This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Pensions in the U.S. Economy

Volume Author/Editor: Zvi Bodie, John B. Shoven, and David A. Wise, eds.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-06285-6

Volume URL: http://www.nber.org/books/bodi88-1

Publication Date: 1988

Chapter Title: The Determinants of IRA Contributions and the Effect of Limit Changes

Chapter Author: Steven F. Venti, David A. Wise

Chapter URL: http://www.nber.org/chapters/c6043

Chapter pages in book: (p. 9 - 52)

1 The Determinants of IRA Contributions and the Effect of Limit Changes

Steven F. Venti and David A. Wise

To encourage employees not covered by private pension plans to save for retirement, individual retirement accounts (IRAs) were established in 1974 as part of the Employee Retirement Income Security Act (ERISA). Emphasizing the need to enhance economic well-being of future retirees and the need to increase national saving, the Economic Recovery Tax Act of 1981 extended the availability of IRAs to all employees and raised the contribution limit. Now (1985) any employee with earnings in excess of $2000 can contribute up to $2000 to an IRA account each year, with tax deferred on the principle and interest until money is withdrawn from the account. The combined contribution of an employee and a nonworking spouse can be as high as $2250. A married couple who are both working can contribute $2000 each. A proposed change in the law contemplates raising the individual IRA limit to $2500 and the (nonworking) spousal IRA limit from $2250 to $2500.

Tax-deferred saving is potentially an important component of saving for retirement and could represent a very substantial increase in tax-free saving for many employees. Indeed, a $2000 contribution to a retirement account represents a future pension benefit greater than many employer-provided private pension plans. The availability of IRAs may also have a substantial effect on national saving. According to IRS data, total IRA contributions in 1982 were over $29 billion.

Steven F. Venti is an assistant professor of economics at Dartmouth College and a faculty research fellow of the National Bureau of Economic Research. David A. Wise is the John F. Stambaugh Professor of Political Economy at the John F. Kennedy School of Government, Harvard University, and a research associate of the National Bureau of Economic Research.

The final version of this paper has benefited from the comments of Gary Burtless. The research was funded in part by the Department of Health and Human Services (HHS) grant #67A-83 and in part by HHS grant #84ASPE130A

Despite the program's size and potential significance, surprisingly little is known about the determinants of IRA contributions. Thus the goals of this paper are: (1) to analyze the effect of individual attributes on whether a person contributes to IRAs, (2) to determine the effect of individual attributes on how much is contributed, and (3) to simulate the effect of potential changes in contribution limits on the amount contributed to IRAs. The results can be used to judge whether the goals that justified introducing the program are being realized. Obviously, persons who do not contribute to IRA accounts will not benefit from them. With national concern about the federal deficit, the short-run tax cost of the program is of substantial interest. The simulations suggest what this cost is and what the cost of proposed changes in the program would be. A fourth issue, the effect of tax-deferred saving on net individual saving, is not addressed in this paper but will be analyzed in future work.

This analysis is based on data obtained through a special supplement of the May 1983 Current Population Survey (CPS). Subsequent analysis will be based on the 1983 Survey of Consumer Finances and a special Carnegie Commission Survey of college and university employees.

Descriptive statistics on contributions to IRA accounts are presented in section 1.1. The statistical model used in the analysis is described in section 1.2. The results are presented in section 1.3, and in section 1.4, results of a similar analysis based on Canadian data (Wise 1985) are compared with results for the United States. The Registered Retirement Savings Plan (RRSP) in Canada is a tax-deferred program that incorporates the characteristics of both IRA and Keogh-like plans in the United States, although the contribution limits are quite different in the two countries. The same statistical model has been estimated on data from both countries.

The major empirical findings may be summarized briefly: tax-deferred saving plans are unlikely to be used by low-income persons. Thus they do not in general substitute for private pension plans, since higher-income persons are more likely than those with lower incomes to be covered by private plans. Given income and other individual characteristics, persons with private pension plans are no less likely than those without such plans to contribute to an IRA. The findings for Canada are very similar to those for the United States. Since the contribution limits are very different, the similar findings support the statistical specification.

Simulations based on the parameter estimates for the United States indicate that if the limits were increased in accordance with the recently proposed Treasury Department changes to the tax system, contributions would increase by about 30 percent.

1.1 Descriptive Statistics

Since model parameter estimates for the United States will ultimately be compared with those for Canada, descriptive statistics for both countries are presented in this section. For several reasons the data for the two countries are not strictly comparable, but they allow rough comparisons.

Most contributions are made by middle-income employees. Although nearly 32 percent of employed persons in the United States have incomes below $10,000, this group is responsible for only about 10 percent of total IRA contributions. Approximately 80 percent of contributions are made by persons with incomes between $10,000 and $50,000. Persons with incomes greater than $50,000 contribute only about 10 percent of total contributions. In Canada, about 82 percent of contributions are made by individuals with incomes between $10,000 and $50,000, and about 15 percent by persons with incomes above $50,000. Only 3 percent of contributions are made by those with incomes below $10,000, compared with 10 percent in the United States (see table 1.1).

As shown in table 1.2, only 5 percent of persons with incomes less than $10,000 made an IRA contribution in 1982 in the United States, and only about 2 percent in Canada. The proportions of higher-income groups making contributions are similar in the United States and Canada, although in general the proportions are lower in the United States than in Canada. Whereas the IRA program is new for most people in the United States, the Canadian RRSP plan was started in 1957.

Only 11 percent of all contributors in the United States have incomes less than $10,000, 80 percent have incomes between $10,000 and $50,000, and about 9 percent have incomes greater than $50,000. Again, the percentages in Canada are very similar to those in the United States. Over 6 percent of contributors have incomes less than $10,000, about 87.5 percent have incomes between $10,000 and $50,000, and 7 percent have incomes greater than $50,000.

Proportions of individuals that contribute to the contribution limits in the two countries are shown in table 1.3. Because the contribution limits vary substantially between the two countries, the numbers must be viewed accordingly.[1] In neither country does the proportion contributing to the limit in any income group exceed 60 percent. In addition, women are apparently more likely than men to contribute to the limit in the United States, whereas in Canada the difference seems less apparent, although at least for persons with incomes below $50,000 the proportion for women is greater than for men, with the exception of the $0–$10,000 income group.

Table 1.1 **Percent Distribution of Individuals and of Contributions, by Income Interval[a]**

	United States		Canada	
Income Interval[b]	Percent of Employed Individuals	Percent of IRA Contributions[c]	Percent of Tax Filers	Percent of RRSP Contributions
0–10	31.7	9.9	46.3	2.9
10–20	35.8	26.1	31.0	21.8
20–30	19.8	26.5	15.4	32.0
30–40	7.7	18.1	4.3	18.2
40–50	2.6	9.1	1.5	9.9
50–60	—	—	0.6	5.3
50–70	1.6	6.5	—	—
60–70	—	—	0.3	3.0
70–80	—	—	0.2	2.0
70+	0.8	3.8	—	—
80–90	—	—	0.1	1.2
90–100	—	—	0.1	0.8
100+	—	—	0.3	3.0

[a]The Canadian data pertain to 1980 and the U.S. data to 1982. Tabulations for the United States are in U.S. dollars and those for Canada in Canadian dollars. Data for the United States are from the May 1983 CPS and supplemental Survey of Pension and Retirement Plan Coverage. The data are weighted to represent the employed population, ages 18–65, excluding the self-employed. The Canadian data are based on a random sample of tax filers and are weighted to represent all tax filers.

[b]In thousands.

[c]Calculations are based on midpoints of reported IRA contribution intervals (see appendix B).

Average contributions in the United States range from $75 for the lowest income group to $1116 for those with incomes greater than $70,000; while the average contribution of contributors ranges from $1517 to $1883 (see table 1.4). This suggests that among those who contribute, a large proportion in each income group contributes at the limit. Unreported tabulations indicate that at very high income levels 85–90 percent of all contributions are at the limit. The percentage of employees with contributions at the limit ranges from about 3 percent for low-income to about 50 percent for high-income employees. The figures for Canada are comparable, but the average contribution levels are considerably higher, reflecting the higher limits. In addition, the Canadian data pertain to both employees and self-employed persons,

Table 1.2 **Percent with Contributions Greater than Zero and Percent of Total Contributors, by Income Interval[a]**

Income Interval[b]	Percent with Contribution > 0		Percent of Total Contributors	
	United States	Canada	United States	Canada
0–10	5.0	1.9	10.9	6.6
10–20	11.3	13.4	28.0	31.8
20–30	19.2	28.0	26.5	33.1
30–40	32.4	45.1	17.2	14.9
40–50	44.9	56.9	8.2	7.7
50–60	—	59.5	—	2.9
50–70	53.5	—	5.8	—
60–70	—	58.5	—	1.4
70–80	—	63.0	—	0.8
70+	59.3	—	3.4	—
80–90	—	63.0	—	0.5
90–100	—	62.6	—	0.4
100+	—	53.6	—	1.0

[a]The Canadian data pertain to 1980 and the U.S. data to 1982. Tabulations for the United States are in U.S. dollars and those for Canada in Canadian dollars. Data for the United States are from the May 1983 CPS and supplemental Survey of Pension and Retirement Plan Coverage. The data are weighted to represent the employed population, ages 18–65, excluding the self-employed. The Canadian data are based on a random sample of tax filers and are weighted to represent all tax filers.

[b]In thousands.

[c]Calculations are based on midpoints of reported IRA contribution intervals (see appendix B).

while the U.S. data pertain only to employees and thus exclude contributions to Keogh plans.

Individuals covered by private pension plans in the United States tend to make somewhat larger contributions than those who are not, and they are also somewhat more likely to make contributions at the limit, as shown in table 1.5. In Canada, the limit on RRSP contributions increases with income and the maximum is higher for persons without than for those with a private plan. Thus for high-income persons, contributions are higher for those without private plans. Nonetheless, for most income intervals those *with* a private pension plan are more likely to contribute at the limit.

In summary, the descriptive data indicate that IRAs are typically not used by low-income employees, and that they do not in general serve as a substitute for private pension plans.

Table 1.3 **Percent with Contributions at the Limit, by Income Interval and Sex**[a]

Income Interval[b]	United States		Canada	
	Men	Women	Men	Women
0–10	1.0	3.7	0.7	0.6
10–20	3.8	9.2	2.8	4.1
20–30	10.5	19.5	6.3	12.9
30–40	21.8	33.4	17.3	25.1
40–50	35.5	41.0	34.0	36.7
50–60	—	—	38.8	33.3
50–70	44.4	58.1	—	—
60–70	—	—	45.6	29.9
70–80	—	—	49.4	31.0
70+	51.0	30.7	—	—
80–90	—	—	51.9	30.5
90–100	—	—	51.3	24.7
100+	—	—	45.7	19.0

[a]The Canadian data pertain to 1980 and the U.S. data to 1982. Tabulations for the United States are in U.S. dollars and those for Canada in Canadian dollars. Data for the United States are from the May 1983 CPS and supplemental Survey of Pension and Retirement Plan Coverage. The data are weighted to represent the employed population, ages 18–65, excluding the self-employed. The Canadian data are based on a random sample of tax filers and are weighted to represent all tax filers.

[b]In thousands.

1.2 The Statistical Model

The results suggest that relatively unambiguous answers can be provided to the three questions addressed in this paper. On the other hand, an analysis of the effect of tax-deferred accounts on net saving will require related but new and somewhat more complicated statistical procedures, and it seems apparent that this question will be answered with more ambiguity and less confidence than the first three. Thus it is important to set forth the analysis so that questions that can be answered relatively precisely can be distinguished from those that inherently leave room for doubt. To put the analysis conducted in this paper in perspective, it may be useful to illustrate how it is related to a more general analysis designed to estimate the net effect of tax-deferred accounts on individual saving. With this goal in mind, a simple but general illustrative model is described first. It serves to motivate statistical analysis of each of the three questions discussed in this paper, while providing a framework within which the fourth question can be

Table 1.4 Average Contribution, by Income Interval[a]

	United States			Canada		
Income Interval[b]	Average[c]	Average, Given Contribution > 0[c]	Percent with Contribution at Limit	Average	Average, Given Contribution > 0	Percent with Contribution at Limit[d]
0–10	$ 75	$1517	2.8	$ 16	$ 834	0.7
10–20	176	1564	6.5	176	1315	3.3
20–30	324	1685	12.9	520	1858	7.6
30–40	571	1762	23.3	1059	2346	18.0
40–50	838	1865	35.8	1637	2877	34.3
50–60	—	—	—	2078	3493	38.2
50–70	1010	1887	45.4	—	—	—
60–70	—	—	—	2489	4181	43.5
70+	1116	1883	49.6	—	—	—
70–80	—	—	—	2899	4604	47.4
80–90	—	—	—	2951	4687	49.2
90–100	—	—	—	2960	4731	48.4
100+	—	—	—	2843	5306	41.8

NOTE: The figures for the United States are *not* comparable becuase the contribution limits are different in the two countries.

[a]The Canadian data pertain to 1980 and the U.S. data to 1982. Tabulations for the United States are in U.S. dollars and those for Canada in Canadian dollars. Data for the United States are from the May 1983 CPS and supplemental Survey of Pension and Retirement Plan Coverage. The data are weighted to represent the employed population, ages 18–65, excluding the self-employed. The Canadian data are based on a random sample of tax filers and are weighted to represent all tax filers.

[b]In thousands.

[c]Calculations are based on midpoints of reported IRA contribution intervals (see appendix B).

[d]Taken to be greater than or equal to 95 percent of actual limit.

addressed. It demonstrates succinctly how the first three questions are related to the fourth. The illustrative model also provides motivation for treating the first three separately from the fourth, although in principle, one general model could be used to address all four questions jointly. Estimation procedures designed to answer the first three questions are then considered, with particular attention given to whether a correctly specified, single behavioral equation can be used to describe both zero and positive levels of tax-deferred saving, or whether two behavioral relationships—one describing whether a person is a potential contributor and the other the desired IRA contribution—are required.

Table 1.5 Average Contributions and Percent with Contribution at the Limit, by Income Interval and Private Pension Coverage[a]

	United States				Canada			
	Employees with Private Pension		Employees without Private Pension		Employees with RPP Contribution[d]		Employees without RPP Contribution[e]	
Income Interval[b]	$[c]	% at L	$[c]	% at L	$	% at L	$	% at L
0–10	138	5.3	61	2.2	54	1.9	17	0.7
10–20	190	7.2	161	5.7	188	3.5	161	3.0
20–30	342	13.7	275	10.7	494	9.1	568	5.1
30–40	588	24.1	516	20.9	830	22.8	1305	12.4
40–50	883	38.2	650	25.9	983	39.3	2429	30.2
50–60	—	—	—	—	1199	45.0	2654	31.9
50–70	1073	48.5	809	35.7	—	—	—	—
60–70	—	—	—	—	1381	47.9	2968	38.1
70+	1170	51.7	978	44.0	—	—	—	—
70–80	—	—	—	—	1355	40.0	3655	50.0
80–90	—	—	—	—	1724	44.6	3396	50.1
90–100	—	—	—	—	1397	41.7	3646	51.3
100+	—	—	—	—	1503	37.8	3641	47.1

[a]The Canadian data pertain to 1980 and the U.S. data to 1982. Tabulations for the United States are in U.S. dollars and those for Canada in Canadian dollars. Data for the United States are from the May 1983 CPS and supplemental Survey of Pension and Retirement Plan Coverage. The data are weighted to represent the employed population, ages 18–65, excluding the self-employed. The Canadian data are based on a random sample of tax filers and are weighted to represent all tax filers.

[b]In thousands.

[c]Calculations are based on midpoints of reported IRA contribution intervals (see appendix B).

[d]Contributes to a Registered Pension Plan (RPP).

[e]The vast majority of this group do not have a pension plan.

1.2.1 A General Illustrative Model

Decisions about the amount to save in various forms are undoubtedly made jointly so that one decision cannot be considered fixed while the other is made. In addition, unmeasured individual attributes are likely to affect decisions about saving in each of two or more different forms. Thus persons who are observed to save more in one form are also likely to save more in another, not because saving in one form induces them to save more in another but rather because they are more inclined to save in any form. This means that one must disentangle the effects of individual-specific attributes from the effect that saving in one form has on saving in another.

The procedure outlined here addresses these problems by considering an individual's preferred allocation of current income to current consumption, tax-deferred saving, and other forms of saving, and then by considering how observed choices are affected by the limit on the tax-deferred saving alternative. Based on such a model, it would be possible to simulate, for example, how total saving would be changed if the limit on tax-deferred saving were raised or lowered. The procedure relies heavily on the fact that the optimal saving behavior of individuals who are not constrained by the limit differs from the behavior of those who are, with a statistical correction for the fact that persons who are at the limit, everything else equal, are likely to have a greater preference for saving than those who are not constrained; they are likely to save more in any form. In practice, the idea is to estimate the parameters of a "preference" function whose primary arguments are IRA contributions, at least one other form of saving, and current consumption. Associated with the preference function are optimal IRA contributions and optimal saving in other forms. In practice, it is necessary to choose these "demand" functions to fit the observable data and then to choose the preference function consistent with them. The procedure can be illustrated by a simple preference function.

Suppose that preferences for consumption and saving out of current income may be described by the simple form:

$$V = (Y - S_1 - S_2)^{1-\beta_1-\beta_2} S_1^{\beta_1} S_2^{\beta_2}, \tag{1}$$

where Y is income, S_1 and S_2 are tax-deferred saving and other saving, respectively, and β_1 and β_2 are parameters to be estimated. This function is intended to represent preferences over possible allocations of current income conditional on individual attributes like income and age, and on individual perceptions of the riskiness of different forms of saving.[2] This approach allows inferences about the relationship of income allocation to age without constraining the functional form to correspond to a particular life-cycle hypothesis. In practice, the parameters would depend upon measured individual attributes and would be allowed to vary randomly among individuals to capture unmeasured variation in individual preferences for current versus future consumption as well as different perceptions of risk, and so forth. In this simple case, the unconstrained optimal saving choices are

$$S_1 = \beta_1 Y, \text{ and} \tag{2}$$

$$S_2 = \beta_2 Y.$$

But in fact, the optimal choice is subject to a constraint; S_1 contributions cannot be greater than the IRA limit L. Until this limit is reached, contributions obey the equations above. But more generally the S_1 and S_2 functions are of the form:

(3)
$$S_1 = \begin{cases} \beta_1 Y & \text{if } \beta_1 Y < L, \\ L & \text{if } \beta_1 Y \geq L, \end{cases}$$

$$S_2 = \begin{cases} \beta_2 Y & \text{if } \beta_1 Y < L, \\ \dfrac{\beta_2}{1-\beta_1}(Y - L) & \text{if } \beta_1 Y \geq L. \end{cases}$$

Thus there are two S_2 saving functions. As long as the IRA limit has not been reached, saving obeys the optimizing rule $\beta_2 Y$. But after the limit is reached, the saving function is of the form $[\beta_2/(1-\beta_1)](Y-L)$. This illustration ignores the tax deferment that makes IRAs more attractive than alternative forms of retirement saving. Introducing the tax rate in the example changes the utility function to

(4)
$$V = [Y(1 - t) - S_1(1 - t) - S_2]^{1-\beta_1-\beta_2} S_1^{\beta_1} S_2^{\beta_2},$$

where, assuming that saving is small relative to income, t is the marginal tax rate. The optimal saving choices then become

$$S_1 = \begin{cases} \beta_1 Y & \text{if } \beta_1 Y < L, \\ L & \text{if } \beta_1 Y \geq L, \end{cases}$$

$$S_2 = \begin{cases} \beta_2 Y(1 - t) & \text{if } \beta_1 Y < L, \\ \dfrac{\beta_2}{1 - \beta_1}(Y - L)(1 - t) & \text{if } \beta_1 Y \geq L. \end{cases}$$

In this formulation, the marginal tax rate does not affect S_1 (IRA) saving, unless it affects preferences for current versus future consumption through the parameters β_1 and β_2. The empirical findings reported below suggest an uncertain effect of the marginal tax rate independent of income, even though the rate of return on IRAs does depend on the marginal tax rate.

In practice, the parameters β_1 and β_2 would be made functions of individual characteristics like age, occupation, possibly income itself, the tax rate, and other conditioning variables that would likely determine individual preferences of possible allocations of current income. To estimate the model, it is also necessary to choose a stochastic specification for the βs. One also needs to choose a specification that allows optimal, or "desired", values S_1 and S_2 to be negative, since many individuals will not save in any form and indeed will borrow.[3]

As emphasized above, this particular functional form is only for illustrative purposes; the form that is ultimately chosen must be determined by the data. But this simple example demonstrates how changes in the limit may affect behavior. In particular, explicit reference to a preference function assures a specification of saving S_2 after the limit L on S_1 is reached that is internally consistent with the function that applies before the limit is reached. And, in a fully specified model, estimates of the βs could be used to simulate the effects of changes in the limit L on total saving, not just the effect on tax-deferred saving.[4] Estimation of only the S_I function is treated in this paper.[5]

1.2.2 Independent Analysis of Contributions to Tax-Deferred Saving Accounts

Within the general framework described above, one can treat the tax-deferred saving equation separately. At least two important issues must be addressed in order to analyze determinants of IRA contributions. The first is simply that in addition to the upper limit on contributions, many individuals, indeed the majority, do not contribute anything at all to IRA (or Keogh) accounts. The standard way to conduct an analysis in this situation would be to use a Tobit model with a lower truncation point at zero and an upper truncation point at the contribution limit. The second issue, however, is that the determinants of whether one contributes at all may be different from the determinants of how much one contributes once an account is established. While it is true that the short-run effect of changes in contribution limits on total contributions is determined only by initial contributors, there may be considerably more room to change total contributions through increasing the number of persons who contribute than by increasing the contributions of current contributors. It is important, therefore, to understand the determinants of the contributor status decision.

Key Issues

To provide accurate predictions of the determinants of IRA contributions and of the effect on contributions of changes in the existing contribution limits, the most important consideration in estimation is to account for the existing limit on observed contributions. Thus, an intuitive discussion of the effect of the limit on estimation, together with procedures that can be used to correct for it, helps to put the important ideas in perspective, although part of the discussion will be familiar to many readers. Mathematical details of the estimation procedure are presented in appendix A.

Consider first figure 1.1. Suppose that the relationship between income and IRA contributions if there were no contribution limits would be represented by the solid line $a + bY$. That is, given Y the average (expected) contribution would be $a + bY$. Of course, for any level of income Y there would be a distribution of contribution levels, represented by the vertical lines; not everyone with income level Y would contribute the same amount. Now suppose that everyone in the sample faced a contribution limit L. We would now observe no contributions above L, and presumably individuals who otherwise would contribute

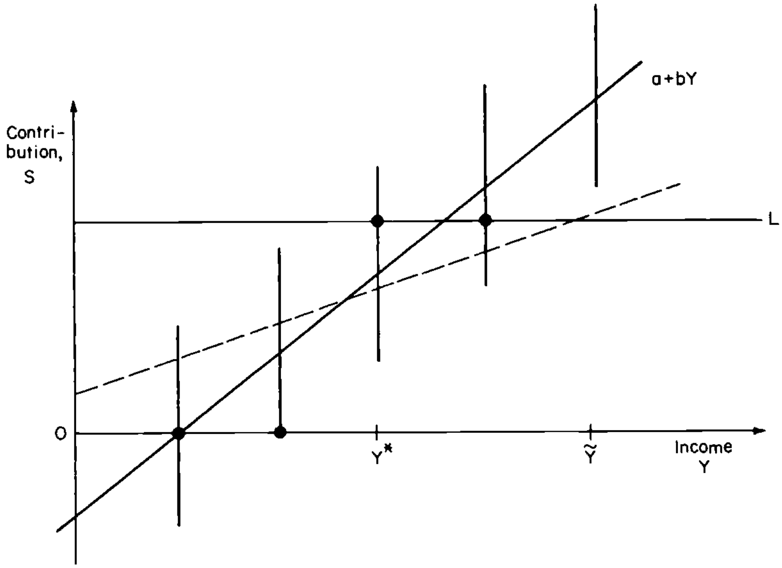

Fig. 1.1 Effect of contribution limits on estimation.

above this limit would in the face of the limit contribute at the limit. This would give rise to a concentration of contributions at the limit, indicated by the heavy dots at that level. In addition, it is not possible to contribute less than zero; we would observe a concentration of points at zero, indicated by the heavy dots along the horizontal axis. If we think of fitting a line, say by least squares, to the data points that are actually observed, we would obtain a fitted line something like the dashed line in the figure.

Suppose that from this fitted line we attempted to predict the relationship between income and contributions S. It is easy to see that this estimate would be a very substantial underprediction. Thus it is clear that standard estimation procedures will not lead to plausible conclusions in this case. And it should also be clear from the figure that the reason is that observed contributions do not represent the contributions that individuals would like to make were they not constrained by the limit.

It is also useful to consider the distribution of contributions for persons with a particular level of income, say Y^*. An illustration of such a distribution is show in figure 1.2. If it were not for the limit at L and the limit at zero, the distribution of contributions S would look something like a bell-shaped curve. But as demonstrated in figure 1.1, we know that we will not observe contributions greater than L, and we will observe no contributions less than zero. The distribution of observed contributions between zero and L would look just like the underlying curve. But instead of a distribution tapering off smoothly to the right and to the left, there would be concentrations of contributions at L and at zero.

The standard Tobit maximum likelihood estimation procedure that takes account of this truncation effect is based on an assumed underlying relationship like $a + bY$, as show in figure 1.1, together with a distribution of contributions around this relationship. In this case there are three possible outcomes: the contribution is zero, it is between zero and L, or it is at L. The values of a, b, and σ that maximize the likelihood of observing the sample values yield estimates of the relationship labeled $a + bY$ in figure 1.1, as well as the dispersion of underlying observations around this expected value. Thus the estimates that are obtained need to be interpreted as pertaining to this underlying relationship. For example, b indicates the relationship between Y and S if there were no limit on contributions. Or, it tells us how an increase in income would affect contributions as long as the contribution limit was not reached; after that, contributions would be observed at the limit L but desired contributions would be above L.

It is also important to realize that only persons who are constrained by the current limit will be affected by a new higher limit. (If the limit is lowered, of course, increasingly large numbers of people will be constrained by it.) With the help of figure 1.2, it is easy to determine the effect of small changes in L on contributions. Consider first an individual whose observed contribution is less than L. Such an individual could contribute more but chooses not to; he is not constrained by the limit. His desired contribution level is less than L, so raising the limit would not increase his contribution level. Consider, on the other hand, a person who is observed to contribute at the limit L. If L were raised, this person would likely contribute more. Thus the effect

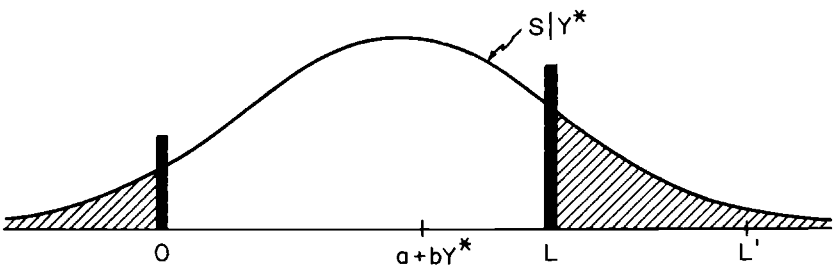

Fig. 1.2 Distribution of contributions, given income level Y^*.

of raising the limit by 1 is just 1 times the probability that the individual is constrained by the limit. Somewhat more formally, we can write the derivative as

(5)
$$\frac{dC}{dL} = \begin{cases} 0 \text{ if } C < L \\ 1 \text{ if } C > L \end{cases}$$
$$= 0 \cdot \Pr[C < L] + 1 \cdot \Pr[C > L]$$
$$= \Pr[C > L].$$

Thus for any individual the expected change in the contribution level is equal to the probability that the underlying desired contribution is greater than L.

It is also important to realize that this derivative depends upon the level of L. Suppose that L were farther to the right than is shown in figure 1.2, say at L'. The effect on contributions of an increase in the limit from L' would be much smaller than the effect of an increase from L because the likelihood that an individual with income Y^* would like to contribute more than L' is much lower than the likelihood that he would like to contribute more than L. While all people, or almost all people, with observed contributions at L would increase their contributions if the limit were raised, very few would increase their contributions to the level L'. Thus to infer the effect of an increase in the limit on contributions, it is necessary to have an estimate of the underlying distribution of desired contribution levels. With an estimate of the distribution of S given Y^*, it is possible to predict the expected contribution given Y^* for any level of L.

Estimation Possibilities

In practice, estimates that address the issues discussed above can be developed in several ways, depending on the hypothesized underlying process that leads to observed contributions. There are two basic possibilities. The first possibility assumes, as in the discussion above, that zero contributions can be thought of simply as a special case of a single underlying preferred contribution behavioral relationship. That is, one could think of a preferred contribution level that declines continuously with decreases in income until the zero contribution level is reached. The second general possibility is that there are two underlying behavioral relationships that determine observed contributions: one relationship describes the likelihood that a person will be a contributor and the second relationship describes the desired contribution, should one be a contributor.

If only one behavioral relationship is assumed, there are at least three ways to obtain estimates. The alternative procedures allow a test of the underlying assumption itself. One procedure uses all observations including those with zero contributions; the second uses only observations with positive contributions; and the third uses only the information on whether a person contributes without using the amount of a positive contribution.[6] If there is, in fact, only one underlying relationship that determines observed contributions, then each of these methods yields consistent estimates of the parameters of this single relationship (except for the third which yields estimates up to a variance scaling factor). If the estimates based on the different groups of observations lead to different estimates, then it is likely that the underlying process should be described by two relationships.

If the goal of the analysis were only to predict the effect of changes in the limits, it is reasonable to concentrate on those who contribute and to allow the parameter estimates to be determined by this group, since noncontributors are not initially affected by changes in the limit. It is at least as important, however, to understand the factors that determine whether a person is a contributor. As emphasized above, changes in the number of contributors at any limit could have a much greater effect on national saving than changes in the limit. To the extent that the determinants of whether one is a contributor are different from the determinants of the amount of the contribution, it is important to consider both of these relationships. The formal details of a two-equation model, together with details of the single-equation estimation possibilities and related tests of behavioral assumptions, are presented in appendix A.

1.2.3 The Empirical Specification

In the illustrative specification in section 1.2.1, desired contributions to tax-deferred saving are of the form $S = \beta Y$. A direct statistical counterpart of this specification is

$$(6) \qquad S = \beta Y e^{\epsilon} = aY^{b} \cdot Y \cdot e^{\epsilon} = \alpha_0 X_1^{\alpha_1} X_2^{\alpha_2} \ldots Y^{b} \cdot Y \cdot e^{\epsilon}.$$

Based on estimates for Canada, this specification fits the observations on positive contributions extremely well. Note that the specification implies that given Y (and the other variables X) the variance of S increases with Y; the disturbance term is heteroscedastic. The specification also leads to a constant income elasticity and is conveniently log-linear. However, this specification is not appropriate if we incorporate contributions at zero and, in the abstract, the possibility of desired contributions less than zero. To consider whether the determinants of contributor status are different from the determinants of the amount of positive contributions, a specification that in principle allows negative as well as positive values and that also fits the observations on positive contributions must be used. Such a specification, and one that in practice fits the observed data well, is of the form

$$(7) \qquad S = a + (Y^b + \eta)Y + \epsilon$$
$$= \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + Y^{1+b} + \eta Y + \epsilon,$$

where η and ϵ are disturbance terms and the variance of $\eta Y + \epsilon$ is given by

$$(8) \qquad V(\eta Y + \epsilon) = \sigma_\eta^2 Y^2 + \sigma_\epsilon^2 = \sigma^2$$

Thus the specification incorporates the property that the variance of S increases with income, and it also allows for "desired" contributions less than zero. The elasticity of desired contributions with respect to income is given by $(1 + b)/[1 + (a/Y^{1+b})]$ and thus approaches $1 + b$ as income increases.[7]

For simplification, appendix A, which describes the details of the alternative estimation procedures, is written in terms of the specification $S = X\beta + \epsilon$, where $V(\epsilon) = \sigma^2$. Development in terms of the above specification may be obtained by replacing $X\beta$ by $a + Y^{1+b}$ and σ^2 by $(\sigma_\eta^2 Y^2 + \sigma_\epsilon^2)$. Recall that the three single-equation approaches use: (a) all observations including those with zero contributions (two-limit Tobit); (b) only observations with positive contributions (one-limit Tobit); and (c) only information on contributor status (probit). A two-equation model that jointly estimates contribu*tor* and contribu*tion* outcomes is also described in appendix A. This model permits the determinants of whether a person contributes to differ from the determinants of the desired level of a positive contribution and allows the stochastic components of the two choices to be correlated.

In addition, the CPS data on IRA contributions are reported only by intervals—\$0, \$0–\$100, \$100–\$500, \$500–\$1000, \$1000–\$2000. Thus the probabilities of positive contributions are of the form

$$(9) \qquad \Phi[(u - X\beta)/\sigma] - \Phi[(l - X\beta)/\sigma],$$

where u and l are the upper and lower bounds of an interval, and $\Phi[\cdot]$ denotes the standard normal distribution function. Thus the likelihood function in this case includes no density function terms; it is composed only of normal cumulative distribution functions. This may be contrasted with the Canadian data that record exact positive contributions.

1.3 Results

1.3.1 Data

The data were obtained through a special supplement to the May 1983 Current Population Survey (CPS). The data on IRA contributions pertain to the 1982 tax year. No information is provided on 1982 contributions to Keogh plans, thus self-employed persons have been excluded from this analysis. In addition, the raw data pertain to individuals, not families. Some of the estimates reported below are based on the individual data, with indicator variables for marital status and sex. Since it is not known from the individual data whether an individual's spouse works, the actual upper limit on family contributions cannot be determined from the individual data alone. Some individuals reported contributions greater than $2000—primarily at $2250 and at $4000, apparently confusing individual contributions with the family total. When the individual data are used, contributions above $2000 are not explicitly recorded at that level, rather any reported contribution above $2000 is treated as a contribution at the $2000 limit. Under the model assumption, this procedure still yields unbiased parameter estimates; it simply does not use all the information.

Family data were created by matching and combining information for individuals in the same household. This allows estimation of family income and of family IRA contribution limits based on the employment status of the husband and wife. Estimated marginal tax rates were also calculated for the family. The estimates were based on average marginal tax rates by income and family status reported by the IRS. As mentioned above, IRA contributions for each family member are reported only by intervals. The intervals for a family were obtained by inferring the possible family intervals from the possible individual reporting intervals. There are twelve possible family intervals in total. Details of the procedures used to create the family data and the tax rates are reported in appendix B.

1.3.2 Parameter Estimates: Single-Equation Models

Individual Data

Estimates by three methods of estimation are reported in tables 1.6 and 1.7. Summary statistics for the variables included in each equation are presented in appendix C. Table 1.7 includes variables indicating whether a person was covered by a private pension plan and whether the worker participated in a salary reduction plan (401[K] or 403[B] plans which permit workers to defer compensation); these variables are not included in table 1.6. Column 1 in each table presents estimates based on the two-limit Tobit specification. Column 2 shows probit estimates where the standard error of ϵ is set at the two-limit Tobit estimate (e.g., 5622 in table 1.6). This allows easy comparison of the two sets of estimates. It may be seen that the parameter estimates are virtually the same. Whether there is a difference between the determinants of contribution status and the determinants of the desired level

Table 1.6 **Parameter Estimates, by Method of Estimation, Individual Data[a]**

Variable	2-Limit Tobit	Probit, σ_ϵ from 2-Limit Tobit	1-Limit Tobit	Probit, σ_ϵ from 1-Limit Tobit
	Method of Estimation			
σ_η	.124(.012)	.120(.016)	.051(.012)	.039(.006)
σ_ϵ	5622(212)	*5622*	2015(210)	*2015*
Constant	−29712(1039)	−29196(758)	−6608(1276)	−10527(266)
Income	.839(.006)	.839(.004)	.753(.013)	.749(.004)
MTR[b]	—	—	—	—
Age	240(9)	228(8)	110(16)	81(3)
Unmarried Women	56(244)	17(247)	211(261)	5(87)
Unmarried Men	7(277)	−208(281)	1353(357)	−74(99)
Married Women	2869(209)	2768(208)	1073(238)	986(73)
Education	650(37)	653(35)	109(35)	230(12)
Private Pension	—	—	—	—
Salary Reduc. Plan	—	—	—	—
LF[c]	−9548.4	−6745.3	−2745.2	−6745.3
N	20513	20513	2999	20513
< 0	17514	17514	—	17514
$> 0, < L$	1003	2999	1003	2999
$= L$	1996		1996	

[a]Standard errors are in parentheses.
[b]Marginal tax rate.
[c]Likelihood function.

of IRA contributions may not be revealed by this comparison, however, since the preponderance of individuals make no contribution and thus the contribution status (the probit portion) will dominate the two-limit estimates. The two-limit Tobit estimates will therefore tend to look like the probit estimates.

A better way to reveal differences in the two relationships is to separate analysis of contribution amounts from the analysis of contributor status. The 1-limit Tobit estimates in column 3 of the tables are based only on the contributions of contributors, and the probit estimates of contributor status in column 4 are obtained by setting the standard error of ϵ equal to the one-limit estimate (e.g., 2015 in table 1.6). These last two columns reveal that the two sets of coefficients are quite similar. The reported coefficient on income is the estimate of $(1 + b)$. It is virtually the same in each of the alternative methods, and the estimated parameters on age seem not to be significantly different in the two cases. The estimated sex effects are also very close in the two cases, with one exception. The estimates suggest that unmarried men contribute more than married men (the omitted category)

Table 1.7 Parameter Estimates, by Method of Estimation, Individual Data[a]

	Method of Estimation			
Variable	2-Limit Tobit	Probit, σ_ϵ from 2-Limit Tobit	1-Limit Tobit	Probit, σ_ϵ from 1-Limit Tobit
σ_η	.124(.012)	.121(.016)	.051(.012)	.040(.006)
σ_ϵ	5621(213)	*5621*	2028(214)	*2028*
Constant	−29608(1039)	−29119(761)	−6713(1308)	−10576(269)
Income	.838(.006)	.838(.004)	.749(.014)	.749(.004)
MTR[b]	—	—	—	—
Age	239(9)	228(8)	110(16)	81(3)
Unmarried Women	42(244)	6(247)	205(264)	1(88)
Unmarried Men	6(277)	−208(282)	1350(359)	−75(100)
Married Women	2856(210)	2759(209)	1079(241)	990(74)
Education	644(37)	648(36)	108(35)	230(13)
Private Pension	23(162)	−18(165)	221(171)	−13(59)
Salary Reduc. Plan	789(346)	751(352)	239(326)	262(125)
LF[c]	−9546.0	−6743.2	−2743.9	−6743.2
N	20513	20513	2999	20513
< 0	17514	17514	—	17514
$> 0, < L$	1003	2999	1003	2999
$= L$	1996		1996	

[a]Standard errors are in parentheses.
[b]Marginal tax rate.
[c]Likelihood function.

but are apparently no more likely than married men to contribute. The constant terms in the two equations differ, although given the estimated standard errors, the difference may not be as great as the estimated values suggest.

A more formal test is to compare the sum of the likelihood values from columns 3 and 4 with the likelihood value in column 1. Under the null hypothesis is that one behavioral relationship is sufficient to describe both contributor status and the amount of contributions, the sum of the likelihoods in columns 3 and 4 will not be statistically different from the likelihood value in column 1. (Minus 2 times the difference will be distributed chi-square with 7 degrees of freedom, with a .05 level of 14.1.) Thus the hypothesis would be rejected in this case. However, the very large sample size will reveal differences even if they have rather small practical importance.

The coefficient on income of .753 implies an elasticity of desired contribution with respect to income of .63, evaluated at the mean of the data for contributors. The desired contribution increases by about \$110 with each year of age according to the estimates for contributors, while the comparable estimate from the probit equation is $81. Given other variables, married women would choose to contribute about $1000 more than married men and the more educated would contribute more than those with less education. The estimated unmarried women effect is not statistically significant.

Summary statistics presented in table 1.5 suggested that employees covered by a private pension plan were more likely to contribute to an IRA. Parameter estimates in table 1.7, however, suggest that the association between pension coverage and IRA contributions can be attributed to other differences in the individual characteristics of those covered and not covered by a pension plan. After controlling for other characteristics, pension plan coverage is not significantly associated with desired contributions. Participation in a salary reduction plan (less than 4 percent of the sample) is positively associated with IRA contributions.

Both the two-limit and one-limit models fit the data rather well. This is demonstrated in table 1.8. Based on the estimates in table 1.6, the

Table 1.8 Model Fit; Actual versus Predicted Proportions by Income Interval, Contribution Interval, and Method of Estimation[a]

	Contribution Interval											
	Zero		Between		At Limit		\$0–\$500		\$500–\$1000		\$1000–\$2000	
Income Interval[b]	A	P	A	P	A	P	A	P	A	P	A	P
	2-Limit Tobit											
0–10	.95	.94	.02	.03	.03	.03	.01	.01	.01	.01	.01	.01
10–20	.89	.89	.05	.04	.06	.06	.01	.01	.01	.01	.02	.02
20–30	.81	.80	.06	.07	.13	.13	.02	.02	.02	.02	.03	.03
30–40	.68	.69	.08	.09	.23	.22	.01	.02	.02	.02	.05	.04
40–50	.56	.57	.09	.09	.35	.33	.01	.02	.01	.02	.07	.04
50–75	.47	.45	.10	.08	.43	.47	.01	.02	.02	.02	.07	.04
75+	.38	.33	.08	.06	.54	.60	.00	.01	.02	.02	.06	.03
	1-Limit Tobit											
0–10	—	—	.43	.48	.57	.52	—	—	—	—	—	—
10–20	—	—	.43	.41	.57	.59	—	—	—	—	—	—
20–30	—	—	.33	.35	.67	.65	—	—	—	—	—	—
30–40	—	—	.27	.27	.73	.73	—	—	—	—	—	—
40–50	—	—	.19	.22	.81	.78	—	—	—	—	—	—
50–75	—	—	.18	.17	.82	.83	—	—	—	—	—	—
75+	—	—	.13	.13	.88	.87	—	—	—	—	—	—

[a]A = actual; P = predicted.

[b]In thousands.

predicted proportion of individuals with contributions at zero, at the upper limit, and within selected intervals are compared with the actual proportions, by income interval. It is important in interpreting these results to realize that gross misspecification of the functional form that relates contributions to income would be revealed in the comparisons by income level. The comparisons indicate close correspondence between predicted and actual proportions. The only apparent discrepancy is that the two-limit Tobit specification underpredicts the proportion of contributions in the $1000–$2000 range and correspondingly overpredicts the proportion of contributions at the limit. Given the differences between a few of the two-limit and one-limit parameter estimates, the similarity of the predictions may be surprising. However, the major difference in parameter estimates is a larger negative constant term in the two-limit than in the one-limit specification, which is offset by a larger disturbance term variance. The likelihood function is quite flat with respect to these two parameters. Thus the sum of the last two likelihoods is not so different in magnitude from the two-limit Tobit likelihood.

Family Data

Parameter estimates based on the family data are reported in table 1.9. The variable specification is identical to that used for the individual data with two exceptions: the marginal tax rate (MTR) has been added and the dummy variable for "married women" has been deleted, with non-single-person families the norm group. Married men and women appear together in the family data, but as two separate observations in the individual data.

Where the variables are the same, the parameter estimates are very similar to those based on the individual data. For example, the estimated income coefficient based on the two-limit Tobit model is .78 using family data and .84 using individual data. The effects of age and education are also quite close.

The results suggest that the marginal tax rate has no effect on the amount of contributions but a positive effect on contributor status. The coefficient on the estimated marginal tax rate in the two-limit specification is 200 with a standard error of 39. This would suggest that an increase of ten percentage points in the marginal tax rate would increase desired IRA contributions by about $2000. On the other hand, the one-limit and probit estimates in columns 3 and 4 suggest that the tax rate has no effect on the level of contributions (column 3) but a positive effect on contributor status (column 4). The latter estimate implies that if all marginal tax rates were increased by ten percentage points—on average from about 24 percent to 34 percent—the proportion of persons who contribute would increase from .134 to .193, or by 44 percent.

Table 1.9 Parameter Estimates, by Method of Estimation, Family Data[a]

	Method of Estimation			
Variable	2-Limit Tobit	Probit, σ_ϵ from 2-Limit Tobit	1-Limit Tobit	Probit, σ_ϵ from 1-Limit Tobit
σ_η	.096(.014)	.084(.023)	.076(.017)	.042(.011)
σ_ϵ	6367(325)	*6367*	3219(548)	*3219*
Constant	−32111(1428)	−31232(1076)	−15224(4284)	−15860(548)
Income	.776(.027)	.756(.039)	.810(.033)	.701(.038)
MTR[b]	200(39)	219(38)	−15(70)	109(20)
Age	211(11)	195(12)	219(51)	98(6)
Unmarried Women	401(384)	207(421)	781(724)	118(216)
Unmarried Men	−273(436)	−645(454)	2724(1016)	−313(233)
Married Women	—	—	—	—
Education	550(45)	546(47)	142(82)	276(24)
Private Pension	—	—	—	—
Salary Reduc. Plan	—	—	—	—
LF[c]	−6727.1	−4601.9	−2089.3	−4601.9
N	15149	15149	2030	15149
< 0	13119	13119	—	13119
> 0, < *L*	756	2030	756	2030
= *L*	1274		1274	

[a]Standard errors are in parentheses.
[b]Marginal tax rate.
[c]Likelihood function.

Canadian estimates for 1981 (see section 1.4) show a much smaller statistically significant effect of the marginal tax rate on contributor status, with a smaller and not statistically significant effect on the amount of contributions. (But the difference between the two estimates is also not significantly different from zero.) Canadian estimates for 1976 show no effect of the marginal tax rate in either equation.[8] An alternative log-linear model for contributors only shows a precisely estimated zero effect of the marginal tax rate on the amount of contributions in both 1976 and 1980. Thus the estimated effect seems quite sensitive to the statistical specification.

Table 1.10 includes indicators of pension coverage and participation in a salary reduction plan. These estimates indicate that if at least one member of a family is covered by a pension plan, the likelihood of contributing to an IRA is higher. The individual data suggested essentially no relationship. A possible explanation is that married persons without pensions, but whose spouses are covered by a pension, have a high likelihood of contributing to an IRA. In the individual data, these people would be treated as not having a private pension.

Table 1.10 Parameter Estimates, by Method of Estimation, Family Data[a]

Variable	2-Limit Tobit	Probit, σ_ϵ from 2-Limit Tobit	1-Limit Tobit	Probit, σ_ϵ from 1-Limit Tobit
	Method of Estimation			
σ_η	.088(.014)	.063(.025)	.078(.017)	.030(.012)
σ_ϵ	6477(327)	*6477*	3198(550)	*3198*
Constant	−32687(1444)	−31470(1032)	−15522(4394)	−15607(514)
Income	.779(.025)	.745(.042)	.813(.033)	.687(.042)
MTR[b]	180(38)	781(423)	−29(71)	104(19)
Age	211(109)	192(11)	221(52)	95(5)
Unmarried Women	520(381)	196(407)	914(740)	106(205)
Unmarried Men	−98(432)	−607(444)	2824(1032)	−293(222)
Married Women	—	—	—	—
Education	538(45)	524(45)	143(83)	258(22)
Private Pension	1626(236)	1599(227)	617(513)	787(112)
Salary Reduc. Plan	791(435)	781(423)	−114(704)	385(208)
LF[c]	−6699.1	−4574.0	−2088.5	−4574.0
N	15149	15149	2030	15149
< 0	13119	13119	—	13119
$> 0, < L$	756	2030	756	2030
$= L$	1274		1274	

[a]Standard errors are in parentheses.
[b]Marginal tax rate.
[c]Likelihood function.

1.3.3 Simulations: Single-Equation Models

Simulations are obtained under three policy assumptions: the existing IRA program, the proposal contained in the administration's recent tax reform proposal (U.S. Department of Treasury 1984), and a modification of the Treasury proposal that restricts spousal IRAs. The Treasury proposal increases the limits to $2500 for both employed persons and nonemployed spouses. The modified Treasury proposal also increases the limit for employed persons to $2500 but sets the spousal limit at $500, instead of $2500.

Simulations based on the family data are presented in table 1.11.[9] To serve as a basis for comparison, the average IRA contribution under the current plan has been simulated for several demographic groups. The model yields an average predicted contribution for all families under the current plan of $312.[10] The simulations indicate that the Treasury plan would increase 1982 contributions by 30 percent to $405 per family. The largest increases are for married, one-earner families whose limit is increased by the Treasury proposal from $2250 to $5000.

Table 1.11 Simulated IRA Contributions, by Plan and Family Type, Based on Family Data[a]

Family Type	Current Plan ($2000/$250)	Treasury Plan ($2500/$2500)	Mod. Treas. Plan ($2500/$500)
All Families			
Avg. Contribution	$312	$405	$370
% Change	—	30	18
Unmarried Head			
Avg. Contribution	136	162	162
% Change	—	19	19
Married, 1 Earner			
Avg. Contribution	267	475	335
% Change	—	78	25
Married, 2 Earners			
Avg. Contribution	536	620	620
% Change	—	16	16

[a]These estimates are unweighted, since it was not clear what weights should be used for the "created" families.

The predicted average contribution for this group would increase from $267 to $475, about 78 percent. The smallest increase, about 16 percent, is for married, two-earner families whose limit increased only from $4000 to $5000.

The modified Treasury plan yields an overall increase of about 18 percent. The limit changes, and thus contributor responses, for unmarried heads and married, two-earner families are the same as in the unmodified Treasury proposal. The modified Treasury plan increases the limit faced by married, one-earner families by only $750, from $2250 to $3000, instead of $5000. The simulated increase in average contributions by this group is 25 percent, about a third as large as the simulated increase under the Treasury plan.

Simulations based on the individual data are shown in table 1.12. Unlike the family data, the individual data do not provide enough information to completely specify the limit faced by each person. Employed single persons face a limit of $2000. For married couples the limits are $2000 per person if both work and $2250 if only one works. If both work, then both will appear in the sample and the appropriate limit for each is $2000. If only one is employed, however, the nonemployed spouse will not be present in the sample, since only employed persons received the CPS pension supplement questionnaire. The appropriate limit for the employed spouse is $2250. The problem is to assign the "correct" limit ($2000 or $2250) to each married person in the sample, given that we do not know if the spouse is employed. [11] If the married person is a woman, a limit of $2000 is assigned, assuming that her spouse also works. If the married person is a man, the limit

Table 1.12 Simulated IRA Contributions, by Plan Based on Individual Data[a]

Individual Type	Current Plan ($2000/$250)	Treasury Plan ($2500/$2500)	Mod. Treas. Plan ($2500/$500)
All Persons			
Avg. Contribution	$246	$326	$296
% Change	—	33	20
Unmarried Males			
Avg. Contribution	120	142	142
% Change	—	18	18
Unmarried Females			
Avg. Contribution	134	158	158
% Change	—	18	18
Married Males			
Avg. Contribution	323	469	395
% Change	—	45	22
Married Females			
Avg. Contribution	280	332	332
% Change	—	18	18

[a]Weighted to reflect national population.

is randomly assigned. With probability P it is set at $2000 and with probability $1-P$ at $2250, where P is the proportion of wives of working husbands that are employed.

The individual data simulations based on this procedure are quite close to those obtained using the family data. For all persons, the simulations indicate that the Treasury plan would increase 1982 contributions by about 33 percent and that the modified Treasury plan would increase contributions by about 20 percent. The largest effects are for married men, the group facing the largest change in limits.

1.3.4 Parameter Estimates: Two-Equation Model

The above results suggest that the observed outcomes can in general be described well with a single behavioral relationship. If two relationships are required, the one-limit Tobit and the probit models together, even if estimated independently, should provide a reasonably accurate description of the determination of contributions. The two-equation model described in appendix A, however, distinguishes between a "potential" contributor behavioral relationship and the level of desired contributions, were one to contribute. Under this representation, a potential contributor could be observed with zero contributions not because the person was a noncontributor, but rather, say, because income was too low for the person to devote current income to future consumption. To the extent that the parameters in the two relationships differ, the single-equation probit estimates, for example, will not provide accurate estimates of potential contributor status. As the parameters in the two relationships become close, and the correlation between them approaches 1, however, the two-equation model approaches the two-limit Tobit specification. If only the variable coefficients were the same, the results could differ if the correlation between the disturbance terms in the two relationships were not unity.

Estimation of several two-equation models indicated only minor differences between parameters based on the single-equation models and those derived from two relationships estimated jointly. The two-equation model is of the form

$$S = a_s + Y^{b_s} + \eta_s Y + \epsilon_s \qquad \text{Contributor Status}$$

$$C = a_c + Y^{b_c} + \eta_c Y + \epsilon_c \qquad \text{Contribution Amount}$$

The details of the specification and estimation procedure are described in section 2 of appendix A. The key distinction between the specification used here and the common sample-selection specification is that even potential contributors can have zero contributions, while others would not contribute under any circumstances. Only the latter are "noncontributors" in the strict sense.

Illustrative estimates for this model are presented in table 1.13. In this specification, $V(\eta_c) = V(\eta_s)$, $V(\epsilon_c) = V(\epsilon_s)$, and all covariances

Table 1.13 Parameter Estimates, Two-Equation Model, Individual Data[a]

Variable	Level of Contribution (C)		Contributor Status (S)
σ_η		.037(.005)	
σ_ϵ		2034(115)	
$\rho(\eta_c, \eta_s)$		.095(.094)	
Constant	−4293(463)		−9547(464)
Income	.758(.009)		.744(.007)
MTR[b]	—		—
Age	92(8)		65(4)
Unmarried Women	423(177)		−111(50)
Unmarried Men	1347(248)		−410(110)
Married Women	998(168)		825(92)
Education	—		250(19)
Private Pension	—		—
Salary Reduc. Plan	—		—
LF[c]		−9497.8	
N		20513	
< 0		17514	
$> 0, < L$		1003	
$= L$		1996	

[a]Standard errors are in parentheses.
[b]Marginal tax rate.
[c]Likelihood function.

other than $\text{cov}(\eta_c, \eta_s)$ are set to zero. Education is excluded from the contributions equation. In practice, covariance or exclusion restrictions were required for identification of key parameters. Because the likelihood function was so flat, more restrictions were necessary than were in principle required.

The parameter estimates indicate that the correlation between η_c and η_s is not significantly different from zero.[12] This suggests independence of the contributor and contributions relationships, given measured individual characteristics. Thus these estimates are very close to the single-equation results presented in columns 3 and 4 of table 1.6.[13]

In principle, however, this specification allows estimation of the proportion of persons who are potential contributors but, because of a liquidity constraint, for example, are observed not to contribute. The probability that a person does not contribute is given in this specification by $1 - \Pr[S < 0] + \Pr[S > 0 \text{ but } C < 0]$. Averaged over all observations in the sample, the proportion of noncontributors is .854, the same as in the probit estimates based on individual data. A proportion of .146 contribute. The proportion of potential contributors, $\Pr[S > 0]$, is estimated to be .182. Thus, the proportion of potential contributors who do not contribute, $\Pr[S > 0 \text{ but } C < 0]$, is estimated to be .036, about 20 percent of potential contributors.

1.4 Comparison of Results for the United States and Canada

Since the Canadian and the U.S. systems are very similar in their general outlines, it is informative to compare the model estimates for the two countries. The Canadian equivalent of IRA and Keogh plans is the Registered Retirement Savings Plan (RRSP). RRSP contributions are also tax-deferred and have upper limits determined both by income and by a maximum level. The Canadian rules also provide for different limits depending on whether a person is a member of a private pension plan.

Since the Canadian tax system is on an individual basis, the most appropriate comparison is with the individual estimates for the United States. Estimates analogous to those in table 1.6 for the United States are shown in table 1.14 for Canada. While the general model specification is identical in the two countries, the specific variables do not correspond precisely. In particular, the variables for women, married, and education are not included in the Canadian version, and there is no marginal tax rate variable in the U.S. version. The comparable parameter estimates, however, are surprisingly similar, based on a comparison of the one-limit Tobit and the corresponding probit estimates in the two countries. The coefficient on income is .75 in the United States, while it is approximately .81 in Canada. The estimated effect

Table 1.14 RRSP Contribution Parameter Estimates, by Method of Estimation, Totals, 1981[a]

Group and Variable	2-Limit Tobit	Probit, σ_ϵ from 2-Limit Tobit	1-Limit Tobit	Probit, σ_ϵ from 1-Limit Tobit
σ_η	.106(.006)	.096(.014)	.125(.032)	.103(.015)
σ_ϵ	2999(183)	—	3199(632)	—
Constant	−9151(561)	−8951(466)	−11657(4690)	−9539(498)
Income	.794(.008)	.789(.009)	.807(.036)	.795(.009)
MTR[b]	37(8)	39(8)	28(40)	42(8)
Age	62(9)	61(9)	79(47)	65(10)
Govt. Employee	—	—	—	—
Employee w/RPP	—	—	—	—
Self-Employed	—	—	—	—
Professional	—	—	—	—
Farmer/Fisherman	—	—	—	—
LF[c]	−6583.8	−1763.9	−4818.9	−1763.9
N	4038	4038	1083	4038
< 0	2955	2955	—	2955
$> 0, < L$	516	1083	516	1083
$= L$	567		567	

[a]Standard errors are in parentheses.
[b]Marginal tax rate.
[c]Likelihood function.

of age is approximately $80 to $110 in the United States, while it is $65 to $80 in Canada. The estimates for Canada also show a close correspondence between the one-limit Tobit and the probit estimates, indicating that a single behavioral relationship apparently describes the observations rather well. Indeed, for Canada the estimates in the two equations are not statistically different, based on the chi-squared test described above.

As already discussed, the estimated effect of the marginal tax rate in Canada is not statistically different from zero in the contributions equation; the estimate for the United States, reported for the family data in table 1.9, is also not statistically different from zero. In both countries the effect of the tax rate on contribution status is positive, although it is much smaller in Canada. As emphasized earlier, these results are very sensitive to model specifications and it may not be possible to distinguish the effect of the marginal tax rate from a nonlinear effect of income.[14]

The parameter estimates from a more highly parameterized model for Canada are shown in table 1.15. The variable "employee w/RPP" indicates individuals in Canada with a private pension plan. Neither

Table 1.15 RRSP Contribution Parameter Estimates, by Method of Estimation, Grouped, 1981[a]

Group and Variable	2-Limit Tobit	Probit, σ_ϵ from 2-Limit Tobit	1-Limit Tobit	Probit, σ_ϵ from 1-Limit Tobit
σ_η	.113(.004)	.105(.009)	.080(.082)	.088
σ_ϵ	2978(125)	—	2523(227)	—
Constant	−8918(371)	−8702(300)	−6527(1209)	−7390(253)
Income	.797(.005)	.790(.006)	.812(.082)	.776(.006)
MTR[b]	33(5)	34(5)	16(13)	29(4)
Age	60(6)	59(6)	42(14)	50(5)
Sex	−45(138)	24(139)	−478(385)	19(117)
Govt. Employee	−693(220)	−796(219)	322(602)	−672(185)
Employee w/RPP	−33(171)	81(170)	−721(403)	68(144)
Self-Employed	31(163)	−101(164)	724(405)	−86(139)
Professional	3995(255)	3133(303)	5312(710)	2644(256)
Farmer/Fisherman	−347(255)	−582(268)	2067(641)	−490(226)
LF[c]	−17566.7	−4996.0	−12513.9	−4995.3
N	11019	11019	3169	11019
< 0	7850	7856	0	7856
$> 0, < L$	1339	3169	1339	3169
$= L$	1830		1830	

[a]Standard errors are in parentheses.
[b]Marginal tax rate.
[c]Likelihood function.

estimate is statistically different from zero, although the one-limit estimate is quite negative. In the United States, there appears to be no relationship between pension coverage and IRA contributions based on individual data, although there is some evidence of a positive pension relationship in the family data.

Because data in Canada are available for several consecutive years, it is possible to check the validity of the model specifications for that country. Between 1976 and 1981, the Canadian Consumer Price Index increased by about 60 percent, but RRSP limits did not change over the period. Thus in real terms the limits declined very substantially between those years. Thus a good external check of the predictive validity of the model is to use estimates for one of the years to predict contributions in the other, when the limit was either considerably higher or much lower. Such predictions, using the two-limit and the one-limit estimates reported in table 1.15, are shown in table 1.16. In general, the predicted values are very close to the actual ones. For example, one-limit estimates for 1981 underpredict 1976 contributions by only 1.7 percent, and one-limit estimates for 1976 underpredict 1981 contributions by only 1.1 percent. Estimates based on the two-

Table 1.16 **Predicted Total Contributions for 1976 Based on 1981 Estimates and Predicted Total Contributions for 1981 Based on 1976 Estimates, by Estimation Method, Using Estimation Files[a]**

For 1976 Based on 1981 Estimates		
Actual:	2149	
Predicted:		
		Difference
Two-Limit	1920	−10.7%
One-Limit	2133	−1.7%
For 1981 Based on 1976 Estimates		
Actual:	4810	
Predicted:		
		Difference
Two-Limit	4940	+2.7%
One-Limit	4754	−1.1%

[a]Estimates are based on parameter estimates in table 1.15.

limit model also yield predicted values very close to actual values. Since the parameter estimates in the two countries are rather close, this suggests that the model also should predict rather well in the United States.

1.5 Conclusions

Persons with low incomes are unlikely to have IRA accounts. In addition, after controlling for income, age, and other variables, persons without private pension plans are no more likely than those with them to contribute to an IRA. Indeed, if anything, those with private plans contribute more than those without them. Both contributor status and the amount of positive contributions are determined in large part by income and, to a lesser extent, by demographic characteristics. The marginal tax rate may have a positive effect on whether one contributes, but it does not appear to influence the contribution amount. Results based on different specifications suggest that it may be difficult to distinguish the effect of the marginal tax rate from a nonlinear income effect. Simulations based on the estimates suggest that the current U.S. Treasury Department proposal would lead to about a 30 percent increase in IRA contributions.[15] Model estimates based on Canadian data for RRSPs are very similar to those for the United States. External checks of the predictive validity of the model for Canada indicate that predictions of the effects of limit changes are quite accurate.

Appendix A
Estimation of IRA Contributions

To estimate IRA contributions, there are two possible methods. The first is to assume that one underlying behavioral relationship leads to all of the observed outcomes. In this case there are three ways to estimate the same parameters, and the difference between the estimates can serve as a basis for a test of the assumption that one behavioral relationship is sufficient. If it is not, the second method is to assume that observed behavior results from two behavioral relationships, one pertaining to the decision to be a contributor and the other describing the desired amount of contributions, were one a contributor. These two methods are described in turn. The first is familiar to many readers, and the goal here is simply to make clear that, under the maintained hypothesis, the three approaches all yield estimates of the same parameters. The second method is not as familiar but is a generalization of a similar procedure in Deaton and Irish (1984).

1. A Single Behavioral Relationship

Assume the following notation:

$s \equiv$ Observed contribution,
$S \equiv$ Latent contribution "propensity,"
$X \equiv$ Vector of individual attributes,
$L \equiv$ Contribution limit,
$\epsilon \equiv$ Random disturbance term,
$i \equiv$ Individual indexes.

Latent contributions are specified as

$$S_i = X_i\beta_1 + \epsilon_i,$$

where ϵ_i is assumed to be distributed normal with mean zero and variance σ^2. Precise amounts contributed by each individual to an IRA are not reported. Instead, we know if the individual did or did not contribute and, if a contribution was made, the interval in which the contribution falls. Details on the intervals are presented in appendix B. The intervals[16] may be summarized as

$$s_i \leq L_0,$$

$$L_{t-1} < s_i < L_t, \qquad t = 1, T,$$

$$L_T \leq s_i,$$

where, in the individual data $T = 3$ and $L_0 = \$0$, $L_1 = \$500$, $L_2 = \$1000$, and $L_3 = \$2000$. In the family data $T = 12$.

Two-Limit Tobit

When all of the observations are used, there are $T + 2$ possible outcomes—contributions at zero, contributions within each of the T closed intervals, and contributions at the upper limit. These outcomes and associated likelihoods [17] are

(1) (i) $s_i = 0$, $\Phi\left[\frac{-X_i\beta}{\sigma}\right]$;

(ii) $L_{t-1} < s_i < L_t$, $\Phi\left(\frac{L_t - X_i\beta}{\sigma}\right) - \Phi\left(\frac{L_{t-1} - X_i\beta}{\sigma}\right), t = 1, T$;

(iii) $s_i = L_T$, $1 - \Phi\left[\frac{L_T - X_i\beta}{\sigma}\right]$;

where $\Phi[\cdot]$ denotes the standard normal distribution function.

One-Limit Tobit with Zeros Excluded

When the zero values are excluded, there are $T + 1$ possible outcomes: the contribution lies within one of the T intervals or at the upper limit. These outcomes and associated likelihoods are

(2) (i) $L_{t-1} < s_i < L_t$, $\left\{\Phi\left(\frac{L_t - X_i\beta}{\sigma}\right) - \Phi\left(\frac{L_{t-1} - X_i\beta}{\sigma}\right)\right\}/\Phi\left[\frac{X_i\beta}{\sigma}\right]$, $t = 1, T$;

(ii) $s_i = L_T$, $\left\{1 - \Phi\left[\frac{L_T - X_i\beta}{\sigma}\right]\right\}/\Phi\left[\frac{X_i\beta}{\sigma}\right]$.

The denominator in each expression, the probability of a positive contribution, reflects the fact that noncontributors have been excluded from the analysis. In a single-equation model, the underlying distribution of contributions is truncated at zero when the one-limit Tobit specification is used, while there is a mass point at zero when the two-limit version is used.

Simple Probit

Finally, estimates can also be obtained with a simple probit specification (up to the scale factor σ). In this case the outcomes and associated likelihoods are

(3) (i) $s_i = 0$, $\Phi\left[\frac{-X_i\beta}{\sigma}\right]$;

(ii) $s_i > 0$, $\Phi\left[\frac{X_i\beta}{\sigma}\right]$.

Likelihood Function

If there are N_0 observations at zero, N_t observations in interval t, and N_{T+1} observations at the upper limit, the log- likelihood function in each of the three cases is:

(A)
$$\ln L = \sum^{N_0} \ln\Phi\left[\frac{-X_i\beta}{\sigma}\right] + \sum_{t=1}^{T}\left\{\sum^{N_t}\ln\left[\Phi\left(\frac{L_t - X_i\beta}{\sigma}\right) - \Phi\left(\frac{L_{t-1} - X_i\beta}{\sigma}\right)\right]\right\} + \sum^{N_{T+1}}\ln\left[1 - \Phi\left(\frac{L_T - X_i\beta}{\sigma}\right)\right].$$

(B)
$$\ln L = \sum_{t=1}^{T}\left(\sum^{N_t}\ln\left\{\left[\Phi\left(\frac{L_t - X_i\beta}{\sigma}\right) - \Phi\left(\frac{L_{t-1} - X_i\beta}{\sigma}\right)\right]/\Phi\left(\frac{X_i\beta}{\sigma}\right)\right\}\right) + \sum^{N_{T+1}}\ln\left\{\left[1 - \Phi\left(\frac{L_T - X_i\beta}{\sigma}\right)\right]/\Phi\left(\frac{X_i\beta}{\sigma}\right)\right\}.$$

(C)
$$\ln L = \sum^{N_0}\ln\Phi\left[\frac{-X_i\beta}{\sigma}\right] + \sum_{t=1}^{T+1}\left[\sum^{N_t}\ln\Phi\left(\frac{X_i\beta}{\sigma}\right)\right].$$

It is clear that (B) + (C) = (A) under the one-equation assumption. If this assumption is inconsistent with the data, the β estimated from (B) will differ from the β estimated from (C). In addition, to the extent that they differ, the sum of the likelihood values from (B) and (C) will be greater than the value from (A) since the "separated" models allow a better fit to the data. Thus a test of the one-equation behavioral assumption can be based either on a comparison of the estimated βs or on the likelihood values. If they differ, a specification with two behavioral equations may be indicated.

2. IRA Contributors and Contributions: A Two-Equation Model

The purpose of this section is to describe a procedure that can be used to relax the one-equation constraint. It is assumed that two behavioral relationships determine the contributions that we observe. One is a relationship between individual attributes and the likelihood that a person is a potential IRA contributor. The other is a relationship between individual attributes and the level of desired contributions, were one to contribute. Of course, both of the outcome variables should be thought of initially as latent variables. In particular, if the latent contribution variable is less than zero, we shall assume that we observe no contribution, even if the contributor latent variable is greater than zero. A desirable property of the model is that it encompasses as a limiting case the standard Tobit model described in section 1 of this appendix.

The model is described by:

$$C = X\beta + \epsilon ,$$

$$S = X\alpha + \eta ,$$

where C is the latent contribu*tion* variable and S the latent contribu*tor* variable, X is a vector of individual attributes, β and α are vectors of parameters to be estimated, and ϵ and η are disturbance terms.[18] We assume that, given X, C and S obey the covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma^2 & \rho\sigma \\ & 1 \end{bmatrix} ,$$

where ρ is the correlation between C and S, given X.

The bivariate distribution between S and C is represented graphically in figure 1.3. The figure includes the limit on IRA contributions L. That is, as usual, we will not observe contributions above L but will observe a concentration at the level L. We shall assume that a person is a contributor if the latent contributor variable S is greater than zero, *and* if the desired contribution amount is greater than zero. Thus when S and C are both greater than zero we observe IRA contributions greater than zero but less than or equal to L.

As in the usual Tobit case, there are three observable outcomes: IRA contributions are zero, contributions are at some level C where C is

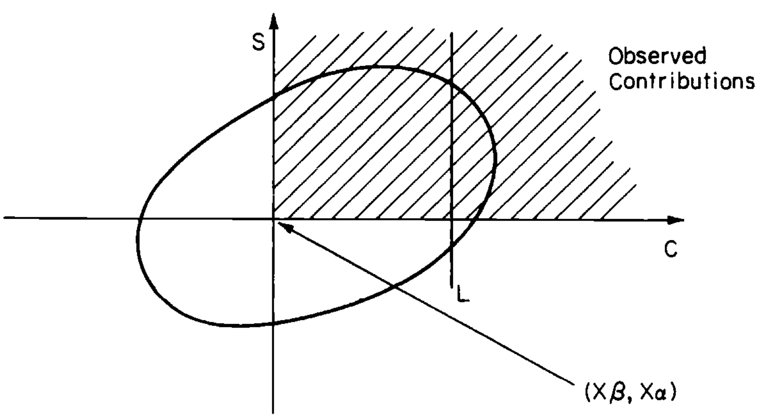

Fig. 1.3 The bivariate distribution between S, the latent contributor variable, and C, the latent contribution variable.

greater than zero but less than L, or we observe C at the limit L. The likelihoods associated with these three outcomes are now

Outcome	Likelihood
$C = 0$	$\Pr[S < 0] + \Pr[S > 0 \text{ and } C < 0]$
$L_{t-1} < C < L_t$	$\Pr[S > 0 \text{ and } L_{t-1} < C < L_t],\ t = 1, T$
$C = L_T$	$\Pr[S > 0 \text{ and } C > L]$

They are described in somewhat more detail by

Outcome	Likelihood
$C = 0$	$\Phi[-X\alpha] + \Phi_2\left[X\alpha, \frac{-X\beta}{\sigma}; -\rho\right]$
$L_{t-1} < C < L_t$	$\Phi_2\left[X\alpha, \frac{L_t - X\beta}{\sigma}; -\rho\right] - \Phi_2\left[X\alpha, \frac{L_{t-1} - X\beta}{\sigma}; -\rho\right],\ t = 1, T$
$C = L_T$	$\Phi_2\left[X\alpha, \frac{-(L_T - X\beta)}{\sigma}; \rho\right]$

where $\Phi_2[\cdot]$ indicates the bivariate normal distribution function.

If indeed C and S are the same underlying stochastic variable, as in the Tobit case, β goes to α and ρ goes to 1.[19] Thus in this case, the two-equation description of IRA contributions reduces to the Tobit specification. By comparing the likelihood values in the two models, one can test explicitly whether the single behavioral equation version can be rejected. The difference between this test and those mentioned in section 1 of this appendix is that the two equations are allowed to be correlated.

Appendix B
U.S. Data Sources

All data for the United States are from the May 1983 Current Population Survey and Supplemental Survey of Pension and Retirement Plan Coverage. Two data sets were created: individual and family.

Individual Data

The CPS data are arranged by individual. The sample used includes all individuals meeting the following criteria:

1. Included in the supplement (working for pay).
2. Between the ages 16 and 65.
3. Not self-employed.
4. Containing valid responses for each of the variables used in table 1.7.

All summary tables and simulations (but not the estimated models) using individual data use the CPS weights designed to represent the nation as a whole. No adjustment to these weights was made for exclusion of observations due to invalid responses.

Several problems arose with the way the IRA variables were coded. Employed persons were asked, "Do you have an IRA?" Those answering in the affirmative were then asked, "Approximately how much of your own IRA did you credit to your 1982 Federal taxes?" Responses are coded in the intervals:

under $100
$100–$499
$500–$999
$1000–$1999
$2000–$2499
$2500

This categorization led to two problems. First, a surprisingly large number of persons reported IRA contributions in the first (under $100) category. Most of these responses probably indicate persons establishing IRAs prior to 1982 and making no contribution in 1982. We have thus interpreted the first category to indicate zero contributions in 1982.

Second, a small number of respondents (186) indicated a contribution exceeding $2500. These responses presumably reflect family rather than individual contributions. These observations have been deleted from our sample.

Family Data

For tax status the family is a more appropriate unit than the individual in the United States. Using relationship codes, ages, and marital status we have converted the CPS data to a family basis. The incidence of unclassifiable persons or otherwise inconsistent units was rather high. In such cases, the observations were deleted from the sample.[20] One consequence of this data conversion is that using the CPS weights is no longer appropriate.

There are two important advantages to forming a family-based sample. The first is that the employment status of husband and wife in two-person families can be determined. This permits unambiguous assignment of contribution limits used to simulate policy changes.

The second advantage is that marginal tax rates can be calculated based on the family information. Our calculations are based on U.S. Internal Revenue Service (1984). This source reports average adjustments and deductions by income category. The first step is to convert each family's reported total income to adjusted gross income by accounting for average adjustments (excluding IRAs and Keoghs) by income class. To obtain taxable income, personal exemptions ($1000 each for self, spouse if married, and each child) and the average itemized home mortgage interest deduction (in excess of the standard deduction if one were not to itemize) for families reporting owning a home are subtracted from gross income. Finally, 1982 tax tables by filing status provide the marginal tax rates assigned to each family. These calculated rates span the entire range from zero to 50 percent.

Appendix C
Summary Statistics

Variable	Individual Data		Family Data	
	All	Contributors Only	All	Contributors Only
Total Individual (Family) Income ($)	17403(12214)	26649(16338)	23399(15717)	38709(18985)
Marg. Tax Rate (%)	—	—	24(11)	34(10)
Age[a] (years)	37(12)	46(11)	39(13)	46(11)
Unmarried Women (%)	.17(.38)	.10(.30)	.22(.41)	.14(.35)
Unmarried Men (%)	.15(.36)	.08(.26)	.17(.37)	.10(.30)
Married Women (%)	.28(.45)	.32(.47)	—	—
Education[a] (years)	13(2.7)	14(2.6)	13(2.9)	14(2.7)
Private Pension[b] (%)	.51(.50)	.67(.47)	.53(.50)	.78(.41)
Salary Reduc. Plan[b] (%)	.03(.18)	.07(.26)	.04(.20)	.10(.29)

[a]In the family data the value for this variable pertains to the CPS reference person in the household.

[b]In the family data the value for this variable is one of either member participates; zero otherwise.

Notes

1. In general, under the Canadian plan persons can contribute 20 percent of their income up to a maximum of \$3500 for those with a private pension plan and up to \$5500 for those without a private plan.

2. This may be contrasted with portfolio composition analysis most recently represented in the work by King and Leape (1984) or earlier work by Feldstein (1976), for example.

3. Given current income, contributions to tax-deferred accounts could of course be taken partially or entirely from other existing asset balances. The identification problem is to determine whether this is the case. It is not possible to address this issue with the CPS data, but it will be considered in subsequent analysis based on other data sources.

4. As far as we know, a model like this one has not been estimated. In any event, data limitations and other choices in the empirical implementation would undoubtedly leave uncertainty about the effect of tax-deferred accounts, as well as the effect on tax-deferred contributions of changes in the contribution limits.

5. Answers to these questions are important in their own right and can be answered with considerable confidence.

6. By referring back to figure 1.2, one can see the difference between these procedures. If all the data are used, then there is a concentration of observations at zero and at L. If the zero observations are deleted, there is no concentration of data points at zero, but the distribution is truncated at this point and the concentration at L remains. The third procedure only considers whether contributions are zero or not.

7. The preference function that corresponds to equation (7) is

$$V(Y - S, S) = (S/\beta)e^{[a+\beta(Y-S)]/S},$$

where $a = a_0 + a_1x_1 + \ldots a_kx_k + \epsilon$, and $\beta = Y^b + \eta$.

8. The estimates are in fact negative, but not statistically different from zero.

9. Given program limits, the estimated parameters, and values of X for each member of the sample, the expected contribution of each individual or family is (following the notation of appendix A):

$$\begin{aligned} E(c) &= \Pr[C \le 0] \cdot E[C \mid C \le 0] \\ &+ \Pr[0 < C < L] \cdot E[C \mid 0 < C < \mathrm{L}] \\ &+ \Pr[C \ge L] \cdot E[C \mid C \ge L] \\ &= (X\beta) \cdot \left[\Phi\left(\frac{L - X\beta}{\sigma}\right) - \Phi\left(\frac{-X\beta}{\sigma}\right)\right] \\ &+ \sigma\left[\Phi\left(\frac{L - X\beta}{\sigma}\right) - \Phi\left(\frac{-X\beta}{\sigma}\right)\right] + \left[1 - \Phi\left(\frac{L - X\beta}{\sigma}\right)\right] \cdot L \end{aligned}$$

10. This estimate may be compared to the average IRA deduction based on IRS wage and salary returns which was \$340 for 1982.

11. This lack of information poses a problem for prediction but not for estimation. The statistical model used to obtain parameter estimates assigns all married persons an open-ended upper limit of \$2000 or more.

12. The implied correlation between $(\eta_c Y + \epsilon_c)$ and $(\eta_s Y + \epsilon_s)$ is about -0.01.

13. It may be noticed that the sum of the individual likelihood functions is 9490.5, whereas the likelihood value for the joint specification is somewhat higher, 9497.8. The higher likelihood value results from the equal variance restrictions on η and the exclusion of education from the contribution equation.

14. Indeed, for some years the estimated effect in Canada is in fact negative, although not statistically different from zero.

15. Strictly speaking, the simulations indicate that had the Treasury proposal been implemented in 1982, contributions would have been 30 percent higher than they were.

16. In principle, the open intervals can be treated as closed intervals by setting limits of $-\infty$ or ∞. We treat open and closed intervals separately for expositional purposes only.

17. To simplify matters this appendix derives likelihood functions for a linear specification of S_i and a homoscedastic error structure. The estimated model is based on the parameter and error structure given by equation (5) in the text.

18. In practice, this model is also estimated with a nonlinear specification, with a heteroscedastic error structure, and with equations analogous to those in the text.

19. First consider the $C = 0$ case. Under the limiting case, the probability of S greater than zero and C less than zero goes to zero because this would be an outcome with zero likelihood. Thus, the bivariate distribution function drops out. The $L_{t-1} < C < L_t$ and $C = L_T$ cases can be rewritten as $\Pr[S > 0|L_{t-1} < C < L_t] \cdot \Pr[L_{t-1} < C < L_t]$ and $\Pr[S > 0|C > L_T] \cdot \Pr[C > L_T]$, respectively. The first term in each case goes to unity in the limiting case, since if C is greater than some positive L_t, S must also be greater than zero. Thus in both cases the bivariate distribution in the last term of likelihoods reduces to expressions containing only univariate cumulative distributions.

20. For example, we eliminated persons married but living in a single-person household. Heads living with other relatives but not married were treated as single persons.

References

Deaton, Angus, and Margaret Irish. 1984. Statistical models for zero expenditures in household budgets. *Journal of Public Economics* 23:59–80.

Feldstein, Martin S. 1976. Personal taxation and portfolio composition: An econometric analysis. *Econometrica* 44:631–50.

King, Mervyn A., and Jonathan I. Leape. 1984. Wealth and portfolio composition: Theory and evidence. NBER Working Paper no. 1468, September. Cambridge, Mass.: National Bureau of Economic Research.

U.S. Department of Treasury. 1984. *Tax reform for fairness, simplicity, and economic growth*. Vol.2. Washington, D.C.: GPO.

U.S. Internal Revenue Service. 1984. *Statistics of income, individual income tax returns, 1982*. Washington, D.C.: GPO.

Wise, David A. 1985. Contributors and contributions to Registered Retirement Savings Plans. Paper prepared for the Tax Policy and Legislation Branch of the Canadian Department of Finance.

Comment Gary Burtless

Steven Venti and David Wise have written a very nice paper on the take-up of tax-preferred saving plans in the United States and Canada. The main objectives of the paper are stated at the outset:

- To analyze the effect of individual attributes on whether a person contributes to IRAs;
- To determine the effect of individual attributes on how much is contributed; and
- To simulate the effect of potential changes in contribution limits on the amount contributed to IRAs.

In addition, the authors very briefly consider the more interesting issue of how much IRAs and similar plans contribute to net personal saving, taking into account the fact that IRA contributions may simply substitute for other forms of saving.

The paper begins with a tabular presentation of some basic statistics on the take-up of IRAs in the United States and Registered Retirement Savings Plans (or RRSPs) in Canada. Both programs have a similar design. Wage earners are permitted to contribute designated amounts to qualified saving plans, and neither contributions nor the interest on contributions is subject to income tax until withdrawals begin during retirement. Under both the U.S. and Canadian plans, there is a maximum permissible amount that workers can contribute to tax-preferred accounts in a given year.

The authors' tabulations show, not surprisingly, that the probability of workers contributing to a plan is strongly correlated with their income. In the United States the percentage of workers contributing to an IRA is about 5 percent in the lowest income class (under $10,000 per year in 1982 earnings), but approaches 60 percent in the highest income class (earnings above $70,000). The percentages are quite similar in Canada, though there is evidence that Canadian taxpayers in the highest income class contribute somewhat less than taxpayers in the upper-middle-income classes. Nor surprisingly, the average contribution per worker—among workers making a contribution—tends to be closer to the permitted maximum in the higher income classes.

How much of the strong correlation between contributions and income is due to an income effect and how much is due to a price effect is difficult to say. The price of retirement consumption (or, equivalently, the price of preretirement saving) can be dramatically reduced by an IRA. Because wage-earners temporarily escape income taxes on both the contribution out of current earnings and the interest on the con-

Gary Burtless is a senior fellow at the Brookings Institution.

tribution, they can essentially buy a greater level of retirement consumption from a $1 increase in current saving. However, the amount of the price change is determined by workers' current and expected future income tax rates. The price reduction is thus determined by workers'marginal tax rates, which are in turn determined by their incomes.

One odd aspect of these tabulations is the finding that workers in jobs covered by private pensions are more rather than less likely to contribute to IRAs than are uncovered workers, holding earnings levels constant. Even when the authors use a formal statistical model to control for the effects of other factors, they never find evidence that uncovered workers are any more likely to contribute to IRAs than are covered workers. Using family data, in fact, they find that uncovered workers are *less* likely to contribute to IRAs (see tables 1.7 and 1.10).

Why this should be is difficult to explain; the authors do not attempt an explanation. Under a private pension plan the firm is saving on its workers' behalf. For at least a few of its workers the firm must be "oversaving" for their retirement. These workers would be expected to compensate for the oversaving by saving less outside of their pension plan. Hence, these workers should save less in the form of IRAs, which are in fact no more tax-preferred than are firms' contributions to private pension plans. By contrast, uncovered workers have no employer which saves on their behalf. One would expect them to do their own saving for retirement, and IRAs are the cheapest way for them to do so. Conceivably, workers with strong preferences for retirement saving sort themselves into pension-covered jobs. Their demand for retirement income is not satisfied by their employers' saving in pension plans, so they salt away additional savings in IRAs. This explanation, while somewhat plausible, is nonetheless surprising.

In the following section, Wise and Venti consider the estimation problem confronting them. Their discussion of the statistical issues is a model of lucidity. Essentially, there are three issues to be dealt with:

- How to model the decision to contribute to an IRA;
- How to estimate the demand function for contributions, given that some amount is going to be contributed; and
- How to deal with the limits on contributions set by current law.

Although the authors discuss three or four solutions, and in fact estimate the IRA demand function using more than one of them, the doubly-truncated Tobit model seems to perform about as well as more elaborate alternatives. That is, the more elaborate alternatives do not yield meaningfully different estimates of participation rates in IRAs, demand for IRAs, or the number of persons at the maximum contribution level.

For those not familiar with the Tobit model, the idea is very simple. For any particular individual, an index of the desire to make contributions to an IRA can extend over the entire range of real numbers—including negative numbers (for workers wishing to reduce their IRA holdings) and high positive numbers (for workers wishing to contribute an amount in excess of the legal maximum). However, there is an upper and lower bound on the *observed* distribution of demands because workers are prevented from reporting withdrawals and are not permitted to make contributions in excess of the maximum. Assuming normality in the distribution of the error terms, it is straightforward to obtain maximum likelihood estimates of the parameters that predict both the *desired* and the *observed* demand for IRA contributions. The model here is somewhat more clever than the usual Tobit model because it properly treats the issue of heteroscedasticity.

Venti and Wise obtain maximum likelihood estimates for both U.S. and Canadian workers. The U.S. data are examined in two different ways. The authors consider individual-level data on contributions and then look at family-level data. The main difference between the two statistical specifications is that the marginal tax rate is excluded as an explanatory variable when the authors estimate individuals' demand for IRAs. I have no idea why this critical variable is excluded for individuals. As mentioned earlier, the marginal tax rate is the main determinant of the marginal price of retirement saving in terms of current consumption foregone. To estimate a demand function that includes only income but excludes price seems odd, but I presume this was motivated by some defect in the data set. It seems clear to me that the marginal tax rate is as well defined for an *individual* in a joint filing unit as it is for the full filing unit. (In fact, it may be better defined if two members face differing tax rates on marginal earnings, as they sometimes will.)

The estimates for both the United States and Canada appear plausible. In the United States, the results indicate an elasticity of contributions with respect to income of 0.63—which doesn't seem unreasonable to me. But this estimate is subject to the qualification that it may be capturing some of the effect of varying the *price* of IRA contributions. As I mentioned above, income and price will be highly correlated under a progressive income tax scheme. In the individual-level specification, Venti and Wise exclude the price (or its proxy—MTR), implying that much of the effect of the price variation must be captured by the income term. In the family-level specification, the price term is included, but is so highly correlated with income (and possibly mismeasured because of the imputation procedure) that its coefficient must be at least a bit suspect. I think this has implications for the simulation results.

Essentially, we have two variables—income and the price of saving—which are highly correlated. Both variables probably have an important and independent effect on the demand for and attractiveness of IRA savings. The statistical procedure must somehow divide up the explanation of overall variance into a part explained by income and a part explained by price. Since the two variables are so highly correlated, this may be difficult to do. The highly erratic estimates of the effect of MTR are apparent in tables 1.9 and 1.10. I'm not sure we can trust those estimates, and for that reason, I'm not sure the simulations (especially those for the new Treasury tax plan) are entirely trustworthy.

The authors, however, are extremely careful to show the correspondence between their findings (1) using different models; (2) for the United States and Canada; and (3) from cross-sectional and time-series analysis of Canadian data. Those comparisons add considerably to the believability of the results.

The simulations in the paper show the effect of raising the present limits on IRA contributions. Since limit changes are frequently proposed, I think the results are interesting. Although I know little about past IRA research or simulations, these simulation predictions seem quite reasonable.

In closing, I wish to consider whether the questions dealt with in the simulations lie at the heart of the IRA issue. The critical question about IRAs and other tax-preferred saving plans is whether they contribute to or subtract from net personal saving. Because they reduce the price of saving for retirement, many Congressmen blithely assume that IRAs must eventually raise the amount of private saving that is done. The more sophisticated members of Congress would agree that in the *short run* most of the contributions to IRAs may come out of non-tax-sheltered savings. But in the long run—so the argument goes—workers will salt away something extra because retirement saving is so cheap.

As economists we cannot be so sanguine. The issue is deeper than the average Congressman imagines. If the price of obtaining $1 of consumption during retirement is reduced by enough, workers may actually engage in less not more saving for retirement; less saving is needed to attain a target level of retirement consumption. In fact, if everyone contributes at the maximum level, the marginal price of saving has not even been reduced. The net effect of IRAs on personal saving is the critical issue economists must address. This paper briefly raises the subject, but because of data limitations is prevented from formally addressing it. What would be useful here is a clear exposition of this main issue and an explanation of how the results in this paper—or extensions of those results—contribute to our understanding of the issue.

This paper, however, generally ignores the issue of the price of retirement saving (or retirement consumption, however you wish to view it). It ignores the objectives of retirement saving both from the point of view of the individual utility maximizer and the typical member of Congress.[1] And it does not treat the problem of saving in the relevant framework of lifetime utility maximization within a lifetime budget constraint that is affected by nonlinear taxes.

Can this paper tell us anything about the fundamental question just mentioned? I think it can. If we look over the tabular results, we notice that only a small minority of wage earners are at or near the maximum contribution level for IRAs or Canadian RRSPs. Even at the highest income (and marginal tax rate) levels, only a bare majority of U.S. and Canadian men make contributions equal to the legal maximum. This implies that the marginal price (not just the average price) of retirement saving has unambiguously been reduced for virtually all wage earners. The IRA does not represent a windfall drop in taxes which has no net effect on the marginal price of saving. In the case of the United States, the reduction in marginal price was sharp and discontinuous. If IRAs on balance *encourage* retirement saving, by reducing its marginal price, we should see some immediate population response, discernible in the aggregate statistics on personal saving. In fact, however, the U.S. personal saving rate has fallen precipitously since 1982, when IRAs were first extended to all wage earners, and in 1986 stood at 3.9 percent—less than half the rate of the 1970s.[2] The net impact of IRAs on personal saving, if any, has evidently been small or swamped by other factors.

In sum, I like the empirical work very much, and I admire the exposition and statistical modeling. But I would be interested in learning the authors' views on the relative importance of the issues raised and formally treated in this paper. Without treating the question of the *net* impact of IRAs on personal saving can any research on IRAs hold broad interest for economists? I realize that the U.S. and Canadian tax authorities will be very interested in these findings. But will the findings interest most public finance economists?

1. If the goal of a typical Congressman is to raise the net *national* saving rate, IRAs are an even poorer instrument than suggested in the previous paragraph. Even if an IRA succeeds in raising the private saving rate, it will reduce the public saving rate, at least initially, because it must reduce government tax revenue. Unless the net private saving response is fairly substantial, the loss in government saving can easily exceed the rise in private saving, leading to a reduction in net national saving.

2. U.S. Council of Economic Advisors, *Economic Report of the President, 1987,* (Washington, D.C.: U.S. Government Printing Office) 1987, p. 274.