet values in the SOI and GuideStar datasets differed by three orders of magnitude, were excluded from the analysis.⁹ Total assets for the SOI group amounted to \$1.345 trillion versus \$1.338 trillion for the GuideStar group, a difference of less than 1 percent. When the totals for Other assets were compared, the GuideStar total was \$34.5 billion (or 41 percent) more than SOI. Most of this difference can be attributed to financial items allocated out of Other assets during the course of SOI processing and, as such, is a measure of filer reporting error. A look at the specific asset categories quickly shows where these “other” assets should have been reported. In the SOI dataset, Investments--other totaled \$129.9 billion versus \$106.4 billion in the GuideStar dataset. This disparity of \$23.0 billion represented two-thirds of the difference in Other assets between the two datasets. Only three other specific asset categories showed an aggregate increase of more than 5 percent after SOI editing: Prepaid expenses and Land, buildings, and equipment, both 8 percent, and Cash, 7 percent.

When the universe of GuideStar-transcribed returns was compared to SOI’s weighted population estimates, similar results were seen. The GuideStar sum of Total

assets was \$1.740 trillion, less than 1 percent larger than SOI’s weighted estimate, while the GuideStar sum of Other assets was \$51.5 billion (or 50 percent) more. Again, Investments--other was the largest misreported category, with an SOI-estimated total that was \$23.3 billion larger than the GuideStar population total.

Researchers and analysts studying the endowments of public charities should be aware of the reporting tendencies of these organizations. To the extent possible, SOI tax examiners allocate assets, liabilities, and expenses to the correct line items; however, not all sources of data have this value added. Further, it is a concern that the growth of electronic filing will be accompanied by a reduction in the amount of usable supplemental data, reducing SOI’s ability to correct these types of reporting errors.

Compensation of Executives and Board Members

Nonprofit organizations, which include public charities and private foundations, are legally required to avoid providing “unreasonable compensation” to executives and board members. Recently, Congress and various independent organizations have proposed legislation aimed to further define and limit permitted compensation amounts. As compensation rates for executives and board members differ substantially among organizations of different types and sizes, analyses of compensation data can provide valuable insight into the development of equitable standards. SOI collects a variety of data related to individual compensation amounts paid to executives and board members, which can assist researchers in analysis of such issues.

All nonprofit organizations that file Form 990 or 990-PF are required to provide individual-level compensation data for all paid executives and board members. These amounts are reported in Part V of Form 990 and Part VIII of Form 990-PF for each board member or trustee, foundation manager or organization director, executive, or officer who was paid by the nonprofit organization during the tax year. Nonprofit organizations report compensation paid to executives and board

members not only for their assistance in operating and administering charitable programs, but also for their work in fundraising, investment management, and other activities not directly related to their charitable purposes. Table 3 shows that, for Tax Year 2002, compensation, including benefits, deferred compensation, and allowances, paid by public charities and private foundations to executives and board members totaled \$15.0 billion. For both public charities and private foundations, the highest paid executives or board members received over \$7 million. Most nonprofit organizations did not report compensating executives or board members; less than half of public charities and less than one-quarter of private foundations indicated that they had paid one or more executives or board members during the tax year.

Among organizations that reported executive and board compensation, patterns of such compensation varied greatly for Tax Year 2002, depending on certain organizational characteristics, such as type and size. For example, median compensation for individual executives and board members at public charities was \$45,000, an amount much larger than the median compensation of \$6,000 paid to individuals with similar positions at private foundations. Likewise, organization size, as measured by total assets, significantly affected compensation practices. For all nonprofit organizations, both median and mean executive and board compensation amounts increased measurably with organization size. Additionally, large nonprofit organizations distributed a larger portion of their total executive and board compensation as employee benefits (13 percent) than medium and small organizations (8 percent and 4 percent, respectively).¹⁰

A different pattern emerges when the aggregate compensation of executives and board members paid by an organization is measured as a proportion of the organization's total expenditures. Although large nonprofit organizations clearly spend more in absolute amounts for compensation than smaller organizations, small nonprofit organizations direct a larger percentage of their overall expenditures toward executive and board compensation. The median proportion of aggregate executive and board compensation to total expenses for small public charities was 8 percent for Tax Year 2002. For medium-sized public charities, the median was 2 percent. And for large public charities, the median

proportion of aggregate compensation was less than 1 percent. Median proportions of aggregate compensation of executives and board members to total expenses also decreased with organization size for private foundations. The median proportion of aggregate executive and board compensation to total expenses was 12 percent for small private foundations, 3 percent for medium-sized private foundations, and less than 1 percent for large private foundations.

In addition to individual executives and board members, many nonprofit organizations also report compensation of institutional trustees, such as banks.¹¹ While public charities paid less than one-half of 1 percent of executive and board compensation to institutional trustees, private foundations reported that 16 percent of compensation was paid to these organizations. Additionally, institutional trustees represented 28 percent of all compensated individuals reported by private foundations. For private foundations, the proportion of compensation paid to institutional trustees to total expenses greatly exceeded that paid to individual executives and board members. The median proportion of compensation paid to total expenses for institutional trustees was 15 percent. In contrast, this proportion, when calculated for compensation paid to individual executives or board members by private foundations, was less than 2 percent.

► **Preliminary Research on Taxation of EO Prohibited Activities**

Chapters 41 and 42 of the IRC outline a number of prohibited activities and their associated penalties. Tax-exempt organizations, certain individuals associated with those organizations, and certain nonexempt trusts that engage in such prohibited activities must pay excise taxes for the tax year in which the prohibited activity occurred. Organizations or individuals liable for such excise taxes calculate their total amounts due using Form 4720, *Return of Certain Excise Taxes on Charities and Other Persons Under Chapters 41 and 42 of the Internal Revenue Code*. Excise taxes may be assessed on a number of activities, such as failure by nonoperating private foundations to distribute minimum amounts toward grants, disbursement of excess amounts toward lobbying, participation in illegal political activities, and

Table 3. Nonprofit Organization Board and Executive Compensation, by Type of Organization and Size,¹ Tax Year 2002

[All figures are samples based on estimates]

Public charities					
Type of organization and size	Number of compensated individuals	Total	Median	Mean	Max
	(1)	(2)	(3)	(4)	(5)
All public charities					
Total compensation and benefits.....	202,316	14,218,864,111	45,000	70,280	7,448,233
Compensation.....	194,537	12,806,782,863	45,000	65,832	6,885,926
Employee plans.....	83,045	1,213,267,385	7,503	14,610	4,559,427
Expense accounts and other allowances.....	25,042	201,114,311	3,000	8,031	743,349
Small charities					
Total compensation and benefits.....	108,035	3,723,646,342	28,146	34,467	333,604
Compensation.....	102,263	3,491,258,605	28,800	34,140	303,113
Employee plans.....	23,826	161,443,629	4,443	6,776	81,493
Expense accounts and other allowances.....	11,351	70,944,108	1,445	6,250	51,600
Medium charities					
Total compensation and benefits.....	73,468	6,393,010,502	70,141	87,018	2,646,940
Compensation.....	71,954	5,811,838,637	66,453	80,771	2,646,940
Employee plans.....	42,521	511,513,724	7,276	12,030	634,936
Expense accounts and other allowances.....	8,875	71,495,761	3,211	8,056	305,400
Large charities					
Total compensation and benefits.....	20,813	4,102,207,268	152,729	197,095	7,448,233
Compensation.....	20,320	3,503,685,622	137,249	172,422	6,885,926
Employee plans.....	16,698	540,310,032	18,338	32,357	4,559,427
Expense accounts and other allowances.....	4,816	58,674,442	5,341	12,183	743,349
Private foundations					
All private foundations					
Total compensation and benefits.....	29,921	743,675,862	6,000	24,855	7,182,301
Compensation.....	29,086	684,732,874	6,000	23,542	7,182,301
Employee plans.....	2,566	51,084,960	11,000	19,909	1,450,943
Expense accounts and other allowances.....	1,563	7,858,028	960	5,026	497,605
Small foundations					
Total compensation and benefits.....	11,767	76,585,846	2,644	6,509	79,102
Compensation.....	11,340	74,440,810	2,684	6,564	63,360
Employee plans.....	388	1,984,176	147	5,108	15,742
Expense accounts and other allowances.....	550	160,860	99	292	960
Medium foundations					
Total compensation and benefits.....	14,411	336,743,345	10,000	23,367	1,472,583
Compensation.....	14,100	320,619,761	10,022	22,739	974,978
Employee plans.....	1,003	12,420,032	6,315	12,377	627,370
Expense accounts and other allowances.....	547	3,703,552	1,600	6,767	497,605
Large foundations					
Total compensation and benefits.....	3,743	330,346,671	29,829	88,257	7,182,301
Compensation.....	3,646	289,672,303	30,000	79,449	7,182,301
Employee plans.....	1,174	36,680,752	20,140	31,244	1,450,943
Expense accounts and other allowances.....	466	3,993,616	3,004	8,570	230,452

¹ For the purpose of analysis, "small" charities hold less than \$1 million in book value of total assets; "small" foundations hold less than \$1 million in fair market value of total assets; "medium" charities hold from \$1 million to less than \$50 million in book value of total assets; "medium" foundations hold from \$1 million to less than \$50 million in fair market value of total assets; "large" charities hold \$50 million or more in book value of total assets; and "large" foundations hold \$50 million or more in fair market value of total assets.

excess benefit transactions or self-dealing activities that benefit individuals associated with public charities or private foundations, respectively.

SOI recently began collecting data from Forms 4720 filed by organizations and individuals. To date, data collection for Calendar Years 2003 and 2004 has been completed. Statistics derived from the population of Forms 4720 received by IRS during those years include data from returns filed for various tax years. For Calendar Year 2004, some 65 percent of the returns included in the population represented Tax Year 2003, and 27 percent represented Tax Year 2002. The additional 8 percent of the Calendar Year 2004 population comprised returns filed for various earlier tax years. While Form 4720 may be filed by a variety of organizations, Form 990-PF filers accounted for more than 95 percent of the return population in each of Calendar Years 2003 and 2004.¹² For Calendar Years 2003 and 2004, approximately 2 percent of all Form 990-PF filers filed Form 4720.

This paper marks the first publication of data collected for the Form 4720 study. Table 4 shows Calendar Year 2003 and 2004 data from Form 4720. Clearly, the excise tax paid on undistributed income is the largest and most commonly reported excise tax. This tax appeared on 85 percent of returns filed and accounted for more than 70 percent of total taxes reported for both Calendar Years 2003 and 2004. After taxes on undistributed income, the most commonly reported taxes were on self-dealing and excess benefit transactions, which are generally prohibited transactions between nonprofit organizations and associated individuals. Examples of

excess benefit transactions include excess compensation to executives or board members and loans made to officers, directors, and trustees. Taxes on self-dealing and excess benefit transactions appeared on 9 percent of returns included in the Calendar Year 2003 study and 10 percent of returns included in the Calendar Year 2004 study. These taxes represented 15 percent of total tax reported for Calendar Year 2003 and 9 percent of total tax reported for Calendar Year 2004.

Data collected from Form 4720 provide additional insight into the types of prohibited activities that occur most commonly and the degree to which such violations occur. However, statistics derived from this information may be limited by both the reliability of nonprofit organizations in reporting prohibited activities and the effectiveness of IRS audit procedures and oversight. For example, a steady annual increase in the percentage of organizations using Form 4720 each year could indicate improved reporting compliance among nonprofit organizations, or increased involvement in prohibited activities. Nevertheless, the statistics may prove helpful in measuring the effectiveness of this oversight. In the future, data from Form 4720 may help determine the impact and effectiveness of any changes made or additions to the regulations that govern the activities of nonprofit organizations.

► **Summary**

The information obtained from SOI statistics, microdata, and research projects can be used in analyses that illuminate a variety of issues faced by legislators,

Table 4. Excise Taxes Reported by Charities, Private Foundations, and Certain Trusts on Form 4720, Calendar Years 2003 and 2004

Internal Revenue Code Section	Item	Calendar Year 2003		Calendar Year 2004	
		Number	Amount	Number	Amount
Section 4942	Tax on Undistributed Income (Schedule B).....	1,551	3,539,633	1,482	5,594,073
Sections 4941 & 4958	Taxes on Self-Dealing and Excess Benefit Transactions (Schedule A).....	170	730,233	170	659,721
Section 4945	Tax on Taxable Expenditures (Schedule E).....	53	277,420	54	1,036,999
Section 4911	Tax on Excess Lobbying Expenditures (Schedule G).....	27	75,255	31	136,033
Sections 4943, 4944, 4912, 4955	Additional Excise Taxes ¹	26	191,318	23	276,670
	Total².....	1,817	4,813,859	1,743	7,703,496

¹ Includes reported taxes on Excess Business Holdings, Investments that Jeopardize Charitable Purposes, Disqualifying Lobbying Expenditures, Political Expenditures, and Personal Benefit Contracts.

² Detail adds to more than total because some organizations reported more than one type of activity subject to excise taxes.

the IRS, and nonprofit practitioners; this paper has highlighted three examples. Several research projects, including an analysis of information derived from the Forms 990/990-T integrated sample and the Forms 990 and 990-T allocation studies, have identified apparent problems with the quality of reporting by tax-exempt organizations. SOI microdata and statistics can be an important asset in research involving information where proper line item allocations are imperative, such as balance sheet or income statement information. Data for individual compensation amounts paid to executives and board members can be employed in a variety of analyses and can provide a glimpse into the compensation habits of nonprofit organizations. The recent introduction of the Form 4720 study provides a new opportunity for research into the degree to which nonprofit organizations deviate from their tax-exempt purposes. Clearly, SOI data can be valuable to researchers and analysts in determining an overall picture of the nonprofit sector, identifying potential problems in tax reporting and compliance, and establishing benchmarks for the administration and operation of nonprofit organizations. Such analyses may provide the framework for future oversight procedures, tax legislation, and self-governance guidelines.

► Endnotes

- ¹ This amount was obtained from the Internal Revenue Service Exempt Organizations Business Master File and includes nonprofit organizations not required to file annual returns with the IRS.
- ² Data indicated as constant dollars were adjusted based on the 2000 chain-type price index for Gross Domestic Product as reported by the U.S. Department of Commerce, Bureau of Economic Analysis. Tax Year 2002 is used as the base year for these adjustments.
- ³ For purposes of analysis, “charitable expenditures” is defined as the sum of program service expenses from Form 990 and disbursements for charitable purposes from Form 990-PF.
- ⁴ Growth rates were derived from the exponential formula for growth, $y=b*m^x$.

- ⁵ For detailed information on Statistics of Income sampling methodology for producing population estimates, see the general appendix, located near the back of the Summer 2005 issue of the *SOI Bulletin*, particularly the Sample Criteria and Selection of Returns section and the Method of Estimation section. The *SOI Bulletin* is available from the Tax Stats section of the IRS Web site, www.irs.gov/taxstats.
- ⁶ A business activity is considered unrelated if it does not contribute importantly (other than the production of funds) to accomplishing an organization’s charitable, educational, or other purpose that is the basis for the organization’s tax exemption. Whether an activity contributes importantly depends in each case on the facts involved. See IRS Publication 598, *Tax on Unrelated Business Income of Exempt Organizations*, for additional information on unrelated business income and tax.
- ⁷ Data collected for the Deductions Allocation Study were controlled to provide statistics solely on amounts of itemized deductions allocated from Other deductions. Any SOI adjustments made for reasons other than allocating, such as correcting math errors, are included in both the SOI adjusted amounts and the taxpayer-reported amounts.
- ⁸ The actual number of Tax Year 2002 large-income Forms 990-T with allocations was 492. Seven returns could not be located for the study, and data on taxpayer entries of itemized deductions were not available from any other source.
- ⁹ Each year, several Form 990 filers report their balance sheet items in thousands of dollars with a note on the return with that information. During IRS Returns Transaction File processing and GuideStar transcription, this note is often missed. SOI processing includes steps to ensure that these returns are transcribed correctly. Consequently, for a certain number of returns each year, SOI balance sheet figures are one thousand times larger than on both the GuideStar file and the Returns Transaction File.

¹⁰ For purposes of analysis, “small” public charities hold less than \$1 million in book value of total assets; “small” private foundations hold less than \$1 million in fair market value of total assets; “medium” public charities hold from \$1 million to less than \$50 million in book value of total assets; “medium” private foundations hold from \$1 million to less than \$50 million in fair value of total assets; “large” public charities hold \$50 million or more in book value of total assets; and “large” private foundations hold \$50 million or more in fair market value of total assets. Of the returns filed by public charities for Tax Year 2002, some 68 percent were filed by small public charities, 30 percent were filed by medium public charities, and 2 percent were filed by large public charities. Small, medium, and

large private foundations represented 70 percent, 29 percent, and 1 percent of returns filed by private foundations for Tax Year 2002, respectively.

¹¹ For additional information on institutional trustees, see Boris, Elizabeth A.; Renz, Loren; and Hager, Mark A (2005), *Foundation Expenses and Compensation: Interim Report, 2005*, The Urban Institute, The Foundation Center, and Philanthropic Research, Inc.

¹² Organizations identified as “Form 990-PF filers” may be private foundations or section 4947(a)(1) charitable trusts that are treated as private foundations for tax purposes. Generally, private foundations represent more than 90 percent of all Form 990-PF filers.

Geographic Variation in Schedule H Filing Rates: Why Should Location Influence the Decision To Report "Nanny" Taxes?

Kim M. Bloomquist, Internal Revenue Service, and Zhiyong An, Department of Economics, University of California, Berkeley Institute

The Schedule H is the Internal Revenue Service (IRS) form used to report Social Security and Medicare taxes on wages of \$1,400 or more paid to household employees. The IRS defines a household employee as someone whose work details are controlled by the employer. A Schedule H is not required to be filed when household work is performed by an agency employee or by a self-employed individual. In the former case, the agency is responsible for work-related details such as who does the work and how it is done. Similarly, a self-employed individual is someone who controls his or her work schedule, provides their own tools or equipment, and offers services to the general public.

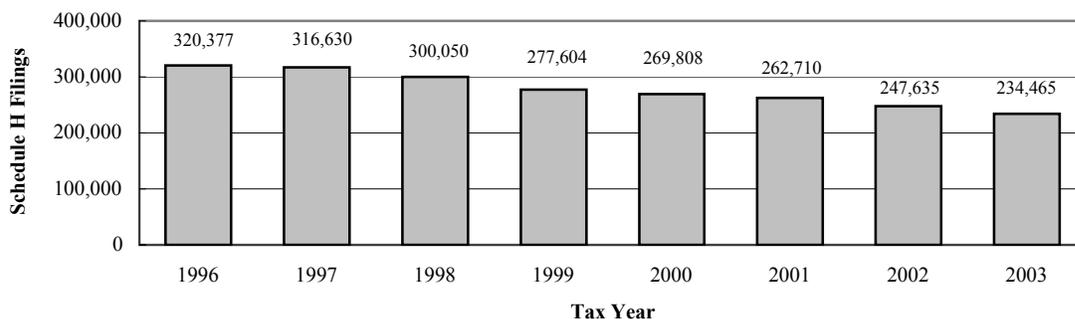
The Schedule H has been referred to as the “nanny tax” form since the early 1990s when several of President Clinton’s political appointees were discovered to have either hired undocumented workers or failed to pay Schedule H employment taxes on former housekeepers. More recently, President George W. Bush’s initial Cabinet head selections for the departments of Homeland Security and Labor were scuttled, in part, for “nanny tax” violations.

These high-profile cases reinforce the commonly-held belief that people perceive little risk in not paying household employment taxes (barring the possibility of being asked to serve as a Cabinet secretary). This

perception is supported by industry experts with first-hand knowledge of compensation practices in this area. Pat Cascio, Board President of the International Nanny Association, recently stated, “A high percentage of nannies are not paid legally. Some people don’t want the extra work or hassle of dealing with taxes. They’d rather pay their nannies out-of-pocket.”¹ If such attitudes are common among people who can afford to hire full-time nannies, it is probably true also for many middle and upper-middle income families who would like to hire someone to provide part-time care for an elderly parent or younger children.

The *Wall Street Journal* recently pointed to the large drop in the number of Schedule H filings (Figure 1) as an indicator of a growing evasion problem.² While this is one possibility, there are other possible explanations for this phenomenon. For example, a decline in Schedule H filings would result if more work in the household sector is being done either by the self-employed or employees of service firms. As noted above, this could relieve the householder of the legal requirement for filing a Schedule H. However, data from the Bureau of Labor Statistics show that between 1999 and 2004 the number of child care workers (i.e., individuals who are not self-employed) grew from 377,110 to 513,110 and the number of personal and home care aides rose from 300,500 to 532,490.³ These figures likely include at least some workers who are non-agency employees and sug-

Figure 1.--Number of Schedule H Filings: TY 1996-2003



Source: Individual Return Transaction File, various years

gest that employment growth in these occupations has been strong even as Schedule H filings have declined.

A second possible explanation for the decline in Schedule H filings not related to evasion could be a fall in demand for the kinds of services offered by household workers. But, the recent strong employment growth for child-care and home health-care aides runs counter to this view. Also, as we shall see in the next section, Schedule H filing is strongly correlated with high-income households. Between TY 1996 and 2003, the number of taxpayers reporting adjusted gross income (AGI) of \$500,000 or more grew from 333,896 to 559,068, an increase of 67 percent. In addition to the jump in number of high-income earners, the Census Bureau reports that the number of family households grew from 69.3 million in 1995 to 75.6 million in 2003. Presumably, at least some of these new families would increase the demand for nannies and other household services.

A third possible explanation for the decline in Schedule H filings is the “outsourcing” of jobs to non-U.S. citizens. One example of this is the growing popularity of au pairs as an alternative to nannies for in-home child care. Au pairs are foreign citizens between 18 and 26 years old and must live with their host U.S. family for a period of not more than two years. The U.S. State Department, which issues J-1 visas to au pairs, reports the number of such visas increased from 11,171 in 2003 to 15,297 in 2004.⁴ However, even if the entire increase in au pair visas displaced an equivalent number of nannies, this could only account for one-third of the drop in Schedule H filings between these two years (see Figure 1).

The use of undocumented workers represents another avenue to outsource jobs in the household sector. When an undocumented worker is hired both the employer and employee have an incentive not to report employment taxes. By evading taxes, employers can pay higher cash wages and workers can stay “invisible” to both tax and immigration authorities. Reports of the growing numbers of undocumented household employees recently prompted even the *Wall Street Journal* to declare, “Nannies are among the most exploited workers in the country.”⁵ As evidence of the growing practice of hiring undocumented workers we need look no further than the aforementioned high profile political appointee

cases, all of whom paid undocumented aliens to work in their homes.

However, it is unclear if the mere presence of a large supply of willing undocumented workers is contributing to the falling trend of Schedule H filing. For example, if the cost of hiring a citizen or documented non-citizen to perform household tasks is prohibitive, households may forgo hiring domestic help altogether and do the work themselves or with other family members. By lowering the cost of labor, a large undocumented workforce may induce demand for household help that wouldn’t otherwise exist. In other words, if all undocumented household workers were somehow removed from the workforce, this would not necessarily produce an increase in Schedule H filing.

The purpose of this paper is to identify factors associated with Schedule H filing and to determine if these factors can account for the recent decline in filing activity. In the next section we examine tax return and other data to identify socioeconomic characteristics of Schedule H filers. The third section presents our analysis of the data using a probit specification of Schedule H filing rates for TY 2003 by 3-digit zip codes and an OLS model of the change in state filing rates between TY 1996 and 2003. The fourth section discusses the implications of our empirical findings and offers several hypotheses to account for the geographic variation in filing behavior that does not appear to be explained by other factors. Finally, we summarize our main findings and briefly outline our plans for future research on this topic.

► **Schedule H Filer Characteristics**

We obtained data for this study from individual tax returns filed between 1997 and 2004 (corresponding to TYs 1996 to 2003). Table 1 displays selected characteristics of TY 2003 taxpayers by Schedule H filing status. The characteristics were chosen based on a priori judgment regarding the types of taxpayers who employ household labor and the kinds of services provided.

Table 1 shows a majority (54 percent) of Schedule H filers reported AGI of \$150,000 or more in TY 2003. Perhaps because married taxpayers also tend to have higher incomes we see that Schedule H filers are more likely

Table 1.--Selected Taxpayer Characteristics: TY 2003

Filed Schedule H?	Taxpayers Count	Reported AGI Over \$150,000 Percent	Married Filing Joint Status Percent	Taxpayer Age 65+ Percent	Children Living at Home Exemptions Average
No	131,792,518	3.47%	41.46%	12.50%	0.612
Yes	234,465	54.18%	68.06%	38.77%	0.914
Total	132,026,983	3.56%	41.51%	12.54%	0.613

Source: Individual Return Transaction File

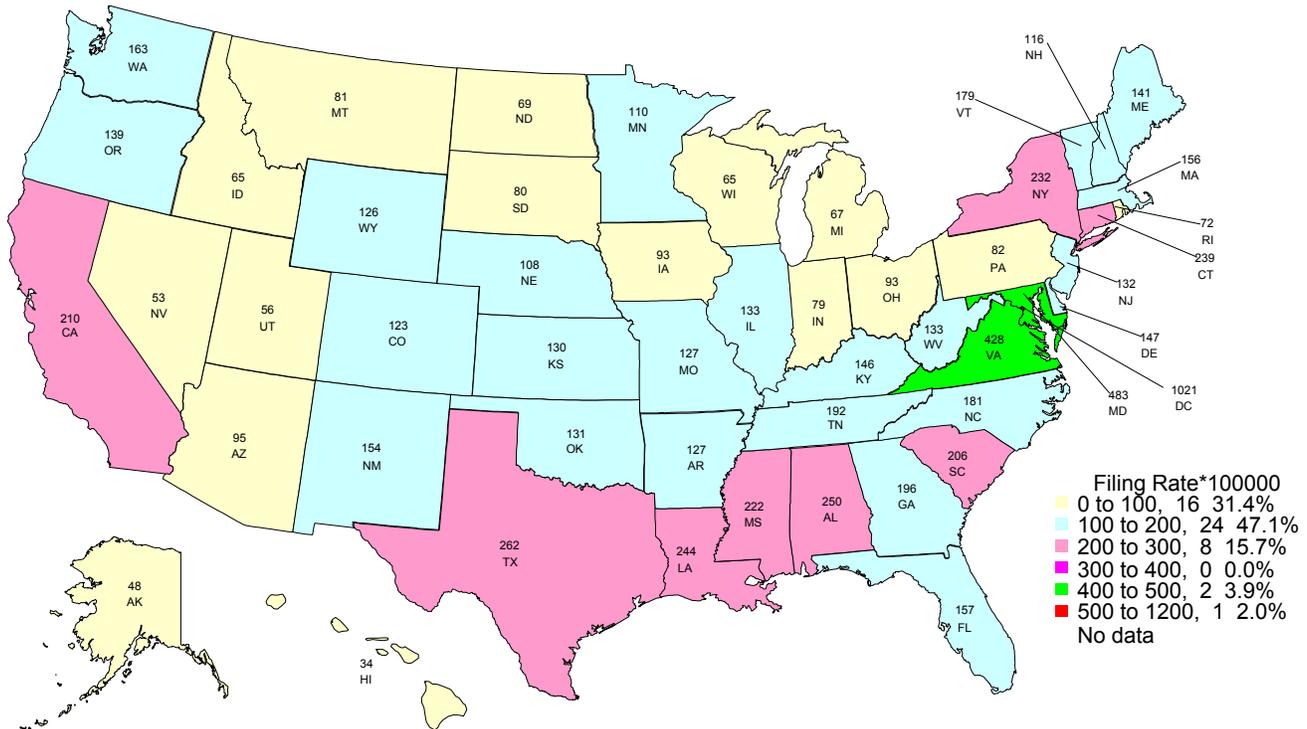
to file jointly than non-Schedule H filers. Persons 65 or more years old accounted for 38.8 percent of all Schedule H filings even though this age group represented only 12.5 percent of all taxpayers. Finally, Schedule H filers also claim more exemptions for children living at home than other filers (an average of 0.914 exemptions versus 0.612 exemptions for non-Schedule H filers).

Figure 2 displays TY 2003 Schedule H filing rates by state. The filing rate (per 100,000 taxpayers) is defined as the number of Schedule H filings divided by the total number of individual income tax filers (including Forms

1040, 1040A, and 1040EZ). From Figure 2, we see that the District of Columbia, Maryland, and Virginia have the nation’s highest filing rates. The three-state combined average of 508 Schedule H filings per 100,000 returns is 3.1 times the national average of 161 filings.⁶ The filing rate for the District of Columbia (1,021 filings per 100,000 returns) is more than six times the national average.

A second feature of Figure 2 appears to show that taxpayers in Southern states are more likely to file a Schedule H than taxpayers in Midwestern and Northern

Figure 2.--Schedule H Filing Rates by State: TY 2003



states. A difference of means test for Schedule H filing rates finds that the average filing rate of 226 filings per 100,000 taxpayers in 11 southern states⁷ is statistically distinct ($p < 0.001$) from the national average. Finally, higher filing rates also occur in the northeastern states of Connecticut and New York and in California.

Spatial Variation in Filing Rates

To examine the spatial variation of Schedule H filing in greater detail, we disaggregated the data by 3-digit zip code. For example, in California the zip codes with the highest filing rates are clustered near Los Angeles and San Francisco. Other major urban areas with high filing rates include New York City, Chicago, and Houston. From the analysis of tax return data we were not surprised to find Schedule H filers concentrated in high-income urban centers. However, we were surprised to find elevated Schedule H filing rates in a number of

small southern cities such as Farmville, VA, Selma, AL, Greenville, MS, and Shreveport, LA. Table 2 lists the 20 zip code areas with the highest filing rates.

The unusually high Schedule H filing rates in and near the nation's capital and, to a lesser extent, in the southern states appear puzzling given relative levels of per capita income (Table 2). In the case of Washington, D.C., we hypothesized that the high Schedule H filing rates could be related to the region's role as the seat of Federal authority and the large population of Federal civilian and military personnel living in the area. There are several reasons why this might be the case. First, due to their choice of career, Federal government workers might identify more with the government obligation to report and pay taxes than non-Federal taxpayers (Akerlof and Kranton, 2000, 2002 and 2005). According to Akerlof and Kranton, the concept of identity implies that if an individual's actual behavior deviates from the

Table 2.--Twenty Zip Code Areas with the Highest Schedule H Filing Rates: TY 2003

Region	State	Zip Codes	Filing Rate (per 100,000)	Percent of National Average Filing Rate	Per Capita Income (1999)
Bethesda/Silverspring	MD	208-209	1,993	1238%	\$35,538
DC	DC	200&202-205	1,841	1144%	\$28,569
New York	NY	100-102	1,265	786%	\$43,077
Greenwich/Norwalk	CT	068-069	822	510%	\$45,815
Alexandria/Fairfax	VA	201&220-223	778	483%	\$34,499
Charlottesville	VA	229	728	452%	\$22,547
Scarsdale/White Plains	NY	105-108	708	440%	\$36,194
Dallas	TX	752-753	694	431%	\$23,489
Morristown	NJ	079	649	403%	\$48,839
Great Neck	NY	110	602	374%	\$35,869
Beverly Hills/Culver City/Torrance	CA	902-905	552	343%	\$24,897
Pasadena	CA	910-912	530	329%	\$27,069
San Francisco/Palo Alto	CA	940-941&943-944	517	321%	\$36,949
Houston	TX	770&772	497	309%	\$20,830
Los Angeles	CA	900-901	472	293%	\$18,041
Mill Valley	CA	949	451	280%	\$38,630
Selma	AL	367	443	275%	\$13,347
Greenville	MS	387	409	254%	\$12,370
Shreveport	LA	710-711	402	250%	\$16,965
Farmville	VA	239	385	239%	\$15,384

Source: Individual Return Transaction File; U.S. Census Bureau (per capita income)

ideal behavior associated with the individual's identification, then the individual experiences a loss of utility. If we apply the concept of identity in the context of tax compliance, the intuition is clear: 1) People are identified with the tax system; 2) The ideal behavior (norms) associated with this identification is that people think they should comply with the tax system and pay the appropriate amount of tax; and 3) If people evade tax and thus their actual behavior departs from the ideal behavior, they will lose utility. Under this interpretation, people would differ by whether they are identified with the tax system or not and to what extent.

A second reason why Federal employees might be motivated to comply is a belief that they would face harsh penalties for modest infractions of the law. For example, Section 1203b of the Revenue Reform Act (RRA) of 1998 requires termination of employment for any IRS employee who fails to timely file a tax return; even if a refund is owed. In addition to potentially career-ending penalties, Federal employees might believe they are subject to a higher level of tax scrutiny than members of the general public – a belief that is not entirely unfounded. In order to allocate its staff to those cases it deems the highest priority, the IRS classifies each new collection case. In recent years, the top three priority categories – in decreasing order of importance – have been: (1) open criminal investigations, (2) IRS employees, and (3) Federal employees and retirees. Other things being equal, collection cases assigned a higher priority are more likely to be worked. Therefore, Federal employees and retirees who fall behind in their tax obligations stand a greater chance of being contacted by the IRS than most other taxpayers.

This explanation is consistent with the standard model on tax compliance (Allingham and Sandmo, 1972). The standard tax compliance model is based on traditional expected utility theory. In this model, a rational individual takes his income (W) that is unknown to the tax authorities, the tax rate (t), the audit probability (p), and the penalty rate (f) as given and chooses his declared income (X). After the individual declares his income, and if his declared income is less than his true income, he faces two possibilities: 1) With probability $(1 - p)$, he will not be audited by the tax

authorities so that he gains by $t(W - X)$; and 2) With probability p , he will be audited and the tax authorities will then know his true income. The consequence is that he will have to pay tax on the undeclared income ($W - X$) at penalty rate (f) that is greater than tax rate (t). In other words, he will lose by $(f - t) * (W - X)$. The individual chooses his optimal declared income (X^*) by maximizing his expected utility function: $E(U) = (1 - p)u(W - tX) + pu(W - tX - f(W - X))$. The model implies that increasing audit probability (p) or penalty rate (f) can reduce tax evasion.

In order to test the hypothesis of higher filing compliance by Federal employees, we compared Schedule H filing rates for IRS employees who reported more than \$150,000 AGI in TY 2003 to non-IRS employee filers in the same income category. [We wanted to use data on all Federal employees but were unable to obtain payroll data from the Office of Personnel Management in time for this study.] Table 3 displays the frequency counts of Schedule H filers by IRS employment status. A Chi-Square value of 16.298 indicates that IRS employees with reported AGI over \$150,000 are more likely to file a Schedule H than non-IRS employees⁸ in the same income group. However, the motive for this behavior (whether identification with government as in Akerlof and Kranton (2000, 2002 and 2005) or fear of detection as in the traditional evasion literature) remains an open question.

Besides Federal employees, other D.C. area residents whose careers are tied directly or indirectly (e.g., lobbyists) to the Federal sector also might be motivated to comply with tax laws covering household employees. Barbara Kline, owner of a nanny placement service in the Washington, D.C. area, observed the following about the Bernard Kerik situation, "Maybe his illegal nanny didn't seem like a problem in New York, but any professionally ambitious Washington parent knows enough by now to play strictly by the rules. They make sure to hire either domestic or documented foreign help, and pay their social security, disability, and unemployment 'nanny' taxes" (Kline, 2005). Another factor enhancing awareness of this issue in the Washington, D.C. area is the prominent press coverage in the Washington Post and other media outlets. Therefore, we believe that the high

Table 3.--Schedule H Filing by IRS Employees and Others with Reported AGI of \$150,000 or More: TY 2003

IRS Employee	Schedule H Filer		Total
	No	Yes	
No	4,744,126	126,850	4,870,976
	97.4%	2.6%	100.0%
Yes	5,246	189	5,435
	96.5%	3.5%	100.0%
Total	4,749,372	127,039	4,876,411
	97.4%	2.6%	100.0%

Source: Individual Return Transaction File

Schedule H filing rates in Washington, D.C. and in the bordering states of Maryland and Virginia, could reflect, in part, a stronger imperative in the minds of taxpayers living in and near the nation's capital of the obligation to report and pay Federal household employment taxes.

Finally, from Table 2 we note that communities such as Greenville, MS and Selma, AL neither have large high-income sub-populations or a significant Federal presence which might account for the higher observed Schedule H filing rates. Therefore, our tentative working hypothesis is that the higher filing rates in the southern states is a relic of historical and cultural factors that have traditionally viewed the hiring of household help as more socially acceptable than in other parts of the nation.⁹ In support of this view, we point out that the combined Schedule H filing rate for high income taxpayers (i.e., with reported AGI of \$150,000 or more) in Puerto Rico and the Virgin Islands is nearly 100 times the U.S. average. Although both Puerto Rico and the Virgin Islands are not included in this study due to their unique taxpayer populations, such large differences in Schedule H filing activity suggest that cultural factors could also be responsible for the higher filing rates in the South.

Temporal Change in Filing Rates

Figure 3 and Table 4 show the change in Schedule H filing rates by state from TY 1996 to 2003. The national trend of declining filing activity is reflected in every state without exception. The states with the largest rate declines are located in the South and in the Washington, D.C. area. However, bear in mind states in these regions

had higher initial levels of filing meaning that a change with the same relative impact on all states would result in disproportionate absolute rate changes in states in the South and in the D.C. area.

This relationship is seen more clearly in Table 4. For example, both Michigan and Alabama experienced a 43.7 percent decline in Schedule H filing rates between 1996 and 2003. However, the filing rate for Alabama fell by 194 Schedule H filings per 100,000 returns whereas for Michigan the equivalent relative change resulted in a decline of only 52 filings per 100,000 tax returns.

However, these regional differences do not explain why Schedule H filing rates fell in all states during this period. To shed some light on this issue we turn to Table 5 which shows the change in Schedule H filing by reported AGI in TY 1996 and 2003. The number of Schedule H filings has declined in all AGI categories except for those households that reported AGI of \$500,000 or more. In TY 1996, households reporting less than \$100,000 AGI accounted for 43 percent of all Schedule H filings, but by 2003 this group's share had fallen to 33 percent of a smaller total. Taxpayers with reported AGI less than \$100,000 accounted for over 70 percent of the total decline of 85,912 Schedule H filings between TY 1996 and 2003. Although the number of Schedule H filings grew among taxpayers with more than \$500,000 in reported AGI, the overall filing rate fell because the number of filers in this income group grew faster than the number of new Schedule H filers.

Although taxpayers with AGI less than \$100,000 account for most of the decline in number of Schedule

Figure 3.--Change in Schedule H Filing Rates: TY1996-2003

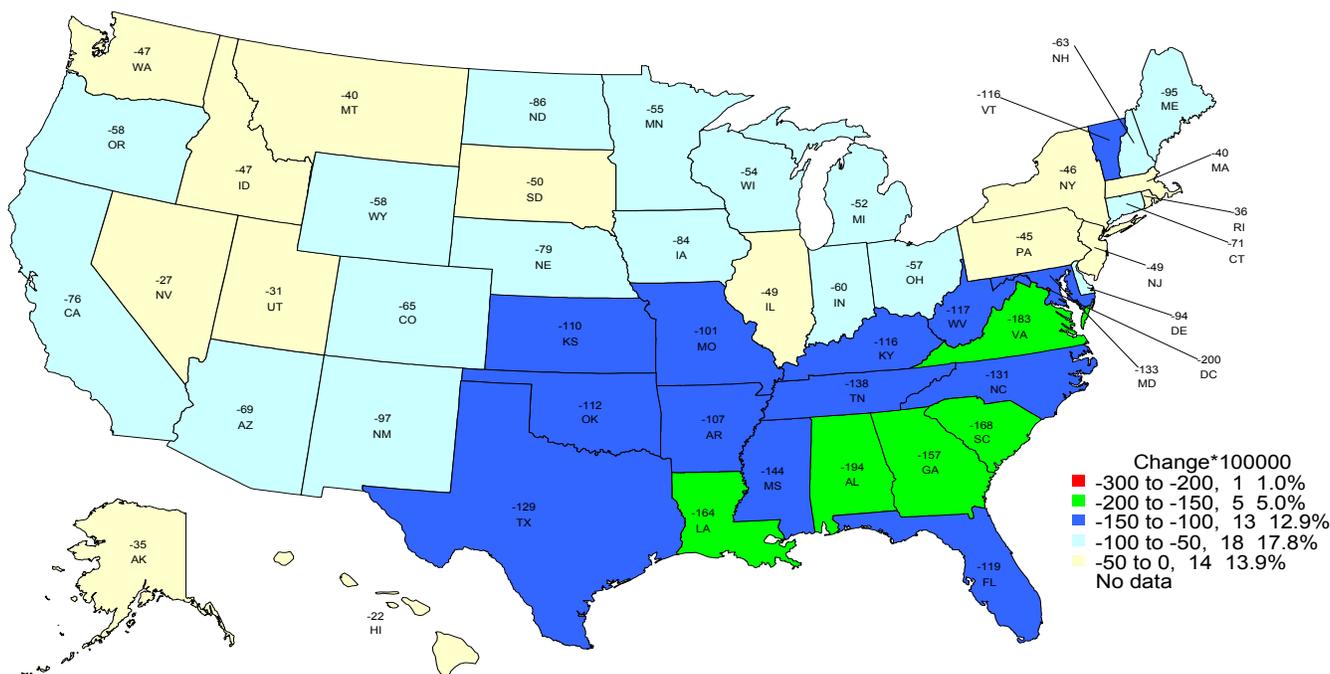


Table 4.--Change in Schedule H Filing Rates per 100,000 Taxpayers: TY 1996-2003

State	Filing Rate Change		State	Filing Rate Change	
	Number	Percent		Number	Percent
North Dakota	-85.7	-55.3%	Delaware	-94.3	-39.1%
Iowa	-83.8	-47.4%	New Mexico	-96.5	-38.5%
West Virginia	-117.4	-46.9%	South Dakota	-49.5	-38.3%
Oklahoma	-111.7	-46.0%	Ohio	-57.2	-38.1%
Kansas	-110.4	-45.9%	Utah	-31.2	-36.0%
Arkansas	-107.2	-45.8%	Pennsylvania	-44.7	-35.3%
Wisconsin	-53.7	-45.1%	New Hampshire	-63.1	-35.2%
South Carolina	-168.3	-45.0%	Colorado	-65.1	-34.6%
Georgia	-156.7	-44.4%	Nevada	-26.7	-33.5%
Kentucky	-115.9	-44.3%	Rhode Island	-35.8	-33.4%
Missouri	-101.0	-44.3%	Minnesota	-54.8	-33.3%
Michigan	-52.2	-43.7%	Texas	-128.9	-33.0%
Alabama	-193.6	-43.7%	Montana	-39.9	-32.9%
Indiana	-60.4	-43.3%	Wyoming	-57.8	-31.4%
Florida	-119.3	-43.1%	Virginia	-182.7	-29.9%
Nebraska	-78.8	-42.1%	Oregon	-57.7	-29.3%
Idaho	-47.5	-42.1%	Illinois	-49.2	-27.0%
Arizona	-68.8	-42.0%	New Jersey	-48.9	-27.0%
Alaska	-34.7	-42.0%	California	-75.7	-26.5%
North Carolina	-131.1	-41.9%	Connecticut	-71.3	-23.0%
Tennessee	-138.1	-41.8%	Washington	-46.8	-22.3%
Maine	-95.2	-40.2%	Maryland	-133.5	-21.6%
Louisiana	-164.0	-40.1%	Massachusetts	-40.5	-20.6%
Mississippi	-144.2	-39.4%	New York	-46.1	-16.6%
Vermont	-116.4	-39.4%	District of Columbia	-200.4	-16.4%
Hawaii	-22.2	-39.4%			

Source: Individual Return Transaction File

H filings, Table 5 also shows that filing rates are lower among all income groups. This could indicate that households are either: (1) no longer reporting to the IRS wages paid to legal or illegal workers, or (2) are changing their lifestyles to reduce their dependence on paid household help, or (3) a combination of the above. As an example of a lifestyle change, the *Wall Street Journal* recently reported that many parents are working flex-time schedules in order to reduce the number of hours needed for a baby-sitter or nanny.¹⁰ In other cases, parents have tried sharing a full-time nanny among several families or enrolling their children in pre-school at an earlier age. Child-care providers involved in such sharing arrangements may be considered self-employed under IRS rules if they control their work conditions (i.e., where and how the work is performed). However, no comprehensive data are available to measure how widespread such practices have become or whether this development alone could account for the large observed drop in Schedule H filings. We suspect that even with these arrangements it is likely that hiring legal domestic help is becoming increasingly a luxury good that is out of reach of most middle and high-middle income households and that the appeal of evasion is growing for many who cannot find legal substitutes among the self-employed or agency employees. As an indicator, the same *Wall Street Journal* article cites hourly rates for part-time nannies from \$13 to \$25, plus benefits such as paid vacations.

► Model Estimation

In this section, we estimate two empirical models of Schedule H filing activity. First, we estimate a pro-

bit model of TY 2003 Schedule H filing rates for 576 3-digit zip code areas. Model specification A includes the four indicators of Schedule H filing propensity identified from tax return data (see Table 1). These are: percentage of taxpayers that report more than \$150,000 AGI (*PctHiInc*), percentage of taxpayers whose filing status is married filing joint (*PctMFJ*), percentage of taxpayers age 65 years or older (*PctAge65+*), and average number of exemptions for children living at home (*AveChHomeEx*). A priori, we expect positive signs on all four variables.

Model specification B adds the percentage of the resident population who are non-citizens (*PctNonCitizen*) and Federal employment as a percentage of total employment (*PctFedEmp*). We include *PctNonCitizen* to account for the possible influence of undocumented workers on the decision to file a Schedule H. Since it is unclear based on the earlier discussion (on page 3) if the mere presence of undocumented workers alone would influence taxpayers' willingness to file a Schedule H, we are uncertain about the sign on *PctNonCitizen*.

We include *PctFedEmp* to represent the hypothesized link (whether due to identification or a heightened sensitivity to the consequences of IRS enforcement actions) between Federal employees and the obligation to pay Federal taxes. Based on the earlier discussion we anticipate a positive sign on this coefficient. We use Census 2000 data as the source for both *PctFedEmp* and *PctNonCitizen*. For this study, we assumed there was no difference within observations on these two variables between 2000 and 2003.

Table 5.--Change in Schedule H Filing by Reported AGI Category: TY 1996 and 2003

Reported AGI Category	All Filers				Schedule H Filers				Schedule H Filing Rate (per 100,000 filers)			
			Change				Change				Change	
	TY 1996	TY 2003	Number	Percentage	TY 1996	TY 2003	Number	Percentage	TY 1996	TY 2003	Number	Percentage
Under \$100K	115,180,718	120,163,036	4,982,318	4.3%	137,097	76,395	-60,702	-44.3%	119	64	-55	-46.6%
\$100-\$200K	4,659,894	9,152,043	4,492,149	96.4%	77,692	52,840	-24,852	-32.0%	1,667	577	-1,090	-65.4%
\$200-\$500K	1,221,645	2,152,836	931,191	76.2%	66,507	60,355	-6,152	-9.3%	5,444	2,804	-2,641	-48.5%
\$500K or More	333,896	559,068	225,172	67.4%	39,081	44,875	5,794	14.8%	11,705	8,027	-3,678	-31.4%
Total	121,396,153	132,026,983	10,630,830	8.8%	320,377	234,465	-85,912	-26.8%	264	178	-86	-32.7%

Source: Individual Return Transaction File

Finally, we also include two regional dummy variables. *South* takes on a value of 1 for 3-digit zip codes located in any of the 11 southern states, 0 otherwise. Again, this variable takes into account any unique cultural or historical factors we believe could be responsible for the higher filing rates in these states. Similarly, *DCRegion* equals 1 for all 3-digit zip codes in D.C., Maryland, and Virginia, else 0. This variable is used to pick up any difference in compliance behavior on the part of non-Federal employee taxpayers living in and near Washington, D.C. We expect positive signs for both *South* and *DCRegion*.

The estimated coefficients for the three models along with the Chi-Squared values are shown in Table 6. The parameter labeled *_C_* in Table 6 is the “natural response” rate which we assumed was equal to 0.0001 in both specifications. In specification A, three of the four tax return variables are statistically significant. The negative sign on *PctMFJ* could indicate, as we mentioned above, that high-income households also tend to be married households and that when these characteristics are entered as independent effects, their influence on Schedule H filing propensity changes. Perhaps among low and middle-income married households, the presence of a second adult in the home means routine domestic chores can be performed largely within the family and not require outside paid assistance.

In specification B, *PctAge65+* is not significant but both regional dummies (*South* and *DCRegion*) are significant and with the predicted sign. *PctFedEmp* and *PctNonCitizen* also are significant. The latter finding could indicate that areas with large non-citizen populations also contain a documented labor force available for employment in the household sector. However, this is only speculation on our part as we have not examined this issue in any detail.

A test for normality of the regression residuals finds that spatial autocorrelation is present and, therefore, it is likely the model has not adequately accounted for all of the factors influencing filing behavior. There are pockets of positive spatial autocorrelation in scattered locations throughout the South, in rural Virginia/West Virginia, and in Southern California. Also present are zones of high negative spatial autocorrelation in New

Jersey, Long Island, southern Connecticut, Atlanta and Dallas. The Virginia suburbs of Washington, D.C. and coastal Virginia appear to have lower than expected filings while the Maryland suburbs of D.C. have higher than expected filings along with D.C. itself. The mixed findings for suburban Washington, D.C. might indicate that the residential location of high-income Federal employees, lobbyists, and officers of corporations with Federal government contracts is more important than the mere presence of Federal employee filers. Another factor possibly influencing Schedule H filing rates is the degree of economic inequality present in an area which could influence the demand and supply for household labor. However, we did not explore this hypothesis in this study.

Using the probit analysis results we estimated an OLS regression model of the percentage change in Schedule H filing rates for the 50 states plus the District of Columbia (right-most column of Table 4). The purpose of this model was to determine if any of the

**Table 6.--Probit Estimation Results: TY 2003
Schedule H Filing Rates**

Parameter	Model Specification		
	A	B	Final
Intercept	-2.5159*** (697.62)	-2.8913*** (1541.81)	-2.8457*** (3312.02)
PctHiInc	5.7906*** (439.42)	5.7937*** (519.86)	5.9590*** (650.67)
PctMFJ	-1.4887*** (91.8)	-1.3152*** (91.41)	-1.2999*** (151.52)
PctAge65+	-0.9272** (4.29)	.3944 (1.74)	
AveChHomeEx	0.0671 (0.43)	-0.0042 (0.00)	
PctNonCitizen		0.6411*** (22.04)	0.5750*** (25.24)
PctFedEmp		1.7650*** (28.44)	1.6835*** (26.35)
DCRegion		0.1389*** (15.37)	0.1409*** (15.95)
South		0.2246*** (218.69)	0.2201*** (216.53)
<i>_C_</i>	0.0001	0.0001	0.0001
N	576	574	574
DF	571	565	567
-Log Likelihood	1,641,266.45	1,624,315.65	1,624,428.68

Chi-Square values in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels respectively. The dependent variable in each regression is the fraction of taxpayers who file a Schedule H.

factors we identified as contributing to the propensity to file a Schedule H could help explain the change in state-level Schedule H filing rates between TY 1996 and 2003. We used state data because we did not have zip code data for non-Census years. For the OLS model, both *South* and *DCRegion* are 0/1 dummy variables for the 11 southern states and the three states (DC, MD, and VA) in the national capital region, respectively. Instead of Census 2000 data for *PctFedEmp*, we use annual Bureau of Economic Analysis (BEA) estimates for state Federal employment to compute the change in percentage of Federal employment (*dPctFedEmp*). Instead of *PctMFJ* (the percentage of married filing joint filers), we calculate the change in percentage of MFJ taxpayers (*dPctMFJ*) from tax return data. Because we did not have non-civilian population data for the beginning and ending years, we used Census Bureau annual estimates to compute the change in percentage of state population from international migration (*dIntMigPctPop*). Finally, we substituted for *PctHiInc* (the percentage of Schedule H files with reported AGI over \$150,000) two variables: (1) *pct96H_AGI150* – the percentage of Schedule H filers with reported income less than \$150,000 in TY 1996 and (2) *dPct_AGI500* – the change in percentage of filers with more than \$500,000 in reported AGI. The variable *pct96H_AGI150* captures the evident change in filing behavior by taxpayers with less than \$150,000 in AGI since TY 1996. The variable *dPctAGI500* is included to account for the ameliorating effects on Schedule H filing associated with growth in the number of taxpayers in the category with highest AGI (see Table 5). We predict all variables will have the same signs as determined from the probit analysis and *dPctAGI500* will have a positive sign. We predict *pct96H_AGI150* will have a negative sign; that is, a larger concentration of TY 1996 Schedule H filers with AGI under \$150,000 will lead to a smaller filing rate in TY 2003. The OLS regression results are shown in Table 7.

► **Discussion**

The results from the OLS regression model in Table 7 show that the two income-based variables are highly significant predictors of the change in Schedule H filing behavior and account for most of the adjusted R Square

value of 0.68. This is a clear indication that the recent decline in Schedule H filing is linked to a shift away from the employment of household workers by middle and upper-middle income taxpayers. However, because the data also show filing rates have decreased for all income groups, we can not rule out the possibility that evasion is increasing, possibly in relation to the steady influx of undocumented workers entering the U.S.

The significance (at the 5% level) of the change in Federal employment on Schedule H filing behavior is interesting and warrants further analysis. Whether this result is due to Federal employees’ identification with the tax system or heightened sensitivity to the consequences of enforcement is unclear. We presented evidence (in Table 3) that high-income IRS employees file the Schedule H more frequently than similarly situated non-IRS employee taxpayers. We will continue efforts to develop a profile of Schedule H filing for all Federal employees. We anticipate this will be accomplished in the near future.

Future research will examine in greater depth the hypothesized relationship between the propensity to file a Schedule H and strength of identification with the

Table 7.--OLS Estimation Results

Parameter	Coefficient
Intercept	-0.0377 (-0.7491)
p96H_AGI150	-0.5350*** (-6.7639)
dPctMFJ	0.7330 (1.1878)
dPctFedEmp	8.2030** (2.0932)
dPct_AGI500	0.0845*** (4.1800)
south	-0.0145 (-0.7894)
dcregion	0.0180 (0.4766)
dIntMigPctPop	-0.0723 (-0.8405)
Adj. R-Square	0.6800

t-values in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels respectively. The dependent variable is the percentage change in Schedule H filing rate from TY 1996-2003.

tax system. Our probit model results indicate this could be a factor in the decision to file a Schedule H for both Federal employees and others living in the national capital region. However, our current research did not yet separate the influence of identification from heightened enforcement environment on Federal employees and retirees and others with ties to the Federal government. One possible approach to tackle this problem might be to combine our data on Schedule H filing with survey data from which we might be able to construct a proxy for taxpayers' identification with tax systems.

In this research, we define the filing rate of Schedule H as the ratio of the number of filers who filed a Schedule H with their tax return over the number of tax filers who filed an individual income tax return. We fully recognize that this definition is less than ideal. One alternative would be to define the filing rate as the ratio of the number of filers who filed a Schedule H divided by the *expected* number of Schedule H filers. Deriving an estimate of the expected number of Schedule H filers is on our research agenda. Large-scale surveys like the Census, the Current Population Survey (CPS), and the Survey of Income and Program Participation (SIPP) might be useful for this purpose. We think that constructing a new measure of Schedule H filing compliance would make an interesting and significant contribution in the area of tax compliance research.

Finally, we will investigate further the role of historical and/or cultural factors in the decision to file the Schedule H. Consultation with industry experts may help in this regard.

► **Summary**

Our analysis of tax return, Census, and other data has determined the following about Schedule H filers and the recent decline in filing activity:

- 1) Schedule H filers are concentrated among households with more than \$150,000 AGI, who select the married filing joint filing status, whose primary taxpayer is age 65 or older, and who claim more exemptions for children living at home than the average taxpayer.
- 2) The states with the highest Schedule H filing rates are the District of Columbia, Maryland, and Virginia. Taken together, filing rates in the three-state region bordering Washington, D.C. are 3.1 times higher than the rest of the nation. The Schedule H filing rate for the District of Columbia is more than six times the national average of 161 filings per 100,000 tax returns. Schedule H filing also occurs with greater frequency among taxpayers living in the 11 southern states.
- 3) A probit model of Schedule H filing rates by 3-digit zip code finds the percentage of high-income households, percentage of married filing joint returns, percentage of Federal employment, percentage of the population who are non-citizens, and location in the 11 southern states or the three-state national capital region (DC, MD, and VA) are statistically significant predictors of Schedule H filing. However, the regression residuals indicate some remaining spatial autocorrelation. Areas of positive spatial correlation occur in the South, in non-urban zip codes of Virginia and West Virginia, and in Southern California. Areas of possible negative spatial correlation occur in Northern New Jersey, Long Island, Connecticut, Florida, and the Virginia suburbs of Washington, D.C.
- 4) Using state data, an OLS regression of the percentage change in Schedule H filing rates between TY 1996 and 2003 finds positive correlations for the percentage change in high-income (> \$500,000 AGI) filers and percentage change in Federal employment. A negative correlation was found for percentage of TY 1996 Schedule H filers with reported AGI less than \$150,000. Analysis of tax return data finds that over 70 percent of the 85,912 drop in Schedule H filings between TY 1996 and 2003 occurred among taxpayers with less than \$100,000 in reported AGI, confirming that Schedule H filing has become increasingly concentrated among the very wealthy. However, the data also show that Schedule H filing *rates* declined substantially among all income groups during this same period underscoring the existence of a broad-based change in taxpayer behavior.

- 5) The observed geographic variation in Schedule H filing rates--higher in the South and the Washington, D.C. area--int at the possible influence of cultural or behavioral factors on taxpayer filing decisions. In particular, the extreme high filing rates in the national capital region could indicate the influence of identity or heightened sensitivity to enforcement consequences not present in the general population. Further research will examine these issues in greater detail.

► Endnotes

- ¹ See *The Beaumont Enterprise News*, "The Nanny 411," January 30, 2005.
- ² See *The Wall Street Journal*, "The Case for Paying the Nanny Tax: Despite Risks, Families Skirt the Law," March 17, 2005.
- ³ See BLS' Occupational and Employment Statistics website at <http://www.bls.gov/oes/home.htm>.
- ⁴ See *The Wall Street Journal*, "Number of Au Pairs Increases Sharply," March 1, 2005.
- ⁵ Cited in Kline (2005).
- ⁶ This difference is statistically significant at the 0.001 level using a t-test with unequal variance.
- ⁷ The 11 southern states are: Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia.

- ⁸ The category "Non-IRS employees" includes all non-IRS Federal civilian and military employees. Thus, if identification with government is a factor responsible for different filing rates, we may be underestimating the difference between IRS and non-Federal employees.
- ⁹ Although we only show state-level filing rates for TY 2003, the 11 southern states as a group exhibit higher filing rates for every year for which we have data.
- ¹⁰ See *The Wall Street Journal*, "Adventures in Babysitting: How to Hire Part-Time Child Care in a Hot Market," September 22, 2005.

► References

- Akerlof, G. A. and Kranton, R. E., Economics and Identity, *Quarterly Journal of Economics* (August 2000), pp. 715-753.
- Akerlof, G. A. and Kranton, R. E., Identity and Schooling: Some Lessons for the Economics of Education, *Journal of Economic Literature* (December 2002), pp. 1167-1201.
- Akerlof, G. A. and Kranton, R. E., Identity and the Economics of Organizations, *Journal of Economic Perspectives* (2005), pp. 9-32.
- Allingham, M. G. and Sandmo, A., Income Tax Evasion: A Theoretical Analysis, *Journal of Public Economics* (1972), pp. 323-338.
- Kline, B., (2005) *White House Nannies*, Tarcher/Penguin, Inc., New York, NY.

2



**Corporate Tax Issues:
Book-Tax Differences and
Measuring Tax Avoidance**

Boyton ♦ DeFilippes ♦ Legel

Prelude to Schedule M-3: Schedule M-1 Corporate Book-Tax Difference Data, 1990-2003*

by Charles Boynton and Portia DeFilippes, Office of Tax Analysis, U.S. Treasury Department, and Ellen Legel, Internal Revenue Service

For most large corporations, the new Schedule M-3 book-tax reconciliation replaces the 4-decade old Schedule M-1, effective December 2004. The goal of this paper is: (1) to present Schedule M-1 data and other selected tax return data for the immediately preceding 14-year period, 1990-2003; and (2) to address tax policy data interpretation issues related to U.S. intercompany dividends (ICD) improperly included on corporate tax returns by some large taxpayers.¹ First, we review events leading to the replacement of Schedule M-1 with Schedule M-3. We then present Schedule M-1 data and other selected tax data for 1990-2003 for two populations: (1) all corporations normally subject to the U.S. Federal corporate income tax; and (2) the subset that would have filed Schedule M-3 if the 2004-2006 requirements had been effective for the earlier years.² Most corporations with total assets of \$10 million or more are subject to Schedule M-3 starting in December 2004, and others entities (corporations and partnerships) will be subject starting in December 2006; we focus our Schedule M-1 discussion on the 1990-2003 data for such corporations. We conclude by discussing certain tax policy issues in interpreting Schedule M-1 data for 1990-2003 relating to U.S. intercompany dividends (ICD) improperly included on corporate tax returns by some large taxpayers. These issues will likely remain unresolved until Schedule M-3 data replace Schedule M-1 data.

► Dissatisfaction With Schedule M-1

A Treasury report in 1999 and Treasury testimony in 2000 by Assistant Secretary (Tax Policy) Jonathan Talisman noted the growing book-tax gap from 1991 to 1997 between pretax book income on Schedule M-1 and tax net income on page 1 of Form 1120. Both the report and the testimony viewed the 1990s book-tax gap as a possible indicator of corporate tax shelter activity, but also noted the difficulty in interpreting Schedule M-1 book-tax difference data.³ Mills-Plesko (2003) proposed

a redesign of Schedule M-1 to increase the transparency of the corporate tax return book-tax reconciliation and to improve data interpretability.⁴ The Mills-Plesko (2003) Schedule M-1 recommendations are largely reflected in Schedule M-3, particularly in Part I.⁵

► Schedule M-1 Versus Schedule M-3

Exhibit I presents a partial detail of Form 1120, page 1 and Schedule M-1. Schedule M-1 is intended to reconcile book income on Schedule M-1, line 1, with tax net income on Form 1120, page 1, line 28.

Exhibit II presents a partial detail of Schedule M-3 Part I and Part II. Part I reconciles worldwide consolidated financial statement income with income per income statement of includible corporations (members of the tax return consolidation group listed on Form 851). Parts II and III reconcile income per income statement of includible corporations (“book”) with tax net income on Form 1120, page 1, line 28. Differences between book and tax are characterized as temporary or permanent.

Part I of Schedule M-3 is important. It defines the starting point for the book-tax reconciliation for the first time in corporate tax history. On Schedule M-1, we know where the reconciliation ends (tax net income) but not where it begins (book). Taxpayers choose Schedule M-1 line 1 book income to suit them. Schedule M-3, Part I, line 11 is what Schedule M-1, line 1 should have been all along. Schedule M-3 uses many of the Schedule M-1 revisions proposed by Mills-Plesko (2003), in particular, Schedule M-3, Part I.

The goal of Schedule M-3 is greater transparency and uniform organization in book-tax data at the time of return filing so that the data may be used to determine what returns will and will not be audited and to determine what issues will and will not be examined on the returns selected for audit.

► **Schedule M-3 Effective 2004**

Effective for all tax years ending on or after December 31, 2004, U.S. corporations with end-of-year total assets of \$10 million or more filing Form 1120, *U.S. Corporation Income Tax Return*, must complete Schedule M-3, *Net Income (Loss) Reconciliation for Corporations With Total Assets of \$10 Million or More*, in place of Schedule M-1, *Reconciliation of Income (Loss) per Books With Income per Return*. Effective tentatively for all tax years ending on or after December 31, 2006, the requirement to complete Schedule M-3 will be extended to U.S. insurance companies (life insurance companies filing Form 1120-L and property and casualty insurance companies filing Form 1120-PC), to S corporations filing Form 1120-S, and to partnerships filing Form 1065, all with total assets of \$10 million or more.⁶ The January 28, 2004, joint Treasury-IRS announcement of Schedule M-3 indicated that Schedule M-3 would become an important IRS audit selection tool both for the selection of corporate returns for audit and the identification of issues on a return for audit.⁷

► **Source of 1990-2003 Data⁸**

A statistical sample of tax return data is electronically encoded annually by the Statistics of Income Division (SOI), Internal Revenue Service, for the use of the Office of Tax Analysis (OTA), U.S. Department of the Treasury, and the Joint Committee on Taxation (JCT), U.S. Congress. These data include Schedule M-1 data. Selected tax return data for all corporations normally subject to the U.S. Federal corporate income tax are summarized annually by SOI in Table 12 of Publication 16, *Statistics of Income, Corporation Income Tax Returns*. SOI Publication 16 tables do not present Schedule M-1 data. To date, only Plesko (2002) (for 1996-1998) and Plesko-Shumofsky (2005) (for 1995-2001) have presented Schedule M-1 data for the SOI Publication 16 Table 12 population.

► **Discussion of Tables 1-4**

Tables 1 through 4 all have the same standardized format for presenting Schedule M-1 data and selected tax return data for 1990-2003.⁹ The title of the table indicates the population or population split for which the

table aggregates data. For example, Table 1 presents data for all corporations excluding those that file specialized Forms 1120 as S corporations, as regulated investment companies (RIC's), or as real estate investment trusts (REIT's). Table 2 restricts the Table 1 population to domestic corporations with total assets at end of year of \$10 million or more as reported on Form 1120, Schedule L.¹⁰

Each table has three panels. The first row of each panel indicates the weighted number of returns for the year for the panel tabulated (N1, N2, and N3 for the first, second and third panels). Returns are weighted because a statistical sample of firms is used to represent the population. Generally, firms larger than \$10 million in total assets have a weight of 1, that is, they represent only themselves in the sample. Smaller firms generally have weights of greater than 1 (for example, 5), that is, the selected firm represents several similar firms (for example, 5 firms). In preparing the tables, we had a "suppression" program check to see if any year (column) of data for any table panel was based on fewer than 10 weighted returns or fewer than three original records ("unweighted" returns). SOI does not allow reporting of data based on such low counts both for statistical reasons (not less than 10 weighted returns) and to preserve taxpayer confidentiality (not less than three original records, that is, unweighted returns). If our suppression program detects a low count for any "data cell", we must suppress not only that data cell but also an adjacent data cell so that the data cannot be recreated by subtraction using any other totals presented or available elsewhere. In Tables 3 and 4, we have suppressed all data in the second and third panels as an overly cautious and simplified response to the restrictions on low counts for any "data cell."

The first panel of each table is divided into two sections, "Summary" and "Schedule M-1 Detail." In the summary section, we present the weighted number of returns on which our data are based and selected aggregate data from Schedule M-1 or elsewhere in the return. For example, tax net income is from Form 1120, page 1, line 28. In some cases, the data are calculated. For example, pretax book income is the result of adding the amounts for Schedule M-1 line 1 and line 2. Book-tax difference is pretax book income minus tax

Table 1. All Corporations (Excluding S, RIC, and REIT)

Dollar amounts in millions. Table values may not add and may differ from SOI Publication 16 values due to rounding.

Summary	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Returns [N1]	2,136,032	2,098,641	2,077,517	2,055,982	2,310,703	2,312,382	2,317,885	2,248,064	2,249,989	2,198,739	2,172,705	2,136,756	2,100,074	2,047,593
Total assets	16,164,132	16,678,060	17,244,251	18,548,429	19,876,869	21,739,327	23,511,425	26,398,625	29,539,369	32,202,629	36,892,288	39,088,335	40,436,428	42,224,867
Total receipts	9,689,005	9,660,584	9,821,790	10,154,951	11,020,931	11,955,287	12,709,002	13,445,455	13,996,496	15,238,419	16,607,285	16,214,517	15,582,599	16,200,948
Pretax book income	292,375	280,145	249,839	412,078	628,654	628,654	752,693	819,582	816,738	853,663	784,075	221,319	347,735	899,320
Tax net income	270,925	248,113	291,867	368,912	426,082	514,751	574,554	607,541	532,246	535,289	517,937	270,774	258,674	455,434
Book-tax difference	21,450	12,031	-42,028	43,166	119,192	113,903	178,139	212,040	284,492	318,374	266,138	-49,455	89,061	443,886
M-1 explains	-37,786	-67,672	-111,638	-24,115	43,759	29,345	61,640	87,597	130,355	172,188	144,482	-183,229	-44,341	289,386
Estimated ICD adj	65,233	71,308	68,806	66,255	76,959	83,549	113,377	107,778	159,381	154,849	129,975	142,549	154,352	140,399
Tabulated ICD adj	0	0	0	0	0	0	0	0	0	0	133,854	127,359	148,227	135,955
Tax net income > 0	416,617	401,582	426,078	496,152	554,084	641,754	714,272	765,753	736,810	743,499	859,531	709,004	676,337	780,053
NOL deduction	38,399	41,152	38,888	45,090	48,872	57,090	55,019	60,289	52,638	64,781	77,079	60,332	65,667	70,294
Special deductions	12,545	12,668	13,527	14,995	14,001	23,612	20,750	24,258	24,109	30,910	27,060	21,824	16,474	13,760
Taxable income	366,311	349,794	377,723	436,613	493,712	564,346	639,366	683,242	662,258	692,591	758,979	634,366	599,637	698,383
Tax before credits	128,007	120,989	131,154	154,359	172,644	198,578	223,454	239,134	230,912	241,430	265,645	220,496	209,252	243,261
Foreign tax credit	24,990	21,097	21,521	22,896	25,401	30,420	40,244	42,200	37,396	38,390	48,506	41,063	42,022	50,034
Tax after credits	96,224	92,435	101,402	119,848	135,387	156,184	170,362	183,916	186,334	192,473	203,408	166,334	153,173	176,956
Pretax book income > 0	484,694	474,568	506,574	607,242	702,602	794,908	924,167	1,005,115	1,064,057	1,138,311	1,245,562	972,495	1,056,118	1,244,020
Book-tax difference > 0	192,458	195,192	201,339	247,767	288,679	280,171	347,260	386,294	495,001	545,780	611,920	485,453	631,824	689,795
M-1 explains > 0	153,113	134,939	152,642	195,417	211,380	219,577	255,044	297,695	375,394	435,037	509,419	394,379	522,626	590,683
Depreciation explains	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Stock options explains	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Stock options explains	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Schedule M-1 Detail														
[H-1] Book net income	203,250	183,548	184,595	305,013	394,789	455,690	553,497	599,870	600,319	600,127	516,667	59,728	166,377	656,741
[H-2] Book federal tax	89,124	76,597	65,245	107,065	150,485	172,965	199,197	219,712	216,419	253,537	267,408	161,591	181,358	242,580
[=] Pretax book income	292,375	260,145	249,839	412,078	545,274	628,654	752,693	819,582	816,738	853,663	784,075	221,319	347,735	899,320
[H-7] Income for book	0	198,573	218,343	274,197	246,539	278,249	320,284	384,592	472,270	523,506	677,680	522,330	601,810	641,836
[H-8] Deduction for tax	0	395,214	422,021	456,629	479,405	517,522	574,240	656,745	743,899	809,100	944,726	955,829	1,000,989	1,051,470
[H-3] Capital loss limit	0	-7,452	-8,443	-9,429	-10,415	-11,401	-12,387	-13,373	-14,359	-15,345	-16,331	-17,317	-18,303	-19,289
[H-4] Income for tax	0	-224,123	-218,376	-238,425	-228,897	-254,185	-298,656	-348,626	-372,892	-405,478	-450,662	-489,598	-463,934	-468,576
[L-5] Expense for book	0	-429,883	-525,184	-509,586	-445,635	-504,461	-525,582	-590,910	-697,518	-733,155	-958,392	-1,126,614	-1,114,248	-859,172
[=] [M-1] explains	-37,786	-67,672	-111,638	-224,123	-228,897	-254,185	-298,656	-348,626	-372,892	-405,478	-450,662	-489,598	-463,934	-468,576
[H-8a] Depreciation for tax	0	110,842	124,957	131,629	140,239	151,985	154,489	177,232	205,270	212,646	229,469	270,189	314,269	298,132
[L-5a] Depreciation for book	0	-75,398	-81,861	-82,424	-86,584	-98,622	-97,709	-111,705	-125,235	-126,499	-141,009	-156,072	-154,068	-163,398
[=] Depreciation explains	22,043	35,444	43,096	49,205	53,655	53,364	56,781	65,526	80,035	86,147	88,460	114,117	160,200	134,734
[H-7a] Tax-exempt interest	0	20,504	20,415	20,567	21,101	21,010	20,492	20,123	22,455	22,972	22,205	21,112	21,601	22,467
[H-8c] Stock options	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[L-5c] Travel-entertainment	0	-2,579	-2,876	-2,963	-7,842	-8,084	-8,771	-9,127	-9,914	-10,002	-10,747	-9,948	-9,626	-15,178
[L-3] Capital loss limit	0	-7,452	-8,443	-9,429	-10,415	-11,401	-12,387	-13,373	-14,359	-15,345	-16,331	-17,317	-18,303	-19,289
M-1 detail explains [5 above]	22,043	45,917	52,192	59,880	59,261	58,508	59,855	62,319	77,173	77,332	79,049	90,105	129,971	101,659
M-1 other explains [balance]	-59,829	-113,588	-163,830	-83,994	-15,502	-29,163	1,784	25,279	53,182	94,857	62,433	-273,334	-174,312	187,727
Pretax Book Income Only														
Returns [N2]	697,321	735,882	682,766	642,260	673,048	669,826	661,701	630,176	651,937	620,207	616,381	616,536	588,560	558,472
Assets	523,627	581,113	548,239	507,901	494,373	532,796	662,690	772,284	964,253	1,007,857	1,169,540	1,313,572	1,311,099	1,279,301
Receipts	322,070	375,260	336,727	353,948	318,683	322,574	379,168	403,338	438,697	467,937	508,472	511,943	476,754	444,808
Pretax book income	-1,086	8,711	6,953	7,326	3,527	3,423	8,713	11,298	11,975	9,959	6,758	5,496	1,620	12,280
Net-income	-1,608	3,318	3,344	4,915	2,796	4,698	6,870	7,144	10,350	8,946	7,406	6,647	6,622	10,046
BTD	522	5,393	3,609	2,410	730	-1,274	1,843	4,154	1,625	913	-648	-1,151	-5,001	2,234
Tax After Credits	1,467	2,539	2,642	3,316	2,181	2,655	3,160	3,924	4,743	4,070	4,451	4,525	3,033	3,676
No Pretax Book Income														
Returns [N3]	177,400	161,191	194,254	231,359	254,577	252,065	247,354	247,898	243,544	246,668	255,714	253,143	277,716	298,155
Assets	159,399	117,770	63,022	153,342	156,995	115,407	269,671	162,129	415,768	374,267	350,925	496,727	328,039	266,268
Receipts	91,410	112,164	70,548	86,262	124,863	81,403	122,370	98,590	156,836	272,055	208,028	240,755	139,081	161,076
Net-income	-663	892	-272	2,703	5,736	719	1,791	-160	2,451	14,887	1,690	4,487	-1,197	1,363
BTD	663	-892	272	-2,703	-5,736	-719	-1,791	160	-2,451	-14,887	-1,690	-4,487	1,197	-1,363
Tax After Credits	535	621	315	929	820	460	1,038	454	1,117	3,631	2,539	1,079	1,062	1,363

Table 2. U.S. Corporations (Excluding F. S. RIC, and REIT) With Assets of \$10 Million or More

Dollar amounts in millions. Table values may not add and may differ from SOI Publication 16 values due to rounding.

Summary	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Returns [N1]	41,625	41,161	40,815	41,662	43,068	44,063	45,324	45,324	45,068	45,976	47,056	45,339	43,760	43,155
Total assets	15,362,220	15,889,853	16,469,258	17,764,780	19,040,111	20,881,409	22,636,159	25,527,012	28,652,482	31,311,114	36,005,051	38,224,380	39,605,830	41,408,040
Total receipts	7,605,263	7,674,997	7,884,246	8,175,249	8,738,525	9,587,949	10,254,017	10,919,097	11,418,473	12,654,017	13,955,969	13,758,278	13,286,373	13,909,128
Pretax book income	295,142	267,890	250,195	398,154	533,071	619,879	738,348	797,156	710,466	829,575	788,014	255,005	380,674	910,546
Tax net income	267,997	247,142	287,375	361,044	418,265	500,042	561,328	589,627	518,597	529,328	533,547	300,639	301,025	486,457
Book-tax difference	-30,988	-20,748	-37,179	37,110	114,806	119,837	177,020	207,528	271,869	300,247	254,467	-45,634	79,649	424,089
M-1 explains	-30,988	-59,909	-106,505	-29,819	40,586	33,352	61,905	85,755	121,305	162,700	133,846	-174,128	-52,462	270,771
Estimated ICD adj	64,618	70,765	68,622	66,131	76,093	83,290	112,965	106,021	156,082	147,565	162,738	140,163	154,212	140,127
Tabulated ICD adj	0	0	0	0	0	0	0	0	0	137,200	130,263	124,991	148,114	135,697
Tax net income > 0	365,177	351,712	377,460	444,111	494,917	578,353	647,549	689,964	656,620	702,927	784,256	645,974	622,273	727,903
NOL deduction	26,500	27,906	26,500	32,249	34,184	42,394	39,284	44,352	35,137	48,480	59,601	44,506	49,757	54,300
Special deductions	11,850	11,969	12,867	14,310	13,169	22,842	19,808	23,277	23,163	29,822	26,040	20,714	15,864	12,654
Taxable income	327,437	313,845	342,168	398,144	449,967	516,357	589,374	624,309	600,210	629,065	702,012	588,011	561,917	662,875
Tax before credits	117,397	111,377	121,341	143,682	160,344	184,856	209,155	221,547	212,473	222,337	248,760	207,173	198,854	233,676
Foreign tax credit	24,745	21,023	21,474	22,852	25,345	30,235	40,082	41,700	36,500	36,905	47,964	40,510	41,938	49,767
Tax after credits	86,262	83,424	91,878	109,848	123,592	143,033	156,746	167,380	164,054	175,332	187,585	154,080	143,285	167,872
Pretax book income > 0	433,972	425,341	454,650	546,865	636,179	730,604	853,187	923,110	968,813	1,035,447	1,150,095	904,474	991,681	1,173,917
Book-tax difference > 0	175,928	179,394	183,506	223,637	240,340	261,116	321,607	355,742	453,673	496,755	563,462	451,821	584,731	640,388
M-1 explains > 0	139,505	120,816	136,599	173,056	185,600	202,689	232,443	270,526	339,389	396,103	467,366	366,181	479,220	544,300
Depreciation explains	0	34,784	42,721	48,369	52,702	51,748	54,522	63,480	77,702	81,803	86,372	112,451	155,284	131,137
Stock options explains	0	0	0	0	0	0	0	0	0	0	0	0	26,561	35,097
Schedule M-1 Detail														
[+L1] Book net income	214,360	198,752	193,022	299,973	393,003	458,074	551,620	592,455	588,748	592,373	534,295	103,654	206,766	675,567
[+L2] Book federal tax	80,781	69,138	57,173	98,181	140,068	161,805	186,727	204,700	201,718	237,202	253,719	151,350	173,908	234,978
[=] Pretax book income	295,142	267,890	250,195	398,154	533,071	619,879	738,348	797,156	790,466	829,575	788,014	255,005	380,674	910,546
[+L8] Income for book	0	178,215	192,522	241,354	218,986	249,019	282,395	340,561	416,985	476,568	595,416	467,940	555,089	590,905
[+L8] Deduction for tax	0	370,323	400,157	433,394	440,859	486,544	537,754	607,051	681,783	755,569	880,520	904,426	949,084	1,007,719
[+L3] Capital loss limit	0	-6,525	-6,116	-5,870	-6,200	-6,805	-6,325	-12,585	-13,363	-19,541	-16,326	-30,928	-65,136	-71,088
[+L4] Income for tax	0	-210,586	-208,586	-228,115	-219,265	-240,937	-284,552	-331,033	-348,847	-472,970	-478,363	-442,734	-454,734	-454,734
[+L5] Expense for book	0	-391,336	-484,481	-470,582	-393,373	-454,469	-467,367	-518,239	-615,252	-662,404	-852,788	-1,037,201	-1,048,708	-802,031
[=] M-1 explains	-30,988	-59,909	-106,505	-29,819	40,586	33,352	61,905	85,755	121,305	162,700	133,846	-174,128	-52,462	270,771
[+L8a] Depreciation for tax	0	106,732	121,470	127,840	135,689	147,147	148,075	168,756	195,531	202,141	203,676	263,599	304,676	289,620
[+L8a] Depreciation for book	0	-71,948	-78,749	-79,471	-82,987	-95,399	-93,553	-103,276	-117,830	-120,339	-133,675	-151,148	-149,392	-158,484
[+L7a] Depreciation explains	0	34,784	42,721	48,369	52,702	51,748	54,522	63,480	77,702	81,803	86,372	112,451	155,284	131,137
[+L7a] Tax-exempt interest	21,546	20,099	20,014	20,142	20,731	20,485	20,034	19,531	21,783	22,388	21,603	20,760	21,248	22,132
[+L8c] Stock options	0	0	0	0	0	0	0	0	0	0	0	0	26,561	35,097
[+L8c] Travel/entertainment	0	-1,718	-2,010	-2,059	-5,680	-5,619	-6,227	-6,566	-7,220	-7,455	-8,198	-7,481	-7,253	-12,895
[+L3] Capital loss limit	0	-6,525	-6,116	-5,870	-6,200	-6,805	-6,325	-12,585	-13,363	-19,541	-16,326	-30,928	-65,136	-71,088
M-1 detail explains [5 above]	21,546	46,639	54,609	60,582	61,133	59,809	62,004	63,860	78,901	77,195	86,372	104,802	130,703	104,339
M-1 other explains [balance]	-52,534	-106,548	-161,114	-90,401	-20,547	-26,458	-99	21,895	42,404	85,505	50,396	-268,930	-183,165	166,389
Pretax Book Income Only														
Returns [N2]	1,404	1,512	1,489	1,397	1,318	1,346	1,282	1,339	1,400	1,422	1,432	1,478	1,458	1,474
Assets	421,196	446,246	421,133	394,444	365,426	403,163	528,867	638,897	830,200	874,652	1,037,732	1,179,462	1,182,833	1,154,880
Receipts	136,527	135,144	121,149	123,303	85,504	93,568	134,483	162,636	201,801	225,503	248,101	270,657	235,755	198,020
Pretax book income	2,799	9,006	7,852	7,337	4,151	6,004	7,378	10,187	10,219	8,966	6,989	4,254	2,571	11,779
Net-income	2,670	3,623	3,982	5,025	3,424	4,907	6,037	6,132	8,626	8,621	8,332	6,990	7,358	9,573
BTD	129	5,383	3,870	2,313	728	1,098	1,341	4,055	1,593	345	-1,342	-2,736	-4,788	2,206
Tax After Credits	1,024	1,589	1,793	2,410	2,026	2,657	2,026	2,657	3,273	3,008	3,331	3,248	1,993	2,718
No Pretax Book Income														
Returns [N3]	392	386	334	319	329	321	361	385	425	429	471	480	428	420
Assets	148,138	109,238	54,961	143,550	146,980	106,362	260,171	152,046	405,871	364,501	339,672	486,811	319,178	257,376
Receipts	45,049	56,758	18,543	41,847	63,518	18,611	60,891	24,803	82,525	182,916	100,065	151,341	70,176	79,386
Net-income	-550	1,231	228	2,984	5,998	747	2,666	495	3,167	1,935	5,747	687	2,339	2,339
BTD	550	-1,231	-228	-2,984	-5,998	-747	-2,666	-495	-3,167	-1,935	-5,747	-687	-2,339	-2,339
Tax After Credits	125	349	69	730	588	227	781	220	803	3,197	680	2,104	702	698

Table 3. ICD Adjustment Required: All Corporations (Excluding F, S, RIC, and REIT) With Assets of \$10 Million or More

Dollar amounts in millions. Table values may not add and may differ from SOI Publication 16 values due to rounding.

Summary	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Returns (NI)	1,845	1,706	1,647	1,659	1,516	1,457	1,335	1,451	1,333	1,333	1,284	1,194	1,141	1,000
Total assets	4,574,526	4,430,581	4,591,137	5,281,177	5,664,100	5,321,909	6,277,836	8,236,197	8,652,226	11,143,376	13,680,997	14,326,198	15,704,660	15,041,004
Total receipts	1,929,575	1,831,758	1,924,403	2,003,054	2,088,626	1,966,866	2,266,975	2,453,235	2,553,234	3,312,601	4,154,112	3,797,227	4,045,988	3,982,062
Pretax book income	135,048	135,518	128,478	145,605	208,376	222,486	291,485	294,786	313,651	377,314	441,489	284,465	340,224	417,076
Tax net income	97,119	86,740	95,924	115,536	130,667	150,054	169,587	186,439	174,490	181,775	262,592	165,295	151,857	225,703
Book-tax difference	37,929	48,779	32,555	30,068	77,709	72,432	121,898	108,347	139,161	195,540	178,897	119,170	188,367	191,373
M-1 explains	-26,408	-26,295	-36,255	-38,423	-2,361	-12,009	3,792	-4,112	-17,103	46,265	49,621	-23,395	43,756	41,557
Estimated ICD adj	64,618	70,755	68,622	66,131	76,093	83,290	112,965	106,021	156,082	147,565	126,738	140,163	154,212	140,127
Tabulated ICD adj	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tax net income > 0	112,884	102,307	107,989	126,147	138,804	158,267	177,471	196,941	184,349	204,741	284,959	220,697	215,675	258,410
NOL deduction	5,211	5,388	5,079	4,855	4,083	5,209	6,820	9,703	5,762	9,039	15,049	9,039	13,731	14,290
Special deductions	4,078	3,667	3,899	3,781	4,600	11,900	8,872	8,432	7,638	8,096	9,204	8,171	6,232	4,096
Taxable income	104,747	94,094	99,863	118,030	130,624	141,752	164,931	180,012	171,413	189,736	262,367	207,311	199,200	240,403
Tax before credits	36,189	33,593	35,331	42,201	46,608	50,510	58,417	63,824	60,653	67,125	92,978	73,189	70,455	84,708
Foreign tax credit	10,947	10,221	11,098	13,497	13,621	15,259	20,424	19,907	16,476	21,309	31,194	31,194	22,105	28,246
Tax after credits	24,757	22,324	23,268	26,406	30,358	33,247	35,862	41,619	41,439	43,308	57,435	44,416	43,394	51,025
Pretax book income > 0	159,246	154,678	158,476	172,538	222,203	232,071	303,021	305,530	326,952	400,275	484,063	373,267	438,686	451,648
Book-tax difference > 0	71,302	75,031	72,231	81,542	101,358	96,057	144,145	137,701	167,325	227,804	247,962	200,728	276,520	230,103
M-1 explains > 0	33,692	27,866	33,580	39,164	49,587	44,819	59,714	63,717	60,387	138,184	157,656	117,799	172,713	143,718
Depreciation explains	0	8,529	14,574	13,963	18,163	17,177	9,956	12,622	24,761	26,309	26,015	34,486	45,034	35,472
Stock options explains	0	0	0	0	0	0	0	0	0	0	0	0	8,883	7,647
Schedule M-1 Detail														
[+L1] Book net income	108,662	114,913	110,692	124,669	169,443	182,129	241,084	238,661	253,045	305,214	360,828	238,522	285,948	344,668
[+L2] Book federal tax	26,386	20,605	17,787	20,935	38,934	40,356	50,400	56,125	60,806	72,100	80,661	45,943	54,276	72,408
[=] Pretax book income	135,048	135,518	128,479	145,605	208,376	222,486	291,485	294,786	313,651	377,314	441,489	284,465	340,224	417,076
[+L7] Income for book	0	69,625	61,222	77,158	75,876	68,965	91,993	100,039	90,742	162,874	219,536	181,787	221,296	223,530
[+L8] Deduction for tax	0	111,533	120,867	132,323	130,781	118,275	145,046	172,024	199,925	231,771	280,133	300,013	294,388	286,799
[+L3] Capital loss limit	0	-2,484	-3,020	-2,602	-1,388	-2,086	-1,140	-4,439	-4,630	-9,274	-5,022	-8,601	-16,179	-17,700
[+L4] Income for tax	0	-82,621	-82,838	-88,398	-88,770	-91,720	-111,284	-141,031	-150,219	-175,452	-212,996	-207,417	-190,451	-200,130
[+L5] Expense for book	0	-122,349	-132,485	-156,904	-118,860	-105,342	-120,823	-130,706	-152,921	-163,655	-232,030	-289,177	-265,298	-250,942
[=] M-1 explains	-26,408	-26,295	-36,255	-38,423	-2,361	-12,009	3,792	-4,112	-17,103	46,265	49,621	-23,395	43,756	41,557
[+L8a] Depreciation for tax	0	30,684	37,406	40,734	41,113	40,027	35,387	43,085	57,419	56,161	69,414	84,438	90,499	88,341
[+L5a] Depreciation for book	0	-22,155	-22,832	-26,772	-22,950	-22,849	-25,431	-30,463	-32,657	-29,853	-43,399	-49,952	-45,465	-52,869
[=] Depreciation explains	0	8,529	14,574	13,963	18,163	17,177	9,956	12,622	24,761	26,309	26,015	34,486	45,034	35,472
[+L7a] Tax-exempt interest	4,060	3,781	4,381	5,554	5,441	4,177	5,740	5,647	5,381	8,045	8,608	9,087	9,738	7,771
[+L8c] Stock options	0	0	0	0	0	0	0	0	0	0	0	0	8,883	7,647
[+L5c] Travel-entertainment	0	-438	-559	-522	-1,973	-1,216	-1,478	-1,615	-1,642	-2,069	-2,390	-2,148	-2,054	-1,767
[+L3] Capital loss limit	0	-2,484	-3,020	-2,602	-1,388	-2,086	-1,140	-4,439	-4,630	-9,274	-5,022	-8,601	-16,179	-17,700
M-1 detail explains [5 above]	4,060	9,388	15,376	16,392	20,243	18,053	13,078	12,215	23,871	23,010	27,211	32,824	45,423	31,423
M-1 other explains [balance]	-30,468	-35,683	-51,631	-54,816	-22,604	-30,062	-9,286	-16,328	-40,974	23,255	22,410	-56,219	-1,667	-10,134
Data suppressed for "Pretax Book Income Only" and "No Pretax Book Income."														

Table 4. ICD Adjustment Not Required: All Corporations (Excluding F, S, RIC, and REIT) With Assets of \$10 Million or More

Dollar amounts in millions. Table values may not add and may differ from SOI Publication 16 values due to rounding.

Summary	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Returns [N1]	39,780	39,456	39,168	40,003	41,562	42,592	43,862	43,873	43,733	44,643	45,773	44,145	42,619	42,155
Total assets	10,787,694	11,459,273	11,877,773	12,483,583	13,376,011	15,559,500	16,358,323	17,290,815	20,000,256	20,167,738	22,324,054	23,898,161	23,901,170	26,367,037
Total receipts	5,675,689	5,843,240	5,959,844	6,142,195	6,649,899	7,621,083	7,987,765	8,462,862	8,865,239	9,341,416	9,801,857	9,961,051	9,240,385	9,927,066
Pretax book income	160,094	132,372	121,717	252,550	324,695	397,393	446,863	502,370	476,816	452,261	346,525	-29,461	40,450	493,469
Tax net income	170,878	160,403	191,451	245,508	287,988	349,988	391,741	403,188	344,108	347,553	270,955	135,344	149,168	260,754
Book-tax difference	-10,784	-28,031	-69,734	7,042	37,097	47,405	55,122	99,182	132,708	104,708	75,570	-164,804	-108,718	232,716
M-1 explains	-4,579	-33,614	-70,250	8,604	42,948	45,361	58,113	89,867	138,408	116,435	84,225	-150,732	-96,217	229,214
Estimated ICD adj	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tabulated ICD adj	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tax net income > 0	252,294	249,405	269,471	317,964	356,113	420,085	470,079	493,023	472,271	498,186	499,296	425,278	406,598	469,493
NOI deduction	21,290	22,519	21,370	27,394	30,101	37,185	32,464	34,644	29,375	39,086	44,552	35,467	36,026	40,010
Special deductions	7,772	8,303	8,969	10,529	8,669	10,942	13,936	14,844	15,525	21,727	16,836	12,543	9,632	8,557
Taxable income	222,691	219,752	242,306	280,115	319,342	374,605	424,443	444,297	428,797	439,328	439,645	380,700	362,717	422,473
Tax before credits	81,208	77,784	86,010	101,481	113,736	134,346	150,738	157,723	151,820	155,212	155,781	133,984	128,399	148,968
Foreign tax credit	13,798	10,801	10,376	9,355	11,724	14,876	19,658	21,793	20,024	15,596	16,770	15,804	19,833	21,521
Tax after credits	61,505	61,100	68,710	83,442	93,234	109,786	120,885	125,760	122,615	132,024	130,150	109,664	99,891	116,847
Pretax book income > 0	274,726	270,664	296,174	374,327	413,976	498,533	550,166	641,861	648,174	635,174	666,032	531,208	552,995	722,269
Book-tax difference > 0	104,626	104,363	111,275	142,095	138,982	165,059	177,462	218,041	286,951	268,951	315,499	251,093	308,210	410,284
M-1 explains > 0	105,814	92,930	103,019	133,892	136,013	157,870	172,728	206,809	279,002	257,919	309,710	248,383	306,507	400,581
Depreciation explains	0	26,255	28,147	34,406	34,539	34,571	44,566	50,858	52,940	55,494	60,357	77,965	110,250	95,665
Stock options explains	0	0	0	0	0	0	0	0	0	0	0	0	17,678	27,450
Schedule M-1 Detail														
[+L1] Book net income	105,698	83,839	82,330	175,304	223,560	275,945	310,536	353,794	335,703	287,158	173,467	-134,868	-79,182	330,899
[+L2] Book federal tax	54,395	48,533	39,386	77,246	101,134	121,448	136,327	148,576	141,112	165,103	173,058	105,407	119,632	162,570
[=] Pretax book income	160,094	132,372	121,717	252,550	324,695	397,393	446,863	502,370	476,816	452,261	346,525	-29,461	40,450	493,469
[+L7] Income for book	0	108,590	131,300	164,196	143,109	180,154	190,402	240,522	326,242	313,694	375,880	286,153	333,792	367,375
[+L8] Deduction for tax	0	258,790	279,290	301,071	310,078	368,269	392,708	435,027	481,859	523,797	600,387	604,413	654,697	720,921
[+L3] Capital loss limit	0	-4,041	-3,096	-3,268	-5,232	-4,720	-5,185	-8,146	-8,733	-10,267	-11,304	-22,327	-48,957	-53,389
[+L4] Income for tax	0	-127,965	-125,748	-139,717	-130,495	-149,217	-173,268	-190,003	-198,628	-212,039	-259,979	-270,947	-252,339	-254,605
[+L5] Expense for book	0	-268,987	-351,996	-313,678	-274,513	-349,126	-346,544	-387,533	-462,332	-498,750	-620,758	-748,024	-783,410	-551,089
[=] [M-1] explains	-4,579	-33,614	-70,250	8,604	42,948	45,361	58,113	89,867	138,408	116,435	84,225	-150,732	-96,217	229,214
[+L8a] Depreciation for tax	0	76,048	84,064	87,105	94,576	107,121	112,688	125,671	138,112	145,980	150,633	179,161	214,177	201,279
[+L5a] Depreciation for book	0	-49,793	-55,917	-52,699	-60,037	-72,550	-68,122	-74,813	-85,172	-90,486	-90,276	-101,195	-103,927	-105,615
[=] Depreciation explains	0	26,255	28,147	34,406	34,539	34,571	44,566	50,858	52,940	55,494	60,357	77,965	110,250	95,665
[+L7a] Tax-exempt interest	17,486	16,318	15,633	14,588	15,291	16,308	14,294	13,883	16,402	14,344	12,995	11,674	11,510	14,361
[+L8c] Stock options	0	0	0	0	0	0	0	0	0	0	0	0	17,678	27,450
[+L5c] Travel-entertainment	0	-1,281	-1,451	-1,538	-3,708	-4,403	-4,749	-4,951	-5,579	-5,386	-5,808	-5,334	-5,200	-11,127
[+L3] Capital loss limit	0	-4,041	-3,096	-3,268	-5,232	-4,720	-5,185	-8,146	-8,733	-10,267	-11,304	-22,327	-48,957	-53,389
M-1 detail explains [5 above]	17,486	37,251	39,234	44,189	40,890	41,757	48,926	51,645	55,030	54,185	56,239	61,978	85,281	72,959
M-1 other explains [balance]	-22,065	-70,865	-109,483	-35,585	2,057	3,604	9,187	36,222	83,378	62,250	27,986	-212,711	-181,498	156,255

Data suppressed for "Pretax Book Income Only" and "No Pretax Book Income."

net income. We present both the SOI tabulated amount for the U.S. intercompany dividend (ICD) adjustment (available from SOI for 1999 on) and our estimate of that adjustment for all years 1990-2003 (more about this later). We calculate an amount we term “M-1 Explains” which is the net amount of book-tax difference reported by the taxpayer on Schedule M-1.¹¹ We also calculate a net error amount indicating the amount of the book-tax difference not included in either M-1 Explains or our estimate of the ICD adjustment.

In the second section of the first panel of each table (“Schedule M-1 Detail”), we present the aggregate amounts from the Schedule M-1 line items and certain calculated amounts. The sign is shown consistently in terms of the effect on a positive book-tax difference. A positive amount increases the book-tax difference; a negative amount decreases the book-tax difference. Consistent with the literature since Talisman (2000), we treat pretax book greater than tax net income as a positive book-tax difference.

The second panel on each table (unless suppressed) presents aggregate data for those corporations in the first panel that, for some reason, reported only pretax book income, that is, no other data appeared in the body of Schedule M-1.¹²

The third panel on each table (unless suppressed) presents aggregate data for those corporations in the first panel that, for some reason, do not even report amounts for Schedule M-1 line 1 and line 2.¹³

Schedule M-1 data for 1990 are not as complete as for other years. SOI only tabulated: line 1, net income (loss) per books; line 2, Federal income tax per books; line 6, total of lines 1 through 5; line 9, total of lines 7 and 8; and line 10, the reconciliation amount corresponding to unedited tax net income (tax net income before the U.S. intercompany dividend (ICD) adjustment).¹⁴

► **Book-Tax Difference Data 1990-2003**¹⁵

For comparison with Table 12 in Publication 16, *Statistics of Income, Corporation Income Tax Returns*, and with Plesko (2002) and Plesko-Shumofsky (2005),

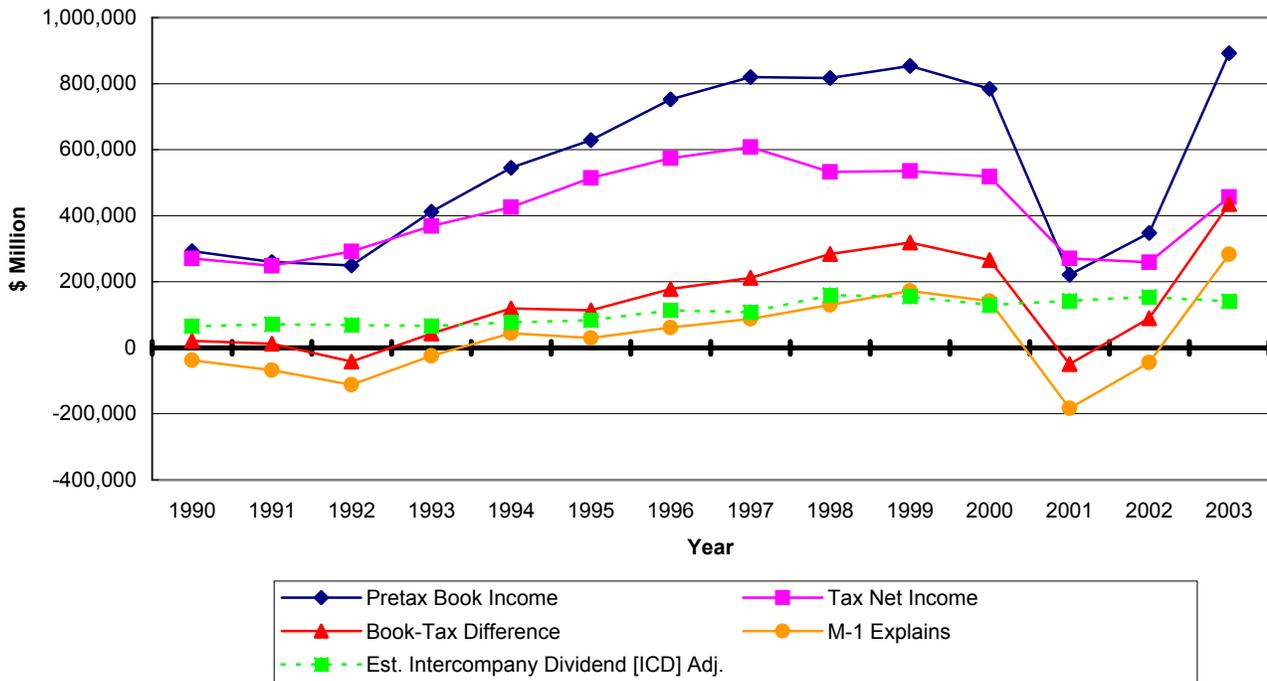
we first present, in this section of the paper, aggregate net data for all corporations normally subject to the U.S. Federal corporate income tax. We then present, in the next section of the paper, the aggregate net data for domestic corporations with assets of \$10 million or more, the corporations that would have been subject to Schedule M-3 if the 2004-2006 requirements had been effective for the earlier years.

Figure 1 based on Table 1 presents aggregate net pretax book income and aggregate tax net income for all corporations for 1990-2003. It also presents the calculated book-tax differences and an amount we term M-1 Explains. Finally, it presents an amount we term “estimated intercompany dividend (ICD) adjustment.”

- Pretax book income is the sum of Schedule M-1, line 1, Net income (loss) per books, and Schedule M-1, line 2, Federal income tax per books.
- Tax net income is Form 1120 line 28 taxable income before net operating loss deduction (line 29a) and special deductions (dividends received deductions) (line 29b).
- Book tax difference is pretax book income minus tax net income. This definition has been in general use since the Talisman (2000) Senate testimony on tax shelters and the possible effect of tax shelters on the corporate tax base.
- M-1 Explains is our term for the book-tax difference actually reported by the taxpayer on Schedule M-1 as originally filed.¹⁶ M-1 Explains and book-tax difference calculated using the Talisman (2000) approach differ by the amount of the U.S. intercompany dividend (ICD) adjustment to tax net income.¹⁷

Some taxpayers improperly include U.S. intercompany dividends (ICD) in tax net income on Form 1120, page 1, line 28, the reconciliation target for Schedule M-1.¹⁸ The taxpayer then removes the same amount as a 100-percent dividends-received deduction on line 29b so that it does not increase final income subject to tax on line 30.

Figure 1. Pretax Book Income, Tax Net Income, Book-Tax Difference, M-1 Explains, and Estimated Intercompany Dividend (ICD) Adjustment For All Corporations (Excluding S, RIC, REIT)



ICD should be eliminated in determining tax net income. SOI removes all ICD amounts that it identifies in tax net income. Taxpayers who include ICD in tax net income must also include it somewhere in Schedule M-1. SOI does not know where in Schedule M-1 the ICD is in general, and, therefore, SOI does not remove ICD from the body of Schedule M-1 but rather, starting in 1999, from Schedule M-1, line 10.¹⁹ The result is that M-1 Explains and book-tax difference as defined by Talisman (2000) differ by the amount of the ICD adjustment to tax net income.

SOI began tabulating the ICD adjustment in 1999, although it made the adjustment without tabulation as a separate file variable starting in 1990. We estimate the ICD adjustment for all years studied: 1990-2003. We estimate the ICD adjustment as unedited Schedule M-1, line 10 minus edited Form 1120, page 1, line 28 (if it is a positive difference) for corporations filing a consolidated return.²⁰ For 1999-2003, we present our estimate and the tabulated ICD. For consistency across years,

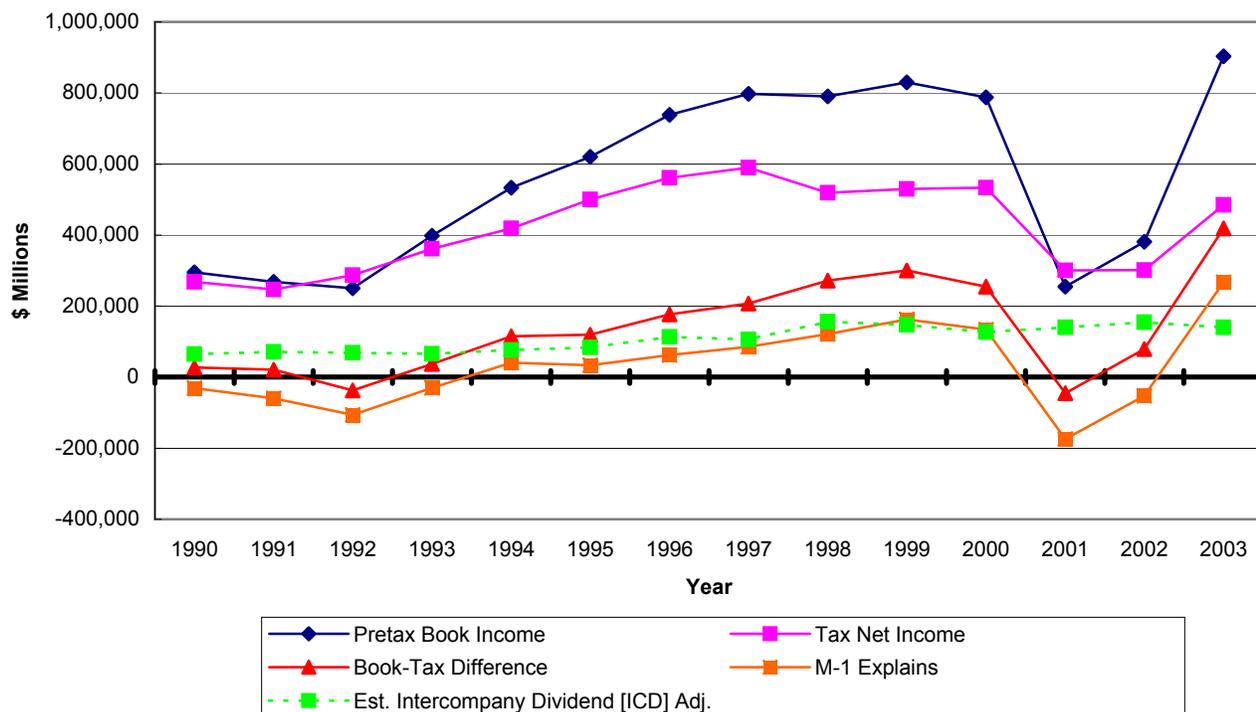
our discussion uses our estimate of the ICD adjustment unless otherwise stated.

► **Assets of \$10 Million or More** ²¹

In this and later sections of the paper, we present the data for domestic corporations with assets of \$10 million or more, the corporations that would have been subject to Schedule M-3 if the 2004-2006 requirements had been effective for the earlier years.

Figure 1 is for all corporations (excluding S, RIC, and REIT). Figure 2 based on Table 2 is for domestic corporations with total assets of \$10 million or more (excluding S, RIC, REIT, and F) and presents a picture of aggregate net pretax book income, tax net income, book-tax difference, M-1 Explains, and ICD adjustment similar to that in Figure 1. This is because most of the aggregate net Schedule M-1 line item amounts (including most of the aggregate net pretax book income, which is the sum of Schedule M-1, line 1 plus line 2), aggregate

Figure 2. Pretax Book Income, Tax Net Income, Book-Tax Difference, M-1 Explains, and Estimated Intercompany Dividend (ICD) Adjustment For U.S. Corporations With Assets >=\$10 Million (Excluding S, RIC, REIT, F)



net tax net income, and aggregate ICD adjustment of all corporations are in fact reported by those domestic corporations with \$10 million or more in assets.

► **What Drives Schedule M-1 Swings?**²²

Schedule M-1 offers detail breakout for depreciation, tax-exempt interest, stock options (starting 2002), travel and entertainment limitations, and capital loss limitation. “M-1 Detail Explains” is our term for the net effect of these items on M-1 Explains. “M-1 Other Explains” is our term for the balance of M-1 Explains not included in M-1 Detail Explains.

Figure 3 presents M-1 Explains, M-1 Detail Explains, M-1 Other Explains, and depreciation explains for corporations with total assets of \$10 million or more. M-1 Detail Explains is essentially depreciation. The other detail items tend to net out. The swings in M-1

Explains are driven by the swings in M-1 Other Explains, that is, by the amounts without detail breakouts. We will not know what is behind M-1 Other Explains until we have the standardized transparent structure of Schedule M-3.²³

► **Issues in Interpreting Schedule M-1 Data**

Figure 4 based on Tables 3 and 4 shows that, for 1993-2000, among corporations with total assets of \$10 million or more, those requiring the U.S. intercompany dividend (ICD) adjustment (to be discussed in Figure 5 under two alternative assumptions labeled Case 1 and Case 2) reported lower net aggregate M-1 Explains than those that did not require the ICD adjustment (to be discussed in Figure 5 as reference Case 3). In particular, the corporations requiring the ICD adjustment appeared to have an aggregate net M-1 Explains of approximately

Figure 3. Schedule M-1 Explains, Schedule M-1 Detail Explains, Schedule M-1 Other Explains, and Depreciation Explains For U.S. Corporations With Asset >=\$10 Million

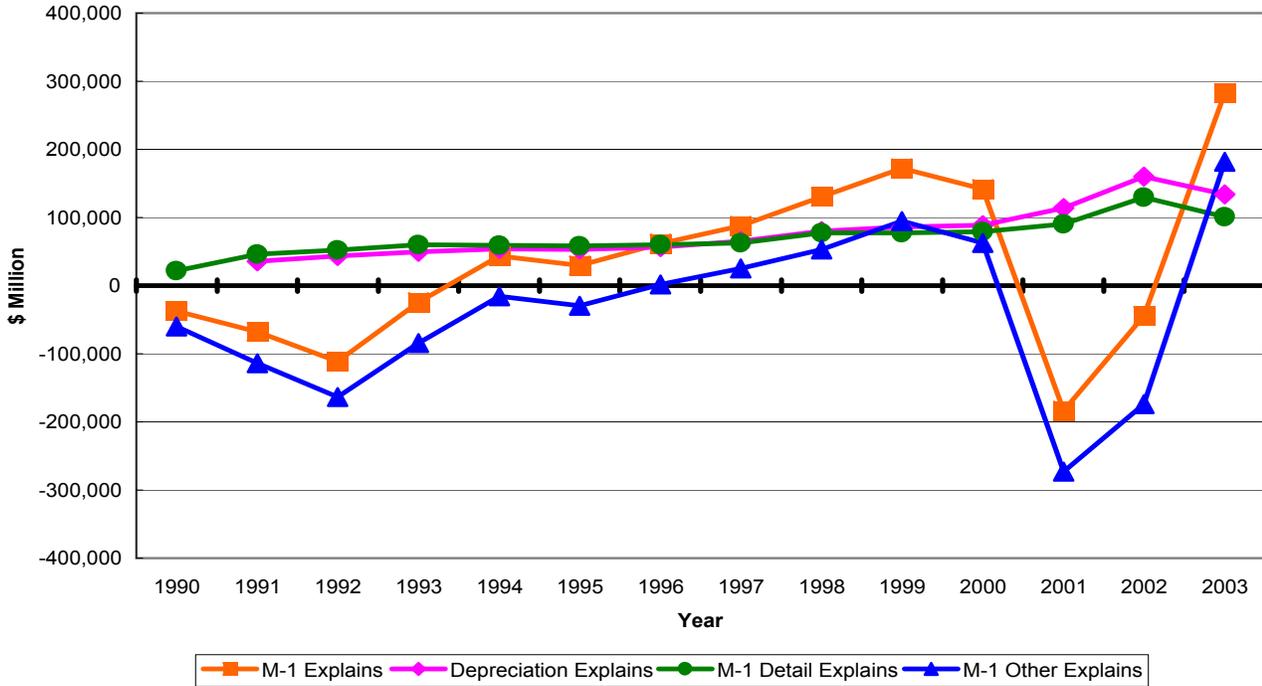
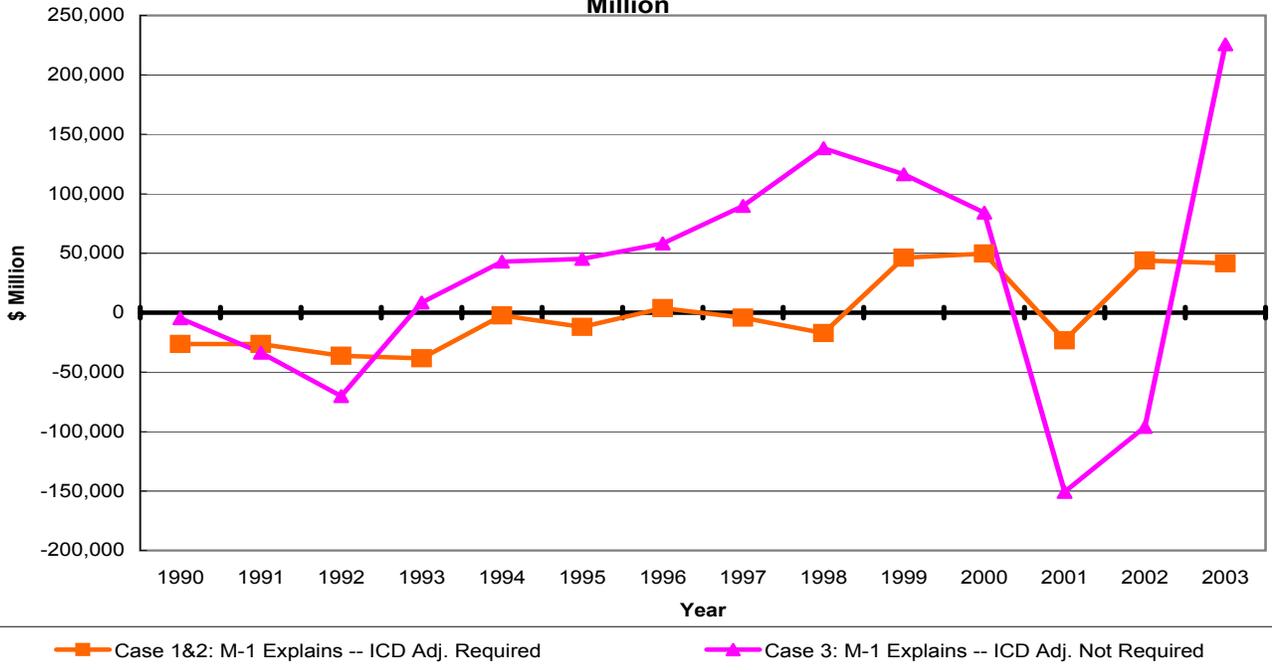


Figure 4. M-1 Explains For Corporations Requiring The Intercompany Dividend (ICD) Adjustment (Cases 1&2) Versus M-1 Explains For Corporations Not Requiring The ICD Adjustment (Case 3) For U.S. Corporations With Assets >=\$10 Million



zero during the boom years of 1994-1998. Corporations not requiring the ICD adjustment had a large aggregate net positive M-1 Explains those years.

► **We Develop “What If” Cases:**

- Case 1: ICD adjustment present, and we back it out of Schedule M-1, line 1.
- Case 2: ICD adjustment present, and we back it out of Schedule M-1, line 4. Here, line 4 is simply a surrogate for any line in the body of Schedule M-1.
- Case 3: ICD adjustment not present. Case 3 is our reference for analysis for Case 1, M-1, line 1 versus Case 2, M-1, line 4. Case 3 controls for changes in the economy across years.

Effect of Case 1: If the ICD adjustment should be removed from Schedule M-1, line 1, pretax book income and book-tax difference will be reduced, and book-tax difference will equal M-1 Explains as observed.

Effect of Case 2: If the ICD adjustment should be removed from the body of Schedule M-1, say, Schedule M-1, line 4, income for tax not for book, M-1 Explains will be increased, and M-1 Explains will equal book-tax difference as calculated using the Talisman (2000) approach that we and others generally follow.

Effect of firm size on our analysis: The approximately 1,100 corporations in 2002 with total assets of \$10 million or more requiring the ICD adjustment are about 25 times larger in mean assets than the approximately 42,000 corporations that year with total assets of \$10 million or more not requiring the ICD adjustment (Cases 1 and 2, \$13.8 billion; Case 3, \$561 million). In the following analysis, we control for the possible effects of size differences by calculating aggregate M-1 Explains as a percentage of aggregate total receipts for the group requiring the ICD adjustment (Cases 1 and 2) and for the group not requiring the ICD adjustment (Case 3).

In Figure 5 based on Tables 3 and 4, the top two lines lie along each other and represent our Case 1 and Case

2 calculated book-tax difference as a percent of total receipts for corporations requiring the ICD adjustment and Case 2 restated M-1 Explains as a percentage of total receipts after the ICD adjustment is removed from Schedule M-3, line 4. In essence, we move Case 2 M-1 Explains up to equal book-tax difference.

In Figure 5, the bottom two lines lie along each other and represent our Case 1 and Case 2 observed M-1 Explains as a percent of total receipts for those requiring the ICD adjustment and the Case 1 recalculated book-tax difference after the ICD adjustment is removed from Schedule M-1, line 1. In essence, we move Case 1 book-tax difference down to equal M-1 Explains.

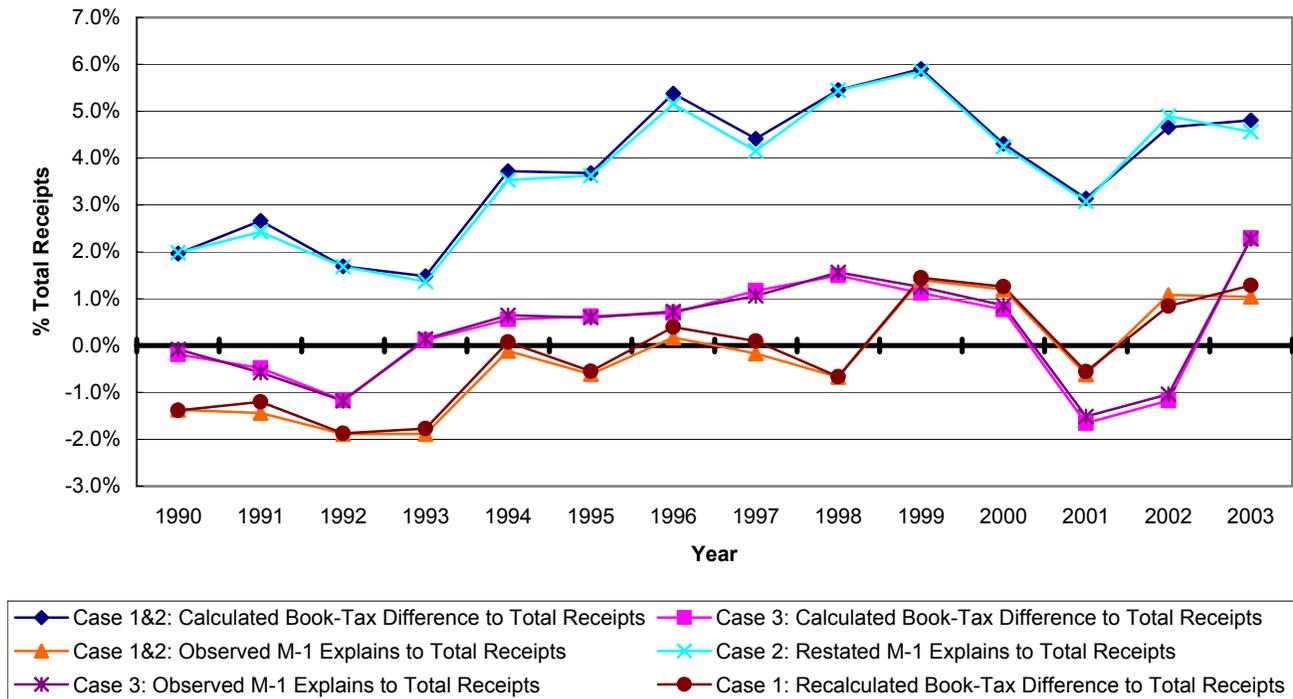
In Figure 5, the middle two lines lie along each other and represent our Case 3 calculated book-tax difference and our Case 3 observed M-1 Explains, each as a percentage of total receipts, for corporations not requiring the ICD adjustment.

In Figure 5, the middle two lines are our reference. If the lower two lines are plausible for corporations requiring the ICD adjustment, then we remove the ICD adjustment from Schedule M-1, line 1, and book-tax difference, effectively recalculating book-tax difference to agree with what taxpayers declared in M-1 Explains. We question whether large corporations would have essential no book-tax difference during the boom years of the 1990's at a time when corporations not requiring the ICD adjustment had a large aggregate net positive book-tax difference and M-1 Explains.²⁴

If the lower two lines are not plausible, or if the upper two lines are more plausible, then we remove the ICD adjustment from Schedule M-1, line 4, accept book-tax difference as calculated under the Talisman (2000) approach, and restate M-1 Explains to agree with our calculated book-tax difference.

The question about where we should remove the ICD adjustment in Schedule M-1 is important. If the ICD adjustment should be removed from Schedule M-1, line 1, book-tax difference as generally calculated involves an overstatement. The worry has been that the ICD adjustment often seemed to be about half of the book-tax gap for the boom years of the 1990's. But we

Figure 5. Book-Tax Difference and M-1 Explains To Total Receipts for U.S. Corporations With Assets >=\$10 Million Requiring ICD Adjustment (Case 1 Assumes in M-1 Line 1, Case 2 Assumes in M-1 Line 4) and Not Requiring ICD Adjustment (Case 3 Reference Case)



show it is often essentially a question of the existence of any book-tax gap for corporations requiring the ICD adjustment.

Figure 6 based on Tables 3 and 4 indicated that the corporations requiring the ICD adjustment generally have more aggregate net positive M-1 Detail Explains (essentially depreciation) as a percentage of total receipts than corporations not requiring the adjustment. We suggest it is not plausible that these corporations would have no other net aggregate book-tax difference.

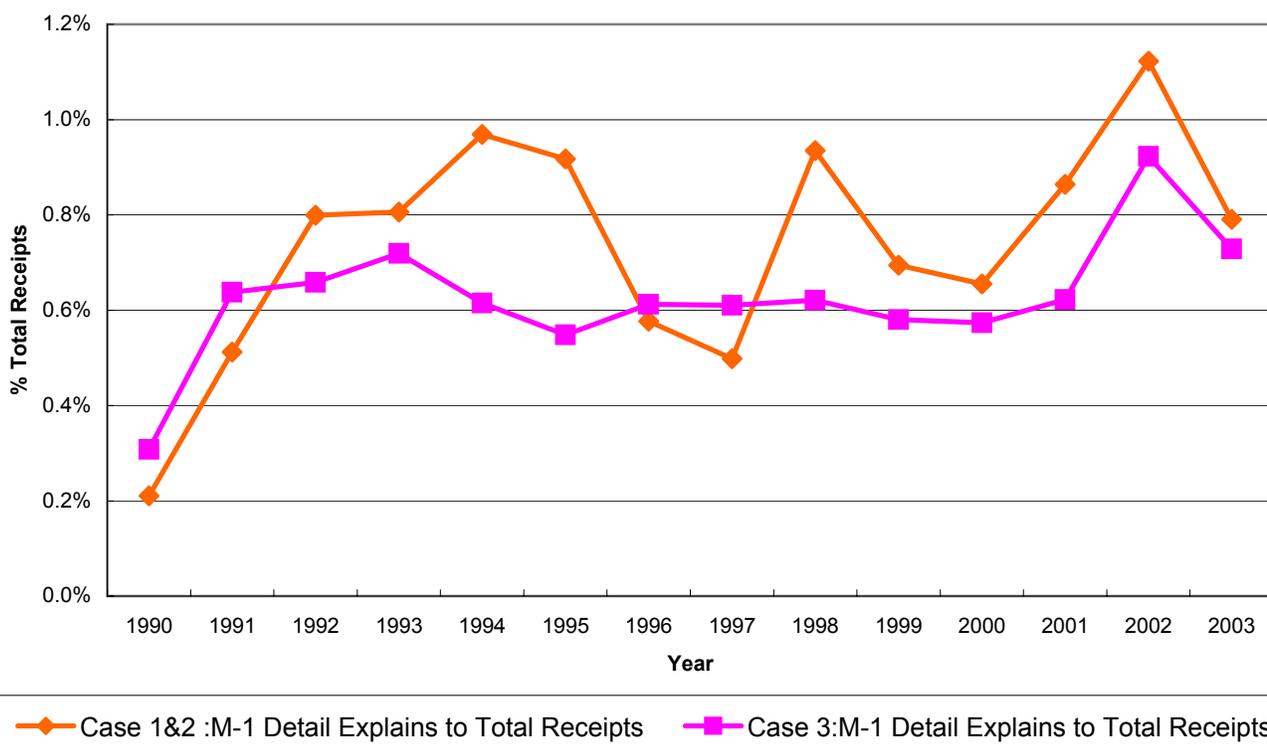
► Evidence From Large Corporations

We also supplemented our analytical research on the ICD adjustment discussed in the prior section with a limited search of large corporation tax returns by SOI. We wished to determine if there was tax return evidence indicating whether Schedule M-1, line 1 or line 4, was

generally used by large corporate taxpayers as the line for inclusion of the matching entry within Schedule M-1 for U.S. intercompany dividends (ICD) improperly included on Form 1120, page 1, line 28 (tax net income), and line 29b (dividends received deduction). In particular, we wished to determine if the relative size of the ICD adjustment compared to the total amount on Schedule M-1, line 4, might function as a flag as to the location of the ICD item within Schedule M-1.²⁵

We first identified all returns for 2003 that involved an ICD adjustment of at least \$1 billion. We then selected for examination five of the returns with an ICD adjustment greater than the total amount on Schedule M-1, line 4, and five of the returns with an ICD adjustment less than the total amount on line 4. One coauthor then searched the supporting detail for these 10 returns for Form 1120, Schedule C (Dividends and Special Deduction) and Schedule M-1, line 4, to identify a caption indi-

Figure 6. M-1 Detail Explains To Total Receipts For Case 1 And Case 2 (ICD Adjustment Required) With Case 3 (No ICD Adjustment) As Reference For U.S. Corporations With Assets >=\$10 Million



cating U.S. dividends included on Form 1120, Schedule C, and, therefore, on Form 1120, page 1, line 28, but not included in book income and an amount similar to the amount of the ICD adjustment.

Note that these returns are each thousands of pages. Searching for a caption and amount in the supporting detail is time-consuming and averaged an hour each even though the coauthor doing the search is very familiar with working with the supporting detail for Form 1120, Schedule C, and Schedule M-1. In the case of all five returns with an ICD adjustment less than the total amount on Schedule M-1, line 4, it was possible to identify an appropriate caption and approximate amount in the supporting detail for line 4. In the case of the five returns with an ICD adjustment greater than the total amount on Schedule M-1, line 4, the pattern was less clear with some support found for the ICD amount being included

on Schedule M-1, line 1, some for line 4, and some totally unclear.

We realize a search on 10 returns out of a much larger number does not prove that the pattern of captions and amounts we found would be found on the returns that were not searched. Further, our search does not prove what would be found if the IRS were to undertake a larger audit of large corporation Schedule M-1 detail. An IRS audit is unlikely because the better-structured Schedule M-3 is replacing the poorer-structured Schedule M-1 for larger corporate taxpayers. We do believe that our search on the 10 returns searched indicates that line 4 of Schedule M-1 is at least a likely location for the matching entry within Schedule M-1 for U.S. intercompany dividends (ICD) improperly included on Form 1120, page 1, line 28 (tax net income), and line 29b (dividends received deduction). We also know from our search that some

corporations do include the ICD amount on Schedule M-1, line 1. If a taxpayer includes the matching ICD amount on line 4 of Schedule M-1, the taxpayer will, either intentionally or innocently, minimize the total book-tax difference reported on Schedule M-1. If the taxpayer includes the matching ICD amount on Schedule M-1, line 1, use of the Talisman (2000) approach will inflate the measure of the taxpayer's book-tax difference by the amount of the ICD adjustment.

We believe that, on balance and given the uncertainties associated with Schedule M-1 data, the Talisman (2000) approach for calculating book-tax differences is the appropriate approach when the goal is the assessment of aggregate compliance risk in the population.

► **Summary and Conclusion**

For most large corporations, the new Schedule M-3 book-tax reconciliation replaces the 4-decade-old Schedule M-1, effective December 2004. The goal of this paper has been: (1) to present Schedule M-1 data and other selected tax return data for the immediately preceding 14-year period, 1990-2003; and (2) to discuss tax policy data interpretation issues related to U.S. intercompany dividends (ICD) improperly included on corporate tax returns by some large taxpayers.

- The method of calculating book-tax differences in general use since Talisman (2000) inflates the reported book-tax gap for the 1990's for those corporations requiring the ICD adjustment that included the matching ICD amount in Schedule M-1, line 1.
- On the other hand, corporations that included the matching ICD amount within the body of Schedule M-1, say on line 4, minimized the total book-tax difference reported on Schedule M-1.
- The authors are aware that some large taxpayers in fact used Schedule M-1, line 1, and some used line 4 for the matching amount to balance the ICD amount improperly included on Form 1120, page 1.
- In light of the ICD interpretation uncertainties, the authors recommend the Talisman (2000) approach

to measuring the book-tax gap of the 1990's for purposes of assessing compliance risk.

- Those issues will likely remain unresolved until Schedule M-3 data replace Schedule M-1 data.

► **Acknowledgment**

We thank Lillian Mills for her many conversations and suggestions as this paper developed. We also thank the participants in the October 5, 2005, OTA Brownbag Seminar, the October 28 SOI Advisory Panel, the November 4 University of North Texas Accounting Workshop, and the November 17 National Tax Association Annual Meeting Concurrent Session on Corporate Book-Tax Differences and Tax Avoidance. Further, we thank each of the following for detailed comments: David Brazell, Curtis Carlson, Geraldine Gerardi, Henry Louie, John McClelland, Susan Nelson, George Plesko, Linden Smith, and Bill Wilson. All errors are ours. Lastly, but not least, we thank Jonathan Mable for his PowerPoint presentation and Erin Sullivan for her assistance.

► **Endnotes**

- * Published on December 19, 2005, in *Tax Notes*, pages 1579-1599. Reprinted with permission of Tax Analysts.
- 1 Our table values may not add and may differ from official Publication 16 *Statistics of Income (SOI), Corporation Income Tax Returns*. values due to rounding. The SOI corporate data file for year t includes all tax years ending between July of Calendar Year and June of Calendar Year $t+1$.
- 2 Corporations normally subject to the U.S. Federal income tax include U.S. corporations filing Form 1120 (no asset limitation) or Form 1120-A (assets of \$500,000 or less), U.S. insurance companies filing Form 1120-L or Form 1120-PC, and foreign corporations with effectively connected U.S. income filing Form 1120-F. Corporations not normally subject to the U.S. Federal income tax include corporations filing Form 1120-S (Subchapter S corporations), Form 1120-REIT (Real Estate Investment Trusts), and Form 1120-RIC

(Regulated Investment Companies) that normally report their incomes proportionately to their owners for taxation imposed on the owners rather than the corporation.

3 See U.S. Department of the Treasury (1999) and Talisman (2000). See also Mills (1998) cited by Treasury (1999, page 32, note 118): “Mills finds evidence that the IRS is more likely to assert deficiencies on firms with large book-tax disparities, indicating that such disparities are correlated with aggressive tax planning.”

4 See Mills and Plesko (2003) for the proposed redesign of Schedule M-1. For discussions of problems in interpreting Schedule M-1 book-tax reconciliation data and problems with the related Schedule L book balance sheet data, see Boynton, Dobbins, DeFilippes, and Cooper (2002), Mills, Newberry, and Trautman (2002), and Boynton, DeFilippes, Lisowsky, and Mills (2005). For discussions of the problems in reconciling financial accounting income and tax income, see McGill and Outslay (2002), Hanlon (2003), McGill and Outslay (2004), Plesko (2004), and Hanlon and Shevlin (2005).

5 For a discussion of the development of Schedule M-3, see Boynton and Mills (2004).

6 Schedule M-1 will continue to apply to domestic corporations with assets of \$250 thousand to \$10 million of total assets or of less than \$250 thousand in total assets but total receipts of \$250 thousand or more. Schedule M-1 will also continue to apply to foreign corporations filing Form 1120-F.

7 U.S. Department of the Treasury; press release dated January 28, 2004, “Treasury and IRS Propose New Tax Form for Corporate Tax Returns.”

“The new Schedule M-3 would expand the current Schedule M-1, which has not been updated in several decades.

“The proposed Schedule M-3 will make differences between financial accounting net income

and taxable income more transparent. This will help agents determine from the return whether the return should be audited and identify the differences that matter most in the audit of the return. We see benefits to taxpayers and the IRS from the new Schedule: a reduction in unnecessary audits and a swifter focus on those differences that are more likely to arise when taxpayers take aggressive positions or engage in aggressive transactions. In addition, the increased transparency will have a deterrent effect,” stated Treasury Assistant Secretary for Tax Policy Pam Olson.

“The new Schedule will let the IRS sharpen and improve monitoring of corporate compliance,” said IRS Commissioner Mark W. Everson. “Our objective is to identify and resolve potential audit issues promptly. This information will help us do so.”

“These changes will enable us to focus our compliance resources on returns and issues that need to be examined and avoid those that do not,” said Deborah M. Nolan, IRS Large and Mid-Size Business Division Commissioner. “Increasing the transparency of corporate tax returns is critical to our objectives to provide certainty to taxpayers sooner and to improve overall compliance.”

8 Our table values may not add and may differ from official SOI Publication 16 values due to rounding.

9 Our table values may not add and may differ from official SOI Publication 16 values due to rounding.

10 Our Table 1 and SOI Publication 16 Table 12 include data from foreign corporations with effectively connected U.S. income required to file Form 1120-F. Our Tables 2-4 include only domestic corporations with \$10 million or more in assets and exclude data from foreign corporations filing Form 1120-F. Corporations filing Form 1120-F are not subject to Schedule M-3 and will continue to complete Schedule M-1.

11 We calculate “M-1 Explains,” the net book-tax difference reported on Schedule M-1, as (line 7

- plus line 8 minus the sum of lines 3, 4, and 5). This is the amount that must be subtracted from pretax book income, the sum of lines 1 and 2, to obtain line 10, the reconciliation amount corresponding to unedited tax net income, that is, tax net income before any U.S. intercompany dividend adjustment. See below for a discussion of the ICD adjustment.
- 12 This is the normal result for one group of corporations, namely, life insurance companies. Form 1120-L does not have a Schedule M-1. Rather the companies attach a financial statement (Annual Statement) prepared according to statutory accounting principles prescribed by the National Association of Insurance Commissioners. The companies also attach a reconciliation of taxable income with the income in the Annual Statement. There is not a fixed form for the reconciliation. SOI creates a dummy Schedule M-1 for life insurance companies with only line 1 and line 2 amounts derived from the Annual Statement.
- 13 Corporations with total assets of less than \$250 thousand and total receipts of less than \$250 thousand are no longer required to complete Schedule M-1 starting with 2002.
- 14 We infer the 1990 amount of -M-1 Explains,--the net book-tax difference reported by the taxpayer on Schedule M-1, as {M-1 line 9 minus line 6 plus line 1 plus line 2} which equals {[line 7 + line 8] - [line 1 + line 2 + line 3 + line 4 + line 5] + [line 1 + line 2]} which equals {[line 7 + line 8] - [line 3 + line 4 + line 5]} which is our defined -M-1 Explains as stated in footnote 11. See below for a discussion of the ICD adjustment.
- 15 Our table values may not add and may differ from official SOI Publication 16 values due to rounding.
- 16 We calculate M-1 Explains, the net book-tax difference reported on Schedule M-1, as [line 7 plus line 8 minus the sum of lines 3, 4, and 5]. This is the amount that must be subtracted from pretax book income, the sum of lines 1 and 2, to obtain line 10, the reconciliation amount corresponding
- to unedited tax net income, that is, tax net income before any U.S. intercompany dividend adjustment.
- 17 In addition to the ICD adjustment, the difference between M-1 Explains and book-tax difference includes other taxpayer errors, but the amount of other errors is small compared to the ICD adjustment.
- 18 Tax net income on Form 1120, page 1, line 28 is also the reconciliation target for Schedule M-3. See above.
- 19 As discussed later, even an extensive search of Schedule M-1 documentation for evidence of the location of the matching ICD amount may prove inconclusive.
- 20 Starting in 1999, we calculate unedited Schedule M-1 line 10 as edited line 10 plus the ICD adjustment for all corporations with an ICD adjustment.
- 21 Our table values may not add and may differ from official SOI Publication 16 values due to rounding.
- 22 Our table values may not add and may differ from official SOI Publication 16 values due to rounding.
- 23 We note that IRS examiners have always been able to investigate the supporting documentation for the line item amounts on Schedule M-1 not on detail breakout lines on a single-firm basis. However, such Schedule M-1 amounts are not useful in return classification and issue identification because supporting details are not standardized and not available in machine-readable form. See below for a discussion of the difficulties of searching the supporting documentation for Schedule M-1.
- 24 There is a plausible explanation for a large multinational taxpayer having a modest, zero, or even negative book-tax difference reported on Schedule M-1 (modest, zero, or negative M-1 Explains in our terminology). If the taxpayer began the

Schedule M-1 with its U.S. domestic income from its financial statements prepared in accordance with Generally Accepted Accounting Principles (GAAP), then its taxable income would be higher due to foreign dividends and other payments from affiliates included in its tax net income, and these amounts would need to be reflected in Schedule M-1, presumably on line 4. If such a taxpayer also improperly included U.S. intercompany dividends (ICD) on Form 1120, page 1, and on Schedule M-1, line 4, any modest, zero, or slightly negative balance for M-1 Explains would probably become very negative. We would expect such a taxpayer to be consistent and to include the U.S. ICD on line 4 if that is where it included the foreign subsidiary dividends and other income. In that case, backing out the ICD from line 4 would only restore M-1 Explains to a modest, zero, or slightly negative balance. It would not cause the restated balance to exceed our Case 3 reference. If the taxpayer included on Schedule M-1, line 1, the sum of its GAAP domestic income and its foreign subsidiary dividends and other income and any improperly included ICD, the foreign subsidiary dividends and income would have no effect on either M-1 Explains or book-tax difference under the Talisman (2000) approach, but the improperly included ICD would inflate the book-tax difference under the Talisman (2000) approach.

- 25 Negative amount representing accrual reversals may be among the items included on Schedule M-1, line 4, or for that matter, on lines 5, 7, or 8, making simple tests of Schedule M-1 line amounts difficult.

► References

- Boynton, Charles; Dobbins, Paul; DeFilippes, Portia; and Cooper, Michael, Consolidation Issues in SOI 1997 Form 1120 Book Data Compared to Matched COMPUSTAT Data, unpublished working paper, Office of Tax Analysis, U.S. Department of the Treasury, Washington, DC, May 2002.
- Boynton, Charles; DeFilippes, Portia; Lisowsky, Petro; and Mills, Lillian, Consolidation Anomalies in Form 1120 Corporate Tax Return Data, *Tax Notes*, Volume 104, Number 4, 2005, pp. 405-417.
- Boynton, Charles and Mills, Lillian, The Evolving Schedule M-3: A New Era of Corporate Show and Tell?, *National Tax Journal*, Volume 57, Number 3, 2004, pp. 757-772.
- Hanlon, Michelle, What Can We Infer About a Firm's Taxable Income from Its Financial Statements?, *National Tax Journal*, Volume 56, Number 4, 2003, pp. 831-863.
- Hanlon, Michelle and Shevlin, Terry (2005), Book-Tax Conformity for Corporate Income: An Introduction to the Issues, *Tax Policy and the Economy*, Number 19, edited by James M. Poterba, National Bureau of Economic Research, Cambridge, MA, 2005.
- McGill, Gary and Outslay, Edmund, Did Enron Pay Taxes?: Using Accounting Information To Decipher Tax Status, *Tax Notes*, Volume 96, Number 8, 2002, pp. 1125-1136.
- McGill, Gary and Outslay, Edmund, Lost in Translation: Detecting Tax Shelter Activity in Financial Statements, *National Tax Journal*, Volume 57, Number 3, 2004, pp. 739-756.
- Mills, Lillian, Book-Tax Differences and Internal Revenue Service Adjustments, *Journal of Accounting Research*, Volume 36, Number 2, 1998, pp. 343-356.
- Mills, Lillian; Newberry, Kaye; and Trautman, William P., Trends in Book-Tax Income and Balance Sheet Differences, *Tax Notes*, Volume 96, Number 8 (August 19, 2002), pages 1109-1124.
- Mills, Lillian and George Plesko, Bridging the Gap: A Proposal for More Informative Reconciling of Book and Tax Income, *National Tax Journal*, Volume 56, Number 4, 2003, pp. 865-93.
- Plesko, George A., Reconciling Corporate Book and Tax Net Income, Tax Years 1996-1998, *Statistics*

of Income Bulletin, Volume 21, Number 4, 2002, pp. 1-16.

Plesko, George A., Corporate Tax Avoidance and the Properties of Corporate Earnings, *National Tax Journal*, Volume 57, Number 3, 2004, pp. 729-37.

Plesko, George A. and Shumofsky, Nina, Reconciling Corporate Book and Tax Net Income, Tax Years 1995-2001, data release, *Statistics of Income Bulletin*, Volume 24, Number 4, 2005, pp. 103-108.

Talisman, Jonathan, Corporate Tax Shelters and the Corporate Tax Base, *Penalty and Interest Provisions*, *Corporate Tax Shelters*, testimony of Jona-

than Talisman, Assistant Secretary (Tax Policy), U.S. Department of the Treasury, before the U.S. Senate, Committee on Finance, Washington, DC, March 8, 2000, pp. 4-6

U.S. Department of the Treasury, Evidence of Growth in Corporate Tax Shelters, *The Problem of Corporate Tax Shelters: Discussion, Analysis, and Legislative Proposals*, Government Printing Office, Washington, DC, July 1999, pp. 31-33.

U.S. Department of the Treasury, *Treasury and IRS Propose New Tax Form for Corporate Tax Returns*, press release dated January 28, 2004, Washington, DC.

Exhibit I
Partial Detail of 2004 Form 1120 Page 1 and Schedule M-1

Form 1120 Department of the Treasury Internal Revenue Service	U.S. Corporation Income Tax Return For calendar year 2004 or tax year beginning, 2004, ending, 20 ▶ See separate instructions.	OMB No. 1545-0123 <div style="font-size: 2em; font-weight: bold; border: 1px solid black; padding: 5px; display: inline-block;">2004</div>
A Check if: 1 Consolidated return (attach Form 851) <input type="checkbox"/> 2 Personal holding co. (attach Sch. PH) <input type="checkbox"/> 3 Personal service corp. (see instructions) <input type="checkbox"/> 4 Schedule M-3 required (attach Sch. M-3) <input type="checkbox"/>	Use IRS label. Otherwise, print or type. Name Number, street, and room or suite no. If a P.O. box, see page 9 of instructions. City or town, state, and ZIP code	B Employer identification number : C Date incorporated D Total assets (see page 8 of instructions) \$
E Check if: (1) <input type="checkbox"/> Initial return (2) <input type="checkbox"/> Final return (3) <input type="checkbox"/> Name change (4) <input type="checkbox"/> Address change		
Income	1a Gross receipts or sales _____ b Less returns and allowances _____ c Bal ▶ 2 Cost of goods sold (Schedule A, line 8) 3 Gross profit. Subtract line 2 from line 1c 4 Dividends (Schedule C, line 19) 5 Interest 6 Gross rents 7 Gross royalties 8 Capital gain net income (attach Schedule D (Form 1120)) 9 Net gain or (loss) from Form 4797, Part II, line 17 (attach Form 4797) 10 Other income (see page 11 of instructions—attach schedule) 11 Total income. Add lines 3 through 10 ▶	1c 2 3 4 5 6 7 8 9 10 11
Deductions (See instructions for limitations on deductions.)	12 Compensation of officers (Schedule E, line 4) 13 Salaries and wages (less employment credits) 14 Repairs and maintenance 15 Bad debts 16 Rents 17 Taxes and licenses 18 Interest 19 Charitable contributions (see page 14 of instructions for 10% limitation) 20 Depreciation (attach Form 4562) 20 21 Less depreciation claimed on Schedule A and elsewhere on return 21a 22 Depletion 23 Advertising 24 Pension, profit-sharing, etc., plans 25 Employee benefit programs 26 Other deductions (attach schedule) 27 Total deductions. Add lines 12 through 26 ▶ 28 Taxable income before net operating loss deduction and special deductions. Subtract line 27 from line 11 29 Less: a Net operating loss deduction (see page 16 of instructions) 29a b Special deductions (Schedule C, line 20) 29b 30 Taxable income. Subtract line 29c from line 28 (see instructions if Schedule C, line 12, was completed) 31 Total tax (Schedule J, line 11) 31	12 13 14 15 16 17 18 19 20 21b 22 23 24 25 26 27 28 29c 30 31

Schedule M-1 Reconciliation of Income (Loss) per Books With Income per Return (see page 24 of instructions)		
1 Net income (loss) per books 2 Federal income tax per books 3 Excess of capital losses over capital gains 4 Income subject to tax not recorded on books this year (itemize): 5 Expenses recorded on books this year not deducted on this return (itemize): a Depreciation \$ b Charitable contributions \$ c Travel and entertainment \$ 6 Add lines 1 through 5	7 Income recorded on books this year not included on this return (itemize): Tax-exempt interest \$ 8 Deductions on this return not charged against book income this year (itemize): a Depreciation \$ b Charitable contributions \$ 9 Add lines 7 and 8 10 Income (page 1, line 28)—line 6 less line 9	

Exhibit II
Partial detail of 2004 Schedule M-3

SCHEDULE M-3
(Form 1120)

Department of the Treasury
Internal Revenue Service

Net Income (Loss) Reconciliation for Corporations
With Total Assets of \$10 Million or More

▶ Attach to Form 1120.
▶ See separate instructions.

OMB No. 1545-0123

2004

Name of corporation (common parent, if consolidated return)

Employer identification number

Part I Financial Information and Net Income (Loss) Reconciliation

1a Did the corporation file SEC Form 10-K for its income statement period ending with or within this tax year?		
<input type="checkbox"/> Yes. Skip lines 1b and 1c and complete lines 2a through 11 with respect to that SEC Form 10-K.		
<input type="checkbox"/> No. Go to line 1b.		
5a Net income from nonincludible foreign entities (attach schedule)	5a	()
b Net loss from nonincludible foreign entities (attach schedule and enter as a positive amount)	5b	
6a Net income from nonincludible U.S. entities (attach schedule)	6a	()
b Net loss from nonincludible U.S. entities (attach schedule and enter as a positive amount)	6b	
7a Net income of other includible corporations (attach schedule)	7a	
b Net loss of other includible corporations (attach schedule)	7b	()
8 Adjustment to eliminations of transactions between includible corporations and nonincludible entities (attach schedule)	8	
9 Adjustment to reconcile income statement period to tax year (attach schedule)	9	
10 Other adjustments to reconcile to amount on line 11 (attach schedule)	10	
11 Net income (loss) per income statement of includible corporations. Combine lines 4 through 10	11	

Part II Reconciliation of Net Income (Loss) per Income Statement of Includible Corporations With Taxable Income per Return

Income (Loss) Items	(a) Income (Loss) per Income Statement (optional)	(b) Temporary Difference	(c) Permanent Difference	(d) Income (Loss) per Tax Return (optional)
1 Income (loss) from equity method foreign corporations				
2 Gross foreign dividends not previously taxed				
3 Subpart F, QEF, and similar income inclusions				
4 Section 78 gross-up				
5 Gross foreign distributions previously taxed				
6 Income (loss) from equity method U.S. corporations				
7 U.S. dividends not eliminated in tax consolidation				
26 Other income (loss) items with differences (attach schedule)				
27 Total income (loss) items. Combine lines 1 through 26				
28 Total expense/deduction items (from Part III, line 36)				
29 Other income (loss) and expense/deduction items with no differences				
30 Reconciliation totals. Combine lines 27 through 29				

Note. Line 30, column (a), must equal the amount on Part I, line 11, and column (d) must equal Form 1120, page 1, line 28.

► Appendix

There are 34 tables which accompany this article. They may be found on the IRS Web site at <http://www.irs.gov/taxstats/productsandpubs/article/0,,id=141315,00.html>. Select the report for “2005.” The tables may also be found at <http://www.irs.gov/taxstats/productsandpubs/article/0,,id=135621.html>. Select the NTA Conference for “2005.” The first four tables appeared with the paper presented at the National Tax Association November 17, 2005, and in the article published in *Tax Notes* December 19, 2005. The remaining 30 tables were developed by the authors as part of the study and are presented here for other researchers.

The authors of this paper request that the following citation be used if data from the 34 Appendix tables are used by other researchers:

“Data are from the aggregate tables of SOI corporate file data prepared for the studies summarized in Boynton, DeFilippes, and Legel (2005, 2006) and are used with the permission of SOI, of the authors, and of Tax Analysts, publisher of *Tax Notes*. Table values may differ from official SOI Publication 16 values due to rounding.”

Table 7 (Identified as Public), Table 9 (Book-Tax Difference of \$10 Million or More Within 1995-1997), Table 13 (Manufacturing), Table 14 (Finance/Real-Estate/Holding-Companies), Table 15 (Transportation/Utilities/Information), and Table 28 (Assets of \$2.5 Million or More) are discussed in Boynton, DeFilippes, and Legel (2006), “Distribution of Schedule M-1 Corporate Book-Tax Difference Data 1990-2003 for Three Large-Size and Three Large-Industry Subpopulations.”

See Boynton, DeFilippes, and Legel (2005) for a discussion of Tables 1-4. Table 1 presents selected tax return and Schedule M-1 data for the population of all corporations (excluding S, RIC, and REIT). The population for Table 1 is the same as for SOI Publication 16, Table 12. Table 2 presents data for U.S. corporations (excluding F, S, RIC, and REIT) with assets of \$10 million or more. Table 3 presents data for U.S. corporations (excluding F, S, RIC, and REIT) with assets of \$10 mil-

lion or more requiring an adjustment for intercompany dividends (ICD). Table 4 presents data for U.S. corporations (excluding F, S, RIC, and REIT) with assets of \$10 million or more not requiring an ICD adjustment.

Tables 5 and 6 divide the population of all corporations (excluding S, RIC, and REIT) by the sign of Tax Net Income. The population for Table 5 is the same as for SOI Publication 16 Table 13.

Tables 7 and 8 for each year divide the population of all corporations (excluding S, RIC, and REIT) by “Identified as Public” or “Not Identified as Public.” A corporation is “Identified as Public” if we identify the corporation as public for any year within the period 1982-2005. Our method classifies a firm as “Identified as Public” for every SOI year in which it is present regardless of whether it was in fact public that year. The COMPUSTAT database prepared by Standards and Poor (S&P) reports Employer Identification Numbers (EIN) reported by firms on their most recent SEC Form 10-K. The COMPUSTAT record covers financial statements for public firms for the most recent 20 years as of the monthly release of a COMPUSTAT database. Data including the most recently reported EIN is reported for a firm by COMPUSTAT in each database release to the extent that the firm had any publicly available financial statements during the 20-year period then ending. We pool the COMPUSTAT EIN data from one database release selected from each of five release years, 2001 through 2005. The first year of a 20-year record for the 2001 release is 1982. The last year for the 2005 release is 2005. If we were able to identify the EIN for a corporation on a SOI annual corporate file as belonging to our pool of COMPUSTAT EIN data, we classify the corporation “Identified as Public.” COMPUSTAT has two files of companies, “active” and “research.” Active companies are currently filing public financial statements (SEC Form 10-K). Research companies are not currently filing public financial statements but have done so in one or more prior years. The research companies may have either ceased to exist through bankruptcy, dissolution, or merger, or have gone private. Early years on the 20-year COMPUSTAT record may be missing for both active and research companies. We use both the active and research files in order to be as inclusive as possible. EIN data on COMPUSTAT may include errors. We cannot ascertain

if the EIN errors are made by the corporation on the SEC Form 10-K or by COMPUSTAT in reporting the data. The following is the breakout of our EIN data for 2003 reflected in Table 7. The number of weighted returns we report in Table 7 for 2003 is 7,702 and corresponds to (3) below in the first column.

COMPUSTAT EIN Counts:

Five-Year Pool	2005 Release	Not 2005 Release	All unique EIN count
17,331	10,624	6,707	(1) Unique EIN count [unweighted count]
6,691	6,165	526	(2) Unique EIN count matched to 2003 SOI corporate file [unweighted count] (excluding S, RIC, and REIT)
7,702	7,004	698	(3) Unique EIN count matched to 2003 SOI corporate file [weighted count] (excluding S, RIC, and REIT)
5,550	5,550	0	(4) Unique EIN count matched to 2003 SOI corporate file and with a 2003 COMPUSTAT non-missing, non-zero financial statement [unweighted count] (excluding S, RIC, and REIT)

Tables 9 and 10 divide the population of all corporations (excluding S, RIC, and REIT) by “Book-Tax Difference of \$10 Million or More Within 1995-1999” or “No Book-Tax Difference of \$10 Million or More Within 1995-1999.” If we were able to identify a book-tax difference of \$10 million or more within 1995-1999 for the corporation, we labeled the corporation “Book-Tax Difference of \$10 Million or More Within 1995-1999.”

Tables 11 and 12 divide the population of all corporations (excluding S, RIC, and REIT) by “Stock Option Expense on Schedule M-1 Within 2002-2003” or “No Stock Option Expense on Schedule M-1 Within 2002-2003.” Stock option expense is tabulated on Schedule M-1 only for 2002 and 2003. If we were able to identify stock option expense on Schedule M-1 within 2002-2003 for the corporation, we labeled the corporation “Stock Option Expense on Schedule M-1 Within 2002-2003.”

Tables 13 through 20 divide the population of all corporations (excluding S, RIC, and REIT) by SOI

major industry code. For 1990-1997, the population for each of Tables 13-20 is the same as for one of the major industry total columns in SOI Publication 16, Table 12. For 1998-2003 we have combined the revised industry codes to approximate the 1990-1997 divisions. For 1998-2003, the population for each of Tables 13-20 is the same as for one of the major industry total columns in SOI Publication 16, Table 12, or is the sum of two or more columns. We indicate the SOI major industry codes involved for each period in the table heading.

Tables 21 through 28 divide the population of all corporations (excluding S, RIC, and REIT) by reported asset size for the given year.

Tables 29 and 30 divide the population of Table 28, U.S. corporations (excluding F, S, RIC, and REIT) with assets of \$2.5 billion or more by whether the corporation required an ICD adjustment for the given year. This division is similar to the ICD division of Table 2, U.S. corporations (excluding F, S, RIC, and REIT) with assets of \$10 million or more by ICD in Tables 3 and 4.

Table 31 is the sum of Tables 26 through 28.

Tables 32 through 34 are the component SOI major industries for 1998-2003 that comprise Table 15.

► References

Boynton, Charles; DeFilippes, Portia; and Legel, Ellen. “Distribution of Schedule M-1 Corporate Book-Tax Difference Data 1990-2003, for Three Large-Size and Three Large-Industry Subpopulations,” Tax Notes 111, No. 2 (April 10, 2006), pages 177-212.

Boynton, Charles; DeFilippes, Portia; and Legel, Ellen. “Prelude to Schedule M-3: Schedule M-1 Corporate Book-Tax Difference Data, 1990-2003.” Tax Notes 109, No. 12 (December 19, 2005), pages 1579-1599.

3



Behavioral Responses to Corporate Taxation

Contos

An Essay on the Effects of Taxation on the Corporate Financial Policy

George Contos, Internal Revenue Service

The taxation of corporate profits in the United States has been one of the most widely discussed issues in the area of public finance. Corporate revenues are currently subject to double taxation. Profits are taxed first at the corporate level and then, when distributed as dividends or when capital gains are realized, taxed a second time at the individual level. The share of tax revenues from corporate profits has been decreasing steadily over the past four decades. In 1962, corporate tax receipts accounted for 21 percent of all tax revenues, but, by 2003, their share dropped to 7.5 percent.¹ In 2003, a proposal by the Bush Administration brought corporate tax integration back to the front pages. The final legislation, the Jobs and Growth Tax Relief Reconciliation Act of 2003, did not eliminate double taxation, but it did reduce the taxation of corporate profits at the individual level.² Double taxation is still a reality; so, the discussion for corporate integration is clearly not over.

In understanding why corporate taxation is such a highly contested issue, critics argue that the current tax system discourages business entities from organizing as taxable corporations and encourages corporations to veer from socially efficient decisions (Scholes et al. (2005), p. 336). Those critics believe that the losses to the U.S. economy caused by the current tax system far exceed the gains from the revenues raised. They call for a neutral tax system that does not enter into the decisionmaking process of firms and does not distort economic efficiency. Supporters of corporate taxation reply to those allegations by saying that corporations are distinct entities and should be taxed separately from their shareholders; that corporations should pay a fee, tax, for the special privileges they enjoy; and that corporate taxation prevents the sheltering of individual income from taxation (Rosen (2002), p. 399).

A large body of research has tested for the effects of corporate taxation. Although the results of empirical models vary significantly, all models agree that, to some degree, corporate taxation affects a broad range of the

decisions made by taxable corporations. The magnitude of those effects and their overall impact on the economy are still under debate. Jane Gravelle (1995) divides the debate on corporate taxation into three key issues. "First who carries the burden of corporate tax--capital, labor, or consumers, and does it play a role in a progressive tax system? Second, how significant are the distortions caused by the excess corporate tax? And third, how can the revenues raised from corporate tax be replaced?" This paper focuses on the second question and more specifically on how the deductibility of interest affects the capital structure of taxable corporations. I test the hypothesis that taxable corporations have a tax incentive to use debt financing versus equity financing because interest paid is tax-deductible while dividends paid to shareholders are not. Measuring the excess debt that corporations carry due to the tax incentive is important because the excessive use of debt may lead to financial distress and even bankruptcy.

This paper extends the work of Gordon and Lee (2001). They use an aggregate data time-series, Tax Years 1950 to 1995, to test for the effects of corporate taxation on the financial policy of firms of different sizes. They found that taxes have a large effect on the use of debt for the smallest and the largest firms. In this paper, I first estimated the Gordon and Lee (G&L) model using the same aggregate Statistics of Income (SOI) data but for a different time period, Tax Years 1993 to 2000, and my findings were qualitatively similar to those of G&L. Next, I introduced a confidential SOI firm-level dataset for the 8-year period, and found an unexpected negative relation between tax rates and debt. However, using a marginal tax rate constructed from taxable income before the interest deduction and the panel dataset, I found, as expected, a positive relation between tax rates and debt. Finally, I divided my panel dataset into small, intermediate, and large size firms, and I found a positive relationship between tax rates and debt for all three firm sizes.

► Corporate Taxation

Before discussing existing research on how taxes affect the corporate capital structure, it is useful to review how double taxation affects the decisionmaking process of firms. Business entities have a financial incentive to organize as “C corporations,” where the term C corporation comes from the subchapter of the Tax Code defining their structure. Corporations are legal entities that can have multiple owners and separate management. The ability to attract multiple investors through the sale of shares or bonds gives corporations broad access to capital and greater potential for growth. The shares of corporations can be easily transferred to other investors without disrupting the operations of the companies. The owners of corporations also enjoy limited liability since, in case of default, their liability is limited to the amount they have invested. Because, in the United States, corporate profits are subject to double taxation, corporations in essence pay a fee for the right to incorporate. Corporate revenues are taxed first on the corporate level and then, when distributed as dividends or when capital gains are realized, taxed a second time on the individual level. Business entities can avoid double taxation but in the process lose some of the special privileges mentioned earlier, if they organize as passthrough entities. Passthrough entities, such as sole proprietorships, partnerships, and subchapter S corporations, avoid double taxation by passing all profits and losses onto their shareholders (Brealey and Myers, 2000).

The firm can finance its investments using equity or debt. Equity is either cash available to the firm or funds raised by issuing stock, primarily common stock. Dividends paid to stockholders are not tax-deductible; thus, dividends are paid from after-tax income. A firm raises debt by borrowing from its shareholders, from financial institutions, or from the public. All interest paid by a corporation to its lenders is tax-deductible, thus generating a tax shield. Clearly, there is a tax incentive for a taxable corporation to use debt instead of equity. So, double taxation directly affects the corporate capital structure.

Since all interest paid is tax-deductible, one would expect that taxable corporations would rely heavily on

debt to finance their investments, but empirical evidence shows that they use significant amounts of equity capital.³ Why is this so? There can be significant nontax costs involved with debt financing. These costs include both the standard costs of borrowing and risks of financial distress that fixed liabilities imply. Firms fall into financial distress when they have difficulty making their debt payments. Extended periods of financial distress can lead to bankruptcy. The higher the debt payment levels, the higher the probability that the firm could fall into financial distress. As the probability of distress increases the risk for the firm’s debtor increases, so they demand higher return for their investments. Consequently, the value of debt tax shields decreases as these forms of nontax costs increase.

The value of tax shields also depends on the marginal tax rate of the firm, and the availability of nondebt tax shields⁴ and tax credits. The marginal tax rate is the tax liability generated, today and in the future, by an additional dollar of income earned today. Estimating the marginal tax rate is not straightforward because of the uncertainty of future earnings, the carryback and the carryforward provisions of the tax law, and the alternative minimum tax (AMT). Corporations can “carry back” and “carry forward” operating losses and tax credits—meaning they can apply them to reduce tax liabilities incurred in past or future years. As Graham (1996) explains, the relationship among operating losses, marginal tax rates, and the value of tax shields is not always obvious. For example, tax shields have very low, if no, value to corporations that expect operating losses in the future. Such firms will have very low marginal tax rates because they can use those net operating loss deductions (NOL’s) in the future to refund any taxes paid today. Firms that experienced losses in the past and expect moderate profits in the future can also use NOL’s to reduce future tax liabilities. However, if that same firm carries back its current-year NOL and the NOL is less than or equal to its past liabilities, then the marginal tax rate of any additional income earned today will be equal to the applicable statutory tax rate. From these examples, it is easy to see that the NOL deduction makes estimating the marginal tax rate of a corporation complex.

The value of debt tax shields also depends on the availability of nondebt tax shields⁴ and tax credits. As

DeAngelo and Masulis (1980) explain, one can make the case of a tax shield substitution effect since the availability of nondebt tax shields may crowd out debt tax shields. Finally, it has been shown that the foreign tax credit limitations do not just reduce the value of debt tax shields, but actually influence U.S. multinationals to decrease their domestic debts by substituting them with equity financing.

In this paper, the corporate marginal tax rate proxies are constructed by selecting the marginal statutory rate that applies to the highest dollar of the current-year taxable income, or taxable income before interest deduction, reported on the tax return. Such proxies have been used successfully in earlier research and can be applied to both the aggregate and firm-level datasets used. Upcoming research by the author explores the effects of the NOL deduction and the various tax credits on the corporate capital structure.

► **Prior Empirical Research**

Modigliani and Miller (1963) were the first to introduce the idea that corporate taxation affects the capital structure of firms. As Scholes et. al. (2005) discuss, Modigliani and Miller showed that if the only imperfection of the capital markets is corporate taxation, the deductibility of interest generates a debt tax shield that increases the value of corporations. When comparing debt and equity financing, Modigliani and Miller explain that borrowing is beneficial to corporations because the cost of debt, interest paid, is tax-deductible while the cost of equity, dividends, is not. In a later paper, Miller (1977) pointed out that, if one takes into account the tax status of corporate investors, equity financing can be a competitive alternative to debt financing. If the interest earned by the debt holders is taxed at a higher rate than the dividends paid to stockholders, then the corporation's tax incentive is the difference between the sum of the corporate tax rate plus the dividend rate, and the individual tax rate of the bondholders. The work of Modigliani and Miller was advanced by DeAngelo and Masulis (1980), who introduced the idea of tax shield substitution. Firms can substitute nondebt tax shields, like the depreciation deduction, for debt tax shields. The work of DeAngelo and Masulis is important because it led to a hypothesis

that can be empirically tested; firms with large amounts of nondebt tax shields will have lower levels of debt than firms with small amounts of nondebt tax shields (Scholes et al. (2005) p. 344).

Since the works of Modigliani and Miller (1963) and DeAngelo and Masulis (1980), a number of empirical studies have examined the impact taxes have on the financial structure of corporations. As Ayers, Cloyd, and Robinson (2001) explain, the capital structure literature can be divided into two streams. The first stream of works compares taxable corporations that have different tax incentives, hypothesizing that firms with greater tax incentives will have higher levels of debt. The second stream of works compares taxable corporations to passthrough entities that are not subject to corporate taxation because, by law, they have to pass all income to their shareholders. Their hypothesis is that taxable corporations will have higher levels of debt than passthrough entities.

The earlier articles of the first stream do cross-section analysis of taxable corporations but do not find convincing evidence that taxation affects the financial policy of firms (Bradley, Jarrell, and Kim, 1984; and Gaver and Gaver, 1985). The more recent articles of the first stream are more successful in finding evidence of a significant positive relationship between debt financing and marginal tax rates. These articles introduce several improvements over earlier work: They examine incremental financing decisions instead of debt levels (MacKie-Mason (1990); Graham (1996); Gropp (1997)); they develop better proxies for marginal tax rates (Graham (1996); Graham, Lemmon, and Schallheim (1998)); they use the ratio of interest expense to gross profit rather than the debt-to-equity ratio as the dependent variable (Cloyd, Limberg, and Robinson (1997); and they research the debt policies of corporations of different sizes (Gordon and Lee (1999)). Here, I briefly present an overview of this work, focusing on the data, the marginal tax rate proxies used, and their key findings.

Bradley, Jarrell, and Kim (1984) use data from 851 large firms to estimate a general equilibrium model. Although they have multiyear data for each firm, in order to avoid business cycle variations or different

adjustment periods, they calculate a 20-year average or "permanent" leverage ratio for each firm. They examine how these ratios vary with the industry of the firm, the volatility in the firm's earnings, the availability of nondebt tax shields, and the expenditures on research and development and advertising. They do not find concrete evidence that taxation affects the firm's leverage ratios, but they find evidence that the leverage ratios are strongly influenced by the firm's industry. They also find that firms with volatile earnings have lower levels of debt, suggesting that the risk of bankruptcy has a negative effect on the amount a firm borrows. Finally, they find that firms with higher levels of nondebt tax shields borrow more, a finding that contradicts the findings of the earlier literature. Bradley, Jarrell, and Kim offer as a possible explanation for this last finding that firms with large amounts of assets have more collateral and thus can borrow more.

The Gaver and Gaver (1985) article does not test directly for the relationship between taxes and debt ratios but rather tests the hypothesis that there is a systematic relationship between the firm's investment opportunity set and its corporate policy decisions. Using longitudinal data from 237 new and 237 established firms, they find evidence that growth firms have significantly lower debt-to-equity ratios than established firms. This is an interesting result that could explain the differences in the debt levels across firms.

The MacKie-Mason (1990) article uses the Compustat data on large publicly traded companies to examine the relationship between nondebt and debt tax shields to measure the firm's tax incentive, using a dummy variable for the net operating loss deduction. Instead of using the aggregate debt over total assets ratio as the dependent variable, he uses the annual change in the total debt levels scaled by the firm's total assets. He finds evidence of substantial tax effects on the choice between issuing debt or equity; that firms with net operating loss carry-forwards are much less likely to use debt; and that the existence of investment tax credits reduces the probability of debt issues only when the firm's marginal tax rate is near zero. His findings support a significant relationship between corporate taxation and the financial decisions of a firm.

Graham (1996) follows MacKie-Mason's incremental choice approach, using a simulated firm-specific marginal tax rate as a proxy for the firm's tax incentives. The data used are a pooled cross-section of differenced time series from about 10,000 Compustat firms from 1980 to 1992. Although he finds a strong positive relation between tax status and incremental debt policy, he is puzzled by the low R-squared of about 5 percent that his regressions produce. He states that "future researchers should study why, given the strong tax incentives firms have to issue debt, taxes do not explain a larger portion of debt policy." Finally, he tests the effectiveness of the tax status proxies used by earlier papers and finds that only the net operating loss dummy variable is a reasonable proxy.⁵

Gropp's (1997) paper builds on the work done by MacKie-Mason and Graham, but, instead of using proxies for expected marginal tax rates, he uses a simple rational expectations approach to estimate the expected effective corporate tax rates of firms. He finds "that current average effective tax rates have substantial predictive power for the estimation of expected corporate tax rates." Controlling for other theories of capital structure choices, he finds that corporate taxation affects the financial policy of firms using a balanced panel from Compustat of 929 publicly traded manufacturing U. S. firms from 1979 to 1991.

Graham, Lemmon, and Schallheim (1998) is the first paper to find a positive relationship between the tax incentive and debt financing using debt levels. They provide evidence that the corporate tax status is endogenous to financing decisions, producing a spurious relationship between the debt ratio and the marginal tax rate of the firm; in other words, the estimated effects of tax status on the debt levels will be biased because companies that have high levels of debt also have low marginal tax rates. To solve this problem, they propose a direct measure of the corporate marginal tax rate using taxable income before the interest deduction as a measure of the firm profits. Using a balance panel from Compustat of 18,193 observations from 1981 to 1992, they find a positive relationship between tax rates and the usage of debt.

Gordon and Lee (2001) is the first paper to research the debt policies of corporations of all sizes and to find a positive relationship between debt levels and after-financing tax rates. They create a dataset from the aggregate data on corporations published by SOI and test for the effects of taxation by comparing the ratios of debt-to-assets of firms in different asset size-classes. Over the 46-year period covered by their data, the corporate tax rates varied significantly,⁶ giving them adequate variation both across time and across firms for a difference-in-difference procedure. This procedure compares the changes in the debt-to-assets ratios for small versus large firms with the changes in the relative tax rates they face. They find that taxes have a large effect on the use of debt for the smallest and the largest firms. For intermediate-sized firms, they estimate a much lower effect, but they provide indirect evidence that this finding is a result of measurement error in the tax variable. Since the SOI data are grouped in asset classes, they only have information on the average rate of return for firms in each asset class, taxable income divided by assets; so, they calculate the average marginal tax rate for firms in each asset class. Due to this limitation, "they are not able to capture the effects of heterogeneity in rates of return across firms on the expected marginal tax rate, arising from the nonlinearity in the tax structure." The effects of heterogeneity in rates of return are more important for intermediate firms since their "taxable incomes are near the point where tax rates change dramatically."

To avoid such problems, I introduced a confidential firm-level dataset of taxable corporations of all sizes, for Tax Years 1993 to 2000. This dataset allowed studying the effects of taxation on firms of all sizes, while capturing the heterogeneity in rates of return across firms. I found an unexpected negative relation between tax rates and debt. However, using a marginal tax rate constructed from taxable income before the interest deduction, I found the expected positive relation between tax rates and debt. Next, I took advantage of the panel aspects of the microdataset; by using fixed effects models, I controlled for the unobserved firm-specific effects and found again a positive relation between taxation and debt. Finally, I divided the panel dataset into small, intermediate, and large size firms, and I found a positive relationship between tax rates and debt for all three firm sizes.

► Empirical Research

The data sample

The data used for this study are the firm-level data collected by SOI and published on an aggregate basis in the annual Corporate Source Book.⁷ The data come from the tax returns of domestic corporations and foreign corporations with U.S. business activities.⁸ The firm-level data are confidential, although SOI employees--like my self--can conduct analyses of the data and share the results with outsiders subject to disclosure review by the Internal Revenue Service (IRS).

I began my analysis with Tax Year 1993 since it is the first year that three new tax brackets, for returns with taxable income greater than 10 million dollars, came into effect. The three brackets were introduced by the Tax Relief Act of 1993 and give my time series additional variation across firms compared to earlier years. I ended my analysis with Tax Year 2000 because it is the last full year before the recession that started in March of 2001.⁹ Tax receipts in Tax Year 2001 decreased significantly; so, including these data would complicate the analysis of my findings.¹⁰ During the 1993 to 2000 time period, the corporate tax schedule remained unchanged; so, the dataset provides significant variation across firms but limited variation across time.

To create the panel, I limited my sample to companies that filed tax returns under the same Employer Identification Number (EIN) and were selected by the SOI sampling process every tax year from 1993 to 2000.¹¹ To confine the data to nonfinancial firms with appreciable business operations, I excluded all financial returns because they follow different tax rules: 1120F filers because SOI does not collect balance sheet information from them; part-year returns which have tax periods of 6 months or less; and all returns with total assets of \$10,000 or less because such firms are too small to help the explanatory power of the empirical model. After these exclusions, the panel consisted of 10,552 firms.

Constructing a "true" balanced panel of corporations is complicated by the need to account and adjust for mergers, acquisitions, and other changes to the structure

of each corporation in the sample. Given the difficulty of this undertaking, and of analyzing firms undergoing major changes, I decided to exclude from the panel all companies for which total assets increased by more than tenfold in a single year and all companies for which total assets decreased by more than 90 percent between 1999 and 2000. The first criterion eliminates from the panel corporations that have merged with or acquired another business entity. The second criterion eliminates from the panel corporations that are in financial distress and will be going out of business in the near future.¹² A total of 60 records were dropped for these reasons, leaving a “final” panel of 10,492 firms.

Apart from the large number of observations, the SOI data offer several advantages over the financial data used in the prior literature. The data collected by SOI are reported by firms to the IRS when financial (book) data are reported by corporations to their shareholders.¹³ As George Plesko (2004) points out, “differences in accounting rules for book and tax reporting purposes can lead to differences in the amount of income reported to shareholders and to the IRS.” Mills, Newberry, and Trautman (2002) find that book-tax income differences grew throughout the 1990’s so that tax rates estimated from book income will be wrong.¹⁴

Financial and tax data may also differ when a parent corporation reports with its subsidiaries. For financial purposes, a parent company must include in the consolidation all domestic and foreign subsidiaries which it owns by 50 percent or more. Under tax rules, however, domestic subsidiaries must be 80-percent or more owned to be included in the parent’s tax return, and foreign subsidiaries cannot be consolidated. Since the Compustat dataset reports financial consolidations and does not separate foreign and domestic income, taxable income could be inflated. The amount of debt reported by some companies in their tax returns could be inflated because they do not eliminate intercompany payables and receivables. Mills, Newberry, and Trautman (2002) report anecdotal feedback of such reporting, but, since the dependent and the control variables of the empirical model are ratios, the effects should be minimal.

Finally, another reason financial and tax data may differ is off-balance sheet financing. Firms in the 1990’s

used special purpose entities to keep debt outside their consolidated financial statements. Mills and Newberry (2004) find “that these financial reporting effects occurred primarily during 1994-1999.” So the financial statements of large firms for that period could under-report both interest expense and debt and inflate taxable income. I believe that, overall, the use of tax data improves the accuracy of my empirical work.

► Summary Statistics

In order to be able to compare my results using the firm-level data with G&L results based on aggregate data, I first present summary information of all variables from the G&L sample and the present sample. As shown in Table 1, the summary statistics of the two samples match very well. The mean total debt-to-assets ratio is about four percentage points higher in the present sample compared to that of G&L, reflecting greater long-term borrowing over prior decades. Looking at the asset side of their balance sheets, firms in the two samples own comparable amounts of depreciable property and land, but firms in the present sample have higher amounts of intangible assets.¹⁵ Finally, although the ratio of accounts receivable to assets dropped by a little bit more than 3 percentage points, cash holdings increased by about 2 percentage points. In comparing the mean marginal rates of the two datasets, it is obvious that, in recent years, corporations have faced significantly lower statutory corporate tax rates: Companies in the 1950 to 1995 period faced higher tax scales with top statutory rates as high as 52 percent, while those in the 1993 to 2000 period faced significantly lower tax scales that topped at 39 percent. The mean marginal tax rate (**mrt**) has decreased from 37.6 percent to 26.5 percent.¹⁶ In contrast, the average yearly individual tax rate on interest faced by individual taxpayers (**ifmr**) in the same two periods was much more stable, slipping from 24.5 to 22.3.¹⁷ It is clear that firms in the 1993 to 2000 period have considerably lower tax incentive (**dmr**) than firms in the 1950 to 1995 period.¹⁸

► Empirical Findings and Sensitivity Analysis

I begin my empirical analysis by regressing the present aggregate sample. The first equation of the Gordon

Table 1
Aggregate Data ¹
Sample Means and Standard Deviations of Variables

Variables	Notation	Gordon & Lee 1950 - 1995 ²		Present Study 1993 - 2000 ³	
		Mean	Standard Deviation	Mean	Standard Deviation
Corporate debt-asset ratios					
Total debt-to-assets	Tdr	25.18	8.05	29.12	6.83
Short-term debt-to-assets	Sdr	9.45	4.07	10.33	3.22
Long-term debt-to-assets	Ldr	15.73	4.36	18.78	4.62
Tax rates					
Marginal tax rate-taxable income	Mrt	37.57	13.15	26.48	9.74
Marginal tax rate-taxable income plus interest paid	Mrtint	37.97	12.81	27.80	9.86
Individual tax rate	lfmr	24.49	2.36	22.26	1.00
Marginal tax rate minus individual tax rate	Dmr	13.04	12.72	4.22	9.75
Corporate assets					
Depreciable assets-to-assets	Dpr	20.79	6.32	21.17	7.09
Land-to-assets	Landr	3.66	2.46	3.51	2.06
Cash-to-assets	Car	9.5	4.00	11.37	6.58
Intangible assets-to-assets	Intr	1.12	1.08	2.45	0.84
Accounts receivable-to-assets	Arr	22.83	4.53	19.01	4.70

¹ Source: SOI Source Book, amounts are in dollars.

² From Gordon and Lee (1999)

³ Author's tabulations

and Lee empirical model measures the effects of tax incentive (**dmr**), nontax factors, firm unique characteristics, and the business environment on the firm's total debt-to-assets ratios.¹⁹ To simplify the model, G&L assume that all nontax factors that affect the corporate financial policy do not change over time or change in a way that is uncorrelated with relative tax rates. To account for those nontax factors, they use an "arbitrary function that measures desired debt-to-assets ratios ignoring tax incentives." In estimation, this arbitrary function is a seventh-order polynomial function of logged real assets.²⁰ The unique characteristics of the firms in each asset class are measured by the composition of the assets of those firms. Finally, the business environment is captured by a set of Tax Year dummies. Thus, the equation estimated is:

$$\begin{aligned}
 tdr_{st} = & \sum_{i=1}^7 \alpha_i \log(rassts_{st})^i + \beta dmr_{st} \\
 & + \gamma X_{st} + \sum_{t=1}^7 \delta_t d_t + \varepsilon_{st} \quad (1)
 \end{aligned}$$

where **tdr** is the debt over asset ratio for firms in asset class *s* at year *t*, **rassts_{st}** are the inflation-adjusted total assets of firms in asset class *s* at year *t*, **log(rassts)ⁱ** is the *i*th order polynomial function of logged **rassts**, **dmr** is the tax incentive of firms in asset class *s* at year *t*, **X_{st}** is a matrix of the composition of the assets of firms in asset class *s* at year *t*, and **d_t** are Tax Year dummies. The main hypothesis is that the coefficient of the tax incentive is positive. For the asset composition variables, I expect that firms with higher depreciable assets, land, and intangibles asset ratios will have higher debt-to-asset ratios when firms with higher cash balances and trade notes and accounts receivable will have lower debt-to-asset ratios. A complete listing of the variables is included in the appendix.

Gordon & Lee use OLS to estimate the first equation, finding the effects of taxes on debt to be modest. Because the marginal tax rate proxy is based on taxable income, they are concerned with possible endogeneity bias: a firm's debt levels through the interest deduction directly affect its taxable income. To correct this bias, they construct an exogenous instrument, based on the findings of Graham, Lemmon, and Schallheim (1998) and re-estimate the model using Instrumental Variable (IV). The instrument is the average tax rate faced by all firms in each time period if the interest deduction is added back to taxable income. Their IV coefficients are not significantly different from their OLS, which G&L attribute to high correlation of the instrument with the marginal tax rate proxy.

The results of the OLS regressions for the present and G&L samples are shown in Table 2. Like Gordon and Lee, I find an unexpected negative relation between tax rates and debt. I next controlled for the firms' size and asset composition by regressing the first equation, resulting as expected in a positive tax coefficient. The coefficients of the control variables, except for the ratio of land-to-assets, had the expected signs and are significant at the 1-percent level. So, I found that the 1990's aggregate data produce the same results as the aggregate data from 1950 to 1995.

G&L also estimate the effects on financial policy of any factors that change over time. These factors are the business cycle, the nominal interest rates, and the tax en-

Table 2
Aggregate Data
Regression Results

Variables	G&L Tdr	Present Tdr	G&L Tdr	Present Tdr	Present Sdr	Present Ldr
Dmr	-0.393** (0.020)	-0.384 (0.065)	0.079** (0.019)	0.078** (0.038)	0.127** (0.027)	-0.048** (0.028)
Log(rassts)			1.853** (0.355)	0.034** (0.007)	0.021** (0.005)	0.013** (0.005)
Log(rassts) ²			-0.641** (0.135)	-0.015** (0.003)	-0.012** (0.002)	-0.003** (0.002)
Log(rassts) ³			-0.568** (0.068)	-0.002** (0.0002)	-0.002** (0.0002)	-0.0002** (0.0001)
Log(rassts) ⁴			0.085** (0.009)	0.0006** (0.0001)	0.0005** (0.00007)	0.0002 (0.00007)
Log(rassts) ⁵			0.019** (0.004)	-0.00003** (0.00009)	-0.0002* (0.00004)	-0.0002 (0.00004)
Log(rassts) ⁶			-0.004** (0.001)	-	-	-
Log(rassts) ⁷			0.002** (0.00038)	-	-	-
Dpr			0.320** (0.058)	0.663** (0.122)	0.096** (0.083)	0.567** (0.092)
Landr			0.317 (0.254)	-1.271** (0.307)	-1.606** (0.208)	-0.335* (0.231)
Car			-0.437** (0.087)	-0.223 (0.225)	-0.394** (0.152)	0.171 (0.169)
Intr			1.447** (0.341)	0.578* (0.409)	0.251 (0.276)	0.326* (0.307)
Arr			-0.027 (0.040)	-0.823** (0.166)	-0.630** (0.112)	-0.193** (0.124)
Constant	25.572** (1.289)	0.311 (0.018)	20.992** (2.187)	0.433** (0.062)	0.370** (0.042)	0.063** (0.047)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	434	88	434	88	88	88
Adj R-squared	0.433	0.246	0.972	0.98	0.974	0.988

* and ** indicate significance levels at 5 percent and 1 percent. Standard errors in parenthesis.
Note: Following G&L, I stopped adding powers to the polynomial when the next higher power was statistically insignificant.

environment. The dependent variable for the second equation is the coefficients of the time dummies estimated on the first equation. Having already controlled for the tax incentives, size of firm, and asset composition, the coefficients of the time dummies capture the effects on financial policy of these nontax factors. In addition, by including in the second equation a yearly measure of the tax incentive (**dmr**), G&L also test if they have adequately controlled for taxes on the first equation. If they have done so, then the coefficient of the tax incentive must be equal to zero. Thus, the equation estimated is:

$$\hat{\delta}_t = \alpha_0 + \alpha_1 ydmr_t + \alpha_2 tb_t + \alpha_3 dj_t + \alpha_4 d > 86 + v_t \quad (2)$$

where $\hat{\delta}_t$ are the coefficients of the Tax Year dummies estimated by the first equation, **dmr** is the average tax

incentive faced by corporations at year t , **tb** is the nominal interest rate measured by the 3-year Treasury bond rate, **dj** is a business cycle proxy equal to the ratio of the Dow Jones index over Gross Domestic Product, and **d > 86** is a dummy capturing any omitted aspects of the Tax Reform Act of 1986.

Table 3 reports both the unexplained yearly variation reported by the G&L and the present samples. According to G&L, if the first equation fully accounts for the effects of taxation on the corporate financial policy, then the tax coefficient of the second equation should be zero; they find that the tax coefficient is positive, large in magnitude, and statistically significant. Because the dependent variable of the second equation is measured net of the estimated effects of taxes estimated in the first equation, to get the complete effect of taxation, they combine the two IV tax coefficients. They find that large firms in the 1970's would finance 9.2 percent of their assets with debt relative to the smaller firms. Using seven annual observations, my replication of the time-series aggregate model showed no unexplained yearly variation. So, for the present sample, the first equation seems to capture the tax incentive in its entirety. This is not totally unexpected since, in the 8 years of my time series, both business cycle and the nominal interest rate variables remained fairly constant when their sample

Table 3
Aggregate Data
Unexplained yearly variation
OLS Regression Results

Variables	G&L	Present
Dmrt	0.264** (0.094)	-0.232 (0.291)
Mrt		
Ifmr		
TB	0.504** (0.148)	0.001 (0.003)
DJ	-4.546** (1.485)	0.015 (0.020)
Dummy for post 1986	3.313** (0.692)	
Constant	0.191 (1.978)	-0.004 (0.044)
Obs.	37	7
Adj R-squared	0.84	0.90

*and ** indicate significance levels at 5 percent and 1 percent. Standard errors in parenthesis.

period permits 37 annual observations and gains power from a structural change in 1986, as well as several economic cycle changes.

I now turn my attention to the balanced panel of firm-level microdata. I began by regressing the first equation on the final panel using OLS. The results of these regressions are reported in the first two columns of Table 4. The tax coefficient is significant at the 1-percent level but negative, and it stayed negative even after I controlled for the size of the firm and asset composition. The asset composition variables had the expected signs, and their magnitudes are consistent with my expectations and were statistically significant. Firms with higher depreciable or intangible asset ratios have higher debt-to-asset ratios, and firms with higher levels of cash at hand and accounts and trade notes receivable have lower debt-to-asset ratios. Finally, the land coefficient was again negative but significantly lower. The

adjusted R-squared of the regression is 0.14 percent. So, my model provides a better fit than earlier firm-level studies but is still unexpectedly poor.

Still not satisfied with the goodness of fit of the liner model, I estimated a log-linear model,²¹ and the OLS regression results are shown in the two last columns of Table 4. The adjusted R-squared of the log-linear regression was higher than the linear model, while the sum of square errors was lower, suggesting a better fit. In particular, the adjusted R-squared was now 0.2 percent, considerably higher than the ones reported by similar firm-level studies. The tax coefficient was again negative, and the asset composition variables had the expected signs.

I next took advantage of the panel aspects of my dataset by using fixed effects.²² Fixed effects allow us to isolate the unobserved firm-specific effects and get a better measure of the true effects of taxation on the financial policy of firms. By unobserved firm-specific effects, I refer to all those firm-unique characteristics that do not change from year to year and help shape the firm's financial policy and capital structure. As shown in Table 5, the relationship between the tax incentive and debt-to-asset ratios is again negative. The tax coefficient when total debt is the dependent variable was -0.115, while the coefficients of the asset composition variables have the expected signs and, except for the ratio of land-to-assets, were statistically significant. The tax coefficient was negative even when I divided debt into short-term and long-term, -0.057 and -0.065, respectively. The overall R-squared of the total, short, and long-term debt regressions were 0.14 percent, 0.016 percent, and .2 percent, respectively.

To test whether the tax coefficients are driven by the presence in my sample of a significant number of firms with no taxable income, I regressed the first equation using two subsets of the final panel. In the first, the sample was limited to 8,900 firms that had a positive marginal tax rate for at least 1 year. Here again, the fixed effects tax coefficient was negative and significant. Next, the sample is further restricted to the 3,100 companies that had a positive marginal tax rate every year; the coefficient remained negative and significant. Both datasets produced the expected signs for all control variables,

Table 4
OLS Regression Results

Variables	Tdr	Tdr	Log(tdr)	Log(tdr)
Dnr	-0.821** (0.009)	-0.381** (0.008)	-0.581** (0.006)	-0.240** (0.005)
Log(rassts)		-8.079** (0.735)		-4.417** (0.436)
Log(rassts) ²		0.960** (0.093)		0.532** (0.055)
Log(rassts) ³		-0.055** (0.006)		-0.031** (0.003)
Log(rassts) ⁴		0.002** (0.0001)		0.0009** (0.0001)
Log(rassts) ⁵		-0.00002** (0.000002)		-0.000005 (0.000001)
Dprr		0.272** (0.005)		0.263** (0.004)
Landr		-0.028** (0.010)		-0.038** (0.007)
Car		-0.384** (0.008)		-0.411** (0.006)
Intr		0.363** (0.020)		0.304** (0.014)
Arr		-0.087** (0.006)		-0.098** (0.005)
Constant		26.654** (2.294)		14.353 (1.362)
Year Dummies	No	Yes	No	Yes
Obs.	83,936	83,936	83,936	83,936
R-squared	0.09	0.14	0.09	0.20

*and ** indicate significance levels at 5 percent and 1 percent. Standard errors in parenthesis.

Note: The final panel includes 10,492 nonfinancial companies that filed tax returns under the same EIN and were selected by the SOI sampling process every tax year from 1993 to 2000 and their total assets did not increase by more than 10 times from one period to the next and did not file final returns in Tax Year 2000. Following G&L, I stopped adding powers to the polynomial when the next higher power was statistically insignificant.

Table 5
Fixed Effects Regression Results

	Log(tdr)	Log(sdr)	Log(ldr)
Variables			
Log(dmr)	-0.115** (0.004)	-0.057** (0.003)	-0.065** (0.003)
Log(rassts)	-2.432** (0.526)	-2.202** (0.409)	-0.503** (0.474)
Log(rassts) ²	0.285** (0.067)	0.242** (0.052)	0.073** (0.060)
Log(rassts) ³	-0.016** (0.004)	-0.013** (0.003)	-0.005* (0.004)
Log(rassts) ⁴	0.0005** (0.0002)	0.0003** (0.0001)	0.0002* (0.0002)
Log(rassts) ⁵	-0.00001** (0.000003)	-0.00001** (0.000002)	-0.000003* (0.000002)
Log(dpr)	0.267** (0.007)	0.034** (0.005)	0.251** (0.006)
Log(landr)	0.145** (0.013)	-0.005 (0.011)	0.154** (0.012)
Log(car)	-0.108** (0.006)	-0.076** (0.005)	-0.038** (0.005)
Log(intr)	0.310** (0.015)	-0.018** (0.012)	0.344** (0.014)
Log(arr)	-0.058** (0.007)	-0.021** (0.005)	-0.040** (0.005)
Constant	8.148** (1.621)	7.850** (1.260)	1.249 (1.461)
Year Dummies	Yes	Yes	Yes
Obs.	83,936	83,936	83,936
R-squared	0.14	0.014	0.20

*and ** indicate significance levels at 5 percent and 1 percent. Standard errors in parenthesis.

Note: The final panel includes 10,492 nonfinancial companies that filed tax returns under the same EIN and were selected by the SOI sampling process every tax year from 1993 to 2000 and their total assets did not increase by more than 10 times from one period to the next and did not file final returns in Tax Year 2000.

Following G&L, I stopped adding powers to the polynomial when the next higher power was statistically insignificant.

and the same or higher overall R-squared as the final panel did.²³

To test whether the negative tax coefficient related to the companies with extreme observations, I excluded from my sample firms that had total debt greater than 80 percent of total assets or firms that had any single asset equal to or greater than total assets. After these restrictions, my sample was reduced down to about 9,000 records. The tax coefficient was again negative and significant, with the rest of the control variables having the expected signs. Excluding those extreme observations reduced significantly the unobserved firm-specific error and raised the overall R-squared to 0.2 percent.

Since the negative relationship between taxes and capital structure seemed to be independent of the dependent variable and the sample, I turned my attention to the possibility of endogeneity bias between the dependent variable and the main regressor.²⁴To correct the possible bias, I constructed an exogenous instrument. The

instrument is the average tax rate faced by all firms in each time period if the interest deduction is added back to taxable income but the instrumental variable tax coefficient is again negative.

Since the instrument does not seem to correct the bias, I followed the example of Graham, Lemmon, and Schallheim and generated a second marginal tax rate proxy (**mrntint**) using taxable income before the interest deduction as a measure of the profits. I proceeded to estimate the log-linear models using fixed effects. Table 6 reports the results of these regressions. The fixed effects tax coefficients of all three regressions are positive and significant at the 1-percent level. The tax coefficient, for the total debt regression, was equal to 0.06. So, after using a modified measure of revenue, one that includes the interest deduction, I found a significant distortion on the corporate financial policy caused by taxation. I estimated that firms in the 39-percent tax bracket are

Table 6
Fixed Effects Regression Results

	Log(tdr)	Log(sdr)	Log(ldr)
Variables			
Log(dmrntint)	0.058** (0.006)	0.014** (0.004)	0.049** (0.005)
Log(rassts)	-1.831** (0.530)	-1.974** (0.410)	-0.344* (0.116)
Log(rassts) ²	0.213** (0.067)	0.215** (0.052)	-0.032* (0.011)
Log(rassts) ³	-0.012** (0.004)	-0.011** (0.003)	0.001* (0.0004)
Log(rassts) ⁴	0.0003** (0.0001)	0.0003** (0.00009)	-0.00002* (0.000007)
Log(rassts) ⁵	-0.000003 (0.000002)	-0.000003 (0.000002)	- (0.000002)
Log(dpr)	0.274** (0.007)	0.038** (0.005)	0.256** (0.006)
Log(landr)	0.156** (0.014)	0.010* (0.011)	0.160** (0.013)
Log(car)	-0.130** (0.006)	-0.086** (0.005)	-0.051** (0.005)
Log(intr)	0.320** (0.016)	-0.013** (0.012)	0.350** (0.014)
Log(arr)	-0.069** (0.007)	-0.027** (0.005)	-0.047** (0.006)
Constant	6.269** (1.633)	7.139** (1.265)	-0.043 (1.467)
Year Dummies	Yes	Yes	Yes
Obs.	83,936	83,936	83,936
R-squared	0.13	0.01	0.20

*and ** indicate significance levels at 5 percent and 1 percent. Standard errors in parenthesis.

Note: The final panel includes 10,492 nonfinancial companies that filed tax returns under the same EIN and were selected by the SOI sampling process every tax year from 1993 to 2000 and their total assets did not increase by more than 10 times from one period to the next and did not file final returns in Tax Year 2000.

Following G&L, I stopped adding powers to the polynomial when the next higher power was statistically insignificant.

forecasted to finance 1.5 percent more of their assets with debt than firms in the 15-percent tax bracket. Firms in the top tax bracket, large firms, are forecasted to finance 1.2 percent more of their assets with debt than small firms. The coefficients of the asset composition variables have the expected signs and are significant at the 1-percent level.

Dividing debt into short-term and long-term also produces very interesting results. The tax coefficient of the long-term debt regression is greater than the tax coefficient of the short-term regression, 0.049 compared to 0.013. These coefficients are drastically different from the aggregate data coefficients presented in Table 2. The coefficients of the asset composition variables for both the short-term and long-term regressions have the expected signs and are statistically significant, except for the land and intangible assets coefficients of the short-term regression that are statistically insignificant.²⁵ Firms with higher depreciable assets have higher long-term debt-to-assets ratios compared to their short-term debt ratios. Firms with higher ratios of cash-to-assets have higher short-term debt-to-assets ratios compared to their long-term debt ratios.

To get a better understanding of the effects of taxation on the financial policy of firms of different size, I divide my sample into small, intermediate, and large firms.²⁶ Small firms have lower debt-to-asset ratios than the rest of the firms, 26 percent of total assets compared to 31 percent for intermediate and large firms. The majority of that debt for all three categories is long-term debt, but, for small firms, long-term debt is a lower percentage of total debt. Large firms have the highest combined ratio of depreciable and intangible assets, with intermediate firms being a close second. The amount of cash firms hold is inversely related to their sizes. Firms in the lowest asset class hold more than one fifth of their assets in cash, while firms in the highest asset class hold only about 6 percent of their assets in cash. The progressiveness of the tax system is evident in both marginal tax rate proxies. The average marginal tax rates, for both proxies, increase as the asset classes rise. An additional dollar of taxable income increases the tax liability of large firms by more than 7 cents, 22.7 percent, whereas an additional dollar of taxable income increases that of small firms by 15.8 percent. The interest paid deduction has the highest

impact on the tax liability of the larger firms. If interest paid was not tax-deductible, then the 7 cents of additional tax liability for large firms would have been 10 cents. These findings are not surprising, since large firms hold more debt, but they give us a measure of the importance of the interest deduction as a tax shield.

The fixed effects regression results of the log-linear model for separate asset-sized classes are reported in Table 7. The dependent variable for the fixed effects regression is the marginal tax rate based on taxable income before the interest deduction (**mr_{int}**).²⁷ The estimated tax coefficients are: 0.057 for small firms, 0.055 for intermediate firms, and 0.085 for large firms. So, I found evidence of a positive relationship between taxation and corporate debt for all three types of firms. Contrary to the G&L findings, taxes had the largest effect on the use of debt for the largest firms, and the tax effect for intermediate firms is comparable to the tax effect for small firms. The coefficients of the majority of the control variables had the anticipated sign and were statistically significant.

Table 7
Fixed Effects Regression Results

	\$1 under \$10,000,000	\$10,000,000 under \$100,000,000	\$100,000,000 or more
	Log(tdr)	Log(tdr)	Log(tdr)
Variables			
Log(drtint)	0.057** (0.007)	0.055** (0.014)	0.085** (0.036)
Log(rassts)	-0.422** (0.101)	-2.807** (0.514)	-0.826** (0.159)
Log(rassts) ²	0.029** (0.007)	0.158** (0.031)	0.042** (0.009)
Log(rassts) ³	-0.0006** (0.0002)	-0.003** (0.0006)	-0.0007** (0.0001)
Log(dppr)	0.292** (0.008)	0.268** (0.013)	0.144** (0.021)
Log(landr)	0.156** (0.016)	0.192** (0.031)	0.118** (0.058)
Log(car)	-0.134** (0.007)	-0.108** (0.012)	-0.190** (0.022)
Log(intr)	0.378** (0.024)	0.307** (0.026)	0.232** (0.027)
Log(arr)	-0.095** (0.008)	0.037** (0.014)	-0.050** (0.021)
Constant	2.113 (0.444)	16.564** (2.883)	-5.479** (1.008)
Obs.	54,024	21,360	8,552
R-squared	0.17	0.09	0.10

*and ** indicate significance levels at 5 percent and 1 percent. Standard errors in parenthesis.

Note: Following G&L, I stopped adding powers to the polynomial when the next higher power was statistically insignificant.

Next, I divided debt into short-term and long-term, and I re-estimated the model. All tax coefficients were positive and statistically significant. The effect of taxation on the long-term debt of small firms was large when the effect on short-term debt was very small. The opposite was true for large firms, where the effect of taxation on short-term debt was approximately two times the effect on long-term debt. Finally, the effects of taxation on short-term and long-term debt for intermediate firms were approximately the same. I believe that these finding can be supported by intuition. Although small firms have relatively less long-term debt than intermediate and large firms, this debt doubles as debt tax shield. Large firms have more mature capital structures; they follow debt target level for their long-term borrowing and use short-term borrowing to create tax shields as needed. Summarizing my findings, I found evidence of a positive relationship between corporate taxation and the total debt ratios of small, intermediate, and large firms.

► Conclusion

Past empirical research on the effects of taxation on corporate financial policy has been limited, due to lack of data, to large publicly-traded firms or small closely-held partnerships. The more recent studies of the capital structure literature find a positive relationship between taxation and the debt levels of those firms. The only work that looks at the entire corporate population is a study by Gordon and Lee. They utilized an aggregate time-series dataset from 1950 to 1995 to find evidence that taxation increases the use of debt. In this study, I used the SOI aggregate and microdata files to research the effects of taxation on the corporate financial policy from Tax Years 1993 to 2000.

When using the aggregate dataset, my findings suggest that taxation in the 1990's still affected the financial policy of firms but to a somewhat lesser extent. I found that large firms in the 1990's finance 1.4 percent more of their assets with debt relative to the smaller firms. That it is a significant decrease compared to the 9.2 percent estimated by G&L. I believe that this decrease is in its entirety due to the lower tax rates faced by all firms and by the reduction in the gap between the tax rates faced by small versus large firms.

When using a firm-level dataset, and after isolating the unobserved firm-specific effects and using a modified measure of revenue, my findings suggest that there is a positive relationship between taxation and the use of corporate debt. Contrary to the G&L findings, taxes have the largest effect on the use of debt for the largest firms and a positive effect on the use of debt for intermediate firms.

Appendix

Definitions of Variables and Expected Signs

Dependent Variables

- Tdr** Ratio of total debt to total assets. Measures total debt as a percentage of total assets. Total debt is equal to the sum of mortgages, notes, bonds payable (Form 1120, page 4 balance sheet, lines 17 and 20).
- Sdr** Ratio of short-term to total assets. Measures short-term debt as a percentage of total assets. Short-term debt is equal to the sum of mortgages, notes, bonds payable in less than 1 year (Form 1120, page 4 balance sheet, line 17).
- Ldr** Ratio of long-term to total assets. Measures long-term debt as a percentage of total assets. Long-term debt is equal to the sum of mortgages, notes, bonds payable in 1 year or more (Form 1120, page 4 balance sheet, line 20).

Tax Variables

- Dmr** Equal to $mrt - ifmr$. Measures the tax incentive the firm has to use debt. (+)
- Mrt** Proxy for marginal rate using taxable income. The rate is set equal to the marginal statutory rate that applies to the highest dollar of taxable income (Form 1120, page 1, line 30). The rate is set to zero when taxable income is zero. (+)

Dmrtint Equal to mrtint minus ifmr. Measures the tax incentive the firm has to use debt. (+)

Mrtint Proxy for marginal rate using taxable income before the interest deduction. The rate is set equal to the marginal statutory rate that applies to the highest dollar of taxable income before interest deduction (Form 1120, page 1, lines 30 and 18). The rate is set to zero when taxable income before interest deduction is zero. (+)

Ifmr Proxy for yearly individual tax rate on interest income multiplied by the fraction of household assets held outside of pensions and life insurance. The yearly rate is the weighted average marginal tax rate reported in the SOI individual returns publication. (-)

Control Variables

Rassts Total assts (Form 1120, page 4 balance sheet, line 15d) deflated by CPI. Real total assets.

Dprrr Ratio of net depreciable assets to total assets. Net depreciable assets are equal to buildings and other depreciable assets less accumulated depreciation (Form 1120, page 4 balance sheet, lines 10 a (c) and b (c)). (+)

Landr Ratio of land to total assets. Land is equal to land net of any amortization (Form 1120, page 4 balance sheet, line 12). (+)

Car Ratio of cash to total assets (Form 1120, page 4 balance sheet, line 1(d)). (-)

Arr Ratio of trade notes and accounts receivable to total assets. Trade notes and accounts receivable are equal to trade notes and accounts receivable less allowance for bad debts (Form 1120, page 4 balance sheet, lines 2 a (c) and b (c)). (-)

Intr Ratio of intangible assets to total assets. Intangible assets are equal to intangible

assets (amortizable only) less accumulated amortization (Form 1120, page 4 balance sheet, lines 13 a (c) and b (c)). (+)

Yearly Variables

Ydmr Yearly average of dmr.

Imr Proxy personal marginal tax rate.

Tb Three-year Treasury Bill rate. Proxy for nominal interest rate.

Dj Average Dow Jones index deflated by GDP. Proxy for the business cycle.

► Endnotes

- 1 Source: Congressional Budget Office Web site; Table 3 Revenues by Major Source, 1962-2003.
- 2 Beginning in 2003, the maximum tax rates on qualified dividends have been lowered to 15 percent from 39.6 percent. For sales and other dispositions of property after May 5, 2003, the maximum tax rates on net capital gains have been lowered to 15 percent from 20 percent.
- 3 Although the ratios fluctuate from year to year, firms rely primarily on internal generated cash (retained earning plus depreciation) to finance new investments. Industry averages show that the ratio can range from 40 percent to 85 percent (Brealey and Myers, 2000).
- 4 The most widely used nondebt tax shields in Tax Year 2000 were: depreciation, compensation of officers, employee benefit programs, advertising, and contributions to pensions and profit-sharing plans.
- 5 In a later paper (1996), he adds two more acceptable marginal tax rate proxies, a trichotomous variable and the statutory marginal tax rate.
- 6 The top corporate tax rate for that time period ranged from a high of 52 percent, from 1952 to 1963, to a low of 34 percent, from 1988 to 1992.

- ⁷ The data are aggregated based on the end-of-year total assets reported in the balance sheet by each firm. For the studies used by Gordon and Lee, the number of asset classes ranged between ten and fourteen. For my dataset, there are eleven asset classes. The breakdown of the asset classes is: (1 under 0.1m), (0.1m under 0.25m), (0.25m under 0.5m), (0.5m under 1m), (1m under 5m), (5m under 10m), (10m under 25m), (25m under 50m), (50m under 100m), (100m under 250m), (250m or more), and (zero assets). The last asset class groups returns that had no ending assets, and was not used in my analysis.
- ⁸ The term domestic corporation refers to companies incorporated in the United States but does not necessarily imply that all their activities are domestic. For foreign corporations engaged in trade or business in the United States, only income that was considered effectively connected with the conduct of a trade or business in the United States was included in the statistics.
- ⁹ The Business Cycle Dating Committee of the National Bureau of Economic Research, November 26, 2001, reports that the longest expansion in the NBER chronology reached its peak in March of 2001.
- ¹⁰ Tax receipts are total income tax after credits reported on Table 1 of the Corporate Income Tax Returns Publication..
- ¹¹ The sample selection process is set up in such a manner that any firms selected into the sample in a given year will be selected again the next year, providing that the firm files a return using the same employer identification number (EIN) in the two years and that it falls into a stratum with the same or higher sampling rate. Note that a firm will usually change its EIN when it merges with another firm. For more detailed explanation of the sampling process, see Section 3 of the Corporate Income Tax Returns Publication.
- ¹² Such firms have unusually large amounts of debt and no taxable income.
- ¹³ Financial reporting usually follows the generally accepted accounting principles (GAAP) rules issued by the Financial Accounting Standards Board (FASB).
- ¹⁴ The use of book data is an issue for all prior literature, Auerbach and Poterba (1987) review pre TRA86 data and they report that the differences between the tax and book amounts reported by firms can be significant.
- ¹⁵ The intangible assets number maybe inflated by the Internet bubble.
- ¹⁶ My findings are in line with the historical marginal tax rates reported at the Tax Policy Center's Web site.
- ¹⁷ Proxy for yearly individual tax rate multiplied by the fraction of household assets held outside of pensions and life insurance. The yearly rate is the weighted average marginal tax rate reported in the SOI individual returns publication.
- ¹⁸ I set the tax incentive as the simple difference between the corporate marginal tax rate and the individual tax rate on interest income. Other literature is investigating the tradeoff and how the individual tax rate differences (dividends versus interest versus capital gain rates) are affecting capital structure, but this issue is beyond the scope of this paper.
- ¹⁹ The total debt is the sum of mortgages, notes bonds payable in less than 1 year and mortgages, notes bonds payable in 1 year or more.
- ²⁰ This is the only variable deflated using the Consumer Price Index (CPI); the rest of variables are in current dollars.
- ²¹ To estimate the model, following the work of Gentry (1994), I transformed all dependent, tax, and control variables by adding one to all observations. I did so because those variables have observations that are equal to zero. I also tried another model with the log of the total debt ratio as the dependent variable, but the log-liner model consistently produced the highest adjusted R-squared.

22 Originally, I thought that, due to the large number of observations in our panel, random effects may be the better choice than fixed effects, but the Hausman test rejected the random coefficients as inconsistent.

23 Because for these regressions I dropped observations based on the magnitude of the dependent variable, these results may be spuriously induced.

24 I also allowed for the possibility of dynamics of adjustment of the debt-over-asset ratio by including in the right-hand side of the empirical model a one-period lag of the ratios and estimating the model using the method of Arellano and Bond. The one-period lag coefficient was both positive and significant with the tax incentive still having a negative effect, but I found that the instrument variables, *dmr* and *dpr*, were correlated to some set of residuals and are not acceptable, and the model failed the Sargan test of overidentifying restrictions.

25 The time dummy coefficients for these regressions were statistically insignificant; so, I did not estimate the second equation.

26 I decided against using the thirteen SOI asset classes because their breakouts were too detailed. My breakouts, based on yearend total assets are: small firms, less than \$10,000,000; intermediate firms, \$10,000,000 less than \$100,000,000; and large firms, \$100,000,000 or more.

27 In order to retain the panel aspects of my datasets and because firms over the eight years time-series moved in and out of asset classes I assigned to all eight observation of each firm the same asset class based on the firms' 1996 year-end total assets.

► References

Ayers, Benjamin C.; Cloyd, C. Bryan; and Robinson, John R. (2001), "The Influence of Income Taxes on the Use of Inside and Outside Debt By Small Businesses," *National Tax Journal*, Volume 54, Number 1, pp. 27-55.

Auerbach, A. (1985), "Real Determinants of Corporate Leverage," in *Corporate Capital Structures in the United States*, edited by B. Friedman, University of Chicago Press, Chicago, pp. 125-143.

Auerbach, A. and Poterba, J. (1987), "Tax Loss Carryforwards and Corporate Tax Incentives," in *The Effects of Taxation on Capital Accumulation*, edited by Martin S. Feldstein, University of Chicago Press, Chicago, pp. 305-337.

Bradley, M.; Jarrell, G.; and Kim, E. (1984), "On the Existence of an Optimal Capital Structure: Theory and Evidence," *Journal of Finance*, Volume 39, pp. 857-878.

Brealey, Richard A. and Myers, Stewart C. (2000), "Principles of Corporate Finance, Sixth Edition," McGraw-Hill Irwin.

Carroll, Robert; Hassett, Kevin; Mackie III, James B., "The Effect of Dividend Tax Relief on Investment Incentives," *National Tax Journal*, Volume 56, Number 3, pp. 629-651.

Cloyd, C. Bryan; Limberg, Steven T.; Robinson, John R. (1997), *The Impact of Federal Taxes on the Debt-Equity Structure of Closely-Held Corporations*, *National Tax Journal*, Volume 50, Number 2, pp. 261-277.

DeAngelo, Harry and Masulis, Ronald W. (1980), "Optimal Capital Structure Under Corporate and Personal Taxation," *Journal of Financial Economics*, Volume 8, Number 1, pp. 3-29.

Feldstein, M. (1995), "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *The Journal of Political Economy*, Volume 103, Number 3, pp. 551-572.

Gaver, Jennifer J. and Gaver, Kenneth M. (1993), "Additional Evidence on the Association between the Investment Opportunity Set and Corporate Financing, Dividend, and Compensation Policies," *Journal of Accounting and Economics*, Volume 16, Number 1/2/3, pp. 125-160.

- Gentry, William M. (1994), "Taxes, Financial Decisions and Organizational Form," *Journal of Public Economics*, Volume 53, Number 2, pp. 223-244.
- Gordon, Roger H., and Young Lee (2001), "Do Taxes Affect Corporate Debt Policy? Evidence from U.S. Corporate Tax Return Data," *Journal of Public Economics*, Volume 82, pp. 195-224.
- Graham, John R. (1996), "Debt and the Marginal Tax Rate," *Journal of Financial Economics*, Volume 41, Number 1, pp. 41-73.
- Graham, John R. (1996), "Proxies for the Corporate Marginal Tax Rate," *Journal of Financial Economics*, Volume 42, Number 1, pp. 187-221.
- Graham, John R. (1998), "Do Personal Taxes affect Corporate Financing Decisions?"
- Graham, John R. (2000), "How Big are the Tax Benefits of Debt," *Journal of Financial Economics*, Volume 55, Number 5, pp. 1901-1941.
- Graham, John R.; Lemmon, Michael L.; Schallheim, James S. (1998), "Debt, Leases, Taxes, and the Endogeneity of Corporate Tax Status," *The Journal of Finance*, Volume 53, Number 1, pp. 131-162.
- Gravelle, Jane and Kotlikoff, L. (1989), "The Incidence and Efficiency Costs of Corporate Taxation When Corporate and Noncorporate Firms Produce the Same Good," *The Journal of Political Economy*, Volume 97, Number 4, pp. 749-780.
- Gravelle, Jane (1995), "The Corporate Income Tax: Economic Issues and Policy Options," *The National Tax Journal*, Volume 48, Number 2, pp. 267-277.
- Gropp, Reint E. (1997), "The Effect of Expected Effective Corporate Tax Rates on Incremental Financing Decisions," *International Monetary Fund Staff Paper*, Volume 44, Number 4, pp. 485-509.
- Hadi, Ali S. (1992), "Identifying Multiple Outliers in Multivariate Data," *Journal of the Royal Statistical Society. Series B (Methodological)*, Volume 54, Number 3, pp. 761-771.
- Hinkins, Susan; Jones, Homer; and Scheuren, Fritz (1988), "Design Modifications for the SOI Corporate Sample: Balancing Multiple Objectives," 1988 Proceedings of the Survey Research Methods Section, American Statistical Association.
- Internal Revenue Service, *Statistics of Income, Corporate Source Book*, Washington DC, various issues.
- Internal Revenue Service, *Statistics of Income, Corporation Income Tax Returns*, Washington DC, various issues.
- Internal Revenue Service, *Statistics of Income, Individual Income Tax Returns*, Washington DC, various issues.
- MacKie-Mason, Jeffrey K. (1990), "Do Taxes Affect Corporate Financing Decisions?," *The Journal of Finance*, Volume 45, Number 5, pp. 1471-1493.
- MacKie-Mason, Jeffrey K. and Gordon, Roger H. (1997), "How Much do Taxes Discourage Incorporation," *The Journal of Finance*, Volume 52, Number 2, pp. 477-505.
- Miller, Merton H. (1977), "Debt and Taxes," *Journal of Finance*, Volume 32, pp. 261-275.
- Mills, Lillian F.; Newberry, Kaye J.; Trautman, William B. (2002), "Trends in Book-Tax Income and Balance Sheet Differences," *Tax Notes*, August 19, pp. 1109-1124.
- Mills, Lillian F. and Newberry, Kaye J. (2004), "Firms' Off-Balance Sheet Financing: Evidence from Their Book-Tax Reporting Differences," working paper.
- Modigliani, Franco and Miller, Merton H. (1963), "Corporate Income Taxes and the Cost of Capital: A Correction," *American Economic Review*, Volume 53, Number 3, pp. 433-443.

- Newberry, Kaye J. (1998), "Foreign Tax Credit Limitations and Capital Structure Decisions," *Journal of Accounting Research* Volume 36, Number 1, pp. 157-166.
- Omer, Thomas C. and Terrando, William (1999), "The Effect of Risk and Tax Differences on Corporate and Limited Partnership Capital Structure," *National Tax Journal*, Volume 52, Number 4, pp. 699-715.
- Plesko, George A. (2002), "Reconciling Corporate Book and Tax Net Income, Tax Years 1996-1998," *SOI Bulletin* (Spring 2002), pp.1-16.
- Plesko, George A. (2004), "Reconciling Corporate Book and Tax Net Income, Tax Years 1995-2001," Data Release
- Rosen, Harvey S. (2002), "Public Finance, Sixth Edition," McGraw-Hill Irwin.
- Scholes, Myron S. and Wolfson, Mark A. (1992), "Taxes and Business Strategy," Prentice-Hall, Englewood Cliffs, NJ.
- Scholes, Myron S.; Wolfson, Mark A.; Erickson, Merle; Maydew, Edward L.; and Shevlin, Terry (2005), "Taxes and Business Strategy, A Planning Approach," Prentice-Hall, Upper Saddle River, NJ.

4



Issues in Quantifying Measurement Error

Scali ♦ Testa ♦ Kahr ♦ Strudler
Davitian
Singmaster ♦ Redmiles

Measuring Nonsampling Error in the Statistics of Income Individual Tax Return Study

Jana Scali, Valerie Testa, Maureen Kahr, and Michael Strudler, Internal Revenue Service

Data collection for the SOI Individual Study begins with a sample of administrative tax records. While the sample is being transcribed, small subsamples of returns are randomly chosen and independently transcribed and processed for a quality evaluation. The IRS Statistics of Income (SOI) Division has an Individual Systematic Improvement (ISI) System which is the tool used to create the quality review sample and improve the Individual Tax Return Study data. The purpose of this paper is to estimate a component of nonsampling error in the SOI Individual Study. The data from the quality review process is used for this purpose.

The paper is organized as follows. We describe SOI's Individual sample design along with some sources of nonsampling error. We describe the editing process and the Individual Systematic Improvement (ISI) System used by SOI to evaluate and improve the quality of the Individual 1040 Program. We describe the study and its limitations. We explain the model used to estimate nonsampling error. We show the Index of Inconsistency. We cover the Intra-Editor Correlation Coefficient and Design Effect by element followed by conclusions.

► Individual Sample and Nonsampling Error Description

The statistics for the SOI Individual Study are estimates from a probability sample of unaudited Individual Income Tax Returns filed by U.S. citizens and residents during Calendar Year 2004. The estimates represent all returns filed for Tax Year 2003 with a small number representing prior years. For Tax Year 2003, some 184,988 returns were sampled from a population of 131,291,334.

The sample consists of two parts. The first part is a stratified probability sample, in which the population of tax returns is classified into subpopulations, called strata, and a sample is randomly selected independently

from each stratum. Strata are defined by the type of return submitted by the taxpayer. A Bernoulli sample is independently selected from each stratum with rates ranging from .05 percent to 100 percent. The second part of the sample is a random sample based on the primary taxpayer's Social Security number. If the last four digits of the primary taxpayer's Social Security number listed on the tax return equals one of five predetermined endings, then the tax return is included in the sample.

The quality of a sample estimator is a function of both sampling and nonsampling errors. Sampling errors arise due to drawing a probability sample rather than conducting a census. Nonsampling errors are due to data collection and processing procedures. They can be the result of misleading definitions and concepts or defective methods of data collection, tabulation, and coding. Nonsampling errors may increase with sample size, and, if not properly controlled, they can be more damaging to a study than sampling errors.

There are four components of nonsampling error. Coverage or frame errors occur when someone does not file a tax return. Nonresponse errors (missing data) arise when the Statistics of Income Division is unable to obtain the tax return because another function within the Internal Revenue Service has the return. Measurement errors are differences in the reported and the actual values. These errors are taxpayer errors. Processing errors occur at the data processing stage. They include editing, coding, data entry, and programming errors. This paper will describe and measure processing errors, which arise due to the following factors:

1. Lack of trained and experienced editors including quality supervisors.
2. Errors in data processing operations such as coding, keying, verification, and tabulation.
3. Procedural, Systemic, or Organizational Defects such as improper instructions, in-

adequate training, and insufficient time to complete a return.

Nonsampling errors are very important to measure because they can cause large biases and produce unreliable estimates if not controlled. By following the correct procedures during sample selection through the analysis of results, nonsampling errors can be controlled and dramatically decreased.

► SOI Editing and Quality Review Processes

For SOI purposes, when we mention editing, it refers to the process of an individual transcribing data items or elements from the tax return into our database. An element is a specific line item from a tax return. The individual transcribing the data is referred to as an editor. For the SOI Individual Study, 97 editors at four IRS Submission Processing Centers edited data from Individual income tax returns selected for the 2003 SOI sample. The data extracted come from Forms 1040, 1040A, and 1040EZ individual income tax returns and approximately 45 associated forms and schedules.

To assist the editors in this process, SOI’s National Office analysts in Washington, DC, implement various procedures to make the edited data adhere to individual tax standards and to try to keep the editing process as consistent as possible across the four centers. For example, the editors receive extensive training on the data editing process and correction procedures before they begin editing individual tax return data for the SOI sample. Then, as data are edited, numerous computerized tests are performed on the extracted data to ensure that certain accounting conditions are satisfied and that data are consistent across forms. All of these computerized tests are reviewed and tested by National Office staff prior to data extraction in a process called Systems Acceptability Testing. Various utilities and help features to aid in the edit process are also built into the computer edit system. For instance, there are utilities that list valid codes and definitions for a particular item. In addition, there is a feature that allows data from the previous year’s tax return to be viewed. There is also a comprehensive editing manual that contains detailed instructions and procedures that editors are expected to

follow while transcribing and correcting the tax return data. The editing manual for the 2003 sample was just over 600 pages.

During data editing, a simple random sample of one or two returns each week is selected for each editor for regular quality review. The goal is to have approximately 50 returns per editor selected for quality review over the course of the editing of the sample. The purpose of the quality review is to assess the accuracy of the data, evaluate the work of the editor, and look for improvement opportunities in the editing process. When an editor’s return is randomly selected for quality review, a different editor from the same team independently re-edits the return. The two edits of the return are then compared line by line, and discrepancies between the two edits, above a certain tolerance, are stored in the SOI database. For money amount fields, the tolerance is \$10; so, money amount fields that differ by \$10 or less are not included. However, there is no tolerance for character and code fields. The next step is for a lead editor to review the discrepancies and determine the correct value: the first editor’s value, the second editor’s value, both, or neither. During the process of reviewing discrepancies, if the first editor value is determined to be incorrect, it is corrected, and the error is charged to the first editor. Then, the reason for the error is determined and coded. There are 32 types of errors; the six most common are shown below.

Table 1.--Types of Errors

Type of Error	Description
Affected Entry	Item was incorrect due to an incorrect related item.
Improper Allocation	An amount that should have been allocated to another item was not moved or was moved incorrectly.
Incorrect Amount	An incorrect amount was entered.
Entry on Omitted Form	An item was not edited because the form or schedule was not edited.
Omitted Entry	A blank or zero item should have had an entry.
Interpretation	Item was edited incorrectly due to being interpreted in a different way than expected.

Affected entries were the most frequent type of error. These types of error occur when multiple errors are the result of one line item being incorrect. For example, if one line item on Form 1040, such as Salaries, Wages, and Tips, is edited incorrectly, then this causes other line items that use that amount, such as total income, adjusted gross income, and taxable income, to also be incorrect.

Table 2.--Number of Errors, by Element

Element	Number of Errors	Error Rate
Salaries, Wages, and Tips	41	0.014
Other Income	51	0.018
Total Credits	13	0.004
Income Tax After Credits	20	0.007
Balance Due / Overpayment	31	0.011
Total Depreciation Deduction	42	0.038
Net Investment Income ¹	19	0.023
Tentative Alternative Minimum Tax	18	0.014
Rental Real Estate and Other Passive Activity Net Income/Loss	21	0.027
Other Taxes ²	28	0.028
Investment Interest ²	11	0.011
Other Investment Interest ²	11	0.011
Contract Labor Expense ³	24	0.021
Utilities Expense ³	27	0.023
Sole Proprietorship Other Expenses ³	109	0.093
Net Profit/Loss from Business ³	20	0.017
Long-Term Gains/Losses from Sale of Capital Assets	19	0.010
Partnership Nonpassive Income	15	0.008
S Corporation Nonpassive Loss	17	0.009

¹ Reported on Form 4952

² Reported on Schedule A

³ Reported on Schedule C

► Study and Limitations

A total of 2,907 returns was selected for regular quality review. Using data from these quality review returns, variables of interest were chosen for this paper. The variables are Salaries, Wages, and Tips; Other Income; Total Credits; Income Tax After Credits; Balance Due/Overpayment; Total Depreciation Deduction; Net Investment Income; Tentative Alternative Minimum Tax; Rental Real Estate and Other Passive Activity Net Income/Loss; Other Taxes; Investment Interest; Other Investment Interest; Contract Labor Expense; Utilities Expense; Sole Proprietorship Other Expenses; Net Profit/Loss from Business; Long-Term Gains/Losses from Sale of Capital Assets; Partnership Nonpassive Income; and S Corporation Nonpassive Loss. These items were chosen by the subject-matter specialists because of the combination of a high number of editor errors and interest in the items.

All returns sampled for the Statistics of Income Individual Tax Return Study are subject to consistency tests. Subject-matter analysts review any returns that fail the consistency tests before the values are considered final. As a result of this review, some values are adjusted; however, there is no information available on these adjustments. The adjusted values replace the original ones.

Several statistics are presented in this discussion of nonsampling error. Net Difference Rate (NDR), t-test, and Index of Inconsistency (IOI) use only the quality review data, while Design Effect (DEFF) uses the entire sample.

► Simple Response Variance Model

We will consider a simple model that was first proposed by Hansen et al. (1952) and Sukhatme and Seth (1952) for measurement error. Their model specifies that the true value μ_i (the final value) is different from the observed value y_i (the editor's value) by an unobserved additive error term ϵ_i . For unit i ($i = 1, 2, \dots, n$), the assumed model is

$$y_i = \mu_i + \varepsilon_i \quad (5.1)$$

While we did not measure response error, we adopted these models to our data to measure processing error and estimate bias. The distribution of the editor error variable ε_i is conceptual; it could be viewed as sampling from a hypothetical population of errors. Thus, the further assumptions for model (5.1) are

$$\begin{aligned} E[\varepsilon_i | i] &= B_i \neq 0 \\ Var[\varepsilon_i | i] &= \sigma_i^2 \\ E[\sigma_i^2] &= \sigma^2 \\ Cov[\varepsilon_i, \varepsilon_j] &= 0, i \neq j. \end{aligned}$$

In words, a systematic bias exists because the mean of the errors is not zero and the error variances are not equal. Also, all errors are uncorrelated. This means that errors made to a return by the first or second editor do not affect other returns edited in the same edit period.

Following Brick et al. (1996), we will assume that the quality review sample is an unrestricted simple random sample, thus

$$\begin{aligned} E[\mu_i] &= \bar{\mu} \\ V[\mu_i] &= \sigma_\mu^2 \\ Cov[\mu_i, \mu_j] &= 0, i \neq j. \end{aligned}$$

Under model (5.1), we assume that the first editor's error term no longer averages to zero, possibly due to editor bias, defined as

$$B = \sum_{i=1}^N (y_i - \mu_i) \quad (5.2)$$

The bias can be estimated by the *Net Difference Rate* (NDR), which is given by

$$NDR = \bar{y} - \bar{\mu} \quad (5.3)$$

where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, $\bar{\mu} = \frac{1}{n} \sum_{i=1}^n \mu_i$, and n is the sample size. It can be shown that, if μ_i is the true value, then the expected value of the NDR is the bias, and its variance exists (Biemer and Atkinson, 1992). Table 3 shows the estimated NDR and t-test values.

Table 3.--Net Difference Rate and T-Test, by Element

Element	NDR	t-test
Salaries, Wages, and Tips	5,159	0.97
Other Income	-5,895	1.11
Total Credits	3	1.73
Income Tax After Credits	-3	0.76
Balance Due	9	0.45
Overpayment	-19	1.30
Total Depreciation Deduction	-1,016	2.43
Net Investment Income ¹	-2,820	0.88
Tentative Alternative Minimum Tax	-3,144	1.34
Rental Real Estate and Other Passive Activity Net Income/Loss	1,581	1.13
Other Taxes ²	186	1.41
Investment Interest ²	-79	0.61
Other Investment Interest ²	79	0.61
Contract Labor Expense ³	-1,109	1.57
Utilities Expense ³	-43	0.15
Profit/Loss from Business Other Expenses ³	-670	0.18
Net Profit/Loss from Business ³	842	0.59
Long-Term Gains from Sale of Capital Assets	-6,524	0.99
Long-Term Losses from Sale of Capital Assets	-5,828	2.23
Partnership Nonpassive Income	461	1.68
S Corporation Nonpassive Loss	-512	1.82

¹ Reported on Form 4952

² Reported on Schedule A

³ Reported on Schedule C

Since the values for the t-test are greater than 1.96 for Total Depreciation Deduction (2.43) and Long-Term Losses from Sale of Capital Assets (2.23), these items

have significant bias. This means that the editors are editing these fields differently.

► **Index of Inconsistency**

Index of Inconsistency and Design Effect cannot be calculated for those elements with a significant bias because these equations assume the elements have zero bias. For the remaining elements in Table 3 with insignificant bias, we assume the bias is zero, $E[\epsilon_i | i] = B_i = 0$, and calculate the following statistics:

$$\begin{aligned} Var[\bar{y}] &= Var[\bar{\mu}] + \frac{\sigma^2}{n} \\ &= SV + EV. \end{aligned} \tag{6.1}$$

The *sampling variance*, SV, is the ordinary variance with no editor error. The *editor variance*, EV, is the variability of returns averaged over conceptual repetitions of editing under the same conditions.

Table 4.--Index of Inconsistency, by Element

Element	IOI
Salaries, Wages, and Tips	0.00184
Other Income	0.18419
Total Credits	0.00000
Income Tax After Credits	0.00000
Balance Due	0.00000
Overpayment	0.00000
Net Investment Income ¹	0.00014
Tentative Alternative Minimum Tax	0.00086
Rental Real Estate and Other Passive Activity Net Income/Loss	0.00009
Other Taxes ²	0.00034
Investment Interest ²	0.00002
Other Investment Interest ²	0.05339
Contract Labor Expense ³	0.00743
Utilities Expense ³	0.00870
Profit/Loss from Business Other Expenses ³	0.01072
Net Profit/Loss from Business ³	0.00476
Long-Term Gains from Sale of Capital Assets	0.00171
Partnership Nonpassive Income	0.00005
S Corporation Nonpassive Loss	0.00007

¹ Reported on Form 4952

² Reported on Schedule A

³ Reported on Schedule C

Hansen et al. (1964) define the *Index of Inconsistency*

(IOI) as
$$IOI = \frac{EV}{SV + EV}, \tag{6.2}$$

which we use to estimate the proportion of random errors associated with editor error in total variance. The IOI obtains values between 0 and 1.0. Estimated IOI values are shown in the Table 4.

Yu et al. (2000) define that the reliability of the data can be expressed in this equation:

$$r = 1 - IOI \tag{6.3}$$

In other words, the reliability of an element is the information without the inconsistent portion. All of the elements, except for Other Income, have index of inconsistencies less than .01, which means that they are over 99-percent reliable. Other Income, with the highest Index of Inconsistency (0.18419), is the element with the least amount of reliability, 82-percent, and the largest amount of processing errors.

► **Design Effect**

By treating the editors as clusters, the Intra-Editor Correlation Coefficient and Design Effect can be used to measure the editor effect on the variance if the sample was an unrestricted simple random sample.

The Intra-Editor Correlation Coefficient (ρ) measures the correlation between the values that is due to editor error. It is a measure of the similarity of the editors in the way the editors edit a specific element.

Kish (1965) defines the Intra-Editor Correlation Coefficient as

$$\rho_{ed} = \frac{S_{between}^2 - \frac{S_{within}^2}{B}}{S^2} \tag{7.1}$$

The ideal range is 0 to 0.1 which indicates no editor variance.

Once the Intra-Editor Correlation Coefficient is calculated, we can use ρ_{ed} to determine the design effect. Design Effect is a measurement of the degree to which an estimate is affected by editor variance,

$$deff = 1 + (B - 1)\rho_{ed} \quad (7.2)$$

where B is the average editor workload or 1,728 returns.

An Editor Design Effect of 1 indicates no increase in variance resulting from the editors. A value of 2 indicates that the variance is doubled.

As Table 5 shows, Overpayment has the largest intra-editor correlation coefficient (0.0124) and design effect (22.40), but one of the smallest Coefficients of Variation. The design effect represents the inflation of variation of the sample if it were treated as a simple random sample with replacement. The design effect for Overpayment can be reduced if editor workload is reduced, but, because the CV is so low, reducing the editor workload in order to reduce the design effect would not be worth the cost.

Table 5.--Design Effect and Coefficients of Variation, by Element

Element	ρ	Design Effect	CV
Salaries, Wages, and Tips	0.0041	8.16	0.21%
Other Income	0.0000	1.01	3.92%
Balance Due	0.0023	5.04	0.81%
Overpayment	0.0124	22.40	0.38%
Other Taxes ¹	0.0004	1.62	4.46%
Investment Interest ¹	0.0005	1.94	1.73%
Long-Term Gains from Sale of Capital Assets	0.0053	10.22	1.36%

¹ Reported on Schedule A

► Conclusions

This paper was written to estimate the nonsampling error and measure the reliability of the Individual Tax Return Study. Quality Review data were used to measure processing errors and determine how editor error affects the accuracy of specific elements.

From the calculations of Net Difference Rate and Index of Inconsistency, we can conclude that bias can be significantly reduced if we work on the editing procedures for Long-Term Gains/Losses from Sale of Capital Assets, Total Depreciation Deduction, and Other Income. Most of the time, processing errors of several elements can be reduced if the editors concentrate on one element. For example, Other Income has one of the largest Net Difference Rates and the largest Index of Inconsistency, but the smallest Design Effect. In other words, more editors than desired are consistently editing the element incorrectly. Since editors are making similar errors, the data quality can be increased if clearer directions or explanations in the edit manuals are provided. Also, more intense training and examples might lead to smaller processing errors. In addition, this will improve the large positive Net Difference Amount, or overestimate, for Salaries, Wages, and Tips because Other Income allocation is most likely the cause of this problem.

Overall, the editors are producing high-quality work with the exception of specific elements that require more than just transcribing. From the research in this paper, improvement opportunities have become available, and subject-matter analysts can put procedures in place to check the editing quality of specific elements. In addition, editing procedures for elements with high processing errors can be revised and clarified to enhance the accuracy and reliability of the Individual Tax Return Study.

► References

- Biemer, P. and Atkinson, D. (1992), "Estimation of Measurement Bias Using a Model Prediction Approach," *1992 Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 64-73.
- Brick, M.; Kim, K.; Nolin, M.J.; and Collins, M. (1996), "Estimation of Response Bias in the NHES: 95 Adult Education Survey," *Working Paper Series*, National Center for Education Statistics, Washington, DC.

Hansen, M.H.; Hurwitz, W.N.; and Madow, W.G. (1952), *Sample Survey Methods and Theory*, John Wiley and Sons, New York, Volume II.

Hansen, M.H.; Hurwitz, W.N.; and Pritzker, L. (1964), "The estimation and interpretation of gross differences and the simple response variance," *Contribution to statistics*, in C.R. Rao (editor) Pergamon Press, Oxford, and Statistical Publishing Society, Calcutta, pp. 111-136.

Kish, Leslie (1965), *Survey Sampling*, John Wiley and Sons, New York, pp. 164-178.

Sukhatme, P.V. and Seth, G.R. (1952), "Nonsampling errors in surveys," *Journal of Indian Society of Agricultural Statistics*, pp. 5-51.

Yu, C.H.; Ohlund, B.; DiGangi, S.; and Jannasch-Pennell, A. (2000), "Estimating the reliability of self-reporting data for Web-based instruction," Arizona State University, Instruction and Research Support.

Corporation Supercritical Cases: How Do Imputed Returns on the Corporate File Compare to the Actual Returns?

Lucy Davitian, Internal Revenue Service

Statistics of Income (SOI) corporation “supercritical” cases are certain large corporations that SOI samples at the 100-percent rate. These supercritical cases account for 58 percent of the total assets of the corporation study while comprising only .03 percent of the total corporation returns; thus, their absence from the Corporation Study would affect the final statistics. Any unavailable returns must therefore be added to the file to protect the validity of the SOI Corporation Study. One method of adding those missing data is to collect the information through surveys sent directly to the corporations. Data collected are then used to create alternate records in the file through various imputation routines. These alternate records are later replaced with the actual return when that information is secured. This paper will give a brief overview of critical cases and the survey process, compare the data in the alternate records to that of the actual returns, evaluate the accuracy of the imputation routines, and make subsequent recommendations for changes to improve data quality where necessary.

► Background on Critical Cases

The critical case list for each program year is created based on the critical cases in the last two program years of the corporation study.¹ In general, there are three levels of critical case classifications: the top level, or supercritical cases, which are the largest corporations; critical cases that comprise 5 percent or more of the total assets of the industry they are classified in; and all other critical cases. The classifications are made based on three different criteria: type of return filed, industry classification, and corporation total assets.

During SOI’s corporation Advance Data processing (beginning after the critical case list creation in December and running through April), all supercritical cases that are unavailable for statistical processing are searched for. Clerks at the IRS submission processing centers in Ogden and Cincinnati search for information on these critical cases. If the clerks cannot secure these returns,

they provide information to assist National Office (N.O.) analysts with additional research. N.O. analysts then use this information to verify mergers between companies or other reasons why the return may be unavailable for SOI’s processing.

Companies that are found to have no tax liability for the tax year, are liquidated or bankrupt, have changed Employer Identification Numbers (EIN’s), or merged into other companies are suppressed from the study file and will not appear on future critical case lists. Between program years 1997 and 2002, an average of 85 supercritical cases were suppressed (see Table 1), thus reducing the number of critical cases that are researched or included in subsequent studies.

Table 1.--Number of Suppressed Critical Cases

Program Year	Total Super Criticals	Number Suppressed
1997	1,006	55
1998	1,160	70
1999	1,416	93
2000	1,622	95
2001	1,584	109
2002	1,595	85

However, if there is no evidence to conclude that a return does not have a filing requirement for the current tax year, and the returns are not located during this advance data period, alternate records, also called added records, are created as a substitute for the unavailable returns. There are four classifications of added records based on the type of information SOI has available to process the corporation return. The most ideal added record is one that uses data from both the IRS Business Master File (BMF)² and a survey sent to the corporation since it contains the most current information on the corporation return. The next level of preference is the use of BMF information only. Then, there are added records created using only survey information. Lastly,

records created based only on prior-year information are included when no other current information is sufficient to create the added record. For the purposes of this paper, only the added records created from survey information will be discussed and analyzed.

► **Filling in for Missing Information: Overview of the Survey Process**

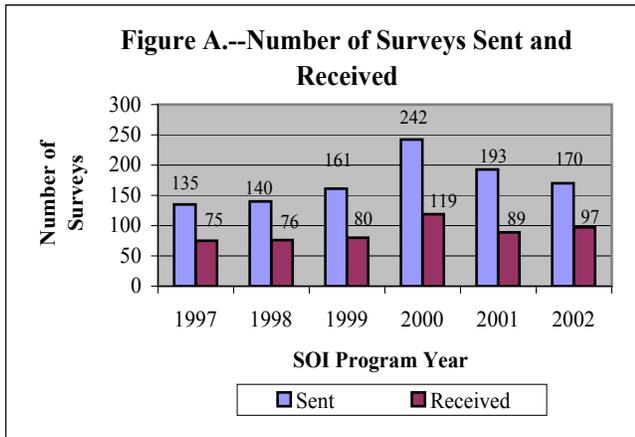
The surveys that are sent to missing corporations initially go through an approval process (renewed every 5 years) through the Office of Management and Budget (OMB). The approval process considers taxpayer burden in filling out and returning the survey, as well as other factors to ensure it meets established OMB guidelines. Once approved for distribution, the survey is sent with an accompanying memorandum signed by the Director of the Statistics of Income Division that states the nature of the survey and informs the corporations that the survey is voluntary. It also notes that the information collected is for statistical use only and not the result of any ongoing or forthcoming examination of the corporation's income tax return. The survey lists approximately sixteen data items from the corporation's tax return relevant to the SOI program year, and asks that the data be returned within 3 weeks of receipt.

Once a survey is returned, SOI processes the data to create an added record, also called a short-edit, in the file until the actual return can be processed. The survey data items are manually typed in, and the program then uses these numbers to calculate the remainder of the current-year amounts (those not included in the survey).³ It does so by using current and prior-year amounts to create ratios that are used to help fill in for the missing data. The returns are then processed through the normal edit function used on all corporate returns to ensure that the total amounts balance and no additional errors are present. Returns created through this short-edit process are then given a weight and included in the study file.

After the close of the Advance Data file and throughout the remainder of the program year (for the 2002 program, file closeout was November 2004), these short-edits (and all types of added records) are replaced once the actual returns are available for SOI processing.

► **Survey Statistics**

Since 1997, an average of 173 surveys have been sent each year to corporations, with average response rates of 51 percent (see Figure A). Over the course of the program years analyzed, many attempts were made to try to increase the response rates. For the 2000 program year, however, there was a higher number of unavailable returns. This was due to the IRS processing center realignments, which resulted in SOI's processing of corporate returns being scaled down from four centers to two. This also created some confusion and resulted in many corporate tax departments still mailing their returns to the same centers as in prior years. This caused a need for the returns to be shipped from these centers to the newly realigned ones. The changes in these processes and the delays they caused directly affected SOI's ability to process the returns for the Advance Data. For the 2001 study, to try to avoid a possible repeat of the prior year, the surveys were mailed earlier. Unfortunately, since many of the corporations were filing extensions, we did not receive as many surveys back until after the extension period was over. Also, in the wake of the September 11 attacks, longer extension periods were granted to corporations that were directly affected by the attacks, and many of these companies were either no longer in business or had portions of their businesses that were dissolved. Since some of the tax departments of these corporations were in New York City, the addresses that the surveys would normally be sent to were no longer valid. This directly attributed to the decline in the number of surveys sent, as well as the number of survey responses. In addition to these challenges with the earlier mailing, we observed the need to call more corporations to obtain the data; they had either misplaced the initial survey or were too busy at the time to fill it out within the 3-week timeframe mentioned in the memo. With that in mind, for the 2002 program, we mailed the surveys a few weeks later than we had for the 2001 study and noticed better response rates and fewer followup calls being necessary to secure the survey data, though, given the circumstances for the prior year files, we will need to evaluate this method further.



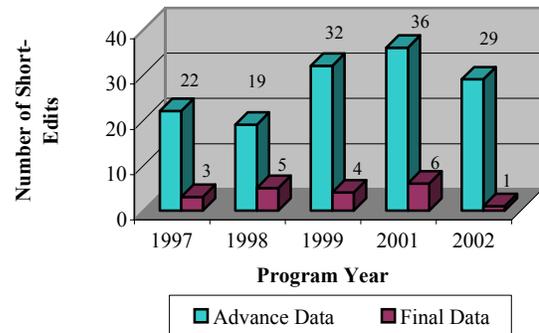
Each year, there is also an attempt to try to increase the number of survey responses and decrease the use of prior-year data. However, despite our efforts, there are still many instances of nonresponse. One reason is that the surveys are voluntary; many corporations do not return the data or do so weeks or months after the specified timeframe. Even though the survey states it has nothing to do with an ongoing or forthcoming investigation of the return, many corporate tax departments are hesitant to submit data that might catch someone's attention--especially if they do not have to. In such nonresponse cases, we attempt to contact the company's tax department directly to see if we can obtain the information we need. This usually causes the corporation to question the need for filling out a survey when it has already filed a return. We explain why the survey is necessary, and that the Statistics of Income Division, while under the IRS, is a statistical organization that uses the data for statistical purposes only and obtains the tax data after the other IRS processing functions. Another reason the survey may not be returned is due to various filing extensions that many corporations file. Depending on the date of the closeout of the Advance Data file, the company might not have enough time to provide the data needed.

The response rates mentioned above also do not consider those corporations that were sent surveys but did not respond because the corporation filed as a subsidiary of another; there are times that our initial research either does not provide all the information about the corporation or it does so after we have already mailed out the survey. In addition, given the time it takes between when

the survey is mailed and returned to SOI, the return may have been selected for processing during subsequent selection cycles and edited before imputation of the survey data is necessary. In such cases, we make no attempt to contact the corporation in nonresponse cases and if the taxpayer calls to ask about the survey, we inform them that the survey information is no longer needed.

Between SOI Program Years 1997 and 2002, of the surveys received, an average of 28 (about 30 percent of all added records) were used in the Advance Data file (see Figure B).⁴ By the end of the Final Data closeout, only an average of 4 remained in the file (19 percent of all added records), the others having been replaced with the actual returns.

Figure B.--Short-Edits Created with Survey Data



► Comparisons of Survey Data to Edited Returns

During Advance Data, the short-edit records accounted for 0.6 percent of the total assets for all corporations in the study file, nearly \$288.7 billion. In addition, all added records comprised 2.7 percent of total assets, or \$1.4 trillion. While the percentages themselves are small, we can see that the missing data could potentially grossly underestimate the total assets in the overall file as well as all the other data items that are collected. To further examine the impact of these variances and see which schedules and forms needed further review, a sample of 50 returns were used to evaluate the trends within the data.⁵ Fields with discrepancies between the added record and actual return were reviewed using a number of different criteria.

Data were first researched by comparing the added record to the actual return for the year studied to view the overall trends within the data. This was then broken into two categories--data that were collected directly from the taxpayer survey, and data that were imputed using the prior-year ratio amount.

Table 2 shows that data items created directly from the information provided by the taxpayer on the survey exhibited little to no change between the added record and the actual return. These small variances may be attributed to differences in taxpayer reporting on the survey and the actual return filed or minor differences in SOI processing of these data items.

Data items for the fields created using the ratio calculations, as exhibited in Table 3, however, showed a much different picture. The largest percent changes were concentrated in the dividends schedule. Using 2002 as an example, for this schedule, dividends from domestic corporations on the added records were \$148.3 million compared to \$0.06 million on the actual returns. This is due to SOI's processing for statistical information purposes where dividend distributions among member corporations electing to file a consolidated return were eliminated from the statistics as part of the consolidated reporting of tax accounts.⁶ The data item, "dividends received deduction," also exhibited similar changes between the added records and actual returns, decreasing from \$129.9 million to \$0.04 million on the actual returns filed. This schedule will need additional review to compensate for these large differences so that amounts imputed on this schedule will more closely match those following SOI's processing of the actual return.

The remaining majority of data items with variances were scattered throughout all parts of the return, and most did not show significant changes between the actual and imputed returns. Many changes, like those on the balance sheet and income and deduction statement of the returns were more susceptible to variances in general. Since the imputations are based on the current-year totals and prior-year data, highly variable data fields like "cash" and "accounts payable" on the balance sheet and "deduction for bad debts" on the deduction statement were susceptible to higher variances from one year to the next. These imputations were not made based on

corporation behavior, and, as such, large accounts payable or receivables, etc. in one year can have an impact (which subsequently disappear once the actual return is filed) on the imputed data items on the added records.

In addition to the above criteria, return types were also evaluated to observe whether a particular return type was susceptible to larger variances. It was observed that, while the type of return filed may contribute to the overall number of variances (especially for larger, more complicated returns), it is not a good indicator of whether or not a data item will change from year to year nor is it a good predictor of trends within the data.

Lastly, companies in the file as added records over multiple years were evaluated to see if they showed distinct trends for the data variation from year to year, and also to see if any one company was driving the changes. For these evaluations, the corporations showed no distinct trends beyond what was observed for the overall sample, other than showing that the same data items changed from year to year.

► **Conclusion and Plans for Future Research**

Critical cases are an integral part of the corporation study and, in some cases, necessary for the statistical validity of the file. This is why studying the alternate records is imperative to ensuring a complete and accurate program file. Reviewing the short-edit records showed the need for further analysis of these returns. While the variances in general are not unreasonably large, there are still some very large changes noticed within the data that could potentially have an impact on the overall corporation file.

The dividends schedule, in particular, is an area that will require further examination for future program years. For the time being, this may involve the manual editing and review of this field by the analyst in charge of the critical case program until additional line items may be added through the OMB authorization process. Once the process is in place for adding the necessary data items, adjustments can be made to the program where necessary to account for the data on this schedule and further improve the data quality.

There are also a number of additional ways to evaluate and hopefully improve the imputation process and, thus, the resulting data that are produced. Such evaluations could decrease the time it takes N.O. staff to incorporate missing data, thereby freeing up resources that can be used on other projects.

One option to do so would be to compile ratios created as an average of the last few years of the return, and subsequently use those in conjunction with the amounts supplied by the taxpayer to create the remainder of the current-year amounts. This might decrease the effect of instances where a company has an unusually large amount one year--thus creating an extremely large ratio that is used to calculate the current-year amounts. Another would be to use the trend within the corporation's industry to calculate the ratios. This would allow the ratios to more closely mirror those of the entire industry and possibly decrease the chances of the corporation being an outlier within the industry.

If these comparisons are done for prior-year returns already in the program file, the accuracy of these proposed options could easily be tracked to determine which would be a more accurate way to add the data.

However, all evaluations aside, the ultimate goal in improving data quality is first and foremost to reduce the number of unavailable records during Advance Data. The lower the number of added records, the better the overall file will be during both phases of the Corporation studies.

► **Acknowledgment**

Thanks go to Patrice Treubert of the Corporation Research Section for her help in creating the SAS data sets that were used in the analysis of the data.

► **Endnotes**

¹ As an example, for the Tax Year 2002 SOI corporation study, which included returns with accounting periods ending July 2002 through June 2003, the

critical case list was finalized in December 2003 and was based on the critical cases in the Tax Year 2000 and 2001 corporation studies. If the returns met the critical case criteria for either of the two prior years, they were classified as critical cases for the 2002 study. Previous and subsequent years also incorporate the same principles for inclusion of returns in the sample.

² All tax data and related information pertaining to individual business income taxpayers are posted to the IRS Business Masterfile (BMF) so that the file reflects a continuously updated and current record of each taxpayer's account. For additional information, please visit: <http://www.irs.gov/privacy/article/0,,id=130752,00.html>.

³ Items from the balance sheet are calculated differently than the remainder of the tax return. Balance sheet items use total assets to impute remaining data items based on ratios of the industry average.

⁴ There were no survey records added for the Tax Year 2000 program so that year was not counted in the survey data comparisons.

⁵ This sample represented 36 percent of all short-edits from Tax Years 1997-2002. Data were selected on a number of factors, mainly, the return type and number of times in the file as an added record. This was done to create a variety of evaluation criteria and ensure that other factors did not influence the data variations. Though the above criterion was used in gathering the sample of returns, the sample was not chosen with the name or size of the corporation as determining factors. The weights for these returns were all the same so that variances were not a result of weighting differences. However, we assumed that the data entered from these returns were free of editor error, that is, the N.O. and field editors entered the amounts in the system correctly for the returns they edited. Since the system is thoroughly tested before program implementation, it is assumed that the program is

also free of error and, therefore, did not contribute to variances in the data.

- ⁶ For tax purposes, dividends reported on these returns represented amounts received from corpo-

rations that were outside the tax-defined affiliated group. See also section on Explanation of Terms, Internal Revenue Service, Statistics of Income, *Corporation Income Tax Returns*, annual publications 1997-2002.

The Impact of the Followup Process on the 2002 Foreign Tax Credit Study Data

Rob Singmaster and Lissa Redmiles, Internal Revenue Service

The followup process is an important step in the data cleansing process of the Foreign Tax Credit study conducted by the Statistics of Income Division of the IRS. The study itself collects data from corporate tax forms and their attached Form 1118's. Analysts review the data, correct anomalies, and disseminate the results. In certain cases, the analysts request additional information beyond what was originally reported by the taxpayer. This paper focuses on the 290 returns selected for additional data requests and the impact of the data received as a result on the study as a whole.

► Overview of the Foreign Tax Credit

The need for a foreign tax credit became apparent with the advent of the modern U.S. income tax in 1913. Since this date, U.S. taxpayers have been subject to taxation on their worldwide incomes. U.S. corporations with international operations or investments may also be taxed on their foreign-source incomes in the country in which the income is earned. The result is double taxation. To correct this problem, the United States passed into law foreign tax credit provisions, beginning with the Revenue Act of 1918. This credit allows U.S. corporations to offset the U.S. tax on their foreign-source taxable incomes with a credit for the foreign taxes that were already paid.

In the close to 90 years that the foreign tax credit has been in existence, the rules and ways in which this credit is reported have undergone many transformations. Perhaps the change that most affected the way the credit is calculated today occurred with the passage of the Revenue Act of 1962. It required corporations to compute a separate limitation for nonbusiness-related interest income. This step prevented corporations from combining foreign-source income from business operations taxed at rates higher than the U.S. rate with interest-bearing investments abroad that was subject to little or no foreign tax.

For Tax Year 2002, taxpayers were required to compute a separate foreign tax credit limitation for each of 11 different income categories. The taxpayer is required to report gross income, various deductions, taxable income, and foreign taxes paid or accrued by country in each appropriate income category. Within each category, taxpayers separate their income, deductions and taxes by type.

The foreign tax credit remains the largest credit that U.S. corporations claim to reduce their U.S. income tax. For Tax Year 2002, 9,383 corporations claimed a total credit of \$42.4 billion. Corporations report the foreign income and taxes related to the credit on Form 1118, *Computation of Foreign Tax Credit--Corporations*, filed with their income tax returns. Gross income, deductions, and taxable income attributed to various countries are reported on Schedule A, while foreign taxes paid or accrued and the foreign tax credit calculation are reported on Schedule B. Schedules C through Schedule J support items on Schedules A and B.

The statistics in this article are based on information reported on Forms 1118 and related corporate returns filed with accounting periods ending between June 30, 2001, and July 3, 2002. The returns in our study were selected after administrative processing but prior to any amendments or audit examination. The estimates are based on a stratified probability sample of 4,157 returns selected from a population of corporations filing a Form 1118 and are subject to sampling error. Each return in the sample is given a distinct weight, calculated by dividing the number of returns in a certain section of the study (industry, accounting period, etc.) by the number of sample returns for the same section. The purpose of these weights is to adjust for the various sampling rates used, relative to the population. For the purposes of this paper, weighted totals are used for all counts and numerical values.

► The Followup Process

During entry of the Form 1118 data, the system performs close to three hundred consistency tests. The data entry personnel resolve some of these tests, and some are shipped to SOI headquarters for further review. If the analysts cannot resolve the remaining errors, and the taxpayer reports a foreign tax credit, a letter may be sent to the taxpayer asking for additional information. (Many corporations with an overall loss file a Form 1118 in order to compute the carryover of taxes available for use in subsequent tax years. Since the form is not required in these cases, we do not typically ask for additional information for these returns.) We ask that the taxpayer respond within 60 days of the original letter but usually grant requests for extensions. If we did not receive a response before the deadline, we phoned the taxpayer. The responses received are used for statistical and analytical purposes only and are not part of tax enforcement or administration.

The most common error that will trigger a letter is missing country detail. We also frequently send letters to those missing Schedule H or Schedule F. Other data requested include explanations for discrepancies between the various schedules on Form 1118 and discrepancies between Form 1120, *Corporation Income Tax Return*, and Form 1118. On Form 1118, the most common discrepancies are between:

- Total not definitely allocable deductions on Schedule A and Schedule H, for the same income type
- Schedule A, total gross income and Schedule F, branch income, for the same country
- Schedule A, definitely allocable deductions and Schedule F, deductions
- Schedule A, total income or loss before adjustments and Schedule B, taxable income
- Total income or loss before adjustments on Schedule A and Schedule J, for the same income type

Between Form 1118 and Form 1120, the most common differences are between:

- total taxable income
- total U.S. income tax against which credit is allowed
- total foreign tax credit
- deemed dividends (subpart F dividends)
- other foreign dividends
- dividend gross-up

By far the most common discrepancy between these two forms is a discrepancy in the dividends and/or dividend gross-up reported on Schedule C of Form 1120 and the sum of the dividends and gross-up reported on Schedule A of Form 1118. This is partly because Schedule C tends to be poorly filed and partly because there are some legitimate reasons for differences in the dividend amounts reported on these forms. In general, we do not ask taxpayers to account for the dividend discrepancies unless we are already requesting other information.

The table below lists the number of requests sent by type. (Since we often requested more than one type of information from one company, the total number of requests exceeds the number of returns in the followup process.)

Number of Requests Sent, by Type

Reason for Followup	Number of Requests
Missing country detail	178
Discrepancies between Form 1120 and Form 1118	84
Schedule F missing	52
Schedule H missing	32
Missing amounts from Sch. H	28
Discrepancy between Sch. A and Sch. F	8
Taxable income discrepancy (Sch. A and J or B and J)	7
Missing Form 1118	7
Other	12

This paper focuses on those returns missing country detail for foreign-source income and/or foreign taxes paid, those missing Schedule F, and those missing Schedule H, because these problems were most likely to be the primary reason for requesting additional information.

► Followup Response

The Foreign Tax Credit study for Tax Year 2002 included data from 4,157 corporate tax returns, representing a population of 9,383. A weighted total of 290 returns were selected for additional data requests. At the end of the study, we had received a response from 206 of these requests, a response rate of 71 percent. Of those that responded, a majority, (166 or 81 percent) provided a fully satisfactory answer to our inquiries and supplied the missing data that they had failed to provide in their original filed tax returns. A smaller group of responses, 31 out of 206 (15 percent), supplied us with at least some information that they had previously withheld. It should be noted that, in many of the cases where we were requesting country detail for either income or taxes paid, the taxpayer was unable to provide this information due to software or time constraints. We chose to rate only 9 out of 206 responses (4.4 percent) as completely unsatisfactory. The remainder of our requests, 84 out of 290 (29 percent), did not respond in any form.

The followup letters sent out for the Tax Year 2002 study represent companies from a wide range of industries. Using NAICS (North American Industry Classification System) to sort these corporations, we discovered that the most well-represented industry in our study was manufacturing, accounting for 121 out of the 290 (41.7 percent) additional data requests. Although manufacturing returns overall accounted for just 18 percent of the total number of returns, they comprised 50 percent of the total foreign-source gross income so that the rate of followup is perhaps slightly lower than expected. The next most populous group was the finance/insurance industry, with 48 out of 290 (16.6 percent). This is as expected, as this industry accounts for about 11 percent of all returns and, more importantly, 16 percent of total foreign-source gross income. The third most populous group was the information industry, with 34 out of the 290 (11.7 percent) total, compared to 6 percent of the

total number of returns and almost 10 percent of the total foreign-source gross income. Although more additional data requests were sent to certain industries than others, we did not find a substantially better or worse response rate when comparing these industries at the end of our study.

► Missing Schedule F

One of the Form 1118 supporting schedules that tends to be missing or poorly filed is Schedule F, *Gross Income and Definitely Allocable Deductions for Foreign Branches*. Amounts from this schedule are included in the total gross income and definitely allocable deductions on Schedule A but are not directly carried forward. The only indication we have that a Schedule F may be missing is if branch taxes were reported on Schedule B, Part I, but no Schedule F was filed and the branch income and branch deductions associated with those taxes are therefore unknown. Sometimes, we can impute a Schedule F using the Schedule A and prior-year data. In other cases, we must write to the taxpayers. Since 261 taxpayers had this condition, we generally limited our requests to those returns that reported over \$1,000,000 of branch taxes or whose branch taxes equaled 25 percent of the total foreign taxes paid or accrued. Of course, if we were sending a letter to a taxpayer due to some other problem, we included a request for the missing Schedule F even if the return did not meet either criterion.

We requested a Schedule F from 52 corporations that reported branch taxes but had not included a completed Schedule F with their Forms 1118. These taxes totaled to about one billion dollars, approximately 20 percent of the total foreign branch taxes reported by all corporations. Of these corporations, 32 or 62 percent, sent in Schedule F data. The total foreign branch gross income reported in response to our letter for these returns was about \$12 billion, 15 percent of the total for all returns. These taxpayers also supplied almost \$7 billion in previously unreported foreign branch definitely allocable deductions, about 17 percent of the total for all returns. By the conclusion of the study, taxpayers had sent in Schedule F's to support a total of \$751 million in branch taxes paid, or about 69 percent of all the unsupported branch taxes from the returns that received letters. Unsupported taxes from all returns then declined from 22

percent of all foreign branch taxes to 6 percent, due to the followup process.

When we examine the ratio of supported taxes, post followup, to the original unsupported tax amounts for those returns selected for followup, by industry, we see most of the major industry groups supplied Schedule F's to support more than 70 percent of the originally unsupported branch taxes. The one exception is the wholesale and retail trade industry group, which provided support for only 29 percent of the taxes missing support from Schedule F.

Followup Returns Missing Schedule F

[Money amounts are in millions of dollars]

Industry	Unsupported Branch Taxes Paid	Taxes supported by Schedule F after Followups	Percent (col. 2/ col. 1)
Manufacturing	\$634	\$453	72%
Wholesale/Retail Trade	13	4	29
Information	30	28	93
Finance/Insurance	97	80	82
Services	230	185	80
Total	\$1,003	\$749	75%

► Schedule H

Another of the supporting schedules included within Form 1118 is the Schedule H, *Apportionment of Deductions Not Definitely Allocable*. This schedule is used to apportion deductions that cannot be definitely allocated to a certain item or class of income. Schedule H is filed only once with each Form 1118 and has two distinct parts. Part I is comprised of research and development deductions, while Part II is a combination of interest deductions and other miscellaneous deductions that do not fit into a specific category. These two parts are then added together to arrive at a total not definitely allocable deduction figure for the schedule. This total figure is also reported on Schedule A, along with the company's definitely allocable deductions.

Every corporation filing a Form 1118 that reports not definitely allocable deductions is required to complete a Schedule H that documents these deductions. We con-

tact taxpayers whose Schedule H is missing and whose not definitely allocable deduction amount exceeds \$10 million.

In Tax Year 2002, taxpayers failed to report a Schedule H to support a total of \$6.8 billion in not definitely allocable deductions. This was approximately 7 percent of the \$100.4 billion in total not allocable deductions from all returns. We wrote followup letters to 32 companies with a request to provide a completed Schedule H. These corporations represented a total of \$4.8 billion in not definitely allocable deductions on Schedule A that were not supported by a Schedule H. This figure accounted for roughly 71 percent of the not definitely allocable deductions not supported by a Schedule H in our study prior to followup. As a result of this process, we received responses from 18 (56 percent) of the companies. They provided supporting Schedule H's that accounted for \$3.18 billion of the \$4.8 billion (66 percent) total represented by the 32 companies. Thus, the followup process decreased the amount of apportioned deductions not supported by a Schedule H from 7 percent to 3.6 percent of the total apportioned deductions.

► Unallocated Income

From a data analysis standpoint, it is desirable for taxpayers to assign as much of foreign income, deductions, and taxes paid total to a specific foreign country as possible. However, they do have the option of categorizing either all or part of their incomes, deductions, or foreign taxes paid or accrued to other or various countries. One of our main goals in sending followup letters is to obtain specific country detail for any large amounts assigned to various countries.

As with the missing schedules, we established criteria for requesting additional country detail when the taxpayer failed to allocate a significant amount of foreign-source gross income to the country or region of source. Generally, we send a letter to those corporations with \$25 million or more of unallocated gross foreign-source income or \$10 million of unallocated foreign-source taxable income. Although we will ask for country detail for the definitely allocable deductions if the return meets the income test and some or all of the deductions have not been sourced, country detail here is not con-

sidered essential to the study. (Many taxpayers prorate their deductions to countries based on each country’s share of foreign gross income, and our system therefore prorates any amounts remaining in “other countries” at the end of the study accordingly.)

We sent followup letters to a total of 160 companies. The unallocated foreign-source gross income for these returns was approximately \$79 billion; about 89 percent of the total unallocated income (\$88.8 billion) and 20 percent of the total foreign-source gross income (\$390 billion). Other income accounted for 42 percent of the unallocated amount, while the next largest category, gross rents, royalties, and license fees, comprised 23 percent. Some of these returns had not allocated any of their incomes, but many had already allocated a considerable portion before we requested additional country detail. Overall, the unallocated amount for these returns was 50 percent of total foreign-source gross income.

A Comparison of Total, Unallocated, and Allocated Income, by Type

[Money amounts are in billions of dollars]

Type of Income	Total FS Gross Income from All Returns	Unallocated Income from Followup Returns	Allocated Income from Followup Returns
Dividends	\$95.4	\$6.6	\$5.5
Interest	55.2	12.4	8.1
Rents	67.1	18.3	5.1
Services	21.8	8.8	2.9
Other	150.8	33.0	21.1
Totals	\$390.3	\$79.0	\$42.7

Of these 160 companies, 88 sent in a satisfactory response, 19 sent in a partial response, 5 included an unsatisfactory response, and the remaining 48 never responded.

By comparing the percentage of total foreign-source income and the percentage of unallocated income from all returns, across industries, we can get an indication of which industries were more or less likely to allocate their incomes to the country of source. Manufacturing companies, for example, earned 50 percent of the total foreign source gross income but accounted for 36 per-

cent of the unallocated income. On the other hand, the information industry comprised just 10 percent of the total but 26 percent of the unallocated income. Finance and insurance companies had only a slightly higher percent of unallocated income than expected based on their percentage of gross income. The other industry groups accounted for about the same fraction of unallocated income as total foreign-source income.

Total Foreign-Source (FS) and Unallocated Income, by Industry Group

[Money amounts are in billions of dollars]

Industry Group	Total Gross FS Income	Percent of Total	Unallocated Income	Percent of Total
Manufacturing	\$194.6	50%	\$32.1	36%
Information	37.2	10%	23.2	26%
Finance/Insurance	60.9	16%	17	19%
Management of Companies	45.2	12%	5.0	6%
Other Industries	52.3	5%	11.6	3%
Totals	\$390.3		\$88.8	

Taxpayers allocated \$42.7 billion of their total gross foreign source incomes to countries and or regions; about 54 percent of the original unallocated amount. They were much more likely to allocate their interest or other income than gross rents, royalties, and license fees or their income from the performance of services. Roughly half of the allocated income was other income, while almost 20 percent was interest income. Most significantly, the total gross foreign-source income attributed to countries or regions as a result of taxpayer correspondence accounted for approximately 11 percent of the total foreign-source gross income for all returns.

The rates of followup response for those corporations missing country detail for gross income and the percentage of foreign source gross income allocated in response to our requests also vary by industry. The professional, technical, and scientific industry group and the management of companies and enterprises group had the highest satisfactory response rates. Manufacturing and the wholesale and retail trade group also had satisfactory

response rates that were well over 50 percent. Rates for transportation and warehousing, information, and the finance and insurance group, however, ranged from 33 percent to 42 percent. A comparison of the original amount not attributable to specific countries or regions to the amount allocated after receiving our requests yields similar results. Top of this list is again the professional, technical, and scientific services industry, with an allocation rate of 81 percent. The management of companies and enterprises industry and the manufacturing industry follow close behind, with 79 percent and 71 percent respectively. Finance and insurance, however, allocated just over half of the amount missing country detail, while the information industry allocated about 37 percent.

A Comparison of Unallocated and Allocated Income for Followup Returns, by Industry

[Money amounts are in billions of dollars]

Industry Group	Income Not Allocated	Allocated Income	Percent Allocated
Manufacturing	\$27	\$19	71%
Wholesale/ Retail Trade	3	1	40%
Transportation/ Warehousing	4	1	13%
Information	22	8	37%
Finance/ Insurance	15	8	53%
Professional/ Scientific/ Technical Services	1	1	81%
Management of companies	4	3	79%
Other industries	2	1	57%
Totals	\$79	\$42.7	54%

While the percentage allocated from the professional, technical, and scientific industries may be impressive, it is important to remember that the total allocated amounts received from this industry group is relatively small. Of the total allocated amount received, manufacturing comprised nearly 45 percent while the finance and

insurance industry group and the information industry each accounted for 19 percent of the data.

► Unallocated Taxes Paid or Accrued

As with the other conditions that cause us to send a followup letter to a certain company, it is necessary to set a minimum threshold for foreign taxes paid amounts for which we want to obtain country detail. After a review of taxpayer reporting trends, we decided to request additional country detail for any unknown foreign tax amount totaling more than \$5 million. Using this number as a guideline, we sent followup letters to 79 U.S. corporations requesting additional taxes paid country detail.

For Tax Year 2002, these companies represented a total of \$5.51 billion in foreign taxes paid, \$2.7 billion (48.5 percent) being attributed to unknown or various countries before followup. This second figure represents 85 percent of the \$3.1 billion total unknown foreign taxes paid amount prior to followup in our study. These totals were broken down by category as follows: \$170.8 million of foreign taxes paid on interest income, \$10.7 million (6.2 percent) for country unknown; \$906.5 million of foreign taxes paid on rents, royalties, and license fees, \$703.3 million (77.6 percent) unknown; \$2.1 billion of foreign taxes paid on foreign branch income, \$905.4 million (43.8 percent) unknown; \$234 million of foreign taxes paid on services, \$219.7 (93.9 percent) unknown; and \$1.8 billion of foreign taxes paid on other income, \$641.2 million (36.2 percent) unknown.[1]

By the conclusion of our Tax Year 2002 study, we received responses from 55 of the 79 companies (69.6 percent) we had contacted to obtain taxes paid country detail for \$2.7 billion of taxes paid attributed to various/unknown countries, approximately 14 percent of the total taxes paid from all returns and roughly 85 percent of the total unallocated taxes from all returns. Taxpayers allocated a majority of their previously unallocated taxes paid on service income, while they provided country detail for about a third of their taxes paid on interest and other income.

A Comparison of Total, Unallocated, and Allocated Taxes, by Type

[Money amounts are in millions of dollars]

Type of Income	Unallocated Taxes from Followup Returns	Allocated Taxes from Followup Returns	Percent Allocated
Interest	\$10.7	\$3.1	29%
Rents	703.3	216.6	31%
Branch Income	905.4	459.5	51%
Services	219.7	206.7	94%
Other	641.2	204.7	32%
Total	\$2,675	\$1,214.9	45%

The additional information we received substantially enhanced the accuracy and usefulness of the study data. Overall, the total amount of taxes paid attributed to various/unknown countries was reduced by \$1.2 billion, from \$2.7 billion to \$1.5 billion, a 45-percent reduction. This \$1.2 billion amounted to almost 7 percent of the total foreign taxes paid.

Taking a closer look at the followup letters we sent for foreign taxes paid country detail, we discovered that the manufacturing industry accounted for the highest percentage of these requests, with 26 out of 79 (32.9 percent) total. The finance/insurance and information industries were also well represented, with 19 (24.1 percent) and 13 (16.5 percent) requests, respectively. Even though the information industry accounted for less overall requests than manufacturing and finance/insurance, it possessed the most foreign taxes paid to unknown countries, with \$976.3 million (36.6 percent) of the total prior to followup. Manufacturing was a close second, with \$943.8 million (35.3 percent) of the total. The finance/insurance industry accounted for only a fraction of these totals prior to followup, with \$221.7 million (8.3 percent). At the end of our study, each of these industries saw a decrease in the amount and percentage of foreign taxes paid to various countries. The most significant drop in unallocated taxes paid was seen in manufacturing, whose unknown foreign taxes paid went from \$943.8 million to \$307.7 million, a

67-percent decrease. The finance and insurance sector experienced the largest percentage decrease in unknown foreign taxes paid of these three industries, going from \$221.7 million to \$91.3 million (59 percent). The information industry showed the smallest change between pre- and post-followup taxes paid data, going from \$976.3 million to \$931 million, a 5-percent reduction.

A Comparison of Unallocated and Allocated Taxes for Followup Returns, by Industry

[Money amounts are in millions of dollars]

Industry Group	Taxes Not Allocated	Allocated Taxes	Percent Allocated
Manufacturing	\$943.8	\$636.1	67%
Wholesale/Retail Trade	86.1	61	71%
Transportation/Warehousing	24.9	24	96%
Information	976.3	45.3	5%
Finance/Insurance	221.7	130.4	59%
Professional/Scientific/Technical services	6.7	3.5	52%
Management of companies	263.4	228.9	87%
Other industries	152.1	85.7	56%
Total	\$2,675	\$1,214.9	45%

► Conclusions

Overall, the response rate for followups was sufficient to make the process worthwhile. Since our data requests covered almost 90 percent of the unallocated income and 87.5 percent of the unallocated taxes, it appears that our thresholds for these data requests are adequate. In future studies, we may want to keep in mind that the information industry is far less likely than the other significant industry groups in our study to provide additional country detail for both foreign-source income and foreign taxes paid. Our criteria for missing Schedule F's also appear adequate, as we sent followups for 92 percent of the unsupported branch taxes. Although we sent followups for a lower percentage of the total unsupported apportioned deductions (71 percent), it is not clear

whether lowering our thresholds for writing to taxpayers to see if we can acquire Schedule H support is justified, since the total unsupported apportioned deductions was just 7 percent of the total.

Reflecting on our results, it appears that the followup process has a substantial impact on the overall quality of our data. By requesting missing Schedule H's, we obtained support for about 3 percent of the total not definitely allocable deductions. Asking for additional country detail enabled us to allocate 11 percent of the total foreign gross income and nearly 7 percent of the total foreign taxes paid or accrued to the source country or region. Although our figures for gross branch income

and deductions are still underreported, without our requests for missing Schedule F's, we would be missing 15 percent of the gross foreign branch income and 17 percent of the foreign branch deductions now reported for this study year. The improvement in the quality of the data as a result of our followup letters more than justifies the effort involved in this process and will be continued in future studies.

► **Endnote**

- [1] For the purposes of this paper we chose not to examine totals for foreign taxes paid on dividends or 863(b) income.

5



Interesting Methodological Topics Related to Internal Revenue Service Tax Statistics

Raub ♦ Chen

Henry ♦ Day

Weber ♦ Bryant

Sailer ♦ Bryant ♦ Holden

A Cluster Analysis Approach To Describing Tax Data

Brian G. Raub and William W. Chen, Internal Revenue Service

The Statistics of Income (SOI) Division of the Internal Revenue Service (IRS) produces data using information reported on tax returns. These administrative data are used by the Department of the Treasury, the Joint Committee on Taxation, and various Federal statistical agencies and are disseminated to the public via the World Wide Web and publications such as the *SOI Bulletin*. The Corporate Foreign Tax Credit (CFTC) study is in many ways typical of SOI studies. Data are collected from tax forms (in this case Form 1118) by SOI field staff and are subjected to error resolution by analysts at National Headquarters. The error-resolved data are used to create statistical tables that are published annually with descriptive text and technical notes. These statistical tables display selected aggregate fields from Form 1118 by industry, type of income, and country to which foreign taxes were paid.

The present paper will describe a population of Form 1118 filers using cluster analysis, with the goal of identifying alternative ways of organizing and analyzing tax data. A second goal is to identify new insights about this population of filers.

► Background

The Corporate Foreign Tax Credit is claimed by U.S. multinational firms to offset some or all of their taxes paid to foreign countries. Under U.S. tax law, U.S. corporations are taxed on income earned both in the U.S. and in foreign countries. Income earned in foreign countries may also be subject to taxation by the authorities in those foreign countries, resulting in double taxation. The foreign tax credit was adopted to alleviate this problem.

To claim the foreign tax credit, U.S. corporations file Form 1118, *Foreign Tax Credit--Corporations*. On this form, taxpayers report their incomes within broad categories such as interest, dividends, services, rents, and other. Deductions and tax liability are also reported.

Further, taxpayers are required to report these items detailed by country.

For 2001, taxpayers were required to segregate their incomes, deductions, and taxes into several limitation categories, or “baskets,” such as the Passive Income basket or the General Limitation Income basket. A separate foreign tax credit was calculated for each basket, with the total foreign tax credit being the sum of the separate foreign tax credits from each basket. The purpose of this provision and related limitations was to prevent taxpayers from using foreign tax credits to offset taxes on U.S.-source income, thus denying the United States tax revenues due on income earned domestically.

For Tax Year 2001, U.S. corporations claimed a combined \$41.1 billion in foreign tax credits. This was the single largest type of tax credit, accounting for 86.7 percent of all credits claimed by corporations in that tax year. This credit is elective, meaning that, if the taxpayer chooses to take the credit, no deductions for those foreign taxes are available. A majority of taxpayers decide to take the credit, since it offsets the U.S. income tax dollar for dollar, unlike a deduction, which may only offset every dollar of U.S. tax by the percentage of the tax rate [1].

► Data Description

The 2001 CFTC study is based on a stratified, weighted sample of corporation income tax returns with a foreign tax credit that were included in the 2001 SOI sample of returns with accounting periods ending between July 2001 and June 2002. These returns were selected after administrative processing but prior to any amendments or audit examination. The corporate tax return forms included in this sample were Forms 1120, 1120S, 1120-L, 1120-PC, 1120-REIT, and 1120-RIC.

The 2001 CFTC data sets contain 2,563 returns claiming foreign tax credits. These returns are weighted

up to a population estimate of 5,478 returns. For the present paper, we used a “defined population” approach by including only those returns with a sample weight of 1. This defined population of 1,075 returns accounted for an estimated 98.3 percent of the total foreign credit claimed on all returns for 2001.

► **Cluster Analysis**

Cluster analysis, or clustering, refers to a set of mathematical techniques for sorting observed data into groups so as to maximize the similarity of observations within the same group and minimize the similarity of observations across different groups. These techniques can be used to discover associations and structures within a data set that may not have been known. Cluster analysis has been widely used in the biological and social sciences to help define classification schemes or taxonomies. It has also been used to suggest new ways of describing a population in business and marketing applications.

Cluster analysis techniques can be broadly separated into two approaches, hierarchical and nonhierarchical. The hierarchical approach builds clusters of successively larger size using some measure of similarity or distance. Typical algorithms used in this approach include single linkage (nearest neighbor), complete linkage (furthest neighbor), and Ward’s Method, which minimizes the mean square distance between the center of a cluster and each member. Nonhierarchical clustering approaches also exist, including the K-means method.

For the present data set, we chose hierarchical clustering since this set of techniques is available in SAS’s PROC CLUSTER. We clustered a sample of our data set using each of the 11 methods available in SAS and ultimately selected Ward’s Method for two main reasons. First is the efficiency of this method, useful given the relatively large number of observations (1,075) and clustering variables (9). Second is the tendency of this method to create clusters of relatively equal size. We noted a strong tendency for other clustering algorithms to create clusters with very few observations. Although the existence of these outliers may be an interesting outcome in a subject-matter sense, allowing very small clusters could create a disclosure problem [2].

In Ward’s Method, the distance between two clusters is defined as

$$D_{KL} = \text{distance between clusters } C_K \text{ and } C_L$$

$$D_{KL} = \left\| \bar{x}_K - \bar{x}_L \right\|^2 \left(1 / N_K + 1 / N_L \right)$$

where

$C_K =$ K th cluster, subset of $\{1,2,\dots,n\}$

$x_i =$ i th observation

$N_K =$ number of observations in C_K

$\bar{x}_K =$ mean vector for cluster C_K

$\|x\| =$ Euclidian length of the vector x , that is, the sum of the squares of the elements of x .

If the distance between observations x and y , $d(x,y) = \|x - y\|^2 / 2$, then the combinatorial formula is

$$D_{JM} = (N_J + N_K)D_K + (N_J + N_L)D_L - N_J D_{KL} / (N_J + J_M)$$

The distance between two clusters is the ANOVA sum of squares between the two clusters added up over all the variables. At each generation, the within-cluster sum of squares is minimized over all partitions obtainable by merging two clusters from the previous generation [3].

To define our clustering variables, we started by considering the main variables in the CFTC study data sets: selected data from Form 1120; gross income and deduction items from Form 1118, Schedule A; foreign tax items from Schedule B, Part I; and foreign tax credit computation items from Schedule B, Parts II and III. The first variable of interest that we identified was the total foreign tax credit, which is calculated on Form 1118, Schedule B, Part III and carried over to Form 1120. One concern that we identified immediately is that the total foreign tax credit amount varies significantly by corporation and is strongly correlated to the overall size of the corporation. Therefore, clustering on this variable

in its original form would tend to create clusters based primarily on the size of the corporation. This clustering would add little to our current knowledge of the filer population and would likely fail to capture relationships between other clustering variables. To overcome this limitation, we standardized this variable by taking the ratio of the total foreign tax credit to the corporation's income tax liability.

Since the types of income, deductions, and taxes reported by taxpayers are important elements of the CFTC study, we chose to use a set of variables that capture these elements. As deductions and taxes for each income type are closely correlated with the gross income for that type, we decided that including deduction and tax variables in our clustering would add little value. Thus, we focused only on gross income for each type--dividends, interest, rents, services, and other. We also standardized each of the gross income variables into a ratio by dividing the total for each type of gross income by the total gross income for the corporation. These ratios became five of our clustering variables.

The final data element of the CFTC data set that we used in our cluster analysis was foreign-source country of the gross income reported by each corporation. Defining clustering elements based on country proved to be somewhat challenging, however, since there are over 300 countries in our system, and it was necessary to limit the number of clustering variables for the sake of efficiency. Ultimately, we decided to create variables for the top three countries as defined by amount of total gross income. These three countries, Canada, Japan, and the United Kingdom, combined for 32.6 percent of the total gross income reported by the firms in our defined population. The corresponding clustering variables were defined as the ratio of gross income allocated to each country to the total amount of gross income for each company. Figure 1 summarizes the clustering variables by description and the names we assigned.

Determining the number of clusters to be used in this cluster analysis was largely a heuristic process.

Figure 1.--Clustering Variables

Variable Name	Variable Description
FTC	Foreign tax credit divided by income tax liability
Dividends	Dividend income divided by total gross income
Interest	Interest income divided by total gross income
Rents	Rents income divided by total gross income
Services	Services income divided by total gross income
Other	Other income divided by total gross income
UK	UK-source income divided by total gross income
Japan	Japan-source income divided by total gross income
Canada	Canada-source income divided by total gross income

From a subject-matter standpoint, we began with the assumption that it made sense to look for at least three clusters but that more than eight clusters would become cumbersome and provide less valuable insight into our defined population. After considering the output from these options, we concluded that viewing our data in four clusters provided the most insight into our data and could be described most effectively. We named these clusters "High Dividend Firms," "Low CFTC/Other Income Firms," "Interest/ Service Firms," and "High CFTC/Manufacturing Firms."

► Clustering Results

Figure 2 displays the number of observations in each cluster.

Figure 2. --Cluster Summary

Cluster	Number of Observations
High Dividend Firms	295
Low CFTC/Other Income Firms	201
Interest/Service Firms	367
High CFTC/Manufacturing Firms	208

The relative similarity in the number of observations in each cluster is consistent with our choice of Ward’s Method for our clustering algorithm, while the absence of very small clusters serves our requirement of protecting taxpayer confidentiality.

In comparing the makeup of the four clusters below, we will use the average of each variable for the firms in the respective cluster, expressed as a percentage rather than a pure ratio for ease of use.

The “High Dividend Firms” cluster is summarized in Figure 3. Dividends is the dominant income variable with an average of 72.0 percent, while the average Interest, Rents, and Services are all below 5.0 percent. The average FTC for “High Dividend Firms” is 16.7 percent, below the overall average of 32.4 percent for companies in our defined population. The UK variable has the highest average value among the four clusters at 15.4 percent, while the average Japan variable is the lowest among the clusters at 0.9 percent.

Figure 3.--“High Dividend Firms” Summary

Variable	Average Percentage Value
FTC	16.7
Dividends	72.0
Interest	3.1
Rents	4.7
Services	1.6
Other	6.7
UK	15.4
Japan	0.9
Canada	18.8

As seen in Figure 4, the average company in “Low CFTC/Other Income Firms” has a significantly different set of characteristics. For this group, the dominant income variable is Other, with an average of 82.8 percent. In contrast, the average Services and FTC values in this cluster are the lowest among the four clusters at 0.6 percent and 8.3 percent, respectively. The average country variables for this cluster are middling--with neither a high nor a low for any country variable among the clusters.

Figure 4.--“Low CFTC/Other Income Firms” Summary

Variable	Average Percentage Value
FTC	8.3
Dividends	4.1
Interest	4.9
Rents	5.7
Services	0.6
Other	82.8
UK	13.5
Japan	4.9
Canada	16.8

Summary statistics for “Interest/Service Firms” appear in Figure 5. For companies in this cluster, Interest, Rents, and Services incomes combine for nearly all of the gross incomes, with an average Interest of 33.4 percent, an average Rents of 31.1 percent, and an average Services of 23.2 percent. The average FTC for companies in this cluster is below the average of all the companies in our defined population at 15.8 percent. Among the country variables, the average Canada and Japan values are the highest of any cluster, 23.1 percent and 8.1 percent, respectively, while the average UK value is the lowest at 9.2 percent.

Figure 5.--“Interest/Service Firms” Summary

Variable	Average Percentage Value
FTC	15.8
Dividends	5.7
Interest	33.4
Rents	31.1
Services	23.2
Other	4.4
UK	9.2
Japan	8.07
Canada	23.1

Figure 6 displays the variable averages for companies in “High CFTC/Manufacturing Firms.” Other is the dominant income variable with an average of 36.0 percent, followed by Dividends and Rents with 28.8 percent and 15.0 percent, respectively. The average FTC

of companies in this cluster is dramatically larger than for any other cluster at 80.2 percent. Among the country variables, the average Canada value is the lowest of the four clusters at 7.1 percent, as is the combined average of the three country variables, 24.6 percent.

Figure 6.--“High CFTC/Manufacturing Firms” Summary

Variable	Average Percentage Value
FTC	80.2
Dividends	28.8
Interest	5.3
Rents	15.0
Services	1.7
Other	36.1
UK	12.4
Japan	5.2
Canada	7.1

► **Industry Analysis**

One additional element of note in the CFTC data is the industry classification of the companies filing Form 1118. Using industry classification in our cluster analysis, however, proved infeasible. Although each corporation in our defined population has a six-digit industry code assigned to it using the North American Industry Classification System (NAICS), this number is of an ordinal, rather than cardinal, nature. Therefore, although the NAICS code could be used as a clustering value, interpreting and describing the meaning of the industry code in the clustering output would be problematic. However, because industry classification is an element of interest, we analyzed the industry breakdown for each cluster ex post facto.

Our industry analysis reveals significant differences between clusters. Although Manufacturing, the largest industry among the firms in our defined population, represents a significant portion of the observations in each cluster, its contribution to the clusters ranged from 26.2 percent of “Interest/Service Firms” to 63.9 percent of “High CFTC/Manufacturing Firms.” Mining, Utilities, and Construction companies are distributed relatively

evenly between the clusters, with a low of 4.0 percent and a high of 7.2 percent. The remaining four industries make up more widely varied portions of the cluster totals. The Finance, Insurance, Real Estate, and Rental and Leasing industry makes up a low of 4.3 percent of “High CFTC/Manufacturing Firms” but a high of 33.6 percent of “High Dividend Firms.” Information companies comprise 3.7 percent of “High Dividend Firms” but 8.2 percent of “High CFTC/Manufacturing Firms.” Services companies make up only 6.0 percent of “Low CFTC/Other Income Firms” but 23.2 percent of “Interest/Service Firms.” Distribution and Transportation companies make up 8.2 percent of “High CFTC/Manufacturing Firms” but 17.4 percent of “Low CFTC/Other Income Firms.”

The industry distribution of “High Dividend Firms,” shown in Figure 7, reveals that Finance, Insurance, Real Estate, Rental, and Leasing is the dominant industry, comprising 33.6 percent of this cluster. This is the highest percentage of firms in this industry among the four clusters. The 13.2 percent of companies in the Services industry was the second highest among the clusters, while the 3.7 percent of companies in the Information industry was the lowest.

Figure 7.--“High Dividend Firms” Selected Industry Breakdown

Industry	Percent of Total
Mining, Utilities, and Construction	6.4
Manufacturing	30.2
Distribution and Transportation	11.9
Information	3.7
Finance, Insurance, Real Estate, Rental and Leasing	33.6
Services	13.2

The industry distribution of “Low CFTC/Other Income Firms,” shown in Figure 8, reveals that companies in the Distribution and Transportation industry represent a larger share than in any other cluster, with 17.4 of the total. In contrast, companies in the Services industry represent a smaller share of the total, 6.0 percent, than in any other cluster.

Figure 8.--“Low CFTC/Other Income Firms” Selected Industry Breakdown

Industry	Percent of Total
Mining, Utilities, and Construction	4.0
Manufacturing	39.8
Distribution and Transportation	17.4
Information	7.5
Finance, Insurance, Real Estate, Rental and Leasing	23.4
Services	6.0

Figure 9 displays the industry distribution of “Interest/Service Firms.” This cluster has the highest concentration of companies in the Services industry, 23.2 percent, and the lowest concentration of companies in the Manufacturing industry, 26.2 percent. “Interest/Service Firms” has 367 members, the most among the four clusters.

Figure 9.--“Interest/Service Firms” Selected Industry Breakdown

Industry	Percent of Total
Mining, Utilities, and Construction	6.0
Manufacturing	26.2
Distribution and Transportation	12.8
Information	6.8
Finance, Insurance, Real Estate, Rental and Leasing	24.0
Services	23.2

As seen in Figure 10, manufacturing firms dominate the “High CFTC/Manufacturing Firms” cluster, with 63.9 percent of the total, while the other industry groups each comprise 8.2 percent or less of the total.

► Implications

To gauge the effectiveness of cluster analysis in gaining insight to our data, we should consider its value to analysts both within SOI and outside. To SOI analysts who work with the CFTC data, some of the output of this cluster analysis may seem relatively obvious and merely confirms prior knowledge about our defined population. An example of this kind of result is that firms

Figure 10.--“High CFTC/Manufacturing Firms” Selected Industry Breakdown

Industry	Percent of Total
Mining, Utilities, and Construction	7.2
Manufacturing	63.9
Distribution and Transportation	8.2
Information	8.2
Finance, Insurance, Real Estate, Rental and Leasing	4.3
Services	8.2

in the “High CFTC/Manufacturing” cluster, dominated by manufacturing companies, claim the highest average foreign tax credit as a percentage of their income tax liabilities. On the other hand, at least one output of our cluster analysis was somewhat surprising: the relationship between reporting primarily Other gross income and offsetting a relatively smaller portion of tax liability with foreign tax, revealed in the “Low CFTC/Other Income Firms” cluster. Although it may have been possible to find this relationship by exhaustively querying our data files, cluster analysis has here served a useful function by pointing us in the right direction for further inquiry.

To those outside SOI who use CFTC data, our cluster analysis may also have value. Because, in most cases, users outside the Department of the Treasury do not have access to our data files, their ability to use our data is limited by what we provide in the published tables or in requested special tabulations. For example, while our published data tables do include summary statistics by industry and by country, they do not capture both relationships together as does our cluster analysis with the ex post facto industry distribution. Here again, the output from our cluster analysis may serve a useful function in revealing areas for further research.

► Limitations

The 2001 Corporate Foreign Tax Credit statistics quoted in this article do not represent the final amounts credited that year. Complete foreign tax credit statistics for 2001 would reflect the results of any audits. Also, some corporations did not file Form 1118 because they did not have a U.S. income tax liability and were, thus, unable to credit any foreign taxes paid, accrued,

or deemed paid for 2001. Finally, other corporations could have deducted their foreign taxes from their gross incomes instead of claiming a foreign tax credit.

As noted above, our analysis used only those firms from our sample with a weight of 1, i.e., those not weighted up to represent a greater part of the population estimates. This group of companies combined to claim 98.3 percent of all CFTC tax credits. Thus, while our analysis includes the large companies that claim an overwhelming majority of the total dollar amount of credits, it excludes many small companies that claim comparatively small CFTC's.

The output of our cluster analysis depended to a significant extent on choices made about our clustering techniques and our selection of clustering variables. As noted above, selecting which clustering algorithm to use and the number of clusters in the output is largely a heuristic process. Our set of clustering variables does not take into account several broad elements of the CFTC data sets, including "limitation baskets," data from Schedules F, G, H, I, and J, and country detail other than for Canada, Japan, and the UK.

► **Conclusion**

Cluster analysis can be a useful set of techniques for exploring and describing data sets, including those produced by SOI based on tax return data. By identifying relationships among the variables that are not immediately obvious to internal or external researchers, clustering can enhance knowledge of the data set and serve as the starting point for further research. The costs of cluster analysis should be manageable in many applications, since widespread software tools such as SAS® include clustering capability.

One challenge in using cluster analysis for data sets like those produced by SOI is that these tools may add the most value for data sets with a very large number

of observations and/or variables where relationships may be more difficult to identify by other techniques. However, these data sets may also be the most difficult to model for efficient clustering. In these cases, an alternative algorithm such as SAS's PROC FASTCLUS may be more appropriate, though at a loss of power and flexibility relative to PROC CLUS.

Another potential challenge in using cluster analysis on data sets like those produced by SOI presents itself for those which use sampling and weighting. Many data sets are significantly less "top-heavy" in dollar terms than the CFTC data set. In these cases, using only returns with a weight of 1 might entail the exclusion of many observations of interest from the clustering analysis. In the alternative, using returns with a weight of greater than 1 would require additional statistical considerations. The tradeoffs between these approaches could be analyzed using a Pareto analysis of the observations in the data set.

Thus, while cluster analysis can be a useful tool for data exploration and description in applications such as SOI's Corporate Foreign Tax Credit project, further study is needed to assess its potential costs and benefits for larger data sets.

► **Endnotes**

- [1] For more background on the Corporate Foreign Tax Credit, see Luttrell, Scott, "Corporate Foreign Tax Credit, 2000," *Statistics of Income Bulletin*, Fall 2004, Volume 24, Number 2.
- [2] The Internal Revenue Code prohibits the IRS from releasing information that could be used to identify specific taxpayers.
- [3] Description of Ward's Method adapted from *SAS/STAT User's Guide, Version 6*.

A Comparison of Income Concepts: IRS Statistics of Income, Census Current Population Survey, and BLS Consumer Expenditure Survey

Eric L. Henry and Charles D. Day, Internal Revenue Service

Several Federal Government agencies produce statistics on individual and household income. Because of the differing purposes to which their data will be put, agencies use different definitions for income (income concepts), as well as different reporting units, sample designs, collection modes, and processing rules. Data users are faced with an array of choices, often without much help to sort out which data series best meets their needs or much guidance to reconcile results based on different sources of data.

In order to help users, a number of papers have been written comparing the Census Bureau's Current Population Survey (CPS) Money Income and Survey of Income and Program Participation concepts, the Bureau of Labor Statistics (BLS) Consumer Expenditure Survey (CE) concept, and the Bureau of Economic Analysis Personal Income concept [1-3]. This paper extends that body of work by first describing the Adjusted Gross Income (AGI) concept, which is used most frequently to define individual income by the Internal Revenue Service (IRS) Statistics of Income (SOI) Division. That description is followed by an explanation of the most important differences between the AGI concept and the definitions of income used in BLS's Consumer Expenditure Survey and the Census Bureau's Current Population Survey. Note that this is a discussion of income concepts only; no attempt is made in this paper to discuss other causes of differences between estimates of income.

The Census Bureau conducts the CPS for BLS. It states that the data are "the primary source of information on the labor force characteristics of the U.S. population. CPS data are [intended for use] by Government policymakers and legislators as important indicators of our nation's economic situation, and for planning and evaluating many Government programs. They are also used by the press, students, academics, and the general public. ... Supplemental questions on ... income ... are often added to the questionnaire." The CPS ques-

tionnaire is administered at the household level, with information being collected for each person living in the household over age 15 [4].

BLS conducts the CE. It is the "basic source of data for revising the items and weights in the market basket of consumer purchases to be priced for the Consumer Price Index." It consists of two components, a quarterly interview survey and a weekly diary survey. The CE targets the entire noninstitutionalized population of the United States [5].

SOI Individual taxpayer data are an administrative data set. The data are collected from a sample of Forms 1040 filed by individual taxpayers [6]. The target population is all individuals required to file a tax return.

The AGI concept is appropriate to administration of the tax laws and thus varies quite a bit from the CPS and CE concepts. In order to make a discussion of those differences tractable and useful to readers, the authors have chosen to discuss those differences of greatest practical significance in comparing the data series, knowing that this will leave out many minor differences.

► The Adjusted Gross Income Concept

This section describes the AGI concept used by IRS's SOI Division. This description includes highlights of changes to the concept over the last 16 years. AGI is the difference between Total Income and Adjustments to Income. A deficit (negative AGI) occurs if Adjustments to Income exceed Total Income.

Total Income includes the following:

Wages, salaries, and tips include compensation for services, including wages, salaries, fees, commissions, tips, taxable fringe benefits, and similar items. AGI does not include money designated for a health flexible spending or health reimbursement arrangement. Similarly, elective contributions and employer matching amounts

for retirement plans, such as 401(k)'s, tax-sheltered annuities, and the Federal Thrift Savings Plan, are not included in salaries and wages for tax purposes. Also excluded from AGI are most forms of armed forces pay earned while in a combat zone or in a hospital recovering from illness or injury suffered in a combat zone. Note that there is a limited exclusion of qualified foreign-earned income.

Taxable interest consists of interest from bonds, savings accounts and certificates of deposit, interest accrued on unpaid amounts due to the taxpayer, and interest on privately held mortgages. Tax-exempt interest, from sources such as tax-free municipal bonds, IRA's, and 401(k) accounts, is excluded from AGI.

Dividends and capital gain distributions do not include the one-time exclusion of part or all of the gain from the sale of principal residence by an individual 55 years of age or older. The words "one-time exclusion of" were deleted in Tax Years 1990 and 1991, brought back in 1992 to 1996, and then incorporated into the current wording, "Exclusion of part or all of the gain from the sale of principal residence up to \$250,000 (\$500,000 on joint returns)," in Tax Year 1997 to the present time.

Refunds of State and local income taxes claimed as itemized deductions in previous years were first included in Tax Year 1990.

Alimony and separate maintenance payments are part of AGI, but child support payments (as IRS defines them) are not.

Net income derived from a business, profession, or farm helps make up AGI. Note that the business must be a "for profit" enterprise. Generation of revenue from a hobby does not qualify an individual to claim all of his or her expenses associated with that hobby.

Net gain from the sale of capital assets or of business property is included in AGI.

Annuities, pensions, individual retirement arrangement (IRA) distributions, and Tier II railroad retirement, reduced by their cost basis, are part of AGI [7].

Rents and royalties, along with **net income from estates and trusts**, help make up AGI.

Partnerships and subchapter S corporations are not taxable entities; therefore, income from these sources is distributed to the partners or owners and is included in individual AGI.

Unemployment compensation is part of AGI, although compensation paid by a union is reduced by the amount of any dues paid.

Taxable amounts of Social Security contribute to AGI. Since the inception of Social Security, railroad employees have had a separate, similar retirement system. **Taxable Tier 1 railroad retirement payments** were added in Tax Year 1990.

Taxable distributions from a Coverdell education savings account were added to AGI in Tax Year 2000.

Among the items of income included in AGI under "Other Income" are **prizes, awards, and gambling winnings, jury duty fees** (started in Tax Year 2000), **amounts received that were claimed as a deduction or credit in a prior year, bartering income, Alaska permanent fund dividends** (started in Tax Year 2000), and **qualified State tuition program earnings** (started in Tax Year 2000).

Statutory adjustments (lines 23 through 32, Form 1040 for Tax Year 2003) are amounts that are subtracted from Total Income to arrive at AGI (line 34, Form 1040 for Tax Year 2003). These include the following:

Reimbursed employee business expenses that were included in reported income (deleted for Tax Year 1990) are used to reduce Total Income.

With some limitations, elementary and secondary educators could deduct up to \$250 in **Educator expenses** (starting in Tax Year 2002) from Total Income for items purchased out-of-pocket for classroom use.

Contributions to self-employed retirement plans (Keogh or simplified employee pension) and **certain**

contributions to IRA's can be deducted when computing AGI.

Up to \$2,500 in **Student loan interest** (started in Tax Year 1998), paid on loans used for tuition, transportation, room and board, books, supplies, and equipment, can be used to reduce AGI by taxpayers with modified AGI under limits based on filing status.

Up to \$4,000 in **Tuition and fees** (started in Tax Year 2002) may be deducted in calculating AGI.

Archer medical savings accounts (started in Tax Year 1997, "Archer" added in TY 2002) are used by employees of small businesses and self-employed persons covered by a high-deductible health plan to save money for paying medical expenses. Contributions to such a plan can be used to reduce AGI.

Moving expenses (started in Tax Year 1994) associated with a move that is closely related to work and covers enough distance may be deducted in calculating AGI.

One-half of self-employment tax (started in Tax Year 1990) can be used to reduce AGI.

Self-employed health insurance expenses may be deducted in computing AGI.

Forfeited interest and penalties incurred by persons who made premature withdrawals of funds from time savings accounts can be used to reduce income in computing AGI.

Alimony payments are deductible for AGI computation purposes. Note that alimony received is considered income.

Forestation or reforestation expenses of up to \$10,000 can be used by owners of qualified timber property to reduce AGI.

The **foreign housing exclusion** is available to reduce AGI for those living abroad whose housing expenses are paid out of amounts provided by their employers.

Repayments of supplemental unemployment compensation from an employer-paid-for fund may be deducted when calculating AGI.

Certain expenses of qualified performing artists, in particular those working for more than one employer and with AGI less than \$16,000 before expenses are deducted, may reduce their AGI by the amount of those expenses, provided they are more than 10 percent of AGI.

Amount of jury duty pay reported on line 21, Form 1040, that was repaid to employers (started in Tax Year 1991).

The **Deduction for clean-fuel vehicles** (started in Tax Year 1999) allows the taxpayer to deduct up to \$2,000 of the cost of a designated clean-fuel vehicle from AGI.

Employee business expenses of fee-basis State or local government officials (started in Tax Year 1999).

SOI uses AGI as its most common measure of income as can be seen in its publications. Many of the components broken out by SOI are then further analyzed by also breaking them out by various sizes of AGI. This is done to compare tax returns to different AGI classes so that economists can easily see counts and money amounts and break out components of the tax return.

► **Comparison of Adjusted Gross Income and the Consumer Expenditure Survey Income Concept**

A description of the Consumer Expenditure Survey reads, "Income is the combined income of all consumer unit members (14 years of age or over) during the 12 months preceding the interview." The income concept includes the following:

Wages and salaries include total money earnings for all consumer unit members (14 years of age and over) from all jobs, including civilian wages and salaries; armed forces pay and allowances; piece-rate payments; commissions; tips; National Guard or Reserve pay (received for training periods); and cash bonuses

before deductions for taxes, pensions, and union dues. This corresponds to Wages, Salaries, and Tips on Form 1040. Portions of income that are nontaxable are the main source of differences between the CE and AGI concepts. AGI does not include money designated for a health flexible spending or health reimbursement arrangement. Also excluded from AGI are most forms of armed forces pay earned while in a combat zone or in a hospital recovering from illness or injury suffered in a combat zone. Note that identifiable amounts as classified under the definition of Salaries and Wages, which may have been reported by taxpayers as “other income,” are treated as salaries and wages for the statistics where possible.

Self-employment income includes net business and farm income, which consists of net income (gross receipts minus operating expenses) from a profession or unincorporated business or from the operation of a farm by an owner, tenant, or sharecropper. If the business or farm is a partnership, only an appropriate share of net income is recorded. Losses are also recorded. This corresponds with net income derived from a business, profession, or farm on the 1040. Rental income taken as crop shares is counted as rental income (line 17) in AGI, not farm income (line 18).

Social Security, private, and Government retirement includes the following: (1) payments by the Federal Government made under retirement, survivors’, and disability insurance programs to retired persons, dependents of deceased insured workers, or disabled workers and (2) private pensions or retirement benefits received by retired persons (or their survivors), either directly or through an insurance company. AGI includes only the taxable portion of Social Security benefits in its AGI computation. At least 15 percent of benefits are not taxable; if income is under \$34,000 (\$44,000 for a married couple filing jointly) and the taxpayer is not married filing separately and living with a spouse, at least 50 percent is not taxable. The CE concept includes income from “companies or unions, Federal Government (Civil Service), military, State or local governments, railroad retirement, annuities or paid-up insurance policies, individual retirement accounts (IRA’s), Keogh, or 401(k) payments.” Note that Tier I railroad retirement

is treated like Social Security for tax purposes. Also, if an employee paid part of the cost of a pension, then payments that represent the return of his or her cost are not included in income.

Interest, dividends, rental income, and other property income include interest income on savings or bonds; payments made by a corporation to its stockholders; periodic receipts from estates or trust funds; net income or loss from the rental of property, real estate, or farms; and net income or loss from roomers or boarders. AGI does not include interest on certain State and municipal bonds, as well as any tax-exempt interest dividends from a mutual fund or other regulated investment company. Dividends do not include nontaxable distributions of stock or stock rights, returns of capital, capital gains, or liquidation distributions. Taxpayers who paid penalties for the premature withdrawal of funds from time savings accounts or deposits could deduct those penalties as an adjustment to total income. Rental income taken as crop shares is counted as rental income in AGI, not farm income.

Unemployment and workers’ compensation and veterans’ benefits include income from unemployment compensation and workers’ compensation and veterans’ payments, including educational benefits but excluding military retirement, which is already included in Government retirement. A minor difference may arise from IRS’s reducing unemployment paid based on regular union dues by the amount of dues paid. Because workers’ compensation benefits paid “under a workers’ compensation act or a statute in the nature of a workers’ compensation act” are not taxable, they are not included in the AGI concept. Veterans’ benefit payments are not included in AGI, since they are not taxable. AGI excludes payments from workers’ compensation or from military or other uniformed services if the payee became entitled to the benefits or was a member before September 25, 1975, or if the payment is due to a combat-related injury. Also, if the payment is from a private disability insurance policy for which the taxpayer paid him- or herself, then the payment is exempt from taxation. Further, railroad retirement disability is treated like Social Security disability for tax purposes.

Public assistance, supplemental security income, and food stamps include public assistance or welfare, including money received from job training grants; supplemental security income paid by Federal, State, and local welfare agencies to low-income persons who are age 65 or over, blind, or disabled; and the value of food stamps obtained. Public assistance, supplemental security income, and food stamps are not included in the AGI concept because they are not taxable.

Regular contributions for support include alimony and child support, as well as any regular contributions from persons outside the consumer unit. Child support, as defined by IRS, is not included in AGI. Regular payments that individuals receive from nonhousehold members are usually not taxable, and thus not included in AGI, although they may be treated as gifts and be taxable to the giver.

Other income includes money income from care of foster children, cash scholarships, fellowships, or stipends not based on working and meals and rent as pay. AGI does not include assistance from friends or relatives. Scholarships and grants that do not represent payment for services, like teaching or research, and which are used for qualified educational expenses, like tuition and books (but not room and board), are not included as they are not taxable. Assistance received from employers can be excluded up to \$5,250.

► **Additional Notes**

Capital gains are not included as income in the CE but are included in AGI. State Tax Refunds are not included in the CE but are included in AGI if the taxes were deducted in the immediate prior year. Also, all lump sum payments like prizes, awards, and gambling winnings are not included in the CE but are included in AGI.

BLS uses income from the CE survey obtained from the interview process as its main component too. These data are then further analyzed by showing income and expenditures by quintiles of income before taxes. This is done to compare both income and expenditure components by varying income classes to more easily see trends in the data.

► **Comparison of Adjusted Gross Income and the Current Population Survey Income Concept**

“Earnings” is a three-part concept in the CPS. The first part includes “wages, salary, armed forces pay, commissions, tips, piece-rate payments, and cash bonuses earned, before deductions are made for items such as taxes, bonds, pensions, and union dues.” This corresponds most closely to Wages, Salaries, and Tips on Form 1040. Portions of income that are nontaxable are the main source of differences between the CPS concept and AGI. AGI does not include money designated for a health flexible spending or health reimbursement arrangement. Similarly, elective contributions and employer-matching amounts for retirement plans, such as 401(k)’s, tax-sheltered annuities, and the Federal Thrift Savings Plan, are not included in salaries and wages for tax purposes. Also excluded from income for purposes of computing AGI are most forms of armed forces pay earned while in a combat zone or in a hospital recovering from illness or injury suffered in a combat zone.

Net income from farm or nonfarm self-employment makes up the other two categories of earnings on the CPS. The CPS concepts are quite close to the AGI concepts; in fact, the CPS accepts replies for these two categories based on the respondent’s tax return. In cases where the respondent does not consult his or her tax return or other official records, differences may arise from change in inventories not being accounted for by the CPS. Also, rental income taken as crop shares is counted as rental income for AGI computation, not farm income.

Unemployment compensation from private or Government sources, as well as strike benefits, are included in both concepts. A small difference may arise from IRS’s reducing unemployment paid based on regular union dues by the amount of dues paid.

Workers’ compensation, defined as “payments people receive periodically from public or private insurance companies for injuries received at work,” is included in the CPS money income concept. Because workers’ compensation benefits paid “under a workers’

compensation act or a statute in the nature of a workers' compensation act" are not taxable, they are not included in the AGI concept.

Social Security pensions are a part of the CPS concept, as well as Social Security survivors' and disability insurance payments. IRS includes only the taxable portion of Social Security benefits in its AGI computation. At least 15 percent of benefits are not taxable; if income is under \$34,000 (\$44,000 for a married couple filing jointly) and the taxpayer is not married filing separately and living with a spouse, at least 50 percent is not taxable.

Supplemental Security Income is included in the CPS concept but not in AGI because it is not taxable.

Public assistance or welfare payments are included in the CPS concept but, again, not in AGI because they are not taxable.

Veterans' payments, under the CPS concept, consist of payments "disabled members of the armed forces or survivors of deceased veterans receive periodically from the Department of Veterans Affairs for education and on-the-job training, and means-tested assistance to veterans." These payments are not part of AGI since they are not taxable.

Survivor benefits include benefits from "private companies or unions, the Federal Government (Civil Service), the military, State or local governments, railroad retirement, workers' compensation, Black Lung payments, estates and trusts, annuities or paid-up insurance policies, and survivor payments." Except for workers' compensation, most survivor benefits are included in AGI. There is an exclusion amount, similar to the Social Security exclusion amount, for railroad retirement survivor benefits. There is also an exclusion amount based on the cost of a private annuity. Also, survivor payments made to families of military personnel who died after September 10, 2001, and payments made to survivors of victims of the 9/11 attacks are nontaxable.

Non-Social Security disability benefits such as disability income from "workers' compensation, companies or unions, Federal Government (Civil Service), military, State or local governments, railroad retirement,

accident or disability insurance, Black Lung payments, State temporary sickness, or other disability payments," are included in the CPS income concept. AGI excludes payments from workers' compensation or from military or other uniformed services if the payee became entitled to the benefits or was a member before September 25, 1975, or if the payment is due to a combat-related injury. Also, if the payment is from a private disability insurance policy for which the taxpayer paid him- or herself, then the payment is exempt from taxation. Further, railroad retirement disability is treated like Social Security disability for tax purposes.

Pension or retirement income is generally included in both concepts. The CPS concept includes income from "companies or unions, Federal Government (Civil Service), military, State or local governments, railroad retirement, annuities or paid-up insurance policies, individual retirement accounts (IRA's), Keogh, or 401(k) payments." Note that part of railroad retirement is treated like Social Security for tax purposes. Also, if an employee paid part of the cost of a pension, then payments that represent the return of his or her cost are not included in income.

Interest income under the CPS concept is made up of all interest income, including interest from "bonds, Treasury notes, IRA's, certificates of deposit, and interest-bearing savings and checking accounts." Some of this income is included in AGI. Other nontaxable interest, from sources such as tax-free municipal bonds, IRA's, and 401(k) accounts, is excluded from AGI.

Dividends received from stock and mutual fund shares are part of the CPS concept. AGI includes these amounts as well, although distributions of stock or options to buy stock (stock dividends or stock options) are usually not taxable, so long as the distribution is made in common stock and in the same way to all common stockholders.

Rents and royalties, net of expenses, and periodic payments from **estates or trusts** are included in both income concepts.

Educational assistance includes Pell grants, other Government assistance, and financial assistance received

from employers, friends, or relatives not residing in the student's household are included in the CPS concept. AGI does not include assistance from friends or relatives. Scholarships and grants that do not represent payment for services, like teaching or research, and which are used for qualified educational expenses, like tuition and books (but not room and board), are not included as they are not taxable. Assistance received from employers can be excluded up to \$5,250.

Alimony is included in both income concepts. Alimony paid is used to reduce the income of the payer in AGI.

Child support makes up part of CPS income but, as defined by IRS, is not included in AGI.

Financial assistance from outside the household that consists of regular payments that individuals receive from nonhousehold members is usually not taxable, and thus not included in AGI, although it may be treated as a gift and be taxable to the giver. This category in the CPS does not include sporadic help or irregular gifts, such as a birthday or holiday present, or educational assistance listed above.

Other income includes all other payments people receive regularly, including foster care payments, military family allotments, and income received from foreign pensions in the CPS concept. AGI includes many types of other income. For example, income from an activity the taxpayer might consider a "business" and might report a net loss for the CPS is included in AGI if the taxpayer did not expect to make a profit. For example, if someone owns two horses and gives a few riding lessons, he or she cannot then treat the upkeep of the horses as a business expense. Rather, the horses would be considered to be kept for personal use, and the income from the lessons would be reported as other income. Alaska permanent fund dividends are reported as other income. This item in AGI also includes some, although not all, foster care payments. Interestingly, the value of found property of which the taxpayer comes into undisputed possession is considered other income. Prizes, gambling winnings,

illegal income, the value of property the taxpayer stole, and rewards all count as other income in AGI.

Capital gains are not included as part of income in the CPS money income concept (although there are several "alternative" concepts for income in CPS that attempt to capture capital gains and other forms of income). AGI includes capital gains except for exclusions enumerated in the AGI definition section of this paper.

State tax refunds that were part of an itemized deduction for State income taxes in the prior year are included in AGI. CPS does not include these amounts.

► **Comparison of Income Data**

Figures A and B present income as measured by the three concepts, along with the capital gains component of AGI. Figure A shows the trend in average income across the agencies. For AGI, this is average income per tax return, and capital gains have been averaged across all tax returns and not just those with capital gains. BLS average income is measured by consumer unit, while Census average income is per household.

Figure B shows the trend in total income across the agencies. Note that, while the definitions on income according to BLS have not changed, the method of collecting income data changed in 2001 with the introduction of brackets. If a respondent reported the receipt of an income component, but refused to answer or did not know the amount, he or she was presented with brackets to select the range that the amount fell into. Prior to the introduction of brackets, these responses were left as invalid blanks. This accounts for the increase in slope for CE average and total income in 2001 [8].

Also worth noting is the acceleration in the rate of increase in AGI starting in the middle 1990's, and the downturn in AGI in 2001. The shape of the trend line for capital gains included in these figures suggests that this behavior may be largely explained by the rapid rise in the value of equities over the last half of the 1990's and the subsequent correction in those values in the early part of the current decade.

Figure A.--Average Income Across Agencies

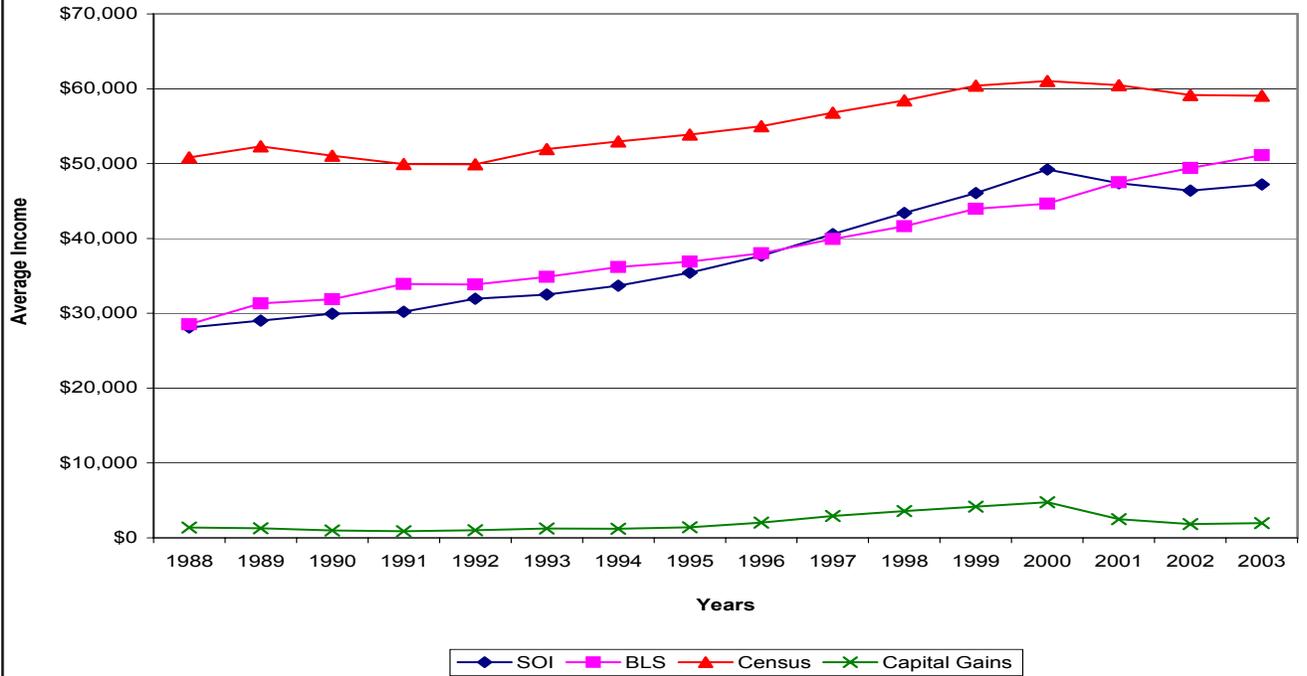
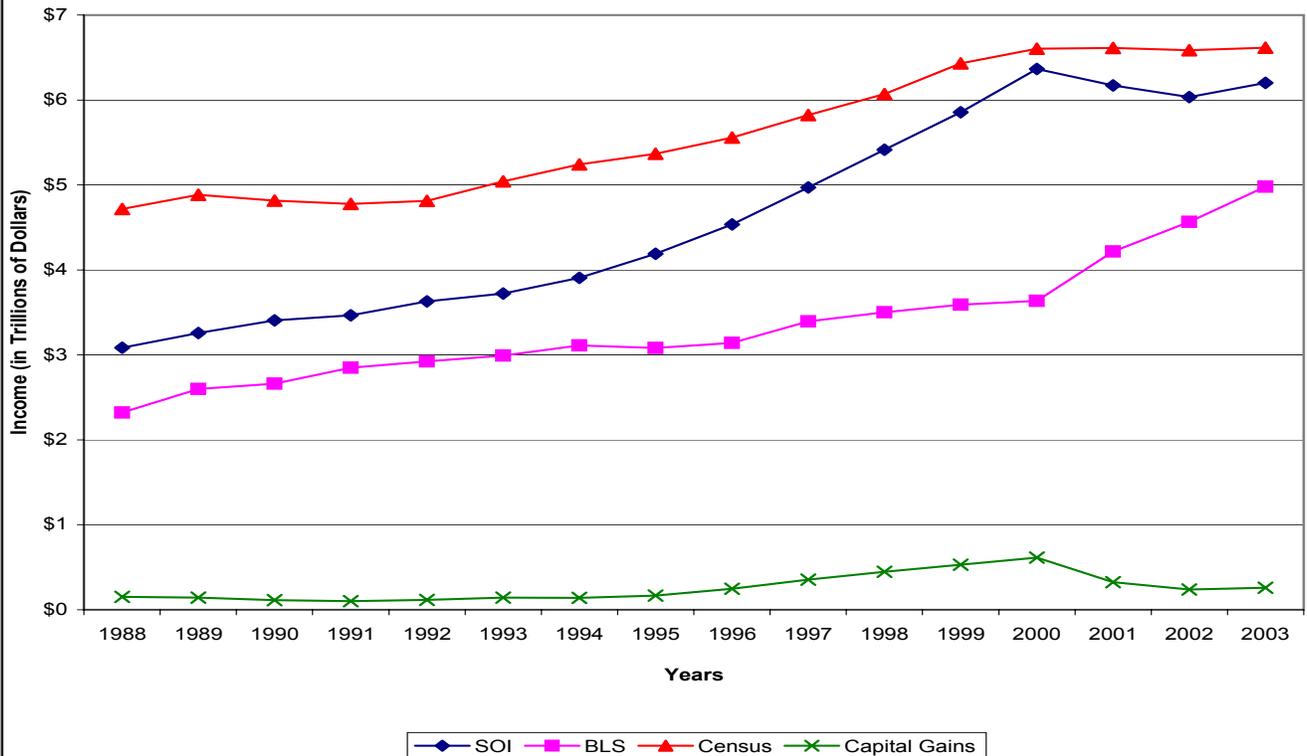


Figure B.--Total Income Across Agencies



► Acknowledgments

The authors wish to thank Laura Paszkiewicz of the Bureau of Labor Statistics and Michael Strudler of the Statistics of Income Division, Internal Revenue Service for their patient efforts in furtherance of the authors' understanding of the concepts described in this paper.

► References

- [1] "Comparability of Current Population Survey Income Data with Other Data," U. S. Census Bureau, <http://www.census.gov/hhes/www/income/compare1.html>.
- [2] Weinberg, D., "Income Data Quality Issues in the Annual Social and Economic Supplement to the Current Population Survey," http://www.welfareacademy.org/pubs/poverty/seminar/2004.10.docs/weinberg_income_data.pdf.
- [3] Ruser, J.; Pilot, A.; and Nelson, C., "Alternative Measures of Household Income: BEA Personal Income, CPS Money Income, and Beyond," <http://www.bea.gov/bea/about/fesac/Alternative-measuresHHincomeFESAC121404.pdf>.
- [4] Material describing the Current Population Survey income concept, along with the data and explanation of terms, can be found at www.bls.census.gov/cps/cpsmain.htm.
- [5] Material describing the Consumer Expenditure Survey income concept, along with the data and explanation of terms, can be found at www.bls.gov/cex.
- [6] Descriptions of AGI are taken from *Individual Income Tax Returns*, Statistics of Income Division, Internal Revenue Service, 1988-2002.
- [7] At the time of the establishment of the Social Security system, a separate system called Railroad Retirement was established for railroad workers. Tier I of the system replaces Social Security for these workers, while Tier II provides a supplemental pension amount.

The 1999 Individual Income Tax Return Edited Panel

Michael E. Weber and Victoria L. Bryant, Internal Revenue Service

The primary product of the Statistics of Income Division's Individual Statistics Branch is an annual cross-sectional sample of individual income tax returns. Some form of this annual cross section, also known as the Individual Complete Report File, has been produced every year since 1916. These annual cross sections provide the basis for most Federal tax policy analysis and research as they are consistently and reliably produced with well-known statistical properties. Longitudinal or panel samples of individual income tax returns, however, have a much shorter history. This has been largely due to their statistical and operational complexity relative to cross-sectional samples, and the added cost of producing panels given limited budgets. SOI produced a few small panels in the mid-to-late seventies and the early eighties, but all of these panels were focused on capital gains and losses. They were not meant to provide longitudinal information on other types of income, deductions, or credits. Beginning with Tax Year 1979, SOI incorporated a few Continuous Work History Sample (CWHS) Social Security Number (SSN) endings as part of the annual Individual Income Tax Return Cross Sectional Sample. These CWHS cross-sectional samples can be used to form a panel as the name implies and have been used for tax policy analysis by researchers both inside and outside the Government.¹ But, while the SOI CWHS has many wonderful longitudinal aspects, it lacks the ability to provide statistically reliable data for high-income taxpayers. For example, in 1999, taxpayers reporting over \$1,000,000 in Adjusted Gross Income (AGI) accounted for 11 percent of all reported AGI and 20 percent of all income taxes. In the annual cross-section file, which utilizes a highly stratified sample design based on income, there were 53,587 returns with \$1,000,000 or more in AGI but only 123 CWHS returns, a statistically inadequate sample for tax policy analysis.²

The first panel that attempted to use a stratified sample design that adequately sampled high-income returns and also represented the underlying annual cross-section or Complete Report File was the 1987-based

Family Panel. This panel followed all of the primary and secondary taxpayers shown on nondependent tax returns found in the 1987 Complete Report. The panel continued until 1996.

► Why the 1987 Family Panel was Terminated

Financial considerations were paramount in the decision to end the panel in 1996. As noted above, the 1987 Family Panel was drawn from the nondependent returns found in the 1987 Complete Report File. So, initially, the Complete Report and the Family Panel samples overlapped. However, since there is great volatility in the reported incomes of taxpayers in the upper income strata, many taxpayers sampled for SOI's Complete Report File at rates of 100 percent in a given year fall into strata with sampling rates of 25 percent or even 10 percent in subsequent years. These original 100-percent strata returns, once selected for the panel, must be processed in subsequent years even though they are not needed for the annual cross-sectional sampling. In addition, in 1991 the Treasury Department's Office of Tax Analysis (OTA) and SOI jointly redesigned the annual cross-sectional sample and thereby shifted the entire underlying sample structure, further reducing the overlap of the two samples. As can be seen from Table 1, in 1988, some 56 percent of the returns sampled for the Complete Report were also used in the 1987 Family Panel. By 1993, that percentage had dropped to 33 percent. If dependent returns, which are usually simple returns, are removed, the comparable figures are 71 percent and 39 percent, respectively (Table 2). If only returns selected for the panel with a 100-percent probability of selection are examined, the comparable figures are 62 percent and 28 percent, respectively (Table 3). This diminishing overlap in the high-income returns is, therefore, very problematic from a cost perspective. In terms of manual processing time, returns in the various 100-percent strata take over 26 minutes on average to process, almost 5 times the amount of time it takes to process returns with AGI under \$100,000. During preparations for processing Tax

Year 1997 returns, it became apparent that, due to the diminishing overlap, SOI would not have enough funds available to complete the processing of both the 1987 Family Panel and the 1997 Complete Report File.

A second reason for ending the 1987 Family Panel was its age. The longer any panel continues, the less its usefulness for the analysis of current issues. For example, assume the 1987 Panel had continued through 2005 and an analysis was performed on the Bush 2001

Table 1.--Overlap between the 1987 Family Panel and the 1987-1993 Complete Reports (CR)

SOIYR	87 Panel	CR	Both	Panel Overlap with CR
1987	86,975	125,788	86,907	99.9%
1988	116,342	110,495	65,385	56.2%
1989	120,803	110,566	59,077	48.9%
1990	124,087	104,277	55,791	45.0%
1991	123,295	125,756	49,494	40.1%
1992	125,228	103,190	45,479	36.3%
1993	132,583	104,357	44,283	33.4%

Table 2.--Overlap between the 1987 Family Panel (nondependent returns) and the 1987-1993 Complete Reports (nondependent returns)

SOIYR	87 Panel	CR	Both	Panel Overlap with CR
1987	86,950	120,520	86,883	99.9%
1988	92,363	106,876	65,109	70.5%
1989	97,207	106,836	58,882	60.6%
1990	101,839	101,512	55,650	54.6%
1991	104,154	123,094	49,385	47.4%
1992	107,917	100,589	45,388	42.1%
1993	112,951	101,779	44,221	39.2%

Table 3.--1987 Panel Returns sampled at 100 percent rate and overlap with SOI cross-section*

SOIYR	Both	Panel overlap with CR
1987 100% panel rate = 12,411		
1987	12,411	100%
1988	7,642	62%
1989	6,301	51%
1990	5,480	44%
1991	4,096	33%
1992	3,571	29%
1993	3,422	28%

* Obtained by matching the 1987 panel 100 percent sample returns in each year with the 100 percent returns in the CR for each year. This is an overestimate as the number of 100 percent records in the panel grows each year due to divorce and dependents filing their own return.

Tax Cuts. The results would not have provided an analysis of how American taxpayers of year 2000 responded to the tax cuts over the next 5 years. It would have provided an analysis of how individual taxpayers who filed a return in the panel base year of 1987 responded to the 2001 tax cuts. Those populations of taxpayers almost certainly were very different. This is not to say that long-lived panels are useless; indeed, long-lived panels are highly valued by researchers, but, as they age, the nature of the analysis that can be performed upon them changes. Given limited resources, there is a tradeoff between the longevity of a panel and the age of its underlying base year data. As any panel ages, it loses its ability to speak to the issues of the current day. Most researchers and analysts find that the most pressing issues, usually defined by their job requirements, are those of the current day.

Thus, given the resource concerns and the age of the panel, a decision was made jointly between SOI and OTA to end the 1987 panel after processing of the 1996 data was complete.

► **The 1999 Edited Panel--The Beginning**

The planning process for the next panel began in the fall of 1997. Consultants from Westat were contracted to moderate the process and to provide statistical guidance and sample design recommendations. Over the next year, Westat met extensively with members of SOI and also moderated several meetings between members of SOI and individuals from OTA.³ The wide-ranging discussions covered such topics as greater utilization of the CWHS concept to completely integrating the cross-section and panel studies into one sample.⁴ In January 1999, Westat produced a report entitled "Issues in the Design of a New Panel of Individual Tax Returns" which provided the basic contours of the sample design for the Tax Year 1999 Edited Panel that was put into operation in May 2001.⁵

► **Basics of the Individual Cross-Section Sample**

Before discussing the specifics of the Edited Panel sample design, the basics of the Complete Report sample design should be discussed. Table 4 shows the final

weighting stratifications for the 1999 Complete Report. The stratifications are based on a tabulated income amount, which is indexed to the GDP each year, and the inclusion of various IRS forms and schedules. For certain income strata, a few additional substrata are created based on a "Degree of Interest" variable. This variable is derived from various components on the tax return such as filing status and the number of dependents.⁶ Prior to the planning and implantation of the 1999 Edited Panel, the prescribed sampling rates ranged from a low of 1 to a high of approximately 1-in-5,000. When ranking the cost of processing returns for the SOI program by stratification, the lower income stratifications (which are dominated by CWHS returns) are the cheapest to process, and the 100-percent stratifications are the most expensive.⁷

► **The 1999 Edited Panel Sample Design**

One of the key Westat panel design recommendations, and one that was readily accepted and implemented, was that the 1999 Edited Panel should make greater use of the CWHS concept and thus contain a larger sample of CWHS returns. This would produce many analytical benefits but would also help SOI to maintain a more constant cost structure over time since CWHS returns could be readily used in the annual cross-sectional file as well as in the 1999 Edited Panel. Consequently, the SOI Complete Report sample design was changed to include five CWHS endings.⁸ Table 5 shows the various Complete Report strata for 1997 and 1999, as well as the percentage of returns found in each stratum that were selected due to their membership in the SOI CWHS sample. As can be seen, some strata now consist entirely of CWHS returns. Indeed, if the "Degree of Interest" stratifications, which require a larger sample size than that generated by five CWHS endings, were eliminated, the CWHS sample would provide all returns required for the Complete Report for returns showing \$120,000 or less of positive income and about one third of the required sample for returns between \$120,001 and \$250,000. In fact, it was decided that the "Degree of Interest" stratifications were not needed for the panel and that a roughly 33-percent subsample of the returns between \$120,000 and \$250,000 of positive income would be adequate as well. Thus, the CWHS sample accounts for all sampled records in the panel with

Table 4.—Number of Individual Income Tax Returns in the Population and Sample by Sampling Strata for 1999

Description of the sample strata	Number of Returns by type of form attached											
	Form 1040, with Form 1116 or Form 2555			Form 1040, with Schedule C but without Form 1116 or Form 2555			Form 1040, with Schedule F but without Schedule C, Form 1116 or Form 2555			All other forms		
	Population counts (2)	Sample counts (3)	Sampling Rate (4)	Population counts (5)	Sample counts (6)	Sampling Rate (7)	Population counts (8)	Sample counts (9)	Sampling Rate (10)	Population counts (11)	Sample counts (12)	Sampling Rate (13)
Total	2,698,596	36,528	100.00	17,272,967	36,746	100.00	1,521,415	4,470	100.00	105,825,250	95,824	100.00
Indexed Negative Income ²												
\$10,000,000 or more	101	101	100.00	504	504	100.00	65	65	100.00	586	586	100.00
\$5,000,000 under \$10,000,000	86	86	100.00	609	609	100.00	121	121	100.00	750	750	100.00
\$2,000,000 under \$5,000,000	346	103	29.77	2,349	741	31.55	533	190	35.65	2,673	862	32.25
\$1,000,000 under \$2,000,000	703	100	14.22	5,188	818	15.77	1,312	214	16.31	5,192	847	16.31
\$500,000 under \$1,000,000	1,472	54	3.67	14,089	498	3.53	3,990	123	3.08	12,007	401	3.34
\$250,000 under \$500,000	3,007	35	1.16	34,810	310	0.89	9,768	78	0.80	27,489	258	0.94
\$120,000 under \$250,000	5,467	34	0.62	75,090	352	0.47	17,257	89	0.52	58,046	287	0.46
\$60,000 under \$120,000	**	**	**	117,062	292	0.25	17,810	36	0.20	87,367	224	0.26
Under \$60,000	**	**	**	321,426	425	0.13	33,741	52	0.15	327,804	446	0.14
Indexed Positive Income ²												
Under \$30,000	143,649	65	0.05	1,874,895	973	0.05	108,513	62	0.06	27,809,524	13,804	0.05
Under \$30,000	199,772	223	0.11	3,464,052	3,586	0.10	172,357	188	0.11	29,242,683	14,749	0.05
\$30,000 under \$60,000	198,137	101	0.05	1,686,282	787	0.05	184,402	83	0.05	6,205,425	6,492	0.10
\$30,000 under \$60,000	314,375	373	0.12	3,351,363	3,562	0.11	281,068	299	0.11	20,613,240	10,179	0.05
\$60,000 under \$120,000	408,896	191	0.05	1,874,804	959	0.05	232,413	120	0.05	5,618,229	6,224	0.11
\$60,000 under \$120,000	350,365	355	0.10	2,274,376	2,361	0.10	190,886	161	0.08	10,025,047	4,905	0.05
\$120,000 under \$250,000	243,101	367	0.15	466,388	680	0.15	106,656	139	0.13	2,374,629	2,408	0.10
\$120,000 under \$250,000	328,531	958	0.29	1,085,930	3,115	0.29	76,074	198	0.26	1,584,226	2,346	0.15
\$250,000 under \$500,000	277,335	1,849	0.67	454,376	3,100	0.68	61,525	371	0.60	567,361	3,727	0.66
\$500,000 under \$1,000,000	128,630	3,105	2.41	125,068	2,979	2.38	16,675	404	2.42	166,746	4,029	2.42
\$1,000,000 under \$2,000,000	54,290	6,581	12.12	31,129	3,767	12.10	4,280	542	12.66	52,437	6,447	12.29
\$2,000,000 under \$5,000,000	27,424	8,938	32.59	10,170	3,321	32.65	1,532	498	32.51	20,333	6,545	32.19
\$5,000,000 under \$10,000,000	7,813	7,813	100.00	2,015	2,015	100.00	302	302	100.00	4,273	4,273	100.00
\$10,000,000 or more	5,096	5,096	100.00	992	992	100.00	135	135	100.00	2,145	2,145	100.00

¹ Each population member is assigned a degree of interest based on how useful it is for tax modeling purposes. Degree of interest ranges from one (1) to four (4), with a one being assigned to returns that are the least interesting, and a four being assigned to those that are the most interesting. 'All' refers to income classes for which returns with all four degrees of interest are assigned.

² Positive and Negative Income classes are divided by a Chain-Type Price Index for the Gross Domestic Product of 1,1480 to represent a base year of 1991.

** Sampling Strata Collapsed.

Table 5.—CWHS Selection as Percentage of Cross-sectional Sample Stratifications, 1997 and 1999 SOI Samples

Description of the sample strata	Degree of interest ³	Stratification by type of form attached							
		Form 1040, with Form 1116 or Form 2555		Form 1040, with Schedule C but without Form 1116 or Form 2555		Form 1040, with Schedule F but without Schedule C, Form 1116 or Form 2555		All other forms	
		1997 CWHS %	1999 CWHS %	1997 CWHS %	1999 CWHS %	1997 CWHS %	1999 CWHS %	1997 CWHS %	1999 CWHS %
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Indexed Negative Income ⁴									
\$10,000,000 or more	All								
\$5,000,000 under \$10,000,000	All								
\$2,000,000 under \$5,000,000	All		0.97%		0.13%				
\$1,000,000 under \$2,000,000	All	1.41%			0.24%		0.93%		
\$500,000 under \$1,000,000	All		1.85%	0.51%	1.00%	0.88%		1.67%	2.24%
\$250,000 under \$500,000	All		11.43%	4.35%	5.16%	2.25%	6.41%	4.95%	6.20%
\$120,000 under \$250,000	All		14.71%	3.70%	11.93%	4.29%	5.62%	5.77%	12.36%
\$60,000 under \$120,000	All	**	**	7.84%	20.21%	5.77%	11.11%	8.76%	18.30%
Under \$60,000	All	**	**	24.47%	35.14%		25.00%	19.52%	32.81%
Indexed Positive Income ⁴									
Under \$30,000	1							90.93%	100.00%
Under \$30,000	2		100.00%	61.42%	100.00%	66.67%	100.00%	61.96%	100.00%
Under \$30,000	3-4	24.14%	53.36%	23.70%	47.52%	23.35%	51.06%	24.73%	48.54%
\$30,000 under \$60,000	1-2	56.76%	100.00%	62.00%	100.00%	59.72%	100.00%	61.79%	100.00%
\$30,000 under \$60,000	3-4	20.59%	46.38%	21.81%	46.50%	20.39%	39.46%	22.96%	45.76%
\$60,000 under \$120,000	1-3	54.08%	100.00%	55.87%	100.00%	52.05%	100.00%	57.05%	100.00%
\$60,000 under \$120,000	4	19.92%	50.70%	19.49%	49.98%	21.88%	50.93%	20.51%	50.00%
\$120,000 under \$250,000	1-3	12.56%	33.79%	16.12%	33.97%	14.09%	28.78%	14.89%	34.65%
\$120,000 under \$250,000	4	6.84%	18.16%	7.04%	16.18%	6.71%	16.67%	7.73%	17.05%
\$250,000 under \$500,000	All	3.84%	7.95%	2.67%	8.10%	2.30%	7.01%	3.09%	8.48%
\$500,000 under \$1,000,000	All	0.93%	2.19%	0.76%	2.32%	1.76%	1.98%	0.76%	1.99%
\$1,000,000 under \$2,000,000	All	0.23%	0.43%	0.10%	0.61%	0.39%	0.74%	0.26%	0.37%
\$2,000,000 under \$5,000,000	All	0.05%	0.13%	0.08%	0.18%	0.00%	0.20%	0.09%	0.15%
\$5,000,000 under \$10,000,000	All	0.04%	0.05%	0%	0.05%	0.00%	0.33%	0.04%	0.07%
\$10,000,000 or more	All	0%	0.04%	0%	0.10%	0.00%	0%	0%	0.00%

positive income up to \$250,000. It was also determined that the additional stratifications by form type would not be needed either. Consequently, the lowest sampling rate in each income strata sampling group (determined by the type of forms and schedules attached to the return) became the maximum sampling rate for that income stratum.

Another recommendation of the Westat consultant's was to design a targeted high-income cohort. The 1987 Family Panel design essentially selected all 1987 cross-section high-income returns for inclusion in the panel, and, in the end, the costs associated with that decision

forced the termination of the panel after 10 years. As a general rule, the larger the selection probability, the more expensive the return is to process; therefore, decisions about sample size for high-income returns, particularly those with over \$2,000,000 of positive income, are crucial in determining project costs. A smaller high-income sample would create the possibility of a longer lived panel and/or the possibility of multiple high-income waves starting perhaps every 5 years. The first step in subsampling high-income returns was to determine how much if any of the 100-percent stratum should be subsampled. A Westat report confirmed OTA's initial opinion that returns above \$20,000,000 of positive in-

come should not be subsampled but rather included in the panel at 100 percent.⁹ Consequently, returns below \$20,000,000 and above \$250,000 would be subjected to subsampling. To that end, analysts from Westat, in conjunction with SOI and OTA, analyzed over 30 potential subsampling schemes using a linked version (or panel) of the 1996 and 1997 Complete Report files.^{10,11} This intensive process required Westat to evaluate each scheme in terms of coefficients of variation (CV) for various items in 1996 and also to compute the CVs for the differences in totals for the various items between 1996 and 1997. To quote from the report: “The primary goal was to select a panel that had acceptably low CV’s for cross-sectional estimates and estimates of change... In addition, a secondary consideration was how the distribution of the sample among income classes would change over time ..(as).. one of OTA’s desires was to avoid allocations that would become too thin at the tails of the income distribution as incomes changed over time.” As various designs were discarded, others were refined, and, in the end, Design 16A was chosen. (See Table 6)

► The Issue of Late Filed Returns

A subtlety of the annual cross-section must be addressed at this point: Not all Tax Year 1999 returns are filed by the end of Calendar Year 2000. A significant portion of Tax Year 1999 returns were filed in Calendar Years 2001 and 2002. Keeping the sample open for an additional 2 years in order to obtain these returns would force policymakers to use outdated data for decision-making. For instance, sampling for the Tax Year 1999 file would not be complete until as late as December 31, 2002. Therefore, in order to provide more timely statistics, SOI produces a sample of tax returns filed during each calendar year. Approximately 97 percent of the returns received in a given calendar year are for the preceding tax year. For example, in Calendar Year 2000, some 97 percent of taxpayers filed their Tax Year 1999 returns. The remaining 3 percent of the returns filed in a given calendar year are generally for the preceding 2 tax years. In our example, these would be Tax Years 1997 and 1998. These “prior year” returns are used as proxies for the Tax Year 1999 returns that were not filed timely during Calendar Year 2000.

When creating panels, however, we have the luxury of time and are thus able to create a sample from a virtually complete set of returns for a given tax year. The Tax Year 1999 Edited Panel is a sample of Tax Year 1999 returns. Since each calendar year was sampled independently, it would be appropriate, when combining all 3 years of Tax Year 1999 sampling, to treat each year as a separate level of stratification. But as can be seen from Table 6, the sample sizes for most of the stratifications for Calendar Years 2001 and 2002 are rather small. This would cause a proliferation of weights. Consequently, a decision was made not to stratify on Tax Year but to treat the 3 years as one sample with one set of stratifications and thereby reduce the variability in the weights.

► Linking Individuals and Tax Returns Over Time

In order to link tax returns and individuals over time, a unique identifier is required. Fortunately, taxpayers are required to provide their Social Security numbers on their tax forms. However, sometimes the SSN’s that are shown on the tax forms are incorrect, and, sometimes IRS transcribes them incorrectly. So, in order to prevent billionaires and millionaires from either disappearing or being linked to Earned Income Tax Credit recipients, SOI performs a review of panel member SSN’s. The 1999 Edited Panel contains 125,108 unique panel member SSN’s. This is simply the number of base year returns in the sample plus the number of spouses on joint returns. Of the 125,108 panel members, only 456 SSN’s (44 for the primary taxpayers and 412 secondary taxpayers) were determined to be incorrect. For 392, a correction was obtained. A total of 29 returns were deleted because the primary SSN’s on these nonjoint returns were determined to be incorrect and no correction could be obtained. Note that this is not a confirmation that the remaining SSN’s are correct. Frequently, invalid SSN’s are not detectable for many years until some point in the future, often when multiple individuals use the same SSN. In addition, many corrections are made to nonpanel member individuals who accidentally, or perhaps intentionally, use an SSN that does not belong to them and thus cause an incorrect linkage to a panel member. While these figures paint a positive picture for the quality of the SSN linkages, one area of concern is

with the use of IRS-generated Taxpayer Identification Numbers or ITIN's which are provided to individuals who are required to file a return but who have not been issued an SSN. Quite often, these individuals will, in time, obtain an SSN from the Social Security Administration and then file using it in subsequent years. This breaks the link to the previous set of returns and, if not caught prior to sampling, will cause the loss of valid sample units.

► Future Plans

The 1999 Individual Income Tax Return Panel is currently being weighted and will include data from 1999 through 2003. Subsequent years of data will be appended to the panel as they become available. Our attention now turns to learning how to use the panel and the publication of tabulations and analysis, hopefully the subject of many future papers.

► Endnotes

- ¹ For more information on the CWHS panel, see Weber, Michael (2004), "The Statistics of Income 1979-2002 Continuous Work History Sample Individual Income Tax Return Panel," 2004 Proceedings of the American Statistical Association, Social Statistics Section.
- ² For example, the estimated amount of AGI, using the full sample of returns with a reported AGI of \$1,000,000 or more, was \$653,184,370,292. The coefficient of variation for this amount is .19. Using the 123 CWHS returns and applying a weight of 2,000 (5 different endings were used in 1999, thus producing a 1-in-2000 sampling rate) produced an estimate of \$696,643,752,000. The specific coefficient of variation for this amount has not been calculated, but can be assumed to be significantly larger than .19.
- ³ Notes from these meetings are found in an unpublished Westat document entitled "Meeting Minutes

For Task Order #13 Under Contract No. TIRNO-96-D-00030.0005."

- ⁴ More information on this topic is found in an unpublished Westat document entitled "Integrated versus Separate Panel and Cross-Sectional Sample Designs," September 1999.
- ⁵ Tax Year 1999 returns were generally filed in Calendar Year 2000. As the Tax Year 1999 Based Edited Panel was defined as a subsample of the 1999 Complete Report File, panel membership did not need to be defined for sampling purposes until Tax Year 2000 returns, which were generally filed in Calendar Year 2001, were received by IRS and ready for SOI sampling in May 2001. As is often the case, final sample decisions were not finalized until the last possible moment.
- ⁶ For additional information on the sample design of the annual Complete Report sample, see Internal Revenue Service, Statistics of Income--Individual Income Tax Returns, Publication 1304, 1999, "Section 2: Description of Sample."
- ⁷ It should be noted that SOI processes many CWHS returns without any manual processing costs.
- ⁸ This change was actually instituted for Tax Year 1998. The sample design for Tax Year 1999 is identical to Tax Year 1998. Consequently, a table showing the Tax Year 1998 stratifications has been omitted.
- ⁹ Westat unpublished memo, "Report on Substrata for Strata 1 and 24," October 9, 2000.
- ¹⁰ Unpublished Westat report "Design of a Panel Sample of Tax Returns--Final Report," May 2001.
- ¹¹ The 1997 file was augmented by data from the IRS Individual Returns Transaction File when a 1996 Complete Report SSN did not appear in the 1997 Complete Report.

Trends in 401(k) and IRA Contribution Activity, 1999-2002--Results from a Panel of Matched Tax Returns and Information Documents

*Peter Sailer and Victoria L. Bryant, Internal Revenue Service,
and Sara Holden, Investment Company Institute*

By combining individual tax returns (Form 1040) and information returns (such as Forms W-2 and 5498) in one panel database, the Statistics of Income (SOI) Division has made it possible to study trends in contributions by individual taxpayers over time to Individual Retirement Arrangements (IRA's), as well as the participation in other types of retirement plans. Using a simple random panel of over 71,000 individual taxpayers who filed for Tax Years 1999 through 2002, this paper will analyze persistency in taxpayers' contribution activities in traditional IRA's and in 401(k) plans. Several possible factors affecting persistency will be considered, including age, marital status, gender, and income.

All of the analysis in this paper is limited to those taxpayers who filed for all 4 years in the study--1999--2002. In the case of joint returns, primary and secondary taxpayers were considered separately. Weighted, the file represents 143.2 million taxpayers, about 81 percent of the original 177.0 million who filed for 1999. Changes in marital status or marriage partners did not affect inclusion in the study--as long as an individual was represented as a taxpayer on a return for all 4 years, he or she could be included in the panel.

► Taxpayers' Use of Traditional IRA's

At yearend 2002, nearly 50 million taxpayers held a total of \$2.5 trillion in IRA assets. The bulk of these were traditional IRA's: 40 million taxpayers with \$2.3 trillion in assets. Traditional IRA's may be contributory and/or the result of rollovers from qualified employer-sponsored retirement plans. This paper focuses on contribution activity among taxpayers in the 1999--2002 panel dataset.

► Definition of Traditional IRA Plans

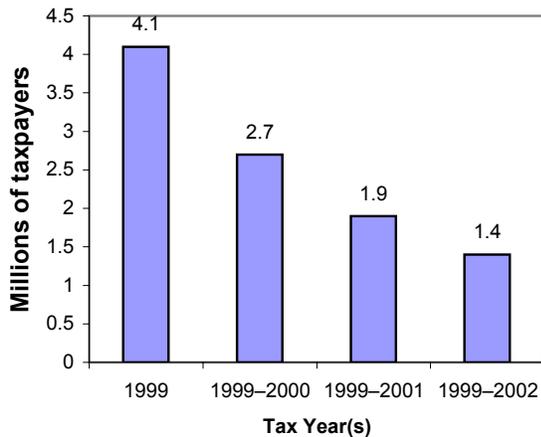
Individual Retirement Arrangements (IRA's) were created by the Employee Retirement Income Security Act (ERISA) of 1974. These first IRA's, termed tradi-

tional IRA's, were still the principal type of IRA's held by most taxpayers in 2002. In general, contributions to traditional IRA's could be made by individuals who received taxable compensation (e.g., wages, salaries, commissions, self-employment income). For 1999 through 2001, the limit was generally the lesser of \$2,000 or the individual's taxable compensation. For 2002, the maximum contribution amount was raised to \$3,000 for taxpayers under age 50, \$3,500 for those age 50 or older (the extra \$500 being a "catch-up" contribution; both catch-up contributions and the higher limits were legislated by the Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA)). Additional restrictions based on age, coverage by a retirement plan at work, income, and filing status limit the amount that could be contributed on a tax-deductible basis in any particular tax year. In general, taxpayers less than 70½ years of age who were not covered by a retirement plan at work could make a traditional IRA contribution that would be deductible on their income tax returns. However, households with an individual covered by a qualified pension plan at work generally found this deduction limited based on income level and filing status (see Internal Revenue Service, Publication 590, for an explanation of the rules).

► Persistency in Traditional IRA Contributions

Figure 1 shows that 4.1 million of the taxpayers represented by the 1999--2002 panel dataset made contributions to traditional IRA plans for Tax Year 1999. Earlier papers explored some of the characteristics of individuals making IRA contributions in a given year (see Sailer, Gurka, and Holden (2003); and Sailer and Holden (2005)). This paper will explore the persistence of the 1999 traditional IRA contributors in following years. Figure 1 shows that, for 2000, only 2.7 million of the 1999 participants made contributions. By 2001, only 1.9 million persisted, and, by 2002, the participation was down to 1.4 million--34.8 percent of the original contributors in 1999.

Figure 1: Taxpayers with Traditional IRA Contributions for 1999 Who Filed for 2000–2002: Persistence of Contributions



One possible reason for dropping out of the IRA savings program could be that some taxpayers lost the immediate tax incentive of being able to deduct the amount contributed. As discussed earlier, taxpayers who were covered by employer-provided pension plans had income limits above which IRA contributions could not be deducted. Contributions could still be made by taxpayers who exceeded the income limitation, but the immediate tax benefit of a deduction would not be available. Nevertheless, income generated by the IRA investment remains nontaxable until it was withdrawn in either event.

Figure 2 divides taxpayers into two groups: Those who were eligible for the deduction in all 4 years, and those who were not eligible in at least 1 year. It shows that the persistency rate for those who were continuously deduction-eligible was higher than for those who were not—42.7 percent versus 27.1 percent in the fourth year, respectively.

Further research showed that marital status and gender were not significant factors in determining persistency of traditional IRA contributions. Age of taxpayer, however, did make a difference. Persistency rose steadily from 19.6 percent for taxpayers under age 25 in the beginning year to 40.4 percent for taxpayers in the 45-to-54 age group, then fell off at higher ages

(Figure 3). Since age 70½ is the cutoff age for making traditional IRA contributions, no persistency was possible above that age.

Figure 2: Taxpayers with Traditional IRA Contributions for 1999 Who Filed for 2000–2002: Persistence of Contributions

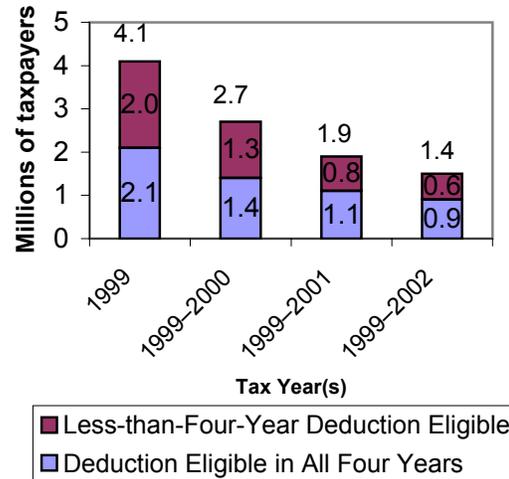
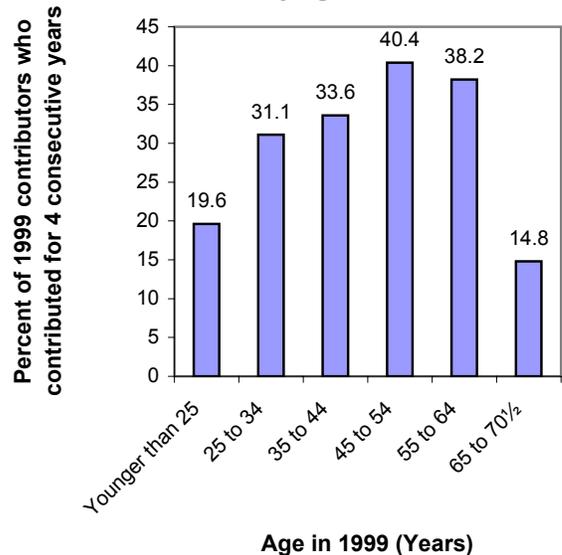
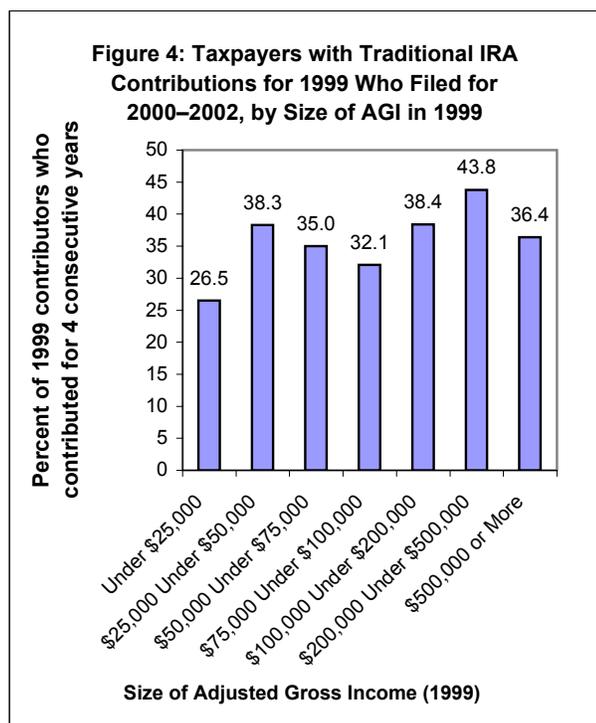


Figure 3: Taxpayers with Traditional IRA Contributions for 1999 Who Filed for 2000–2002, by Age in 1999



Size of adjusted gross income (the best indicator of total household income on the tax return) also made a difference although, somewhat unexpectedly, the distribution proved to be bimodal, with the “Under \$25,000,”

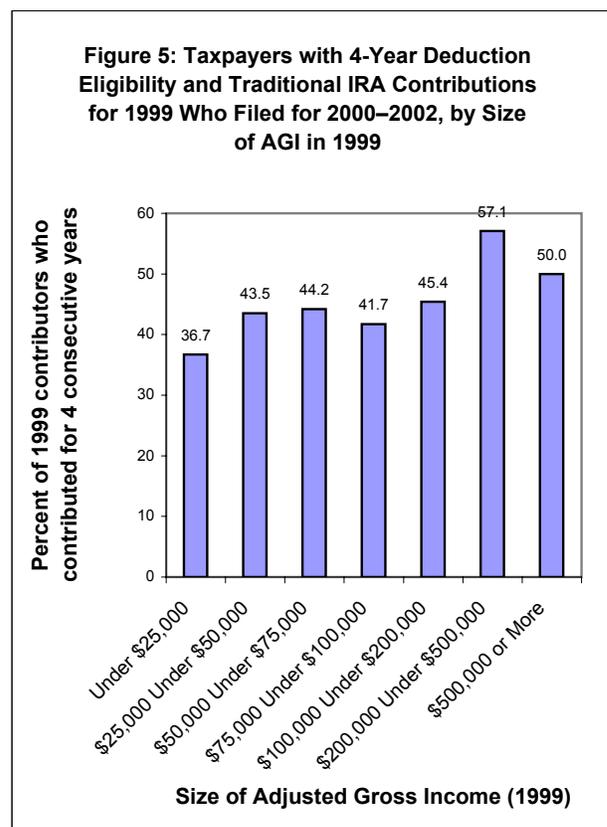
the “\$75,000 under \$100,000,” and the “\$500,000 or more” income classes showing lower persistency rates (Figure 4). Persistency is most difficult for lower-income taxpayers, and, given the many other investment



opportunities for those with high incomes, perhaps not that relevant for higher-income taxpayers. The lower persistency rates in the middle of the distribution may be related to the phaseout of the deductibility of traditional IRA contributions for some taxpayers at those levels. When only taxpayers who were eligible for IRA deductions in all 4 years were considered, persistency was higher across all income groups and did not vary as much among the lower-to-middle income groups (Figure 5).

► Reasons for Leaving the Program

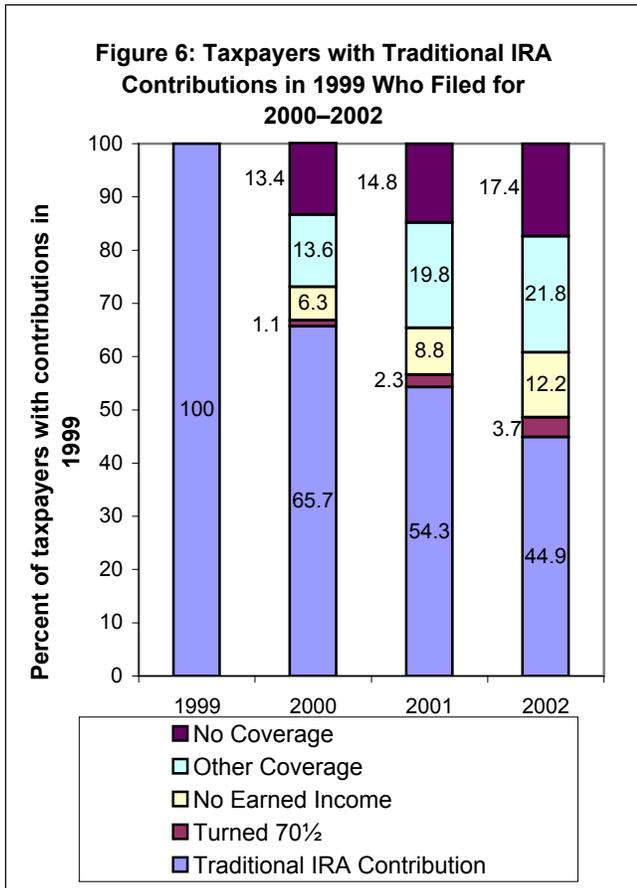
In Figure 6, several factors are considered that may have caused taxpayers who contributed to traditional IRA plans in Tax Year 1999 not to contribute in subsequent years. As mentioned previously, reaching age 70½ disqualifies a taxpayer from making contributions.



By 2002, some 3.7 percent of the 1999 contributors were no longer eligible to contribute due to their ages. Also in 2002, some 12.2 percent no longer had earned income (salaries and wages or self-employment income) and thus were ineligible. A total of 21.8 percent of the 1999 IRA contributors still met the basic age and income requirements, but had opted to save for retirement under different plans--401(k) plans, Roth IRAs, SEP or SIMPLE IRA plans--or had coverage under another employer-sponsored retirement plan. This left 17.4 percent of the 1999 contributors who were not contributing to any pension plan, even though they appeared to be eligible to do so.

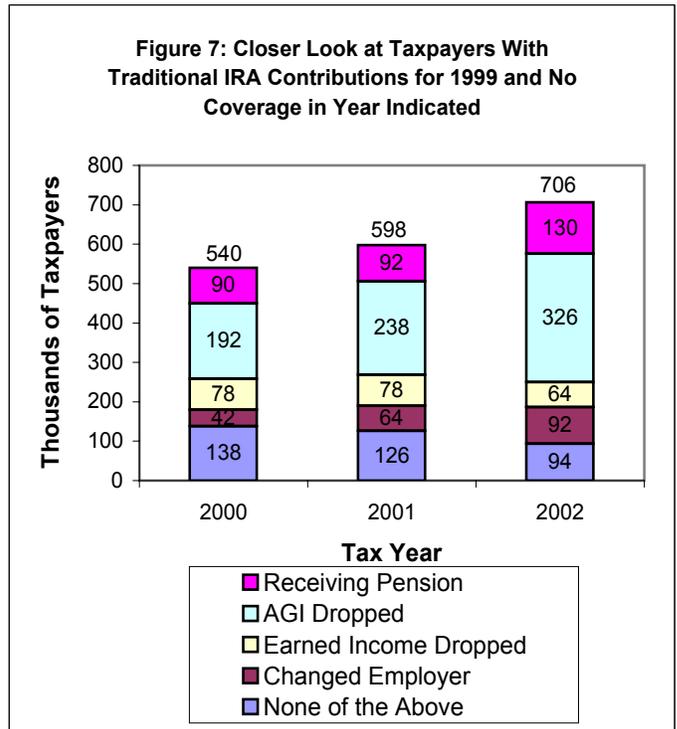
The motivation of taxpayers who stopped contributing to retirement plans is, of course, a matter of speculation. But the matched database of tax returns and information documents does contain information that supports somewhat informed speculation. Looking at the taxpayers who stopped contributing between

Figure 6: Taxpayers with Traditional IRA Contributions in 1999 Who Filed for 2000–2002



1999 and each of the 3 succeeding years, Figure 7 shows that between 90,000 and 130,000 of these individuals, depending on the year, had started making withdrawals from their pension plans (information reported on Form 1099-R). So, while they were still receiving earned income, they were presumably semiretired and no longer felt the need to build up their pension reserves. A substantial number of these taxpayers were not receiving pension income, but had experienced a drop in adjusted gross income since 1999, and may not have felt able to afford pension plan contributions. By 2002, these taxpayers numbered 326,000--over half the individuals who had stopped making pension contributions. A smaller number of taxpayers (64,000 for 2002) did not have a drop in overall income, but did have a drop in salaries and wages (earned income), which may have had a similar effect. And another 92,000 of these taxpayers changed employers between 1999 and 2002, or changed from employee to self-employed individuals--changes which may have disrupted their contribution patterns.

Figure 7: Closer Look at Taxpayers With Traditional IRA Contributions for 1999 and No Coverage in Year Indicated



► **Taxpayers’ Use of 401(k) Plans**

At yearend 2002, nationwide, 401(k) plans had accumulated \$1.5 trillion in assets (see Investment Company Institute (August 2005)). This paper uses information from individuals’ W-2 forms in conjunction with the Individual Tax Return (Form 1040) to analyze taxpayer contributions to 401(k) plans among taxpayers in the 1999–2002 panel dataset.

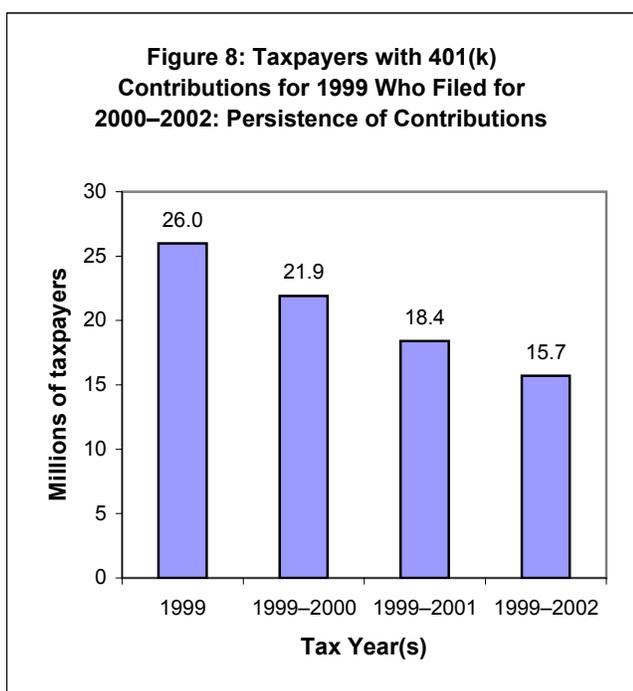
► **Definition of 401(k) Plans**

The key provision of 401(k) plans, which are a type of employer-sponsored defined contribution plan, is the ability to defer salaries by making before-tax contributions (deferrals) to an account maintained in the given participant’s name. In most instances, the participant directs the investment of the account assets, which grow tax-free until they are withdrawn. In many cases, the plan sponsor may make a matching contribution (for example, contributing 50 cents for every dollar the participant contributes up to 6 percent of salary; for a detailed analysis of 401(k) plan participant contribution activity, see Holden and VanDerhei (October 2001)).

Contribution limits in 401(k) plans are higher than in IRA's. In Tax Year 1999, the participant deferral limit in 401(k) plans was \$10,000 (\$10,500 in 2000 and 2001, and \$11,000 in 2002). "Catch-up" contributions were also permitted in 401(k) plans starting in 2002 under EGTRRA.

► Persistency in 401(k) Contributions

A comparison of persistency in 401(k) contributions (Figure 8) to that for the traditional IRA contributions (Figure 1) shows that persistency of contributors to 401(k) plans is much higher. Over 60 percent of contributors to 401(k) plans in 1999 contributed for the following 3 years as well--as compared to 34.8 percent for contributors to IRA plans.



► Future Research

The Statistics of Income Division is developing a larger, stratified panel, which will contain data for over 140,000 individual taxpayers. The data shown in this article will be rerun from this larger panel when it becomes available. In addition, further analysis of taxpayers with 401(k) contributions in 1999 and not in later years will be explored.

► Note

The views in this paper are those of the authors and do not reflect those of the Investment Company Institute or its members, nor are they the official positions of the Internal Revenue Service. Any errors are solely the responsibility of the authors.

► References

- Holden, Sarah; Ireland, Kathy; Leonard-Chambers, Vicky; and Bogdan, Michael. "The Individual Retirement Account at Age 30: A Retrospective," *ICI Perspective*, Volume 11, Number 1, February 2005 (<http://www.ici.org/pdf/per11-01.pdf>).
- Holden, Sarah and VanDerhei, Jack, "Contribution Behavior of 401(k) Plan Participants," *ICI Perspective*, Volume 7, Number 4, and *EBRI Issue Brief*, Number 238, Investment Company Institute and Employee Benefit Research Institute, Washington, DC, October 2001.
- Investment Company Institute, "Mutual Funds and the U.S. Retirement Market in 2004," *ICI Fundamentals*, Volume 14, Number 4, Investment Company Institute, Washington, DC, August 2005. (<http://www.ici.org/pdf/fm-v14n4.pdf>).
- Sailer, Peter; Gurka, Kurt S.; and Holden, Sarah (2003), "Accumulation and Distributions of Retirement Assets, 1996–2000: Results From a Matched File of Tax Returns and Information Returns," *Special Studies in Federal Tax Statistics*, Internal Revenue Service, Statistics of Income Division, Washington, DC, pp. 53-61.
- Sailer, Peter and Holden, Sarah (2005), "Use of Individual Retirement Arrangements To Save for Retirement--Results from a Matched File of Tax Returns and Information Documents for Tax Year 2001," *Special Studies in Federal Tax Statistics: 2004*, Internal Revenue Service, Statistics of Income Division, Washington, DC, pp. 25-32. (<http://www.irs.gov/pub/irs-soi/04saiaasa.pdf>).

Sailer, Peter J. and Nutter, Sarah E., "Accumulation and Distribution of Individual Retirement Arrangements, 2000," *SOI Bulletin*, Internal Revenue Service, Statistics of Income Division, Washington, DC, Spring 2004, pp. 121-134. (<http://www.irs.gov/pub/irs-soi/00retire.pdf>).

Sailer, Peter J.; Weber, Michael E.; and Gurka, Kurt S. (2002). "Are Taxpayers Increasing the Buildup

of Retirement Assets? Preliminary Results from a Matched File of Tax Year 1999 Tax Returns and Information Returns," *National Tax Association Proceedings, 95th Annual Conference on Taxation*, Washington, DC, pp. 364-369.

U.S. Internal Revenue Service, "Individual Retirement Arrangements (IRA's)," Publication 590, for Tax Year 2002. (<http://www.irs.gov/pub/irs-prior/p590--2002.pdf>).

6



Estate and Personal Wealth Sample Design

McMahon

Origins of the Estate and Personal Wealth Sample Design

Paul B. McMahon, Jr., Internal Revenue Service

In Estates and Personal Wealth, we have two studies with different populations under consideration. The Estates Study is concerned with the assets, debts, and taxes left by a decedent who had more than a certain amount of wealth. The Personal Wealth Study, on the other hand, is focused on the wealth holdings of the living. For Estates, essentially all the population appears on a sampling frame, but, to study the living, we must rely on proxies that can be observed for only a portion of the distribution, the portion in the tail.

One set of samples is the source for the data in both series of studies.

We will first briefly describe the interest in these populations. The “questionnaire” in this set of surveys is an administrative record, the Form 706, *Estate Tax Return*, and the sampling frame is a system of electronic records derived from the initial filing. We will provide a bit of background on these as well.

We focus on the studies initiated since 1982, with strata designs that changed somewhat over that time. While some previous papers have addressed certain estimation issues, such as with the Personal Wealth Estimation (Johnson and Woodburn, 1994), there have been only the briefest descriptions of the strata design or concepts.

Our goal, then, is to show how the different requirements for studies of the two populations affect this one sample design, and how that design has evolved in the light of tax law changes.

Finally, we will discuss some future directions for the series, in light of pending legislation.

► Analysts and Uses

The two main sponsors of these studies are the Office of Tax Analysis in the Department of the Treasury and Congress’s Joint Committee on Taxation. Their objective is to gather data for oversight on the opera-

tion of the tax laws and, in this case, on Estate Taxes, and projecting the effects of proposed changes to those laws. This is not limited to the revenue aspects of the tax laws.

That is, this study has to meet two uses. First, the measurement of current law, and second, determining the effect on the living population who have estates large enough for the eventual filings. In order to look at trends in the analysis, we need to be concerned about the effect of economic conditions at the time of the observations (the date of death), the time of life considerations (youthful spenders versus middle-age savers, for example), and what the sociologists call age cohorts, where history affects economic decisions (the Depression generation’s thrift).

There is also an underlying philosophical question: Does the operation of the Estate Tax, in concert with a graduated income tax, prevent the concentration of wealth into few hands? At the beginning of the twentieth century, some politicians, like Theodore Roosevelt argued in favor of the Estate Tax on just this issue. More recently, there have been numerous articles this past spring in the *New York Times* and the *Wall Street Journal*, for example, on the concentrations of incomes. Income is often taken as a proxy for wealth; so, this question is clearly of continued interest.

Indeed, using data from Estate Tax Returns dating back to 1916, the National Bureau for Economic Research (NBER) published a working paper that considers this very concentration issue (Kupczuk and Saez, 2004). Although the data used in that study are from many years in the past, the sample designs for most of those years actually originated in the mid-1980’s and reflect the plans developed for sampling more recent tax filings.

► The Administrative Records

The basic data for these studies use the records that arise from what some have called the “Death Tax.” It is more accurate, though, to call it a transfer tax, as the

change of an asset’s title to some beneficiary or heir is the proximate cause for this tax or its complement, the gift tax. The tax return, which acts as the questionnaire for our studies, is Form 706, *Estate Tax Return*.

The assets that are considered for this tax are everything owned by the decedent: art, bonds, cars, personal effects, through to zoom lenses and beyond. That is, the filing is based on a complete inventory of an individual’s possessions. In this, it is similar to the information that the Federal Reserve attempts to obtain in its Survey of Consumer Finance.

There are major differences between the data collected for the Federal Reserve surveys and the IRS studies, however. First, the tax form also includes insurance payments to the estate and gifts made before the decedent’s death, which would not be included in the Finance Survey. Then, the law permits deductions for the costs of such items as estate administration, the funeral, and legal counsel, as well as exempting the contributions to charities and the spouse of the decedent.

Another difference is that the value of the assets is usually assessed at the time of death, not as of some common reference date for all respondents.

The main difference, though, arises from the populations these two sets of studies targets. The Survey of Consumer Finance seeks to estimate the holdings of all households, while the Estates and Personal Wealth studies are limited to individuals who exceed a certain threshold set by the tax code.

Figure 1.--Estate Tax Return Filing Thresholds for Selected Years

<u>Year of Death</u>	<u>Gross Estate Threshold</u>
1997	\$600,000
1998	\$625,000
1999	\$650,000
2000 & 2001	\$675,000
2002 & 2003	\$1,000,000
2004 & 2005	\$1,500,000
2006 – 2008	\$2,000,000

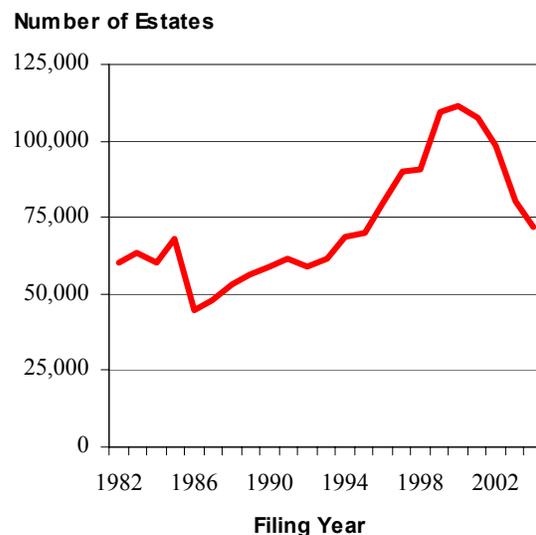
If the value of those possessions at the time of the decedent’s death is below the threshold amount shown

in Figure 1, then there is no estate tax. That threshold varies depending on the year of the decedent’s death. It is currently \$1.5 million, rising to \$2 million in January 2006. These values have been updated in the tax code periodically; in 1977, for example, the threshold was \$60,000.

Filing is not required for smaller estates, though some do if the value is near the boundary. This may be due to the difficulty in itemizing all of an estate’s assets. In those cases, amended returns will be filed, and perhaps a tax assessed, but such cases are outside the scope of this set of studies; we are only concerned with initial filings.

One can see the effect of raising the threshold quite clearly in Figure 2. In 1986, the exclusion was doubled, to \$120,000, with a resultant sharp drop in filings and again, after the 2001 tax bill passed, which raised the limit several times in succession.

Figure 2.--Annual Filings of Estate Tax Returns



While the law and regulations provide one source of limitations on the studies, and thereby the design, another is in the physical properties of the documents and the processing regimen.

The Estate Tax Return is filed on paper as a large package with sections that are partly structured and partly

respondent-created. While Form 706 is, on the surface, highly standardized, the space allowed for some schedules (such as a list of heirs) is sometimes insufficient. This leads the attorney or executor to create substitute schedules of their own design.

The filing regulations also mandate the inclusion of the will, unless the decedent died intestate, appraisals of real property, and the death certificate. While the last may be relatively standardized, the will and appraisal(s) are not.

Moreover, all of these filings are subjected to an audit review, unlike the small proportion of Individual Tax Returns. Such audits keep the return unavailable for considerable lengths of time. Thus, the Statistics of Income studies must capture the return first and cannot wait for the entire population to become available; the sample must be selected as the returns are processed through the administrative pipeline.

The filing deadline for these documents is 9 months after the decedent's death. Extensions to this deadline are often required, because it takes time to locate some financial records, and for some assets to come to light. Since evaluating the effect of changes to the law is an objective, focus on a particular year of death means we must continue the selection over more than 2 years: the focus year and at least the following 15 months.

In practice, given the administrative environment, the minimum effective sampling period is 3 years. The additional months arise from the cycle of updating the computer programs, where the latest versions are introduced each January.

We want to use an electronic record in the sampling of these estates because, while selecting the returns as paper records ensures their retention for statistical purposes, this direct approach is costly and difficult and limits stratification options. The 1977 Study's manually-selected sample was limited to three strata, for example, and required considerable daily coordination with the ten national Service Centers where the returns were filed.

Yet the use of the computer records also gives rise to limitations. Ignoring audit trail codes, tracking data,

and name and address information, there were only 16 amounts available in 1982, less than we can use today, but not by much. Most of those, 13, were involved in the calculation of the tax liability. This left a bare handful as possibly useful for sampling purposes, including some of the "code" fields.

Decedent's Year of Death was available. This was, and is, a tax-related field due to changes in the filing threshold; so, it was an administrative requirement.

For 1982, though, the Statistics of Income Division managed to convince the other interested parties within the Service that the age of the decedent could be useful. Rather than have a clerk calculate the age, though, the Service decided to include the Date of Birth. Gender, which could have been an important stratifier, is not available.

► The Stratifiers

Longitudinal studies in the sociology field have long noted that there are three effects to the group under observation: current events, time of life, and age cohort. We cannot easily address this last effect, that of the age cohort, at least not in the near future, because the observations on this group trickle in over such a long time.

We could address the aspect of current events' effect by focusing on all the decedents in a single year. "Current events," in this context, means not only the operation of economic conditions, but also the tax provisions then in force. Years ending in 2, 6, or 9 were selected; so, the first focus year included in this review is 1982.

Likewise, we could address the "time of life" through the age of the decedent (since we have the dates for both birth and death). This sociological concern has an economic component in the nature of financial holdings. For example, middle-aged people are often counseled to focus their investment strategy on growth, while retirees frequently look to revenue-producing equities. One tax consideration that arises is the unrealized capital gains included in the estate. By considering the age of the decedent, then, we can improve the measures in the composition of estates.

Age can also improve the reliability of the personal wealth estimates, which depend on this factor in the construction of the weighting classes.

Age and a focus year, though, would not aid in reducing the sampling error of the monetary estimates all that much, though. For that, we needed a variable that was reasonably correlated with the key amounts of interest. Given that this is a general sample to support ambiguous analysis (at the time of the design, anyway), that left Total Gross Estate as the monetary stratifier.

► Selection Method

Since the selection process was computerized, we took advantage of a Bernoulli mechanism, the “Transformed Taxpayer Identification Number,” used in selecting other IRS Business Master File samples, such as for the Corporations and Partnership Studies (Harte, 1986). This permanent random number procedure was meant to improve the year-to-year estimates of change by increasing the likelihood of an entity being included in the sample in succeeding years. Clearly, this is not an issue for Estates, but it did reduce the programming burden.

The selection probabilities were set within strata, with those records with a Transformed Taxpayer Identification Number below the designated probability selected for the sample.

In addition to that selection process, a 1-percent Continuous Work History Sample (CWHS) set of ending digits for the Social Security numbers was employed. We felt that, since some of the CWHS digits were in use for the Statistics of Income Individual Study, this might allow a greater overlap between the two studies.

► Strata Boundaries

There are two sets of boundaries that need to be determined: age, and size of Gross Estate. Fortunately, in the later case, our task was simplified by the administrative systems. Each return was assigned a Gross Estate Code, manually, based on the size of the Estate. At the time this design was first implemented, the value itself was not available.

Gross Estate Codes, shown in Figure 3 below, with a value of less than 6 were for returns below the filing threshold in 1982, and thus were not subjected to the Bernoulli sampling. These smaller estates were filing for the record only, though we did sample them using the CWHS digits.

Figure 3.--Defining the Gross Estate Code

<u>Size of Gross Estate</u>	<u>Code</u>
Under \$300,000	1 - 5
\$300,000 under \$500,000	6
\$500,000 under \$1,000,000	7
\$1,000,000 under \$5,000,000	8
\$5,000,000 or More	9

Determining the age groups was a more difficult problem. The sample has to address two populations: the estates affected by the tax law and the living population for the Personal Wealth Estimates. In addition, we made the assumption that the age distributions within the Gross Estate categories would have a significant impact; so, we planned separate age classes for the various Gross Estate Codes. The reasoning was that, as age increases, the opportunity to accumulate wealth also increases. Thus, the median age for the smaller estates’ decedents would be less than that for larger estates.

The data we had available at that time were from the 1977 Estates Study, which as we noted above had but three strata based on the size of Gross Estate. The estimates were tallied into 5-year bands. As one might expect, given the nature of the population under consideration, most of the low age-groups were empty of observations.

Over the years from 1977 to 1982, though, the number of estates in each category grew, even as the total number declined due to a rise in the filing threshold. This growth resulted from both inflation effects and the normal growth of the economy.

That growth adjustment only addresses the expected filing volume, not the population of interest. To address this, we need a further adjustment to predict the population of the living wealthy. That adjustment was the inverse of the mortality rate developed by the National

Center for Health Statistics, NCHS (then, in 1980, the data were in a pamphlet; now, they are available on their Web site).

The main reason for using the estimated wealthy population instead of the expected filings of estate Tax Returns is that we wished to focus on the scarcity of “youthful decedents.” This mortality- weighted set of estimates allowed us to determine, in effect, what age a “youthful decedent” might be.

We used the Dalenius-Hodges’ cumulative square root of the frequency method to find reasonable strata boundaries, with a goal of choosing five groups (Dalenius and Hodges, 1959). In the end, a sixth was added because there were a fair number of cases where there was no age reported. In later years, this “Age Unknown” group was folded into the highest- age category because research showed that these decedents actually were members of that group, and the numbers became quite small.

While the strategy outlined above was applied to the estates within the focus year, some felt that, with appropriate “aging” of assets for decedents from other years, we might be able to create better Personal Wealth estimates. Hence, as is seen in Table 1, some strata are reserved for “young,” nonfocus-year decedents.

The later sample design tables show this strategy was revisited after the first focus year, and the strata for nonfocus-year filings expanded, duplicating the strata outline of the focus year. This revision reflected an increase in funding for this series of projects, as well as better meeting the need for data on the annual processing operations.

► **Sample Allocation**

Weighted strata variances for the value of Gross Estate (the value of all of an estate’s assets) were available from the prior 1977 study. Since the data collection is from administrative records, without any costs related to contacting a taxpayer, we simply assumed that the costs were essentially the same regardless of the stratum. The sample size was set at about 13,000 records per year.

Neyman Allocation (with a set sample size or otherwise) also requires a population estimate. Since we are primarily interested in the effect of the tax law as it is applied in a given year, and that law has effects on the living as well as the estates, the appropriate population was the same as the one used to find the age-strata breaks.

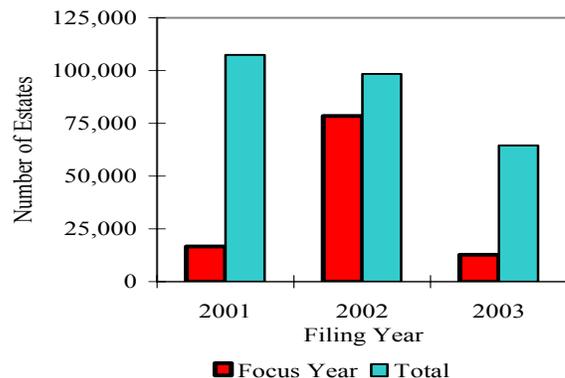
For the initial 1982 study, we allocated sample to strata under the plan for sampling the returns over 3 years, concentrating only on the year of death of the decedent, and ignoring the year of filing the administrative record.

Since the “Personal Wealth” population is more numerous than the Estates population, there were a lot of cases where the allocation prescribed more sample than there were expected estate filings. Thus, the allocation was reiterated several times, removing the certainty strata each time, before the final design’s sample sizes were derived.

These sample sizes, when divided by the expected filing volumes, became the sampling probabilities used in the Bernoulli selection. These are the sampling rates shown in Tables 1 through 5, below, exclusive of the CWHS sample selections.

As a result of the filing pattern, as in the example shown in Figure 4, only about 15 percent of the sample, or about 2,000 records, were to be designated in the first year of the study, and a similar amount in the final year of the set.

Figure 4.--Estates For Decedents Who Died During 2001



Starting with the 1986 Estates Study, while the allocation of the sample to the focus year was set at the target 10,000 to 15,000 records, the difference between the expected sample size in any given filing year and the target was allocated to the nonfocus-year records within a filing year. Thus, using 2005 as an example (Focus Year 2004), while the overall sample size is about 10,000 records, about 3,000 were allocated to estates of decedents who died before 2004 or in 2005.

The allocation for nonfocus-year returns used the expected filing volume of returns, instead of the population of the wealthy used in the allocation for the focus-year strata.

► **Changes--1986 to 2004**

The initial design, in Table 1, shows the result of having age stratification dependent on the Gross Estate class. Although we show a zero probability of selection for the "Under \$300,000" Gross Estate classes and other strata, those records were subjected to the 1-percent CWSH selections.

For the 1986 version of the design, shown in Table 2, the age groups were made independent of Gross Estate and were replicated for the nonfocus-year decedents. This also resulted in new age boundaries.

(Note, in this table and in subsequent ones, we will not show the classes that fall below the filing threshold due to space constraints. We used red to highlight the changes as well.)

The 1989 edition of the design, Table 3, also shows only a minor change: the introduction of an age group "65 under 75."

The next significant change arose for the 1992 study (Table 4). Here, we were finally able to replace the Gross Estate Code with the actual amount and thus expand the stratification. This design outline stood for about a decade.

The anticipated changes to the Estate Tax Law in 2001 left the design, Table 5, in some question. As a result, instead of planning to select the earliest filings

for the Focus Year (2001 decedents) at the same rates as filings in later years, we planned on the initial year's sample to support estimation by itself. The focus-year pattern was also amended; so, the Statistics of Income studies will coincide with the Federal Reserve Board's Survey of Consumer Finance.

As of this writing, the tax law is still subject to change, but at least one update, having the strata boundaries match the filing thresholds, is planned for 2007.

► **Future Research**

The current trend for the tax law suggests that, in a few years, we will be canvassing the entire population, and, under some legislation, this part of the tax code would expire. However, at some future time, there may again be reason to sample a successor tax return, for one lesson from history is certainly that the Estate Tax may someday be revived. We hope that, should that arise, this paper might be of some help to that future statistician.

One more immediate issue that the Estates and Personal Wealth studies have is that the original filings on which they are based may be prone to errors in the reporting, and especially underreporting of financial assets. When such problems are discovered, the executor or lawyer will file amended returns. While such amendments are possible with other types of tax filings, because the sole person knowledgeable about the various holdings for an estate has passed away, it may be that the effect would be more serious. At this time, we simply do not have the data to examine this issue.

However, we are starting to accumulate a database that might permit such research in a few years.

► **References**

Dalenius, T. and Hodges, J.L. Jr. (1959), "Minimum Variance Stratification," *Journal of the American Statistical Association*.

Johnson, Barry W. and Woodburn, R. Louise (1994), "The Estate Multiplier Technique, Recent Improvements for 1989," *Compendium of Federal*

Estate Tax and Personal Wealth Studies, Publication 1773 (4/94), Department of the Treasury, Internal Revenue Service.

Harte, J. M. (1986), "Some Mathematical and Statistical Aspects of the Transformed Taxpayer Identification Number: A Sample Selection Tool Used at IRS," *Proceedings of the Section on*

Survey Research Methods, American Statistical Association.

Kopczuk, Wojciech and Saez, Emmanuel (2004), "Top Wealth Shares in the United States, 1916-2000: Evidence From Estate Tax Returns," Working Paper 10399, National Bureau of Economic Research (<http://www.nber.org/papers/w10399>).

Table 1.--Strata and Selection Probabilities, 1982

*Size of Gross Estate
(Based on Gross Estate Code)*

<i>Age of Decedent</i>	<i>Under \$300,000</i>	<i>\$300,000 under \$500,000</i>	<i>\$500,000 under \$1,000,000</i>	<i>\$1,000,000 under \$5,000,000</i>	<i>\$5,000,000 or More</i>
Decedent Died in 1982					
<i>Under 45</i>	0	1.00	1.00	1.00	1.00
<i>45 under 55</i>		0.50	1.00		
<i>55 under 60</i>		0.35	0.50		
<i>60 under 70</i>					
<i>70 or Older</i>		0.10	0.25		
<i>Unknown</i>		0.10	0.25		
Decedent Died in a Year Other Than 1982					
<i>Under 45</i>	0	1.00	1.00	1.00	1.00
<i>45 or Older, or Unknown</i>	0	0	0	0	1.00

Table 2.--Strata and Selection Probabilities, 1986

*Size of Gross Estate
(Based on Gross Estate Code)*

<i>Age of Decedent</i>	<i>\$500,000 under \$1,000,000</i>	<i>\$1,000,000 under \$5,000,000</i>	<i>\$5,000,000 or More</i>
Decedent Died in 1986			
<i>Under 40</i>	1.00	1.00	1.00
<i>40 under 50</i>	1.00	1.00	1.00
<i>50 under 65</i>	0.35	1.00	1.00
<i>65 or Older, or Unknown</i>	0.07	0.50	1.00
Decedent Died in a Year Other Than 1986			
<i>Under 40</i>	1.00	1.00	1.00
<i>40 under 50</i>	0.25	0.35	1.00
<i>50 under 65</i>	0.04	0.50	1.00
<i>65 or Older, or Unknown</i>	0.01	0.01	1.00

Table 3.--Strata and Selection Probabilities, 1989

*Size of Gross Estate
(Based on Gross Estate Code)*

<i>Age of Decedent</i>	<i>\$500,000 under \$1,000,000</i>	<i>\$1,000,000 under \$5,000,000</i>	<i>\$5,000,000 or More</i>
Decedent Died in 1989			
<i>Under 40</i>	1.00	1.00	1.00
<i>40 under 50</i>	1.00	1.00	1.00
<i>50 under 65</i>	0.50	1.00	1.00
<i>65 under 75</i>	0.12	0.50	1.00
<i>75 or Older, or Unknown</i>	0.12	0.50	1.00
Decedent Died in a Year Other Than 1989			
<i>Under 40</i>	1.00	1.00	1.00
<i>40 under 50</i>	0.25	0.35	1.00
<i>50 under 65</i>	0.05	0.06	1.00
<i>65 under 75</i>	0.03	0.05	1.00
<i>75 or Older, or Unknown</i>	0.03	0.05	1.00

Table 4.--Strata and Selection Probabilities, 1992

Age of Decedent	Size of Gross Estate				
	\$600,000 under \$1,000,000	\$1,000,000 under \$2,000,000	\$2,000,000 under \$5,000,000	\$5,000,000 under \$10,000,000	\$10,000,000 or More
Decedent Died in 1992					
Under 40	1.00	1.00	1.00	1.00	1.00
40 under 50	1.00	1.00	1.00	1.00	1.00
50 under 65	0.22	0.44	1.00	1.00	1.00
65 under 75	0.10	0.20	0.40	1.00	1.00
75 or Older, or Unknown	0.03	0.06	0.18	1.00	1.00
Decedent Died in a Year Other Than 1992					
Under 40	1.00	1.00	1.00	1.00	1.00
40 under 50	0.15	0.20	1.00	1.00	1.00
50 under 65	0.06	0.11	0.33	1.00	1.00
65 under 75	0.06	0.11	0.33	0.45	1.00
75 or Older, or Unknown	0.03	0.05	0.16	0.22	1.00

Table 5.--Strata and Selection Probabilities, 2001

Size of Gross Estate	Age of Decedent			
	Under 40	40 under 50	50 under 65	65 or Older
Decedent Died in 2001				
\$675,000 Under \$1,000,000	1.00	1.00	1.00	0.13
\$1,000,000 under \$1,500,000	1.00	1.00	1.00	0.20
\$1,500,000 under \$2,000,000	1.00	1.00	1.00	0.20
\$2,000,000 under \$3,000,000	1.00	1.00	1.00	0.40
\$3,000,000 under \$5,000,000	1.00	1.00	1.00	0.80
\$5,000,000 under \$10,000,000	1.00	1.00	1.00	1.00
\$10,000,000 or More	1.00	1.00	1.00	1.00
Decedent Died in a Year Other Than 2001				
Under \$1,000,000	1.00	0.01	0.01	0.01
\$1,000,000 under \$1,500,000	1.00	0.01	0.01	0.01
\$1,500,000 under \$2,000,000	1.00	0.01	0.01	0.01
\$2,000,000 under \$3,000,000	1.00	0.02	0.02	0.02
\$3,000,000 under \$5,000,000	1.00	0.04	0.04	0.04
\$5,000,000 under \$10,000,000	1.00	0.11	0.11	0.11
\$10,000,000 or More	1.00	1.00	1.00	1.00

7



IRS Area-To-Area Migration Data

Gross

Internal Revenue Service Area-To-Area Migration Data: Strengths, Limitations, and Current Trends

Emily Gross, Internal Revenue Service

The mobility of Americans has long been a subject of interest for demographers, scholars, and the media. Just a few decades ago, the ultimate success story in this country was home ownership and staying in one neighborhood for all of adulthood. Currently, people and families move many times during their adult lives, with the peak moving years being between 20-24 years of age.¹ To where are these people moving, and from where did they originate? One of the few accurate sources of area-to-area migration data in the United States comes from the Statistics of Income Division (SOI) of the Internal Revenue Service (IRS), which maintains records of all individual income tax forms filed in each year.

This paper will highlight the data IRS has on taxpayer migration, particularly the county-to-county migration data created by U.S. Bureau of the Census analysts using IRS data. First, the paper will discuss the IRS Individual Master File from which these datasets are derived. Then, it will cover how the Census Bureau reviews the file and runs it through a geocoding program. Next, the paper will cover how the dataset returns to the IRS for disclosure proofing and how the data are marketed. The data themselves will be discussed, highlighting strengths and limitations. Finally, some current trends in migration will be examined.

► Statistics of Income (SOI) Division and the Data Source

The Statistics of Income program began in 1916, when Congress passed a revenue act that included a provision requiring the annual compilation of statistics with respect to the operation of the tax law. This requirement has reappeared in each major rewrite of the tax law since then and is currently included as section 6108 of the Internal Revenue Code of 1986.

Besides annual SOI publications, based on individual and corporate income tax returns, other data are also published in the quarterly *Statistics of Income*

Bulletin. The *Bulletin* includes studies on sole proprietorships, partnerships, tax-exempt organizations, estate tax returns, and estimates of personal wealth, as well as studies on “international” tax returns. Most of the SOI publications are available on the “tax stats” portion of the IRS Web site (www.irs.gov), which contains over 3,900 files related to tax statistics.

From time to time, SOI undertakes special reimbursable studies for Government and private users. One customer, the Census Bureau (which is allowed access to tax return data under the Internal Revenue Code but must be able to justify the data items it receives as needed for its own statistical programs) pays IRS for annual data on every entity on the IRS Individual Master File (IMF). (The IRS Master File includes administrative records for every Form 1040, 1040A, and 1040EZ.) The tax and income items that Census receives from the IMF include:

- Tax Filing Units (the filer and spouse of filer, plus all exemptions represented on the forms)
- Mailing address
- Age classification (the filer is classified as “under age 65” if he or she did not mark the age 65+ checkoff box)
- Income data: wages and salaries, interest income, dividend income, gross rents, and royalties
- Adjusted gross income (includes all taxable income, less adjustments to income)
- Total income (a special definition which most closely approximates the Census Bureau’s definition of total income).

The Master File data that Census receives were based on all returns filed by late September of the filing year. This extract is believed to include 95 percent to 98

percent of the individual filing population. The individuals covered by the returns include the filer and the spouse of the filer, as well as any exemptions claimed on the tax return. The Tax Year 2002 file, the most recent data available, contained about 130.5 million returns.²

In addition to using these data for their population estimates, Census also uses them to produce area-to-area migration data for SOI. The tax and income data included in the migration data are Number of Returns, Number of Exemptions, Aggregate Adjusted Gross Income (AGI), and Median AGI.

► **Census Bureau Processing**

In accordance with the agreement mentioned above between the IRS and Census Bureau, the 1040 Individual Master File dataset is provided annually to the Planning, Research, and Evaluation Department at Census. Both the Social Security Number (SSN) and the taxpayer name are stripped from each return. In their place, a special identification number called a Protective Identification Key (PIK) is assigned to each return.

To further prepare the data for its own purposes, as well as to prepare the migration files, the Census Bureau geocodes the IMF data. Geocoding involves assigning a set of codes to each return that represent the residence of the filer. These codes are assigned from the United States Post Office (USPO) ZIP/Sector-to-County Cross Reference (CCRS), which is generally reflected in the “ZIP plus 4” codes. The “plus 4” codes have two characters each—a sector code and a segment code. According to USPO guidelines, each sector code must identify one county only. This is the key to how Census is able to geocode each return by county of origin. From the combination of ZIP sector codes and mailing State code for each individual return, Census is able to assign each record with a State/county code from the CCRS. To prepare the migration data, Census must use 2 consecutive filing years of IMF data. For each set of filing years, a code was given to the current-year return and the prior-year return, using the current-year CCRS. County equivalent codes are assigned to the District of Columbia, the Virgin Islands, Puerto Rico, APO/FPO (military), and “other foreign.”

► **Identifying Migrants**

Once the geographic codes are in place, Census determines who in the file has or has not migrated. The coded returns for 2 consecutive years are then compared to one another for two criteria: (1) the street address and (2) the mailing address State plus ZIP code. If the two are identical, the return is labeled a “nonmigrant.” If any of the above information changed from the first prior year of study to the current year, the return is considered a mover. However, the return is only a “migrant” if the taxpayer’s geographic code changed. If a taxpayer’s address codes change from one year to the next, that taxpayer is an “in-migrant” for the address on the return filed in the second year, and an “out-migrant” for the address on the return filed the first year. If a taxpayer changed streets but stayed in the same county, that taxpayer would not be a migrant for purposes of this dataset.

As previously mentioned, the filer’s return address determines the migration status of the record. There are instances, however, where the taxpayer may not have changed residences but the return address suggests a move. This may happen if: (1) the filing address is that of a financial institution or tax preparer, and not the actual taxpayer; (2) the taxpayer is a college student living away from home who filed with a home address one year and the college address another; (3) the taxpayer puts his or her place of business as the return address; (4) the taxpayer maintains dual residences, primarily residing in one county but having the tax return sent to the other; and (5) the taxpayer uses a post office box for mailing purposes.

► **Tax Year versus Migration Year**

This section distinguishes among what is meant by tax year, filing or calendar year, and migration year. When dealing with income taxes, the year in which a return is filed is the “filing” or calendar year and almost always follows the actual “tax year.” For this reason, clarification of what exactly is meant by the year of migration is necessary. The residence of a taxpayer, for purposes of the Migration data files, is noted at the time the individual income tax return is filed. Because most tax returns are filed the spring after the tax year

has ended, the migration (filing) year coincides with the previous year's tax data. For example, the 2003 migration data cover the place of residence for individuals who were filing their 2002 Forms 1040 in Calendar Year 2003. Furthermore, since the migration data show movement from year to year, the files are expressed in 2-year increments, such as the 2002-2003 migration data. Thus, the file would show actual changes in residence from Calendar Year 2002 to Calendar Year 2003.

► IRS Preparation and Marketing of Migration Products

After Census geocoding and error checking, the Census Bureau maintains a file to supplement its internal population studies.³ A copy is then delivered to the Statistics of Income (SOI) Division of the Internal Revenue

Service. A statistician at the SOI Division checks the data for outliers, adds column headings and labels, and parses the data into Excel spreadsheets. Once SOI is satisfied with the dataset, it authorizes Census to release the file to State demographers. For each State, there is an inflow and an outflow spreadsheet, which shows the following information about the returns in each county: the number of migrant returns (used to estimate households); the number of exemptions attached to these returns (used to estimate individuals); the aggregate adjusted gross income of the migrating returns; and the median adjusted gross income of these returns. There is also a line item for nonmigrants with their relative incomes. An example of a page of the Minnesota inflow file for 2002-2003 follows (Figure A). This example shows the summary information for returns moving into Minnesota between 2002 and 2003, as well as detailed information

Figure A -- Inflow File for Minnesota (MN), 2002-2003

From St Abbr	From County Name	Number Of Returns	Number Of Exemptions	Aggregate Adjusted Gross Income (thousand dollars)	Median Adjusted Gross Income (whole dollars)
MN	Total Mig - US & For	146,999	257,176	5,894,696	25,079
MN	Total Mig - US	144,355	253,910	5,858,968	25,484
MN	Total Mig - US Same St	103,195	179,330	4,075,991	26,690
MN	Total Mig - US Diff St	41,160	74,580	1,782,977	22,294
MN	Total Mig - Foreign	2,644	3,266	35,728	4,877
MN	Aitkin County Tot Mig-US & For	454	875	18,991	28,102
MN	Aitkin County Tot Mig-US	454	875	18,991	28,102
MN	Aitkin County Tot Mig-Same St	393	767	16,643	28,599
MN	Aitkin County Tot Mig-Diff St	61	108	2,348	24,999
MN	Aitkin County Non-Migrants	5,175	11,257	200,253	25,733
MN	Hennepin County	58	105	2,833	38,332
MN	Anoka County	54	116	2,309	36,666
MN	Crow Wing County	47	91	1,627	18,999
MN	Ramsey County	29	52	1,640	45,832
MN	Itasca County	21	30	559	18,124
MN	Mille Lacs County	19	38	932	26,249
MN	Dakota County	18	32	964	37,499
MN	St Louis County	16	35	795	39,999
MN	Washington County	13	21	760	54,999
MN	Cass County	12	23	290	19,999
MN	Scott County	10	16	410	32,499
MN	Wright County	10	23	550	39,999
SS	Other Flows - Same State	86	185	2,974	24,999
DS	Other Flows - Diff State	61	108	2,348	24,999

for the first county of destination, Aitkin County, MN. For more information on interpreting this file, see IRS documentation.⁴

Once the files are prepared, they are announced for sale via the SOI Web site (www.irs.gov/taxstats/index.html), as well as in various SOI publications. The migration data are free to Federal, State, and local government agencies and are among the most popular products distributed through the SOI Division's Statistical Information Services (SIS) Office. In 2004, well over 200 migration data sets were distributed to customers in government, business, and academia. Information on pricing can be found on the Web site (www.irs.gov/taxstats/indtaxstats/article/0,,id=96816,00.html); in the Products and Services Section of each *Statistics of Income Bulletin*, Publication 1136; or by contacting the SIS office at (202) 874-0410.

► **Strengths and Limitations of the Dataset**

The county-to-county migration data may be the largest dataset that tracks movement of both households and people from county to county, including family incomes. Because these data are obtained from income tax records, they are inclusive and reliable. However, the source and design of this dataset have some limitations. As mentioned previously, those who are not required to file United States Federal income tax returns are not included in this file. Because of this, the dataset underrepresents the poor. Also not included is the small percentage of tax returns filed after late September of the filing year. Because the IRS granted most taxpayers who file this late an extension, and because most taxpayers who request an extension are more likely to file high-income tax returns, the migration data set can underrepresent the very wealthy.

The matching process also causes some returns to be missed. When the current-year tax return is compared to the prior-year tax return, only the Social Security number of the primary taxpayer is considered. If a secondary filer exists (as in the case of a married couple filing jointly), that Social Security number is not recorded or compared. If, for example, a husband and wife file a joint return in the prior year but file separately in the current year, only

the husband's current year will have a match with the prior year. The spouse's current-year return becomes a nonmatch and will not be included in the data. This problem not only occurs when couples decide to switch filing status from year to year, but also when marriage or divorce changes an individual from being a primary taxpayer (included in the file) to a secondary taxpayer (not included in the file).

In addition to the dataset not including the entire individual filing population, it also underrepresents the elderly, another large segment of the population which may not be required to file individual tax returns.

► **Uses of the County-to-County Migration Data**

Statistics of Income tax data are mainly used within the Government by the Treasury Department's Office of Tax Analysis (OTA) and by the Congressional Joint Committee on Taxation. Both use the data in tax policy research and in revenue estimating.⁵ The county-to-county migration data, however, are created for users outside the IRS or Treasury Department.

The Census Bureau uses these files to back up its demographic data between Decennial Censuses. Most of the individuals ordering these data are from academia, the media, and the private sector. Academic papers using the data show trends and shifts in demographics. Newspapers often highlight trends showing the fastest growing counties, where the wealthy are moving, and what parts of the country are losing population. Private firms include researchers hired by corporations, developers following movement of housing consumption, and technology companies estimating future demand, to name just a few. The county-to-county migration data are one of the most frequently requested products disseminated by the SOI Division. In Calendar Year 2004, the Statistical Information Services Office of the Division answered 367 requests about its migration data.

► **Current Migration Trends**

The wealth of useful data present in the county-to-county migration files can be illustrated by examining some current demographic trends shown in the data.

This section looks at three regional trends, as well as how customers used SOI data in their work.

► **Loudoun County, Virginia**

A look at inflow and outflow files for the State of Virginia shows that the fastest growing county in the Washington, DC metropolitan area is Loudoun County, Virginia. Loudoun County is situated just to the west of what used to be considered the outer limits of the Washington, DC suburbs as recently as 15 years ago. As the greater DC area continues to grow as a result of a long period of economic growth and small unemployment rate, more and more households have been moving into the area.

Two enormous residential communities, Ashburn and South Riding, evolved in the 1990’s and are continuing to grow and attract affluent professionals by the thousands each year.

Figure B. -- Loudoun County, Virginia

	Number of Returns	Number of Exemptions	Aggregate AGI (thousand dollars)	Median AGI (whole dollars)
Inflows	13,073	27,035	939,231	50,864
Outflows	7,391	14,632	492,439	44,932
Nonmigrants	68,231	166,364	5,987,797	65,184

A look at the 2002-2003 data in Figure B compares the individual income tax return data of those who came into the county and those who exited the county between these 2 years. The Number of Returns column shows that the number of households increased by 7.5 percent between 2002 and 2003. The rise in number of exemptions nearly mirrors this change. A comparison of Adjusted Gross Income (AGI) between in the immigrants and outmigrants is equally striking. The median AGI column shows that the median adjusted gross income of the returns moving into Loudoun County is considerably higher than the median income of those who are leaving. Both are lower than the median income of the nonmigrants (those who resided in Loudoun County for

both years), suggesting that perhaps the immigrants are younger and less-established families than those who have resided there longer.

► **Clark County, Nevada**

Another notable county in the United States in terms of migration is Clark County, Nevada. Clark County is the home of the cities of Las Vegas, North Las Vegas, and Henderson, as well as the unincorporated towns of Paradise (including the Las Vegas strip, the University of Las Vegas, and McCarran International Airport), Sunrise Manor, Spring Valley, and Enterprise. An examination of Figure C shows that, while 28,962 returns left the county from 2002 to 2003, some 44,311 returns came in. Thus, the returns moving into the county outpaced the returns leaving the county by 53 percent in that year. While Clark County is considered an excellent place to retire, data from the Nevada State Demographer’s office show that the percentage of Clark County residents age 65 and older has held steady at approximately 11 percent for the past several years.⁶

Figure C. -- Clark County, Nevada

	Number of Returns	Number of Exemptions	Aggregate AGI (thousand dollars)	Median AGI (whole dollars)
Inflows	44,311	83,219	1,916,647	22,547
Outflows	28,962	54,254	1,028,971	21,010
Nonmigrants	511,010	1,084,081	25,334,202	32,015

The IRS county-to-county migration files also show that, of the top ten counties of origin for those moving into Clark County, none of them originates from the State of Nevada. The top five counties of origin are: Los Angeles, San Diego, Orange, and San Bernadino (all southern California counties), and Maricopa County, Arizona. Further study of the Nevada State Demographer’s published data show that Clark County is projected to double in size between the years 2003 and 2024, accounting for 85 percent of the total expected growth in the State of Nevada for that time period.

► Riverside County, California

The U. S. county with the highest net gain of returns between Calendar Years 2002 and 2003 was Riverside County, California. Riverside County is situated just to the east of Los Angeles and Orange Counties, two of the most populated counties in Southern California. As shown below in Figure D, Riverside had a net gain of 20,404 returns during this time period. Where did these residents come from? According to the IRS data, 10,425 of the 50,843 returns coming in to Riverside County were former residents of Orange County. While having twice the population of Riverside County, Orange County is geographically small: only 789 square miles, compared to Riverside's 7,207 square mileage.

Figure D. -- Riverside County, California

	Number of Returns	Number of Exemptions	Aggregate AGI (thousand dollars)	Median AGI (whole dollars)
Inflows	50,843	114,863	2,282,503	30,189
Outflows	30,439	62,084	1,151,864	23,437
Nonmigrants	488,511	1,204,255	23,218,621	31,618

The second largest source of in-migrants to Riverside County was Los Angeles County, which lost 9,167 residents to this neighbor to the East. This loss may be a drop in the bucket for hugely populated Los Angeles, which has over 3 million residents, but illustrates a national trend: households are leaving the cities and close-in suburbs for more land and more affordable housing. In fact, Los Angeles had a significant net loss of households in the year examined, with 18,432 of its Year 2002 returns calling a different county home in 2003. The top five recipients of Los Angeles outflows were all neighboring Southern California counties.

► Summary

As this paper shows, the migration data contain a wealth of information that can be used to analyze and illustrate major demographic trends. The Census Bureau,

in partnership with the IRS, creates a unique product rich in information yet simple enough to understand for all customers: from demographers, newspapers, and Government agencies to the public at large.

► Endnotes

- 1 U.S. Census Bureau, *Current Population Survey, 2003 Annual Social and Economic Supplement*.
- 2 *Statistics of Income--2002, Individual Income Tax Returns*, Publication 1304, Internal Revenue Service.
- 3 Long, John F, "Postcensal Population Estimates: States, Counties, and Places," presented at the Annual Meeting of the American Statistical Association, San Francisco, CA, August 1993.
- 4 Internal Revenue Service (1999), *Area-to-Area Migration and County Income*, internal documentation, Statistics of Income Division.
- 5 Kozielec, John (1996), "The Tax Return: A Unique Data Source for Tracking Migration," *Turning Administrative Systems Into Information Systems: 1995*, Publication 1299, Internal Revenue Service.
- 6 State of Nevada Demographer: "Nevada's Age, Sex, Race, and Hispanic Origin Estimates for 2003," <http://www.nsbdc.org/demographer/pubs/>.

► References

Sater, Douglas K. (1994), "Geographic Coding of Administrative Records--Current Research in ZIP/Sector-To-County Coding Process," working paper, United States Census Bureau.

United States Census Bureau (1996), *Supplemental Documentation for External Data Products*, internal documentation.

Index

of IRS Methodology Reports on Statistical Uses of Administrative Records

Special Studies in Federal Tax Statistics--2004

Selected papers given primarily at the 2004 Annual Meetings of the American Statistical Association in Toronto, Ontario, Canada, and two other professional conferences--the Luxembourg Wealth Study Workshop in Perugia, Italy, and the Conference on Privacy in Statistical Databases in Barcelona, Spain. The volume is divided into five major sections. It begins with four papers on recent developments in Statistics of Income research. Section 2 includes five papers on quality assessment of administrative records data. Section 3 presents a paper on estimates of income and wealth from survey and tax data. Section 4 contains a paper on disclosure protection techniques. Finally, Section 5 presents a paper on some current theoretical research on multivariate analysis presented in a poster session at ASA.

Special Studies in Federal Tax Statistics--2003

Selected papers given primarily at the 2003 Annual Meetings of the American Statistical Association in San Francisco, CA. The volume is divided into four major sections. It begins with four papers presented in the same session under the topic, "Are the Rich Getting Richer and the Poor Getting Poorer?" Section 2 includes a paper on survey methods. Section 3 presents five papers on new developments in tax statistics and administrative records. Finally, Section 4 contains a paper on survey nonresponse and imputation.

Special Studies in Federal Tax Statistics--2002

Selected papers given primarily at the 2002 Annual Meetings of the American Statistical Association in New York City and at the 2002 National Tax Association Conference in Orlando, FL. The volume is divided into seven major sections. It begins with two papers on recent IRS research. Section 2 includes a group of four papers on methodological and analytical advances in tax statistics. Section 3 presents two papers on statistical uses of administrative records. Section 4 contains a paper on disseminating IRS locality data. Section 5 includes a paper on confidentiality and data access issues. Section 6 presents a paper on measuring the quality of IRS responses to taxpayer inquiries. Finally, Section 7 includes two papers on distributional theory and computation.

Special Studies in Federal Tax Statistics--2000-2001

Selected papers given primarily at the 2000 and 2001 Annual Meetings of the American Statistical Association in Indianapolis, Indiana and Atlanta, Georgia, plus one other paper presented at the International Conference on Establishment Surveys II in Buffalo, New York in 2000. The volume is divided into four major sections. The book begins with five papers on statistical applications. Section 2 presents two papers on confidentiality and data access issues. Section 3 presents two papers on changing industry codes. Finally, Section 4 includes five papers on analyses of Federal tax and information returns.

Turning Administrative Systems Into Information Systems--1999

Selected papers given at the 1999 Annual Meetings of the American Statistical Association (ASA) in Baltimore, MD. In addition, the report includes one paper presented at the 1998 ASA conference in Dallas, TX. The volume is divided into six major sections. The book begins with a complete ASA session analyzing administrative records from the U.S. tax system. It contains four papers, as well as a set of comments on the presentations. Section 2 presents four papers on the statistical uses of administrative records. Section 3 includes two papers, which focus on employee satisfaction and customer satisfaction surveys at the IRS. Section 4 contains two papers, one of which was presented

at the 1998 ASA conference, that provide an update on the Survey of Consumer Finances. Section 5 presents one paper that looks at the feasibility of preparing State corporate data by matching receipts and employment data by State and industry. Finally, the volume concludes with a paper on distributional theory and computation.

Turning Administrative Systems Into Information Systems--1998-1999

Selected papers given at the 1998 Annual Meetings of the American Statistical Association in Dallas, Texas. In addition, the report includes a session of papers presented in 1999 at the Annual Meetings of the American Economic Association (AEA) plus one other paper. The volume is divided into five major sections. The book begins with the AEA session in memory of the late Dr. Daniel B. Radner, Social Security Administration economist. It contains four papers on new empirical findings in the distributions of personal income and wealth, as well as two sets of introductory remarks and two sets of comments on the presentations. Section 2 presents two papers on data measurement and data bases for economic research. Section 3 includes two papers, which focus on sample design, estimation, and imputation research. Section 4 explores issues dealing with public-use files, including the potential for disclosure. Finally, Section 5 concludes the volume with a paper verifying the classification of public charities in the 1994 Statistics of Income Study Sample. (It is the only paper not presented at the ASA or AEA meetings.)

Turning Administrative Systems Into Information Systems--1996-1997

Selected papers given primarily at the 1996 and 1997 Annual Meetings of the American Statistical Association in Chicago, Illinois and Anaheim, California, plus one non-ASA article. The volume is divided into nine major sections. The book begins with a paper originally printed as a textbook article on inheritance and wealth in America. Section 2 presents papers on using administrative records for generating national statistics. Section 3 contains two sets of panel reports on the statistical uses of administrative records. Section 4 focuses on methodological research. Section 5 explores issues dealing with quality improvement in government. Section 6 presents a panel discussion on Customer Satisfaction Surveys. Section 7 focuses on the effect of downsizing on Federal statistics. Section 8 explores the privacy area. Finally, Section 9 concludes with seven papers on statistical disclosure limitation.

Turning Administrative Systems Into Information Systems--1995

Selected papers given primarily at the 1995 Annual Meetings of the American Statistical Association in Orlando, Florida and another conference. The volume is divided into five major sections. The book begins with a paper on SOI migration data, giving an example of how this unique dataset can be used by demographers and policy researchers. Section 2 presents papers on sample designs and redesigns, as well as on SOI efforts in the corporation and partnership areas. Section 3 contains papers on weighting and estimation research. Section 4 focuses on analytical approaches to quality improvement, from graphical techniques to cognitive research. Finally, Section 5 concludes with papers from an invited session on record linkage applications for health care policy, a session organized by SOI in view of its long-term interest in improving matching techniques for administrative and survey data.

Turning Administrative Systems Into Information Systems--1994

Selected papers given primarily at the 1994 Annual Meetings of the American Statistical Association in Toronto, Ontario, Canada. The volume is divided into nine major sections. The book begins with an overview of the Statistics of Income Programs, describing the origins and customers of various SOI data and highlighting our products and services. Section 2 presents the descriptive results from two recent studies--one on sales of capital assets and one on self-employed nonfilers. Section 3 contains papers and discussion from a session on privacy issues involved in using administrative record data. The next two sections are much more methodical in nature: Section 4 focuses on sample design and estimation work in SOI, beginning with a reprint of a 1963 paper by W. Edwards Deming, which presents an evaluation of the SOI sample. Section 5 presents data on record linkage. Section 6 draws together the

papers from a session on nonresponse in Federal surveys. Section 7 is a more statistical section, which contains a collection of papers on imputation methodology in a number of different arenas. Section 8 focuses on another long-time theme of these volumes--quality improvement efforts. Finally, Section 9 presents two unrelated papers on data preparation techniques.

Turning Administrative Systems Into Information Systems--1993

Selected papers given at the 1993 Annual Meetings of the American Statistical Association in San Francisco, California and other related conferences. The volume contains seven major sections, each focusing on a somewhat different area of research. The first section begins with a paper that presents a view for the future of the Federal statistical system. This effort is part of a dialogue with other agency leaders to redefine a cohesive plan for Federal data producers and users. Section 2 contains several descriptive papers based on tax data about individuals, and Section 3 looks at similar uses of tax data for businesses. Section 4 focuses on sample design issues for several SOI projects, while Section 5 presents information on improvements to analytical techniques. Finally, Sections 6 and 7 describe a number of different studies SOI is involved in to improve the quality and productivity of other areas of IRS.

Turning Administrative Systems Into Information Systems--1991-1992

Selected papers given mostly at the 1991 and 1992 Annual meetings of the American Statistical Association, held, respectively, in Atlanta, Georgia and Boston, Massachusetts. Papers chosen for this volume exemplify some of the basic changes that are occurring in the Statistics of Income program during the 1990's, including discussions of methodological improvements and applications currently under way in the U.S. Federal statistical community. The volume contains seven general areas of interest: information from tax return data; the 1989 Survey of Consumer Finances; estimation and methodological research in the SOI business program; sample design and weighting issues in the SOI individual program; some quality improvement applications; some technological innovations for SOI research; and a look to the future data needs for the Federal sector. Previous volumes in the series were called Statistics of Income and Related Administrative Record Research (see below). The title was changed to more clearly reflect how the Internal Revenue Service's Statistics of Income function is adapting to better meet the informational needs of its many customers.

Statistics of Income and Related Administrative Record Research--1990

Selected papers given primarily at the 1990 Annual meeting of the American Statistical Association in Anaheim, California. Papers selected for this volume contain discussions of methodological improvements and applications currently under way in the U.S. Federal statistical community. In particular, the focus is on work being done by the Statistics of Income Division of the Internal Revenue Service (IRS). The volume covers five general areas: longitudinal panel data and estimation issues; analytical research using survey and administrative data; design issues for Federal surveys; information on the conclusions of the Establishment Reporting Unit Match Study; and a look at future data needs for the Federal sector.

Statistics of Income and Related Administrative Record Research--1988-1989

Selected papers given mostly at the 1988 and 1989 Annual Meetings of the American Statistical Association in New Orleans, Louisiana and Washington, D.C., respectively. Papers for the volume focus on perspectives on statistics in government--in celebration of ASA's 150th anniversary; improvements in income and wealth estimation; methodological enhancements to administrative record data; some looks at the effects of tax reform; and technological innovations for statistical use.

Statistics of Income and Related Administrative Record Research--1986-1987

Selected papers given, for the most part, at the 1986 and 1987 Annual Meetings of American Statistical Association

in Chicago and San Francisco, respectively. Papers focus on ongoing wealth estimation research and U.S. and Canadian efforts regarding methodological enhancements to corporate and individual tax data and recent refinements to disclosure avoidance techniques.

Record Linkage Techniques--1985*

The Proceedings of the Workshop on Exact Matching Methodologies held in Arlington, Virginia, May 9-10, 1985. Includes landmark background papers on record linkage use and papers describing methodological enhancements, applications, and technological developments, as well as extensive bibliographic material on exact matching.

Statistical Uses of Administrative Records: Recent Research and Present Prospects*

A two-volume reference handbook on research results involving the use of administrative records for statistical purposes from 1979 through 1982:

- ❑ Volume I (March 1984) focuses on general considerations in administrative record research, applications of income tax data, uses based on data from other major administrative record systems, and enhancements to statistical systems using administrative data.
- ❑ Volume II (July 1984) focuses on comparability and quality issues, access to administrative records for statistical purposes, selected examples of end uses of linked administrative statistical systems, and a status report that sets goals for the future.

Statistics of Income and Related Administrative Record Research--1984*

Selected papers given at the 1984 Annual Meeting of American Statistical Association in Philadelphia. Papers focus on future policy issues, applications, exact matching techniques, quality control, missing data, and sample design issues.

Statistics of Income and Related Administrative Record Research--1983*

Selected papers given at the 1983 Annual Meeting of American Statistical Association in Toronto. Papers focus on use of administrative records in censuses and surveys, applications for epidemiologic research and other statistical purposes, and statistical techniques involving imputation and disclosure and confidentiality

Statistics of Income and Related Administrative Record Research--1982*

Selected papers given at the 1982 Annual Meeting of American Statistical Association in Cincinnati. Papers focus on statistical uses of administrative records, resulting methodologic advances, and estimates and projections for intercensal updates.

Statistics of Income and Related Administrative Record Research*

Selected papers given at the 1981 Annual Meeting of American Statistical Association in Detroit. Papers focus on applications and methodologies with an emphasis on IRS's Statistics of Income Program, the Small Business Data Base, nonprofit and pension data, and on Canada's Generalized Iterative Record Linkage System.

Economic and Demographic Statistics*

Selected papers given at the 1980 Annual Meeting of American Statistical Association in Houston. Papers focus on evaluation of the 1977 Economic Census, CPS hot deck techniques, and efforts to upgrade Social Security's Continuous Work History Sample.

*Out of print--Copies of selected papers can be obtained upon request.

NOTE: The IRS Methodology Reports on statistical uses of administrative records are now being offered free of charge. To obtain copies, write to:

Statistical Information Services (SIS)
Statistics of Income Division (RAS:S:SS:SD)
Internal Revenue Service
P.O. Box 2608
Washington, DC 20013-2608

Phone: (202) 874-0410
FAX: (202) 874-0964
E-mail: sis@irs.gov