

ERIC M. ENGEN

Board of Governors of the Federal Reserve System

WILLIAM G. GALE

The Brookings Institution

JOHN KARL SCHOLZ

University of Wisconsin at Madison

Do Saving Incentives Work?

AMERICAN SAVING RATES have recently fallen to their lowest levels since 1950. After averaging roughly 8 percent in the 1950s, 1960s, and 1970s, the net national saving rate fell to about 4.5 percent in the 1980s and has fallen below 2 percent since 1990.¹ The personal saving rate has also declined, from an average of 7 percent between 1950 and 1980 to an average of 4.6 percent since 1990.² These declines have raised concerns that the economy may be unable to finance investment and sustain growth over the long run and that a significant fraction of the baby-boom generation may not be saving adequately for retirement.³

We thank Zhiqin Zhou for outstanding assistance, and Doug Bernheim, Jim Eisner, Bill Gentry, Jon Gruber, Robert Haveman, Andrea Kusko, Larry Ozanne, Jim Poterba, Andrew Samwick, Louise Sheiner, Dan Sichel, Jon Skinner, Joel Slemrod, Martha Starr-McCluer, Steve Venti, David Wise, and seminar participants at Duke, Wisconsin, and the Tax Economists' Forum for helpful comments. We also acknowledge financial support from the National Science Foundation through grants SES-9220794 and SES-9211553. The opinions expressed in this paper are those of the authors and not necessarily shared by the Board of Governors of the Federal Reserve System, members of its staff, or other institutions.

1. The data are taken from the *Economic Report of the President 1994*, tables B-23 and B-29, and the *Economic Report of the President 1991*, tables B-22 and B-28. The net national saving rate is gross saving less depreciation, all divided by net national product. Congressional Budget Office (1993a) provides an assessment of the saving decline.

2. The data are taken from the *Economic Report of the President 1994*, table B-27, and the *Economic Report of the President 1991*, table B-26. The personal saving rate is defined as personal saving divided by disposable personal income. Bosworth, Burtless, and Sabelhaus (1991) provide an analysis of the decline in personal saving.

3. See Bernheim (1992), Bernheim and Scholz (1993), and Congressional Budget Office (1993a, 1993b).

As a consequence, raising the saving rate has been a frequent focus of policymakers and academics. Over the past 20 years, a popular legislative remedy for stimulating saving has been the development of specially designated personal saving accounts, such as individual retirement accounts (IRAs), 401(k) plans, and Keogh accounts. These voluntary accounts, which we refer to as saving incentive plans, feature preferential tax treatment of contributions and investment earnings, annual contribution limits, and penalties for early withdrawals. The programs have been popular: since 1986, contributions to these saving incentive plans have amounted to about one-third of personal saving as measured by the National Income and Product Accounts.⁴ Yet, despite their popularity, saving incentives may not work—they may not raise national saving.

The effect of saving incentives on national saving is the sum of their effects on public and private saving. To raise private saving, an incentive must reduce consumption or raise labor supply.⁵ Standard economic theory indicates that raising the after-tax rate of return on all private saving has an ambiguous effect because of opposing income and substitution effects. Saving incentive programs that raise the after-tax return on limited amounts of assets placed in designated accounts may be an even less effective way to stimulate private saving than increasing the rate of return on all saving. This is because a household does not need to reduce consumption or raise labor supply to claim the tax advantages of a saving incentive. Instead, the contributions may be financed by transferring existing taxable assets, by increasing debt, or by reallocating current saving that would have been done anyway. Moreover, because the subsidies in saving incentive plans are capped by the annual contribution limit, the saving incentive will not affect marginal returns on saving for some households. Rather, for those households, the incentive will generate only an income effect, which should reduce their overall level of saving.

The effect of saving incentives on public saving, holding other government policies constant, is generally negative in the short run because contributions are typically tax deductible. Thus, to raise national saving in the absence of other policy changes, a saving incentive plan must raise private saving by more than the associated tax loss to the government.

4. See table 1 and *Economic Report of the President 1994*, table B-27.

5. Kotlikoff (1990).

In this paper, we examine the effects of saving incentives on private and national saving using both empirical and simulation analyses. In the next section, we describe and analyze the main features of current saving incentive plans and review trends in aggregate saving and the growth of saving incentive plans. Then, we present new evidence on the effects of 401(k) plans. Our results are consistent with the view that 401(k) plans have not stimulated private saving or wealth. We also present evidence that IRAs, 401(k) plans, and other retirement saving plans are, to some extent, substitutes in households' portfolios. In the third section, we present new evidence on the effects of IRAs. The evidence suggests that a small proportion of IRA contributions has represented new private saving. However, coupled with the revenue losses created by tax-deductible contributions, our results indicate that saving incentives have had little, if any, positive effect on national saving. In the fourth section, we analyze saving incentives in the context of a stochastic, life-cycle simulation model of saving. Our simulation results are consistent with the empirical findings for the short and medium terms (0–30 years). They also suggest that over longer horizons, saving incentives can raise the national saving rate.

An Overview of Saving Incentives

A wide variety of saving incentive options are currently available. Individual retirement accounts were established in 1974 for workers without pensions and featured tax-deductible contributions up to an annual limit, tax-free accrual of interest, and penalties for early withdrawal of funds. In 1981, eligibility was extended to all workers and the limits were raised to their current levels. For a single worker, the annual limit is the lesser of either earnings or \$2,000. The presence of a nonworking spouse raises the limit by \$250. In a two-worker household, the limit is the lesser of earnings or \$2,000 for each worker, so the maximum combined limit is \$4,000.

From 1982 to 1986, eligibility was universal and contributions were tax deductible. The Tax Reform Act of 1986 then restricted the deductibility provisions. For households with a retirement plan, deductibility is now phased out at adjusted gross income levels between \$40,000 and \$50,000 for joint filers and between \$25,000 and \$35,000 for single filers. The act did not restrict IRA eligibility or the tax-free accrual of interest.

Deductible contributions are taxed when withdrawn, and nondeductible contributions are not. In either case, returns on contributions are untaxed until withdrawn, at which point they are subject to income taxes. Withdrawals can be made at any time but are subject to an additional 10 percent penalty if the account holder is younger than 59.5.⁶

IRA contributions may be financed with loans, the interest on which is deductible as investment interest.⁷ However, IRA assets cannot be used as collateral for a loan. IRA funds can be invested in virtually any type of financial asset but not directly in physical assets.

Keogh plans and simplified employee pensions (SEPs) are similar to tax-deductible IRAs, except that they apply to the self-employed and typically have higher contribution limits.⁸

Established in 1978, 401(k) plans are employment-based saving incentives. Important clarifying regulations were issued in 1981.⁹ Like IRAs, 401(k) plans feature tax-deductible contributions, tax-free accrual of interest, annual contribution limits, and restrictions on withdrawals. However, 401(k) plans differ from IRAs in several important respects. First, 401(k) plans are available only to employees of firms or other organizations that choose to sponsor the plans.¹⁰ Second, 401(k) contributions are made through regular payroll deductions, while IRA contributions can be made on an irregular basis. Third, employers may make tax-deductible contributions into the plan.¹¹ These contributions may be unconditional or may depend on employee contributions. Most firms with 401(k)s provide employer contributions, often matching employee contributions at rates ranging from 25 percent to 100 percent, up

6. The penalty does not apply in the event of permanent disability or death. Individuals may initiate penalty-free annuity distributions from IRAs at any age.

7. Restrictions against deducting interest on loans used to make tax-exempt investments (such as municipal bonds) do not apply because the return on IRAs is tax deferred, rather than tax exempt.

8. Keoghs and SEPs may also be set up for employees of small businesses.

9. Andrews (1992).

10. Strictly speaking, 401(k) plans, as authorized by section 401(k) of the Internal Revenue Code, are only available to employees of for-profit firms that offer the plans. Sections 403(b) and 457 authorize salary-reduction plans for employees of nonprofit institutions and state and local governments, respectively. Federal government employees may participate in thrift saving plans. These plans are similar to 401(k) plans in most respects. Throughout this paper, we refer to all of these as 401(k) plans.

11. 401(k) plans differ from other defined contribution pension plans in that the employee's contribution is tax deductible.

to 5 or 6 percent of salary.¹² Fourth, annual contribution limits to 401(k) plans are governed by several factors. The Internal Revenue Service (IRS) limit on employee contributions is \$9,240 in 1994.¹³ However, employers often impose a limit of 10 to 15 percent of salary.¹⁴ More generally, under the Employee Retirement Income Security Act the combined employer and employee contributions, summing 401(k) plans and other defined contribution plans, cannot exceed the lesser of 25 percent of salary or \$30,000. Highly compensated employees may face even tighter limits required by complicated nondiscrimination requirements.¹⁵ Fifth, investment options for 401(k)s are determined by the employer and hence are typically more limited than with an IRA. The assets in 401(k)s often must be allocated among the employer's stock, a guaranteed investment contract, or specific equity, bond, or money-market mutual funds. Sixth, borrowing from a 401(k) is allowed in some plans.¹⁶ The loan is generally secured by funds in the account.¹⁷ Seventh, withdrawal provisions differ from those for IRAs. In some plans, penalty-free withdrawals are permitted in cases of documented financial hardship.¹⁸ Withdrawals upon employee separation from a firm face a 10 percent penalty (plus income tax) unless the funds are rolled over into an-

12. In survey results reported in U.S. General Accounting Office (1988a), 38 percent of plans provided only matching contributions, 24 percent provided only unconditional (or discretionary) contributions, 13 percent provided both, and 25 percent provided neither.

13. In 1982, the limit was \$45,475, the same as for all employer-based defined contribution pension plans. The limit was reduced to \$30,000 starting in 1983. In the 1986 tax reform, the limit was reduced to \$7,000 (1987 dollars) and indexed for inflation (Andrews, 1992).

14. About two-thirds of employers have employee contribution limits between 9 and 16 percent (U.S. General Accounting Office, 1988a). Kusko, Poterba, and Wilcox (1994) analyze 401(k) participation and contribution behavior in a single firm with a limit of 10 percent of salary. In 1990, 23 percent of participants were constrained by this limit, while only 1 percent were constrained by the IRS limit.

15. In 1994, the IRS defined highly compensated employees as those with 1993 salaries above \$64,245. Some company owners and officers may also be considered highly compensated, regardless of salary. Papke (1992) provides an excellent discussion of the nondiscrimination provisions.

16. Survey evidence suggests that borrowing is permitted in two-thirds or more of 401(k) plans (U.S. General Accounting Office, 1988a; Buck Consultants, 1989; Hewitt Associates, 1993).

17. If the loan is secured by elective 401(k) contributions, the interest paid is not tax deductible. However, if the employer permits, the loan may be secured by an employee's house, in which case the interest paid is tax deductible, as with a home equity loan.

18. U.S. General Accounting Office (1988a, 1988b).

other 401(k) or IRA or converted to an annuity, or unless the account holder is over age 59.5, retired and over age 55, or disabled. Otherwise, withdrawals are not allowed.

Some Simple Analytics of Saving Incentive Plans

Tax-free accrual of interest can create significantly higher rates of return. The tax deductibility of contributions to a saving incentive account may also provide a tax advantage. Let r be the interest rate, τ_0 the tax rate at time period 0, τ_1 the tax rate that applies between time period 0 and time period T , and τ_T the tax rate that applies to withdrawals made at time T . The after-tax balance ($B1$) in period T per dollar of pretax income placed in a conventional saving account in time 0 is

$$(1) \quad B1 = (1 - \tau_0) [1 + r(1 - \tau_1)]^T.$$

In a front-loaded saving incentive plan, such as 401(k)s, the contribution is tax deductible, but the contribution and any investment earnings are taxed when withdrawn. The after-tax balance ($B2$) in period T per dollar of pretax income contributed to a front-loaded saving incentive in time 0 is¹⁹

$$(2) \quad B2 = (1 - \tau_T) (1 + r)^T.$$

The difference D equals ($B2 - B1$) and is the increased private return to a saver from a front-loaded saving incentive relative to a conventional saving account, per dollar of pretax income. If the amount placed in the saving incentive account would have been saved even in the absence of this tax-preferred plan, D is also forgone revenue to the government. D is positive (except when all tax rates are zero or all tax rates are one), rises with r and T , and, as shown in figure 1, is a concave function of τ .²⁰

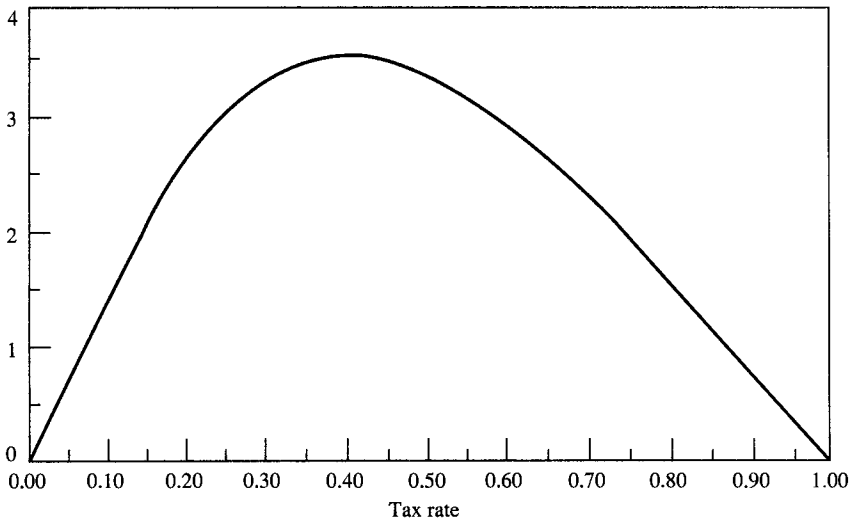
In a back-loaded saving incentive plan, the contribution is not tax deductible, but neither the original contribution nor the earnings are taxed on withdrawal. The after-tax balance ($B3$) in period T per dollar of pretax income contributed to a back-loaded saving incentive in time 0 is

$$(3) \quad B3 = (1 - \tau_0) (1 + r)^T.$$

19. For a 401(k) plan with an employer matching contribution rate of m , $B2 = (1 - \tau_T) (1 + m) (1 + r)^T$.

20. If the tax rate effective at the time of withdrawal from a saving incentive account is less than the tax rate in effect when the contribution is made, the tax savings would be even larger than in figure 1.

Figure 1. Return Arising from Tax-Free Accrual for Each Dollar Invested
Dollars



Source: Authors' calculations in which the return is calculated as $B2$ from equation 2 minus $B1$ from equation 1. The holding period, T , is set equal to 30, and the interest rate, r , is set equal to 0.08.

Referring to equations 2 and 3, if τ_0 equals τ_T , the after-tax balances of front-loaded and back-loaded plans are equal ($B3 = B2$), so the upfront tax deductibility of contributions confers no tax advantage.²¹

If contributions are made during working years when people face relatively high tax rates, and withdrawals are made after retirement when people face lower tax rates, there is a potential tax advantage to front-loaded plans of $\tau_0 - \tau_T$ per dollar contributed. Tax changes, such as the 1986 tax reform, that flatten the profile of marginal tax rates decrease this potential tax advantage. A steep lifetime income profile or policy changes that increase future tax rates could cause τ_T to exceed τ_0 and make upfront deductibility a liability rather than an advantage.²²

Since the Tax Reform Act of 1986, some higher-income families have not been allowed to make tax-deductible IRA contributions but can nevertheless make nondeductible contributions. These nondeductible IRAs are different from back-loaded IRAs since the return on contribu-

21. Even in this case, a taxpayer may still prefer the upfront tax deduction if uncertainty about future income creates a demand for precautionary saving (see Engen and Gale, 1993).

22. For further discussion of these issues, see Burman, Cordes, and Ozanne (1990).

tions (but not the original contribution) are taxed at the end of the holding period. The after-tax balance ($B4$) in period T per dollar of pretax income contributed to a nondeductible IRA in time 0 is

$$(4) \quad B4 = (1 - \tau_0)[(1 - \tau_T)(1 + r)^T + \tau_T].$$

The tax advantages of a nondeductible IRA are smaller than for front-loaded or back-loaded IRAs ($B4 < B3$, $B4 < B2$), as long as tax rates and interest rates are positive. However, a nondeductible IRA still provides a higher after-tax balance than a conventional saving account ($B4 > B1$).²³

Early withdrawal penalties make saving incentive funds illiquid. Account holders who plan to make early withdrawals may still wish to put funds in designated accounts because after a period of time the higher effective interest rate will more than offset the penalty. In a front-loaded IRA, the penalty is assessed on both principal and earnings. The after-tax after-penalty balance ($B5$) in period T per dollar of pretax income contributed to a front-loaded IRA in time 0 is

$$(5) \quad B5 = (1 - \tau_T - p)(1 + r)^T,$$

which is equivalent to $B2$ except it includes the penalty p .

The advantage of contributing to a saving incentive plan relative to conventional saving depends on the holding period. The holding period that makes the benefits of tax-free accrual outweigh the costs of early withdrawal rises with p , falls with r , and is convex in the tax rate. With p equal to 0.10, this holding period is about five years or more for a wide range of parameters.²⁴

23. For example, if $T = 30$, $r = 0.08$, and tax rates are constant at $\tau = 0.40$, then a dollar in pretax income will yield an after-tax balance of $B1 = \$2.45$ in a conventional saving account, $B2 = B3 = \$6.04$ in a front-loaded or back-loaded IRA, and $B4 = \$3.86$ in a nondeductible IRA.

24. For example, with a base case of $p = 0.1$, $\tau = 0.4$, and $r = 0.08$, the breakeven holding period, T^* , is 7 years. Raising r to 0.10 reduces T^* to 5 years, while reducing r to 0.05 raises T^* to 10 years. From the base case, raising the penalty to 0.20 raises T^* to 14 years. From the base case, reducing τ to 0.2 raises T^* to 9 years, but raising τ to 0.7 raises T^* to 8 years.

For a nondeductible IRA, the after-tax balance, if the early withdrawal penalty is incurred, is equal to $(1 - \tau_0)[(1 - \tau_T - p)(1 + r)^T + (\tau_T + p)]$. Note that the nondeductible contribution is not subject to either tax or penalty when withdrawn early, although the earnings are subject to both tax and penalty. With the base case of $p = 0.10$, $\tau = 0.4$ (taxes are constant), and $r = 0.08$, T^* is 12 years.

Matching contributions by employers in 401(k) plans are a part of total employee compensation. If employers adjust wages for the average amount of matching contributions (for the firm as a whole or within subgroups),²⁵ a worker that contributes the average amount has no income effect from the employer's matching. However, there is still a substitution effect since an additional dollar contributed would receive a matching contribution that is not completely offset by a reduction in wages. Above-average 401(k) contributors would be subsidized by below-average 401(k) contributors.

Borrowing to finance contributions to a saving incentive account is profitable if the after-tax return on the contribution exceeds the after-tax cost of borrowing.²⁶ If r_s and r_b are the returns on contributions and borrowing, debt-financed contributions are profitable when²⁷

$$(6) \quad (1 - \tau_T - p)(1 + r_s)^T > [1 + r_b(1 - \tau_1)]^T.$$

This condition is often difficult to achieve in practice because r_s is typically less than r_b .²⁸

However, the employer matching rate for 401(k) plans changes the condition for profitable debt-financed contributions to

$$(7) \quad (1 - \tau_T - p)(1 + m)(1 + r_s)^T > [1 + r_b(1 - \tau_1)]^T,$$

where m is the employer matching rate. If τ is 0.4 (tax rates are constant), p is 0.1, r_s is 0.10, T is 15, and the match rate is 0.5, debt financing is profitable with r_b as high as 13.1 percent. If the match rate is zero, as, for example, in an IRA, borrowing is only profitable if the borrowing rate is below 8.3 percent. These differences suggest that the possibility of borrowing to finance contributions is a more important issue for 401(k) plans than for other saving incentives.

25. Full adjustment of wages for the cost of fringe benefits is consistent with the evidence provided by Gruber and Krueger (1991) for workers' compensation, Gruber (1994) for mandated maternity benefits, and some of the evidence presented in Montgomery, Shaw, and Benedict (1992) for defined benefit pensions.

26. See Gale and Scholz (1994) and Kotlikoff (1990) for discussions of the possibility of borrowing to finance IRA contributions.

27. We assume that interest payments on the borrowed funds are tax deductible, as is the case, for example, with home equity loans.

28. With $r_s = 0.08$ and $r_b = 0.10$, constant tax rates, $\tau = 0.4$, and a penalty of $p = 0.10$, it would take a 38-year holding period for the tax advantages of an IRA to overcome the early withdrawal penalty and borrowing costs (assuming that the interest payments are tax deductible).

Table 1. Saving Incentive Contributions, 1980–90^a

Billions of nominal dollars, except as noted

<i>Year</i>	<i>IRA</i>	<i>Keogh and SEP</i>	<i>401(k)</i>	<i>Total contributions^b</i>	<i>Total contributions (percent of GDP)</i>
1980	3.4	2.0	...	5.4	0.20
1981	4.8	2.0	...	6.8	0.22
1982	28.3	2.5	3.9 ^c	34.7	1.10
1983	32.1	2.9	8.7 ^c	43.7	1.28
1984	35.4	4.1	16.3	55.7	1.48
1985	38.2	5.2	24.3	67.7	1.68
1986	37.8	6.2	29.2	73.2	1.71
1987	14.1	6.2	33.2	53.4	1.18
1988	11.9	6.6	39.4	57.9	1.18
1989	11.0	6.3	46.1	63.4	1.21
1990	9.9	6.8	49.0	65.7	1.19

Sources: U.S. Internal Revenue Service (various issues) and U.S. Pension and Welfare Benefits Administration (1993).

a. The data do not include rollovers or nondeductible contributions and do not account for withdrawals.

b. May not add because of rounding.

c. Authors' estimates.

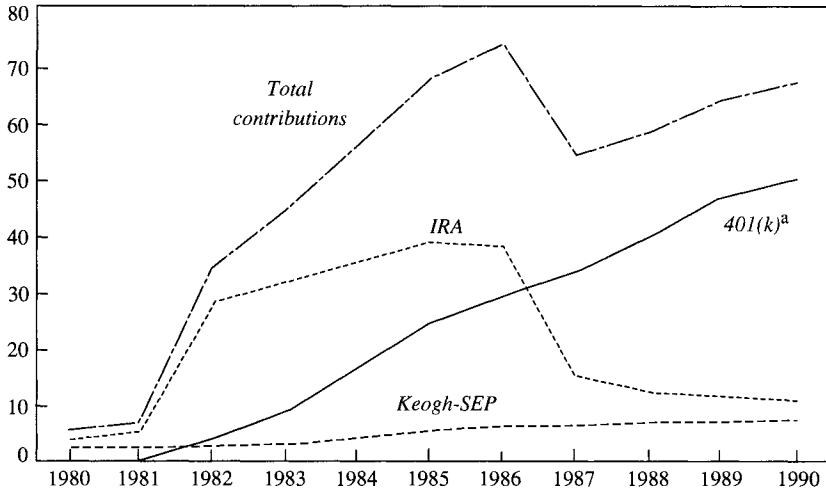
Saving Incentives and Aggregate Measures of Personal Saving

Table 1 and figure 2 show the evolution of IRA, 401(k), and Keogh contributions from 1980 to 1990.²⁹ Contributions to 401(k)s grew throughout the decade and totaled \$49 billion in 1990, or 0.9 percent of gross domestic product. IRA contributions grew dramatically following universal eligibility in 1982 and stayed high through 1986. In 1987, the restricted deductibility and other features of the 1986 tax reform took effect. Deductible IRA contributions fell 63 percent in 1987 and continued to fall thereafter. By 1990, nominal IRA contributions were 26 percent of their 1986 value. Keogh and SEP contributions have been much smaller, but tripled from 1980 to 1986 and then rose slowly through 1990. Overall, nominal contributions to all these accounts grew from about \$5 billion in 1980 to \$73 billion in 1986, fell to \$53 billion in 1987, and rose again to about \$66 billion in 1990. Contributions rose from 0.2 percent of GDP in 1980 to 1.7 percent in 1986. They fell to 1.2 percent of GDP in 1987 and remained roughly at that level through 1990.

29. The data do not include rollovers and nondeductible contributions and do not account for withdrawals. Data on 401(k) contributions were unavailable beyond 1990 at the time this paper was written.

Figure 2. 401(k), IRA, Keogh, and SEP Contributions, 1980–90

Billions of nominal dollars



Sources: IRA and Keogh-SEP data are from U.S. Internal Revenue Service (various issues); 401(k) data for 1984–1990 are from U.S. Department of Labor (1993).

a. 401(k) data for 1982 and 1983 are estimated.

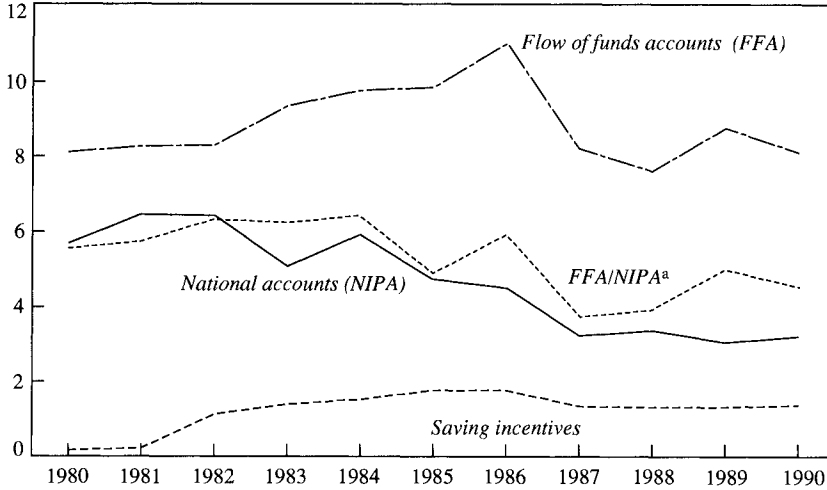
These patterns can be compared with the evolution of several measures of aggregate personal saving. The National Income and Product Accounts (NIPA) measure personal saving as the difference between personal disposable income and personal spending. The flow of funds accounts (FFA) of the Federal Reserve provide an alternative measure, which is the sum of the household sector's net acquisitions of financial assets and net investments in tangible assets less the net change in liabilities. A third measure is the flow of funds calculated on a national accounts basis (FFA-NIPA). This measure removes consumer durables, government insurance and pension reserves, and saving by corporate firms from the flow of funds measure. It thus employs the national accounts concept for saving but uses data from the flow of funds.³⁰

Figure 3 plots each of these measures and contributions to saving incentives as a percent of GDP from 1980 to 1990. Major shifts in the ratio of saving incentive contributions to GDP occurred only twice. From 1981 to 1982, contributions rose by 0.9 percentage point of GDP, follow-

30. Holloway (1989) and Wilson and others (1989) provide extensive discussions of the national accounts and flow of funds measures of personal saving.

Figure 3. Personal Saving Rates and Contributions to Saving Incentives, 1980–90

Percent of GDP



Sources: IRA and Keogh-SEP data are from U.S. Internal Revenue Service (various issues). 401(k) data for 1984–90 are from U.S. Department of Labor (1993). 401(k) data for 1982 and 1983 are estimated. Personal saving data are from the *Economic Report of the President 1994*, table B-27, and the Board of Governors of the Federal Reserve System, flow of funds accounts.

a. Flow of funds is calculated on a national accounts basis by removing consumer durables, government insurance and pension reserves, and saving by corporate farms from the flow of funds measure.

ing the establishment of universal eligibility for IRAs. During this period, the national accounts and flow of funds saving rates were flat, but the FFA-NIPA rate rose by 0.6 percentage point. In the year following the Tax Reform Act of 1986, contributions to saving incentives as a percent of GDP fell by 0.5 percentage point. At the same time, saving in other forms fell by even larger amounts, so that the national accounts rate fell by 1.2 percentage points, the flow of funds rate fell by 2.8 percentage points, and the FFA-NIPA rate fell by 2.2 percentage points. Taking the decade as a whole, contributions to saving incentives rose by 1.0 percentage point of GDP, but the national accounts and FFA-NIPA rates were lower in 1990 than in 1980, while the flow of funds rate was about the same.

It is possible that personal saving in the 1980s could have fallen even further in the absence of saving incentive plans. Nevertheless, it is difficult to conclude from the aggregate data that changes in saving incentives were an important influence on saving behavior in the 1980s.

New Evidence on 401(k) Plans and Private Wealth

This section presents new evidence on the effects of 401(k) plans on private wealth using information from the Survey of Income and Program Participation (SIPP). The SIPP is a nationally representative household survey, conducted by the Bureau of the Census and composed of a series of overlapping panels. Households are interviewed several times over a period of about two and a half years. Every “wave” of the survey collects core data on income, demographics, and other items. Periodic modules collect detailed information on specialized topics.

We use data from topical modules with information on 401(k)s. The 1984 SIPP wave 4 was undertaken between September and December 1984. We refer to this as 1984 data. The 1985 SIPP wave 7 and the 1986 SIPP wave 4 surveys occurred between January and April 1987. Because the relevant variables in these two samples have very similar means and medians and the samples otherwise look very similar, we have combined them to form our 1987 data. Interviews for the 1990 SIPP wave 4 occurred between February and May 1991; we refer to this information as 1991 data. Each topical module contains information about workers’ 401(k) eligibility and participation.³¹ The 1987 and 1991 data also contain information on 401(k) balances.

We focus on families in which the reference person is 25–64 years old, in which at least one person is employed, and in which no one is self-employed.³² Table 2 provides information on 401(k) eligibility and par-

31. As noted above, we treat 403(b) and 457 plans and federal thrift saving plans as the equivalent of 401(k) plans. Skip patterns in the SIPP indicate that 401(k) questions were not asked of workers who said (i) their employer or union did not offer a retirement plan, (ii) they did not know if their employer or union offered a retirement plan, (iii) they were not included in the retirement plan, or (iv) they did not know if they were included in the retirement plan. Hence, eligibility rates could be underreported if workers were not aware of the availability of plans, and eligibility and participation could be underreported if workers did not think of 401(k) plans as retirement plans. Calculations from the 1991 SIPP data reveal 17.2 million 401(k) participants (Yakoboski and Boyce, 1993), while the 1990 Form 5500 reports filed by pension plans count 19.5 million active participants (U.S. Pension and Welfare Benefits Administration, 1993). These differences may be due to the skip patterns noted above, different definitions of active participants in the two samples, or other factors.

32. We focus on this group for several reasons. 401(k) plans are employment based and not available to the self-employed. For people aged 65 and older, issues concerning

Table 2. 401(k) Eligibility and Participation Rates by Family
Percent

<i>Sample category^a</i>	<i>Eligibility^b</i>	<i>Participation^c</i>	<i>Participation of eligible families</i>
<i>All families</i>			
1984	14.7	8.5	57.8
1987	21.9	13.7	62.7
1991	36.4	25.4	69.8
<i>Age of reference person in 1991 sample</i>			
25-34	31.9	23.3	73.2
35-44	40.7	27.2	66.8
45-54	39.4	27.5	69.8
55-64	32.4	23.0	70.9
<i>Income in 1991 sample</i>			
Less than 10,000	10.3	7.1	69.1
10,001 to 20,000	18.8	12.7	67.4
20,001 to 30,000	28.2	16.2	57.5
30,001 to 40,000	37.9	26.2	69.3
40,001 to 50,000	42.8	30.6	71.6
50,001 to 75,000	52.6	37.1	70.5
More than 75,000	54.1	44.5	82.3

Source: Authors' calculations based on data from the Survey of Income and Program Participation (SIPP) as described in the text.

a. Sample sizes are 9,433 in 1984, 10,926 in 1987, and 10,105 in 1991. All figures are weighted to reflect population totals.

b. A family is considered eligible if at least one member of the family is eligible for a 401(k).

c. A family is considered a participant if at least one member of the family has a positive 401(k) balance.

participation in the various years. A family is defined as eligible if anyone in the family is eligible for a 401(k) plan. Because the SIPP waves mentioned above do not provide data on 401(k) contributions, a family is defined as a participant if at least one member has a positive 401(k) balance. Both eligibility and participation rates have grown rapidly. Overall participation rates almost tripled from 8.5 percent in 1984 to 25.4 percent in 1991, mainly because of a rise in eligibility rates but also because of an increase in participation conditional on eligibility. In 1991, 401(k) eligibility was somewhat more common for families headed by a person be-

retirement behavior and saving by the elderly substantially complicate the analysis. SIPP questions about 401(k)s are asked only to people 25 and older. The reference person is the person in whose name the family's home is rented or owned. If jointly owned or rented, either spouse may appear as the reference person.

tween the ages of 35 and 54 and eligibility rates rose with family income.³³

Comparing 401(k) Participants and Nonparticipants

Table 3 presents data on families that did and did not participate in 401(k) plans.³⁴ The table shows that, at any given point in time, 401(k) participants typically have higher overall wealth and higher non-401(k) wealth than nonparticipants. For example, in 1991, median financial assets for participants were \$19,550,³⁵ more than eight times the median level of financial assets for all nonparticipants, which were \$2,249.³⁶ In 1991, the median participant had about 2.7 times the net worth of the median nonparticipant.³⁷

The difference in wealth between participants and nonparticipants at a given point in time cannot be taken as evidence that 401(k) plans raise saving. Part of the difference is due to differences in observable characteristics. For example, the table shows that 401(k) participants tend to

33. Similar patterns hold in 1984 and 1987. Contribution rates on a per worker basis follow similar patterns in the 1988 Employee Benefit Supplement of the Current Population Survey (Andrews, 1992).

34. Income and wealth measures are in real terms (1991 dollars), using the seasonally adjusted CPI for urban households, which was 135.0 in February 1991 and 111.9 in February 1987. These dates approximate the midpoints of the periods during which the 1987 and 1991 SIPP surveys were undertaken.

35. This figure is much higher than comparable figures in Poterba, Venti, and Wise (1993, table 3.1a) because their "families with 401(k)s" do not include families that have 401(k) plans and IRAs, whereas our definition includes all families that have a 401(k).

36. Financial assets are defined as the sum of checking accounts, U.S. saving bonds, other interest-earning accounts in banks and other financial institutions, other interest-earning assets (such as bonds held personally), stocks and mutual funds, and IRA, Keogh, and 401(k) balances.

37. Net worth is defined to include financial assets plus equity in houses, other real estate, businesses and automobiles, and less unsecured debt.

As George Perry has pointed out to us, households may "double count" assets in survey data. For example, an IRA with \$5,000 in equities might be reported as both \$5,000 in stocks and \$5,000 in IRA balances, so that we would calculate financial assets of \$10,000. We do not believe this problem is important for the SIPP. First, questioners are very aware of the issue. Second, in constructing the financial assets measure, the SIPP specifically adds together IRAs and other assets. Third, in the 1991 SIPP, over 60 percent of IRA holders have IRA and 401(k) balances exceeding their other financial assets or report that they have certain types of assets (such as stocks) in their IRA but not outside of their IRA.

Table 3. Median Characteristics of 401(k) Participants and Nonparticipants^a
 1991 dollars, except as noted

<i>Family characteristic</i>	<i>1984</i>	<i>1987</i>	<i>1991</i>	<i>Change, 1987-91</i>
<i>Participants</i>				
401(k) balances ^b	...	4,826	6,000	1,174
Financial assets ^b	...	21,474	19,550	-1,924
Financial assets other than 401(k) balances	10,918	12,575	7,848	-4,727
Net financial assets ^b	...	12,426	10,847	-1,579
Net financial assets other than 401(k) balances	3,686	3,728	1,100	-2,628
Net worth ^b	...	80,469	58,358	-22,111
Net worth other than 401(k) balances	75,141	71,632	46,045	-25,587
Defined benefit pension coverage (percent)	70.2	66.5	63.5	-3.0
Age (years)	42	42	41	-1
Income	50,287	51,014	45,699	-5,315
Education (years)	14	14	14	0
Sample size	810	1,543	2,517	...
<i>Nonparticipants</i>				
Financial assets	2,338	2,328	2,249	-79
Net financial assets	-103	0	0	0
Net worth	31,148	26,624	21,355	-5,269
Defined benefit pension coverage (percent)	46.9	42.5	34.5	-8.0
Age (years)	38	39	40	1
Income	34,011	33,286	31,242	-2,044
Education (years)	12	12	12	0
Sample size	8,623	9,383	7,588	...

Source: Authors' calculations based on data from the SIPP.

a. Table reports median values, except for pension coverage and sample size.

b. Data on 401(k) balances are unavailable for 1984.

be slightly older and have more education and substantially higher incomes than nonparticipants. However, an important part of the differences in wealth holdings in a given year can be explained by heterogeneity: different households have different propensities to save, even after controlling for observable factors.

The notion that heterogeneity arising from unobserved factors plays

an important role in saving is now widely recognized.³⁸ Heterogeneity implies that families with high propensities to save will make up a disproportionate share of those that participate in voluntary plans like IRAs and 401(k)s. Hence, observing in cross-sectional data that participants have higher wealth than nonparticipants, even after controlling for other factors, can lead to misleading inferences about the effects of saving incentives, because unobserved determinants of saving may differ systematically between the groups.

Table 3 can also be used to compare the assets of participants and nonparticipants over time.³⁹ Although participants' median real 401(k) balances grew by \$1,174 from 1987 to 1991, their median real financial assets (including 401(k) balances) fell by \$1,924. Median real financial assets for nonparticipants fell by less than \$100.⁴⁰

However, financial assets may be too narrow a wealth measure for analysis of 401(k) plans. As discussed above, employer matching of employee contributions creates opportunities for financing 401(k) contributions with debt rather than new saving. To allow for this, we define net financial assets as financial assets less nonmortgage debt. Table 3 shows that from 1987 to 1991, median real net financial assets fell for participants and were constant for nonparticipants.

38. See Gale and Scholz (1994) or Poterba, Venti, and Wise (1994). Heterogeneity in unobserved determinants of saving, or tastes for saving, can be demonstrated by running cross-sectional regressions of non-401(k) wealth on observable characteristics and an indicator for whether a family participates in a 401(k) plan. If the coefficient on the 401(k) indicator is positive, it implies that, controlling for observable characteristics, families that save more in 401(k) plans also save more in non-401(k) forms. The coefficient is positive and economically and statistically significant in regressions using each of the 1984, 1987, and 1991 SIPP samples for non-401(k) financial assets, non-401(k) net financial assets, non-401(k) net worth, and defined benefit pension coverage. Hence, participants tend to save more in all forms than nonparticipants, even after controlling for observable characteristics.

39. It is important that these comparisons be made in real rather than nominal terms. Because the level of assets in any year is much higher for participants than nonparticipants, an equal percentage change in asset levels (say, because of inflation) translates into a larger arithmetic change in nominal assets for participants. Hence, changes in real assets will generally be smaller than changes in nominal assets, and can even be in the opposite direction. For example, for 401(k) participants, nominal financial assets rose from 1987 to 1991, even though real financial assets fell.

40. Interpreting trends in these figures would be more difficult if there had been extraordinarily high or low rates of return on existing wealth during this period. But this does not seem to have occurred. For example, the Standard & Poor's 500 rose from 280.9 in February 1987 to 362.3 in February 1991, approximately the dates of the SIPP surveys. This represents an annualized rate of return of about 6.6 percent.

It may also be useful to consider broader measures of wealth. Medians of financial assets and net financial assets are small relative to medians of overall net worth. Therefore, while most households may not have much to substitute away from in terms of net financial assets, they may have other forms of wealth that can be adjusted. Moreover, the debt-financed arbitrage opportunities are even more lucrative if financed by property-backed debt, the interest on which is tax deductible. Table 3 shows that net worth declined more for participants than for nonparticipants from 1987 to 1991.

These data present relevant background information, but descriptive tabulations clearly cannot determine whether 401(k)s increase private wealth. For example, it can be shown that changes in observable characteristics can explain all of the reduction in financial assets and net financial assets, and most of the fall in net worth for 401(k) participants. Stronger evidence on the effects of 401(k) plans requires two items: controlling for changes in the many factors (including income, age, and education) that might influence observed levels of wealth in the cross-sections and finding an appropriate way to identify the effects of 401(k) plans. We discuss our identification strategy below.

Comparing Two Types of "Savers"

Our strategy is to compare two groups: (i) 401(k) participants and (ii) IRA participants who are not eligible for 401(k) plans. An IRA participant is defined as a family with a positive IRA balance.⁴¹ There is no overlap between these two groups. The logic of this comparison is two-fold. First, families in each group, by choosing to participate in voluntary saving plans, have indicated that they are "savers." Second, 401(k) participants had expanded opportunities to use saving incentives between 1986 and 1991 while IRA participants not eligible for 401(k) plans had sharply curtailed opportunities. Not surprisingly, between 1986 and 1991, 401(k) contributions grew rapidly, while IRA contributions plummeted. These changes in the structure and use of saving incentives provide a way of testing the effects of such plans on private wealth.

The basic test is that if 401(k) plans stimulate private saving, then, controlling for other factors, wealth measures for 401(k) participants should have risen from 1987 to 1991 relative to wealth measures for IRA

41. The SIPP does not contain information on IRA contributions, only on balances.

Table 4. Median Characteristics of IRA Participants Not Eligible for 401(k) Plans^a
1991 dollars, except as noted

<i>Characteristic</i>	<i>1984</i>	<i>1987</i>	<i>1991</i>	<i>Change, 1987-91</i>
IRA balances	5,418	8,083	8,200	117
Financial assets	20,551	23,525	27,600	4,075
Financial assets less IRA	12,845	12,426	12,999	573
Net financial assets	13,139	15,683	20,000	4,317
Net financial assets less IRA	6,194	5,911	6,300	389
Net worth	98,944	98,550	98,253	-279
Net worth less IRA	92,726	90,060	83,703	-6,357
Defined benefit pension coverage (percent)	48.8	45.8	41.0	-4.8
Age (years)	49	46	46	0
Income	47,034	47,323	44,151	-3,172
Education (years)	13	14	14	0
Sample size	1,795	2,024	1,257	...

Source: Authors' calculations using data from the SIPP.

a. An IRA participant is defined as a family with a positive IRA balance. Table reports median values, except for pension coverage and sample size.

participants not eligible for 401(k) plans.⁴² The test assumes that, controlling for observable characteristics, the saving behavior of 401(k) participants in a world without 401(k)s would be similar to the saving behavior of IRA participants who are not eligible for 401(k) plans.

Table 4 provides data on IRA participants not eligible for 401(k) plans. Median real IRA balances for this group rose by only \$117 between 1987 and 1991. Hence, 401(k) balances rose for 401(k) participants (as shown in table 3), but IRA balances were roughly constant for IRA participants not eligible for 401(k) plans. Table 4 also shows that IRA participants not eligible for 401(k) plans had increases in median financial assets and had small increases in financial assets excluding IRAs. In 1987 and 1991, IRA participants not eligible for 401(k)s were older than 401(k) participants, had slightly lower income, had the same median educational attainment, and had somewhat higher levels of assets and wealth.

42. The 1987 SIPP data were collected from January to April 1987, and the survey asks for IRA and 401(k) balances at the end of the month preceding the interview. Because contributions to 1986 IRAs could be made until April 15, 1987, the 1987 data refer to asset accumulation that occurred before the provisions of the 1986 tax reform took effect.

TESTING THE EFFECTS OF 401(K) PLANS. To implement the test described above, we use families from the 1987 and 1991 SIPP samples that either (i) had a 401(k) plan or (ii) had an IRA and were not eligible for a 401(k).⁴³ We specify regressions of the following form:

$$(8) \quad W = X\beta + \gamma IN91 + \delta PART + \alpha [(IN91)(PART)] + \epsilon.$$

In this specification, β , γ , δ , and α are parameters to be estimated, and ϵ is a residual assumed to be independently and normally distributed. W is a measure of wealth, discussed below. X is a vector of variables thought to influence wealth accumulation, including age, age-squared, family income, family income-squared, an interaction term between age and income, family size, indicators for educational attainment (12 years, 13–15 years, and 16 years or more), defined benefit pension coverage, marital status, the presence of two earners, occupational classifications, and the race and gender of the reference person. Standard life-cycle considerations suggest that saving is influenced by age, income, and demographic factors, including family size and marital status. Education, pension coverage, and occupation are often thought to influence saving behavior. $IN91$ is an indicator that the observation is in the 1991 sample and captures general differences in wealth between 1987 and 1991. $PART$ is an indicator that at least one member of the family participates in a 401(k) plan; it captures average differences in wealth between 401(k) participants and IRA contributors not eligible for 401(k)s that are not captured by other variables.

The key independent variable is the interaction term $(IN91)(PART)$, which measures the amount by which wealth *increased* for 401(k) participants relative to IRA holders not eligible for 401(k)s from 1987 to 1991, after controlling for family characteristics (through X), general changes between 1987 and 1991 (through $IN91$), and general differences in wealth between the two groups (through $PART$). If 401(k) plans raise private wealth, the coefficient α should be positive and economically and statistically significant.

This equation is estimated for a variety of different wealth measures. We report results for real financial assets and real net financial assets. In each case, we remove IRA balances from the wealth measures of IRA participants not eligible for 401(k) plans.⁴⁴ This adjustment biases the

43. Data from 1984 are excluded because there is no information on 401(k) balances.

44. Thus, the dependent variables are total or net financial assets for 401(k) participants (including their 401(k) and IRA balances) and total or net non-IRA financial assets for IRA contributors not eligible for 401(k)s.

Table 5. Median Regression Explaining Net Financial Assets^a

<i>Independent variable^b</i>	<i>Coefficient</i>	<i>Standard error^c</i>	<i>t-statistic</i>
Age	-0.9776	0.2684	-3.64
Age-squared	1.1738	0.3726	3.15
Income	-0.5410	0.1177	-4.59
Income-squared	0.2351	0.0924	2.54
Age-income interaction	1.5993	0.2115	7.56
12 years of education	0.0213	0.0094	2.24
13-15 years of education	0.0197	0.0085	2.30
16 years or more of education	0.0767	0.0119	6.44
Defined benefit pension coverage	0.0046	0.0052	0.88
Married	0.0272	0.0141	1.92
Male	0.0057	0.0072	0.79
White	0.0502	0.0088	5.69
Two earners	-0.0679	0.0119	-5.68
Family size	-0.1813	0.0330	-5.49
Family member with 401(k) (<i>PART</i>)	0.0850	0.0101	8.39
1991 sample dummy (<i>IN91</i>)	0.0129	0.0088	1.46
Interaction term [<i>IN91</i>] (<i>PART</i>)	-0.0158	0.0126	-1.24
Constant	0.1063	0.0449	2.36

Source: Authors' regressions using data from the SIPP.

a. The dependent variable is net financial assets measured in 1991 dollars. Occupational dummies are omitted from the table. Variables are measured in the following units: net financial assets divided by 100,000, age divided by 100, age-squared divided by 10,000, income divided by 100,000, income-squared divided by 10^{10} , age-income divided by 10^7 , and family size divided by 10. Income and assets are expressed in 1991 dollars.

b. The sample pools 1987 and 1991 families that either (i) have a 401(k) plan or (ii) have an IRA but not eligible for a 401(k) plan. The sample size is 7,341.

c. Standard errors are estimated with bootstrap methods.

tests in favor of finding that 401(k) plans raise saving but is done to ensure that any increase in the comparison group's wealth is not the result of IRAs. This issue arises because some families roll over pension balances into IRAs. These should not be counted as increases in wealth but would be counted as such in the SIPP, because questions are asked directly about IRA balances but not about balances in defined contribution pension plans other than 401(k)s.⁴⁵

Table 5 provides estimates of the median regression for net financial assets.⁴⁶ Controlling for other factors and taking account of the qua-

45. We discuss below the results of adding IRAs back into the asset measures for the comparison group. For analyses and data on rollovers, see Chang (1993) and Fernandez (1992).

46. Because median regressions reduce the effect of outliers they are useful in analyzing wealth data, which are typically highly skewed. Appendix table A1 provides summary statistics on the variables used in the regression.

Table 6. Effects of 401(k) Participation on Private Wealth Accumulation^a

<i>Dependent variable</i>	<i>Quantile regression</i>			<i>Ordinary least squares</i>
	<i>25th percentile</i>	<i>Median</i>	<i>75th percentile</i>	
Financial assets	- 810 (1.67)	279 (0.21)	- 131 (0.05)	- 3,856 (0.99)
Net financial assets	- 1,310 (1.33)	- 1,576 (1.24)	- 3,409 (1.19)	- 2,146 (0.50)

Source: Authors' regressions using data from the SIPP.

a. The table provides estimates of the coefficient on the interaction term $(IN91)(PART)$ in 1991 dollars in a pooled sample of families that either (i) participated in a 401(k) or (ii) participated in an IRA but were not eligible for a 401(k). Quantile regression standard errors are estimated with bootstrap methods; *t*-statistics are shown in parentheses.

dratic and interactive terms, the coefficients imply that the median of net financial assets rises with age, income, and, in general, education; is higher for married couples than for singles; and is lower for families with two earners and for larger families. On average, participants in 401(k)s have higher levels of net financial assets, controlling for other factors.

The key coefficient is that on the interaction term $(IN91)(PART)$. This coefficient should be positive and significant if 401(k) plans raise private saving. Instead, it is negative and insignificant, indicating that the net financial assets of 401(k) participants did not rise relative to the net financial assets of IRA holders not eligible for 401(k) plans, controlling for other factors. This result occurs even though median real 401(k) balances among 401(k) participants rose by more than \$1,100 over this period. The implication is that contributions to 401(k) plans did not raise private saving relative to what it would have been in the absence of 401(k)s.

Similar estimates occur for a variety of additional specifications. Table 6 reports the coefficients on the interaction term for the median regression, the 25th and 75th quantile regressions, and an ordinary least squares (OLS) regression for financial assets and net financial assets. In one case, the interaction term is positive: median financial assets grew by \$279 more for 401(k) participants than for IRA participants not eligible for 401(k)s. This figure is estimated imprecisely (the *t*-statistic is 0.27) but represents about 24 percent of the \$1,174 increase in median 401(k) balances for 401(k) participants.

From this result, it is possible to develop a point estimate for the proportion of 401(k) contributions that represent increases in national sav-

ing. We estimate an average tax rate of 22 percent for 401(k) participants.⁴⁷ This would suggest that 2 percent of 401(k) contributions represent increases in national saving.

In all other cases, controlling for other factors, asset measures for 401(k) participants grew less than for IRA contributors not eligible for 401(k)s by a statistically insignificant amount. Results from other regressions that use net worth as the dependent variable or include IRA balances in the asset measure of IRA participants not eligible for 401(k) plans generate results even less favorable for 401(k) plans.

In each of the quantile regressions, the effects of 401(k) plans on wealth are successively more negative as the definition of the wealth variable is widened from narrow measures, such as financial assets, to broader ones, such as net financial assets or net worth.⁴⁸ This suggests that families may be able to substitute a variety of assets for 401(k) balances and suggests the importance of examining measures of wealth that are sufficiently broad to allow for adjustments in portfolios.

It also suggests that the non-401(k) wealth of 401(k) participants should have *fallen* relative to that of IRA participants not eligible for 401(k)s. To examine this proposition, we reestimated the median regressions in table 6, excluding 401(k) balances from the wealth measures of 401(k) participants. If 401(k)s and other financial assets are substitutes, the coefficient on the interaction term (*IN91*) (*PART*) should be negative and significant in this specification. The results are consistent with this hypothesis. When the dependent variable is financial assets (excluding 401(k) balances), the coefficient on the interaction term is $-\$2,333$ (with a *t*-statistic of 2.87); when the dependent variable is net financial assets, the coefficient on the interaction term is $-\$2,985$ (with a *t*-statistic of 3.53).

These results refer only to private saving, while our original criterion for whether saving incentives worked also focused on national saving. Clearly, if 401(k) plans do not raise private saving, they reduce national saving in the absence of any other policy change.

47. The average marginal tax rate in 1988 was 22 percent for IRA participants—anyone who had contributed to an IRA between 1979 and 1988—using the IRS-Michigan Tax Panel, described later in this paper. Relevant characteristics, such as income, family size, percent married, and homeownership, are very similar for 401(k) participants and IRA participants in the 1987 SIPP.

48. That this relation does not hold for the ordinary least squares (mean) regression suggests the importance of outliers in influencing the OLS relationships.

DISCUSSION. The tests above are based on two key factors: changes in the opportunities for and use of saving incentives for different groups of families, and the sample choice of families such that one group of savers can be compared with another. We discuss each factor further below.

Concerning the shift in opportunities, a potential caveat is that the Internal Revenue Service limit on employee 401(k) contributions fell from \$30,000 in 1986 to \$7,000 in 1987 (and was then indexed for inflation). If a large portion of 401(k) participants were contributing more than \$7,000 a year before 1986, the new limit could affect our results. However, it seems that very few people gave amounts exceeding \$7,000. In the 1987 SIPP, 50 percent of participants had accumulated 401(k) balances of \$4,000 or less, 75 percent had balances of \$10,395 or less, and only 10 percent had balances of more than \$29,000. Andrea Kusko, James Poterba, and David Wilcox show that only about 1 percent of 401(k) participants in a medium-sized manufacturing firm in 1990 were constrained by the IRS limit.⁴⁹ Leslie Papke presents data from Form 5500 reports filed by pension plans indicating that less than 1 percent of contributors made employee contributions of \$5,000 or more in 1987.⁵⁰ Hence, the lower IRS limit is unlikely to have affected the median regressions.

A second issue concerns our choice of comparison groups. At least three issues arise in comparing 401(k) participants and IRA participants not eligible for 401(k)s.⁵¹ Do the groups start out with unequal propensities to save? Does the propensity to save of the 401(k) participant group fall over time? Does the propensity to save of the IRA group rise over time? If the answer to all three questions is no, our choice of comparison groups does not lead to systematic biases in the analysis. Each of these questions lends itself to empirical examination, which we describe below.

First, do the two groups start out with similar underlying propensities to save? If 401(k) participants have systematically lower propensities to save, our test might be biased against finding a favorable effect of 401(k) plans on private saving. One might argue that it is “easy” to save in a

49. Kusko, Poterba, and Wilcox (1994).

50. Papke (1992).

51. We thank Doug Bernheim, Joel Slemrod, and Jon Skinner, among others, for stimulating our thoughts on this topic. As discussed below, related issues arise in the comparison of other groups as well.

401(k) plan: employers set up the plan, provide for automatic payroll deductions, and often provide high rates of matching contributions. One may not need to be much of a “saver” to participate. In contrast, to contribute to an IRA, one must set up an account and arrange to make contributions to it. These considerations suggest that 401(k) participants may have systematically less intense tastes for saving on average than IRA participants not eligible for 401(k) plans. However, non-401(k) pension coverage is higher for 401(k) participants than for IRA participants not eligible for 401(k) plans, as shown in tables 3 and 4. If workers with strong desires to save self-select into firms that offer pensions or 401(k) plans, then 401(k) participants may on average have higher propensities to save than IRA participants not eligible for 401(k) plans.

We examine initial propensities to save by estimating median regressions, using all 401(k) participants and IRA participants not eligible for 401(k) plans in the 1987 SIPP data. The dependent variable is a measure of assets; the independent variables include all of the *X* variables and a dummy variable for whether the household was a 401(k) participant. If, controlling for other factors, the 401(k) group has a smaller overall propensity to save, the coefficient on the dummy variable indicating 401(k) participation should be negative. Our estimates indicate that when the dependent variable is financial assets (net financial assets), the coefficient on 401(k) participation is \$972 (\$1,986) with a *t*-statistic of 1.46 (1.64). Thus, there is no evidence in these regressions that 401(k) participants have smaller propensities to save via total or net financial assets than IRA participants not eligible for 401(k) plans.⁵²

The second issue is whether the propensity to save for 401(k) participants fell. If average tastes for saving among 401(k) participants fell from 1987 to 1991, our tests may be biased against finding a positive effect of 401(k) plans on saving. This possibility is based on the large increase in participation and eligibility between 1987 and 1991, and the idea that there is heterogeneity in propensities to save. For example, suppose that there are “serious” savers and “occasional” savers and that at the beginning of the 401(k) program all of the serious savers enrolled. If an increasing number of occasional savers participate over time, the average taste for saving among participants would fall.

52. The dependent variables in these regressions include the IRA balances of IRA participants not eligible for 401(k) plans. Obviously, excluding IRA balances would raise the coefficient on 401(k) participation and strengthen the conclusion.

But this simple scenario may omit important factors. Many 401(k) plans are cashed in upon employee separation from the firm. Thus, an alternative scenario is that over time, more and more serious savers enrolled (and stayed enrolled) in a 401(k) as the plans proliferated. New occasional savers established 401(k) accounts, but other occasional savers cashed in their 401(k) plans and hence departed from the sample of 401(k) participants.⁵³ The direction of potential shifts in tastes for saving is not obvious. Moreover, as shown in table 3, median 401(k) balances among participants rose from 1987 to 1991; given the large expansion in 401(k) participation from 1987 to 1991, one might have expected balances to have fallen if the sample were becoming more diluted over time.

To examine these ideas, we estimate median regressions, pooling all 401(k) participants in the 1987 and 1991 SIPP data. The dependent variable is a measure of assets; the independent variables include all of the **X** variables and a dummy variable for whether the family was in the 1991 sample. If, controlling for other factors, the 401(k) sample is becoming diluted with families that have low propensities to save, then the coefficient on the 1991 indicator should be negative. But when the dependent variable is financial assets (net financial assets), the coefficient on the 1991 dummy variable is \$1,638 (\$267) with a *t*-statistic of 2.31 (0.38). Thus, for 1987 to 1991 there is no evidence of a reduction in the propensity of 401(k) participants to accumulate total or net financial assets.

The third issue is whether characteristics of IRA participants not eligible for 401(k) plans changed so as to raise their average tastes for saving between 1987 and 1991. If so, our tests would be biased against finding a positive effect of 401(k)s on saving. The concern is that, since contributions to IRAs decreased sharply after 1986, only serious savers may still have had IRAs in 1991, whereas both serious and occasional savers had IRAs in 1987.

To address this conjecture, we estimate median regressions, this time pooling all IRA participants not eligible for 401(k) plans in the 1987 and 1991 SIPP data. The dependent variable is a measure of assets that excludes IRA balances for the reasons given above and for comparability

53. Although most amounts received through lump sum distributions are rolled over into some form of saving, broadly defined, a substantial proportion of the recipients consume all or part of the funds. See Chang (1993) and Fernandez (1992).

with the regressions in tables 5 and 6. The independent variables include all of the X variables and a dummy variable for whether the family was in the 1991 sample. Controlling for other factors, if the overall taste for saving was rising in this group between 1987 and 1991, the coefficient on the 1991 indicator should be positive and significant. When the dependent variable is non-IRA financial assets (non-IRA net financial assets), the coefficient on the 1991 indicator is \$1,209 (\$1,060) with a t -statistic of 1.33 (1.16). Thus, for financial assets, the increase in the taste for saving is smaller for this group than it is for 401(k) participants and is not statistically significant. For net financial assets, the increase in the taste for saving is larger than it is for 401(k) participants but again is not statistically significant. These three sets of results provide no evidence that the samples we have chosen are biased toward finding that 401(k) plans do not raise total or net financial assets.⁵⁴

However, our tests contain two omissions that may bias the results in favor of finding that 401(k) plans *raise* private saving. First, we ignore the distinction between pre- and posttax asset balances. A 401(k) balance at any point in time represents a pretax value; one cannot consume the entire balance because income taxes and perhaps penalties are due upon withdrawal. In contrast, one may consume the entire balance of a taxable asset. Hence, at a point in time, a family with all of its portfolio in fully taxable assets has higher current available wealth than a family with an equal asset balance but all of its portfolio in a 401(k) plan, but the SIPP would record them as having equal wealth. Moreover, if tax rates change by relatively small amounts, as they did from 1987 to 1991, comparisons of increases over time in fully taxable assets and pretax 401(k) balances over time will tend to overstate the increase in wealth arising from 401(k) plans.⁵⁵

Second, a variety of estimates suggest that about one-third or more

54. The tests implicitly assume that the underlying relationship in 1987 between assets and observable family characteristics accurately reflects the relationship in 1991 as well.

55. Footnote 19 implies that, ignoring any penalties, the after-tax balance after T years, resulting from a \$1 contribution to a 401(k) plan with a zero match rate, is $(1 - \tau)(1 + r)^T$. The value reported in the SIPP, which asks about the current balance, would be $(1 + r)^T$. Hence, if the reported value of 401(k) balances rose by an amount, Z , from one survey to the next, the increase in after-tax balances (assuming tax rates were constant) would be only $(1 - \tau)Z$. Our tests do not adjust 401(k) balances downward to account for this effect. Similar problems arise for IRAs, but our tests remove IRA balances for IRA participants not eligible for 401(k) plans.

of 401(k) plans represent conversions of previously existing pensions.⁵⁶ This presents several potential problems for analyses using the SIPP data, because the SIPP provides information on 401(k) balances but not on balances in other pension plans. Consider two observationally equivalent families each with equal amounts of taxable assets. One family also has \$10,000 in a 401(k) while the other family has the same amount in a defined contribution plan. The SIPP data would show the 401(k) family to have higher wealth. Moreover, as more pension plans were converted to 401(k)s between 1984 and 1991, the successive cross-sections of the SIPP would show increases in reported assets associated with becoming a 401(k) participant, even though a real increase may not have occurred.

Comparisons Using Other Groups

The tests above are based on comparing two groups of savers who experienced different changes in their opportunities to use saving incentive plans. There is obviously no single "correct" choice of comparison groups, however, so this subsection examines alternative comparisons.

The underlying question in all of these comparisons is what would have happened to the assets of 401(k) contributors if they had not been eligible to make 401(k) contributions. The difference between this counterfactual scenario and what actually happened is the "effect of 401(k)s on household saving."

Addressing this counterfactual comparison requires some sort of model. One approach is to develop a structural econometric model of household saving behavior.⁵⁷ An alternative is to make less structured comparisons, controlling for a variety of factors. We have chosen the latter approach, which assumes that, after controlling for other characteristics, the saving of IRA participants not eligible for 401(k)s would be equivalent to the saving of 401(k) participants in the absence of 401(k)s.

56. Andrews (1992) reports that 15,689 of the 45,054 401(k) plans in existence in 1987 reported being established before 1984 even though there were only 1,703 actual 401(k) plans in existence in 1984. Most of the discrepancy is thought to be due to preexisting plans reconstituted as 401(k) plans. Papke, Petersen, and Poterba (1993, table 3.1) report that 45 percent of the firms they surveyed, representing 37 percent of the participants covered by the plans, reported that a 401(k) plan replaced a previously existing pension plan. In a 1989 survey by Buck Consultants, 28 percent of 401(k) plans surveyed had been converted from thrift or profit-sharing plans, and 7 percent were converted from defined benefit plans (Buck Consultants, 1989, p. 4).

57. See Gale and Scholz (1994) and Venti and Wise (1986, 1987, 1990, and 1991).

Comparisons using a less structured approach must be made with care. An important cost of this approach is that at least two criteria are needed for the comparisons to be informative. First, the type of household within each comparison group should remain similar over time. For example, we show above that the sample of 401(k) participants did not become diluted over time with families with low tastes for saving and that the sample of IRA participants not eligible for 401(k) plans did not become concentrated with families with high tastes for saving. Second, the two groups should not be too different at a point in time. For example, we show above that 401(k) participants and IRA participants not eligible for 401(k) plans have similar tastes for saving in 1987.

The second criterion arises because in a less structured approach, it is extremely difficult to control for all of the variables that affect wealth accumulation, and groups in very different circumstances may be influenced differently by omitted variables. For example, in the late 1980s, at least two important events might have affected disparate groups in different ways. First, Jerry Hausman and Poterba show that the Tax Reform Act of 1986 had substantially different effects at the top end of the income distribution than at its middle.⁵⁸ Second, the stock market crash in 1987 likely had different effects on households at opposite ends of the income and wealth distribution.

COMPARISONS BASED ON PARTICIPATION. An alternative way to examine the effects of 401(k)s on wealth is to compare families with 401(k)s but without IRAs to families without 401(k)s.⁵⁹ This comparison is one of those presented by Poterba, Steven Venti, and David Wise, who use a different methodology than we do to examine the effects of 401(k) plans on household saving.⁶⁰ If one assumes that the saving behavior of 401(k) participants without IRAs in the absence of 401(k)s would be similar to the saving behavior of households without 401(k)s, controlling for other characteristics, the effects of 401(k)s on household saving can be estimated using a specification similar to that in table 5. To do so, we include all families from the 1987 and 1991 SIPPs that either

58. Hausman and Poterba (1987).

59. Not surprisingly, the median values of economic characteristics of families with 401(k) plans but without IRAs are generally between those of all 401(k) participants and those of nonparticipants. The median real 401(k) balances for families with 401(k)s but not IRAs rose by \$1,234 from 1987 to 1991.

60. Poterba, Venti, and Wise (1992, 1993, and 1994).

(i) had a 401(k) but not an IRA or (ii) did not have a 401(k). This specification drops from the sample those households that have both a 401(k) and an IRA.⁶¹ The independent variables include the same *X* variables as above, an *IN91* indicator, a *PART* indicator, and the (*IN91*) (*PART*) interaction term. As before, the effect of 401(k) plans on wealth should appear in the term (*IN91*) (*PART*).

These regressions show that, controlling for other factors, median real financial assets of families with 401(k)s but not IRAs increased by a point estimate of \$445 relative to that of families without 401(k) plans between 1987 and 1991. For net financial assets the point estimate was \$401. These figures represent 36 percent and 32 percent of the \$1,234 increase in median 401(k) balances for families with 401(k)s but not IRAs. Using an average tax rate of 22 percent for 401(k) participants would suggest that between 14 percent and 10 percent of 401(k) contributions represent net increases in national saving in the short run.⁶²

We note, however, that the point estimates are insignificantly different from zero and do not address the possible upward biases that arise from failing to distinguish between pre- and posttax asset balances and from ignoring the prevalence of 401(k) plans that represent rollovers of previously existing plans. Moreover, when net worth is the dependent variable, the coefficient on the interaction term is negative. Finally, using these comparison groups implicitly assumes that the saving behavior of 401(k) participants, to whom we have referred earlier as “savers,” would, in the absence of 401(k) plans, be the same as nonparticipants, controlling for observable characteristics.⁶³

61. Among 401(k) participants, 46 percent had IRAs in 1987 and 35 percent had IRAs in 1991.

62. Ordinary least squares regressions also generate positive but statistically insignificant coefficients on the interaction term. The point estimate is that mean financial assets of families with 401(k)s but not IRAs increased by \$1,114 relative to that of families without 401(k) plans between 1987 and 1991. For net financial assets the increase was \$2,418. These should be compared with an increase in mean 401(k) assets of \$3,125.

63. To examine differences in initial propensities to save, we ran median regressions using all families in the 1987 SIPP that either (i) had a 401(k) but not an IRA or (ii) did not have a 401(k). The dependent variable was financial assets, and the independent variables were the same *X*s as above and an indicator variable for having a 401(k) plan. The coefficient on the indicator variable was about \$3,700 and estimated precisely. By comparison, the median level of financial assets in the 1987 SIPP sample was less than \$3,000, and the median level of net worth was under \$30,000.

In similar tests comparing families with 401(k)s but not IRAs to IRA participants not eligible for 401(k)s, the coefficient on the interaction term is \$317 when the dependent variable is financial assets and \$43 when the dependent variable is net financial assets. Neither is estimated precisely. These figures represent 26 percent and 3.5 percent of the \$1,234 increase in median 401(k) balances for families with 401(k)s but not IRAs. Adjusting for the tax loss, this comparison also suggests that 401(k)s have had little, if any, effect on national saving.

Another test would compare families with an IRA but not a 401(k) to families without an IRA. These two groups, however, appear to be very different with respect to initial propensities to save and with respect to wealth.⁶⁴ Hence, we believe it is unlikely that differences in tastes for saving or differential changes in the economic environment facing the two groups (such as the 1986 tax reform and the stock market crash) can be adequately captured using the modeling approach above. Another comparison would involve families with both an IRA and a 401(k) and families with neither. Again, however, we believe that these groups are too disparate for the comparison to be informative.⁶⁵

COMPARISONS BASED ON ELIGIBILITY. From an analytical perspective, a potentially advantageous aspect of 401(k) plans is that—unlike IRAs—they are not universally available. If 401(k) eligibility is distributed independently of underlying propensities to save, then differences in eligibility across families could be used to identify the effects of 401(k) plans on saving. In this subsection, we examine the relation between 401(k) eligibility and propensities to save.

Table 7 presents information on households by 401(k) eligibility status in the SIPP. Median non-401(k) financial assets and non-401(k) net worth are several times higher for eligibles than noneligibles. In 1991,

64. To examine these differences, we ran median regressions using all families in the 1987 SIPP that either (i) had an IRA but not a 401(k) or (ii) did not have an IRA. The dependent variable was financial assets, and the independent variables were the same *X*s as above and an indicator variable for having an IRA. The coefficient on the indicator variable was about \$16,000 and estimated precisely.

65. We ran median regressions using all families in the 1987 SIPP that either (i) had both an IRA and a 401(k) or (ii) had neither. The dependent variable was financial assets, and the independent variables were the same *X*s as above and an indicator variable for having both saving incentive plans. The coefficient on the indicator variable was about \$39,000 and estimated precisely.

Table 7. Median Characteristics of 401(k)-Eligible and Noneligible Families^a

1991 dollars, except as noted

<i>Family characteristic</i>	<i>1984</i>	<i>1987</i>	<i>1991</i>
<i>Eligible families</i>			
401(k) balances ^b	. . .	1,206	2,000
Financial assets ^b	. . .	13,422	13,456
Financial assets other than 401(k) balances	7,026	8,445	6,652
Net financial assets ^b	. . .	5,248	6,305
Net financial assets other than 401(k) balances	1,130	1,745	556
Net worth ^b	. . .	62,795	51,339
Net worth other than 401(k) balances	61,573	58,632	42,900
Defined benefit pension coverage (percent)	73.6	69.5	63.5
Age (years)	40	40	41
Income	46,242	48,137	43,605
Education (years)	14	14	14
Sample size	1,410	2,433	3,613
<i>Noneligible families</i>			
Financial assets	2,236	2,051	2,000
Net financial assets	-89	0	0
Net worth	30,299	25,576	19,250
Defined benefit pension coverage (percent)	44.6	39.2	28.0
Age (years)	39	39	39
Income	33,467	32,320	30,000
Education (years)	12	12	12
Sample size	8,023	8,493	6,492

Source: Authors' calculations using data from the SIPP.

a. Table reports median values, except for pension coverage and sample size.

b. Data on 401(k) balances are unavailable for 1984.

401(k)-eligible households were 35.5 percentage points more likely to participate in a defined benefit pension than noneligible families.⁶⁶ Some of these differences are due to differences in observable characteristics: eligibles have higher median income and more years of education. The key issue is whether, controlling for other factors, 401(k)-eligible families tend to have greater propensities to save than noneligible families.

Table 8 addresses this issue. The four panels show estimates of the effects of eligibility on various measures of wealth that *exclude* 401(k)

66. The probability of having a defined benefit plan and being vested is about 20 to 25 percentage points higher for eligibles than noneligibles in the three survey years. Including data on reported participation in defined contribution plans raises these differences, but it is unclear whether respondents included a 401(k) plan as a defined contribution plan in answering the survey, so we exclude data on these plans. 401(k) eligibles are also about 10 percentage points more likely to have employer-provided health insurance.

Table 8. Effects of 401(k) Eligibility on Non-401(k) Wealth

<i>Measure</i>	<i>Year</i>	<i>Controlling for income only</i>	<i>Controlling for income and other factors^a</i>
Probability of having a defined benefit pension ^b	1984	0.26 (16.5)	0.24 (15.2)
	1987	0.26 (21.1)	0.25 (19.8)
	1991	0.33 (31.3)	0.31 (28.7)
Median level of non-401(k) financial assets ^c	1984	737 (3.22)	346 (1.98)
	1987	992 (3.81)	773 (4.96)
	1991	215 (1.28)	280 (2.07)
Median level of non-401(k) net financial assets ^c	1984	372 (1.64)	173 (0.65)
	1987	521 (2.73)	768 (3.07)
	1991	-307 (1.71)	-48 (0.23)
Median level of non-401(k) net worth ^c	1984	4,668 (2.36)	2,500 (1.99)
	1987	2,149 (1.58)	3,291 (3.20)
	1991	-848 (0.71)	1,478 (1.75)

Source: Authors' regressions using data from the SIPP. Numbers in parentheses are *t*-statistics.

a. The other factors are the explanatory variables in X, as described in the text.

b. In the first panel, the entry represents the marginal increase in the probability of being covered by a defined benefit pension plan associated with being eligible for a 401(k), controlling for other factors. The marginal probabilities are calculated from the coefficients of a probit equation (see Greene, 1990, p. 664).

c. The entry represents the coefficient on being eligible in a median regression of non-401(k) wealth on eligibility status and other factors. For these panels, the median regression standard errors are estimated with bootstrap methods.

balances. For each measure and each sample year, two cross-sectional equations are estimated. The first controls for a constant, family income, and 401(k) eligibility. The second controls for all of the X variables in table 5 and for eligibility.⁶⁷ If, controlling for other factors, eligi-

67. The only exception is that the probits for pension coverage do not include pension coverage as an independent variable, for obvious reasons.

bility is positively correlated with measures of non-401(k) wealth, then it is difficult to see how eligibility can be uncorrelated with tastes for saving. In that case, exploiting variations in eligibility will not be a useful way to measure the effects of 401(k) plans on saving.

The first panel shows that 401(k)-eligible families are between 24 and 31 percentage points more likely to be covered by a defined benefit pension plan than are ineligible families, controlling for other factors. Most of the difference in coverage rates reported in table 7 can be explained solely by the eligibility term in table 8. The second panel of table 8 shows the effects of eligibility on median levels of non-401(k) financial assets. Eligibility is typically statistically significant, and the estimated effects are large relative to the median financial assets of noneligibles. The third and fourth panels show smaller effects of eligibility on net financial assets and larger effects on overall net worth (again, excluding 401(k)s).

Table 8 thus shows that 401(k)-eligible families save more in *non-401(k)* assets than observationally equivalent noneligible families, even after controlling for other factors. The results show that 401(k) eligibility is positively correlated with an underlying taste for saving and with unobserved determinants of saving behavior, and in most cases this relationship is economically and statistically significant.⁶⁸ Moreover, the results may understate the positive correlation between tastes for saving and 401(k) eligibility. The appropriate way to measure differences in the saving propensities of eligibles and noneligibles is to examine the actual wealth of noneligibles and the wealth of eligibles in the absence of 401(k) plans. The test in table 8 assumes that all 401(k) saving is new saving, so that no adjustment is made for 401(k) wealth. If, however, x percent of 401(k) saving is thought not to be new saving, then x percent of 401(k) balances should be added to the wealth measured for eligibles in table 8. This would obviously raise the coefficient on eligibility in the regressions.

The implication is that using variations in 401(k) eligibility to identify the effects of 401(k) plans on saving will overstate those effects.⁶⁹ The

68. The equations that control only for income provide particularly strong evidence of this proposition. Eligible families are about the same age as noneligibles but have higher income and education, implying that they have higher permanent income than noneligibles. Controlling for current income, someone with higher permanent income should have lower wealth, yet eligibles have higher wealth, even ignoring 401(k) balances.

69. One could try to use instrumental variables to examine the effect of 401(k) eligibility on household saving. A good instrument needs to be a determinant of eligibility and

finding that 401(k) eligibility is positively correlated with tastes for saving is consistent with theoretical models of pensions⁷⁰ and with survey evidence concerning why 401(k) plans are created.⁷¹

Substitution among 401(k) Plans, IRAs, and Pensions

Our results are consistent with the view that 401(k) plans do not raise private wealth. This implies that contributions to 401(k)s are financed by some combination of saving that would have been done anyway, shifting of other assets, or increasing debt. In particular, funds that would otherwise have been placed in IRAs or private pensions are one possible source of 401(k) contributions. In this subsection, we examine the extent to which IRAs or private pensions act as substitutes for 401(k)s, with particular emphasis on the relationship between IRAs and 401(k) plans.

IRAs and 401(k) plans are unlikely to be perfect substitutes. 401(k) plans are tied to the work place, while IRAs are not. 401(k)s have different matching rates, contribution limits, hardship criteria, loan provisions, and asset allocation choices. 401(k) contributions tend to be regu-

not of saving. But probit regressions with industry dummies, region of the country, year dummies, union status, occupation, and firm size explained very little of the variation in eligibility, and it is difficult to find other possible instruments. Given the econometric problems with using a poor first-stage instrument, we do not believe this approach would generate reliable results. See Bound, Jaeger, and Baker (1993); Nelson and Startz (1990); and Staiger and Stock (1993). Skinner (1994) proposes an alternative way to address this issue.

70. Theoretical models suggest that pensions exist where workers demand them or that workers who demand pensions will be attracted to firms that offer them (Allen, Clark, and McDermed, 1993). In either case, it seems reasonable to believe that pension coverage will be positively correlated with tastes for saving. Ippolito (1993) provides a model explaining the rapid growth of 401(k) plans in terms of their ability to help firms attract more productive workers. These workers will typically have lower discount rates as well (for example, they may value future income more and hence work harder to achieve promotion) but are difficult for the firm to identify. 401(k) plans help firms identify such workers because 401(k)s can be cashed in upon leaving the firm. High-discount rate workers find it relatively more desirable to leave the firm and cash in the 401(k). One implication is that workers that remain with a firm that offers a 401(k) plan will tend to have low discount rates: that is, they will have higher propensities to save.

71. A 1989 report by Buck Consultants, summarizing information obtained from 424 responding companies, representing a broad range of firm sizes, geographical locations, and industry types, reveals that "perceived employee interest" was the second-most-frequently stated reason that a firm installed a 401(k) plan and was noted by 63.5 percent of respondents (Buck Consultants, 1989, p. 4).

lar salary deductions; IRAs can be funded anytime. For example, people may want to put some money into a 401(k) but then keep the flexibility of contributing to an IRA at a later date in the tax year, depending on their income and expenses.

Nevertheless, IRAs and 401(k)s could be good substitutes for some people (particularly those who hold both 401(k)s and IRAs). The plans represent alternative ways to save for retirement, and 401(k)s should be particularly attractive for high-income households after removing the tax deductibility of IRA contributions in 1986. Moreover, a variety of papers show that households often shift the composition of their assets and liabilities in response to tax changes,⁷² even if they do not change the overall level of wealth. Similar statements apply to pensions and 401(k) plans.

PREVIOUS RESEARCH. Other researchers have uncovered some evidence of substitution. James Eisner uses data from the 1986 SIPP to estimate the probability of 401(k) participation (conditional on eligibility).⁷³ He finds that, evaluated at sample means and controlling for other factors, having a defined benefit plan reduces a worker's 401(k) participation rate by 28 percentage points, while having a defined contribution plan reduces the participation rate by 10 percentage points. Andrews, using the Employee Benefits Supplement of the 1988 Current Population Survey, also finds that participation in a preexisting pension plan reduces the probability of participating in a 401(k) plan, conditional on eligibility and other factors.⁷⁴

Another way to test for substitutability is to exploit the removal of the tax deductibility of IRA contributions for some higher-income families in the 1986 tax reform. If households find IRAs and 401(k)s to be substitutes, then, controlling for other factors, (i) the overall decline in the probability of giving to an IRA after 1986 should be higher for 401(k) eligibles than noneligibles; (ii) there should be little difference in the decline in IRA contributions for 401(k) eligibles and noneligibles among income groups whose IRA tax status did not change; and (iii) the decline in IRA participation should be larger for 401(k) eligibles than for noneli-

72. See, for example, Feenberg and Poterba (1993), Feldstein (1976, 1993), Manchester and Poterba (1989), Scholz (1994), Skinner and Feenberg (1990), and Slemrod (1992a).

73. Eisner (1993).

74. Andrews (1992).

gibles among income groups who lost deductibility of IRA contributions.

Poterba, Venti, and Wise present evidence related to this issue using data from the Employee Benefits Supplement of the 1983 and 1988 Current Population Surveys.⁷⁵ Their results are broadly consistent with the patterns noted above. First, during 1982–87, the decline in IRA participation was 10.7 percentage points for 401(k) eligibles compared with 7 percentage points for those not eligible. Second, for low-income groups, whose IRA tax status was not affected, the decline in IRA participation rates was about the same for 401(k) eligibles and noneligibles. Third, for families with incomes above \$75,000, the IRA participation rate dropped substantially more for 401(k) eligibles than for noneligibles. In this group, the IRA participation rate fell by 57 percentage points for eligibles compared with 27 percentage points for noneligibles.

Poterba, Venti, and Wise use the same data to estimate the probability of contributing to an IRA before and after 1986 as a function of year, income class, and 401(k) eligibility.⁷⁶ The point estimates suggest that the fall in IRA contribution rates from 1982 to 1987 was larger for 401(k) eligibles than for noneligibles in all income categories above \$40,000. The direction of this effect is consistent with the presence of substitution, and the coefficient is statistically significant for families with incomes above \$75,000 but not for families with incomes between \$40,000 and \$75,000.

NEW EVIDENCE ON SUBSTITUTABILITY. We now turn to a new test, using the SIPP data described above. The intuition behind our test is that if 401(k)s and IRAs are substitutes, 401(k)-eligible families whose IRA deductibility was restricted in 1986 should have funneled more money into 401(k) plans between 1987 and 1991 than eligible families whose deductibility was not removed, controlling for other factors.

To measure this effect, we pool the 1987 and 1991 samples of 401(k)-eligible households from the SIPP and estimate a tobit model with real 401(k) balances as the dependent variable. The independent variables are the Xs used above, with non-401(k) net worth added as an additional term.⁷⁷ Also included is a dummy variable for being in the 1991 sample,

75. Poterba, Venti, and Wise (1992).

76. Poterba, Venti, and Wise (1994, table 6).

77. Results are very similar if the non-401(k) net worth term is omitted.

Table 9. Tobit Estimates of 401(k) Balances^a

<i>Independent variable</i>	<i>Coefficient^b</i>	<i>Standard error</i>	<i>t-statistic</i>
Age	579.70	249.66	2.32
Age-squared	-4.39	2.92	-1.50
Income	0.16	0.06	2.69
Income-squared	-240.46	232.76	-1.03
Age-income interaction	177.51	108.33	1.63
Non-401(k) net worth	0.02	0.00	8.49
12 years of education	970.89	1,118.78	0.86
13-15 years of education	1,803.67	1,179.46	1.52
16 years or more of education	3,300.42	1,176.64	2.80
Defined benefit pension coverage	-3,048.59	603.58	-5.05
Married	1,531.91	1,081.87	1.41
Male	969.04	786.32	1.23
White	5,873.44	935.96	6.27
Two earners	-1,612.15	750.98	-2.14
Family size	-1,325.96	261.26	-5.07
1991 sample dummy (<i>IN91</i>)	4,256.59	877.50	4.85
Tax change	-2,014.07	1,204.47	-1.67
Interaction term [<i>IN91</i>] (<i>tax change</i>)	2,537.89	1,173.2	2.16
Constant	-30,504.20	5,431.57	-5.61
Standard error of regression	20,636.64	237.94	. . .

Source: Authors' estimates from a tobit model using data from SIPP.

a. The dependent variable is real 401(k) balances (1991 dollars). Occupational dummies are omitted from the table. Assets and income are expressed in 1991 dollars. Variables are measured in the following units: income-squared divided by 10⁹, and age-income interaction divided by 10⁵.

b. The sample pools 1987 and 1991 families that are eligible for a 401(k) plan. Sample size is 6,046. The log likelihood is -47,511.479.

a dummy for having the tax deductibility of IRA contributions removed in 1986 (this is called *tax change*),⁷⁸ and an interaction term between the two dummies. The interaction term shows the effects on 1991 401(k) balances of having tax deductibility removed, after controlling for family characteristics, general differences between the 1987 and 1991 samples, and general differences between the affected group and the rest of the sample.

Table 9 shows the results. Balances in 401(k) accounts rise with income, age, non-401(k) net worth, and education. The coefficient on the interaction term is \$2,538 and statistically significant at conventional

78. This variable equals one if the family is single, had real income above \$35,000, and had a pension, or if the family is married, had real income above \$50,000, and had a pension. These conditions were satisfied by 46 percent of eligibles in 1987 and 39 percent in 1991.

levels. Thus, among 401(k)-eligible families, the prohibition against deducting IRA contributions raised 401(k) balances, even after controlling for other family characteristics such as income, pension coverage, and non-401(k) net worth. These results imply that there is substitution between 401(k) plans and IRAs.

One can gauge the importance of substitution in at least two ways. First, the average real 401(k) balance among participants (eligibles) rose by about \$4,900 (\$4,300) between 1987 and 1991 for the group for whom the tax deductibility of IRA contributions was removed. Hence, the removal of deductibility can account for half or more of the increase in 401(k) balances in this group. Second, calculations using aggregate 401(k) balances in the 1987 and 1991 SIPP and the coefficients and sample in table 9 suggest that the removal of IRA deductibility could account for up to 17 percent of the increase in 401(k) balances over this period.

Table 9 also provides evidence on the substitutability between defined benefit pension plans and 401(k) balances. Having a defined benefit pension reduces 401(k) balances by about \$3,000 and is statistically significant at conventional levels. This implies that there is substitution between 401(k) plans and defined benefit pension plans. Thus, several analyses, using a variety of different tests, show evidence of substitutability between IRAs and 401(k)s and between pensions and 401(k)s.

New Evidence on IRAs and Household Saving

Previous research has generated a variety of results concerning the effects of IRAs on saving.⁷⁹ Glenn Hubbard, using cross-sectional data, finds that IRA contributors have higher ratios of net worth to income than noncontributors, controlling for household characteristics.⁸⁰ Daniel Feenberg and Jonathan Skinner use interest and dividend data from the IRS-Michigan Tax Panel, which is described below, to infer that, in each of several ranges of initial taxable wealth, taxpayers with IRAs raised their taxable financial assets by more than those without IRAs between 1980 and 1984.⁸¹ However, tabulations from the tax panel also

79. In addition to the papers mentioned below, other analyses of IRAs include Andrews and Bradford (1988), Collins and Wyckoff (1988), Feldstein (1992), Galper and Byce (1986), Gravelle (1991), Huizinga (1991), O'Neil and Thompson (1987), and Skinner (1992).

80. Hubbard (1984).

81. Feenberg and Skinner (1989).

show that *incomes* of IRA contributors are considerably higher than those of noncontributors, controlling for initial taxable wealth, so that contributors also consume more than noncontributors with the same initial taxable wealth. In any case, it is unclear whether these papers provide evidence that IRAs increase saving or merely reflect the presence of heterogeneity in saving behavior.

Douglas Joines and James Manegold present regressions from the tax panel data that suggest that reductions in consumption financed less than 1 percent of IRA contributions for the full sample and 30.5 percent of contributions for households with imputed taxable wealth below \$25,000.⁸²

Venti and Wise estimate formal models of IRAs and saving using various data sets. They find that between 3 and 20 percent of the IRA contributions that would have accompanied increases in the IRA contribution limit would be financed by reductions in other saving. They conclude that little substitution of IRAs and other saving occurs and that contributions to IRAs represent substantial net saving increases.⁸³

Gale and Scholz develop a dynamic model of IRAs and saving based on utility maximization in the presence of uncertainty.⁸⁴ Estimates of the model using data from the 1983–86 Survey of Consumer Finances imply that 2 percent or less of the increased IRA contributions that would have resulted from limit increases would represent net additions to national saving. Gale and Scholz also show that IRA holders generally, and contributors who are at the limit particularly, have substantial amounts of taxable assets to shift into IRAs.

Tax Panel Data on IRAs and Saving

The data for our analysis of IRAs come from the 1979–88 IRS-Michigan Tax Panel and include almost all items on taxpayers' 1040 forms as well as many items from supplemental schedules.⁸⁵ Because we analyze IRA contributors over time, we follow Feenberg and Skinner as well as Joines and Manegold in restricting the sample to taxpayers that appear

82. Joines and Manegold (1991).

83. Venti and Wise (1986, 1987, 1990, and 1991).

84. Gale and Scholz (1994).

85. The sampling procedures and data are described in Slemrod (1988, 1990), Christian and Frischmann (1989), and Joines and Manegold (1991).

in the data in every year, which is 5,315 tax filing units in the ten-year panel.⁸⁶

As emphasized by Joines and Manegold, a strength of this data set is that it provides repeated observations of taxpayers' IRA contributions over a period of statutory variation in IRA eligibility and limits. Panel data and variation in statutory limits allow us to examine the effects of IRAs on the saving of the same households over time. Previous research using cross-sectional data sets has inferred the effects of IRAs from variations in contribution limits across households at a given point in time. These variations, arising solely from differences in marital status and earnings, may be correlated with saving behavior, which complicates inferences about the effects of IRAs.

The potential limitations of the data are threefold. First, because data on asset balances are not available, we cannot calculate saving directly. Instead, following previous researchers, we infer asset balances by capitalizing reported taxable interest and dividend receipts using average interest rates and dividend yields for the appropriate year.⁸⁷ This introduces measurement error and may bias estimates to the extent that the measurement error is correlated with other characteristics included in the regression analysis.⁸⁸ To reduce measurement error, we average data over periods when statutory IRA provisions are constant. Second, tax returns do not provide information on several household characteristics, such as age and education, that may affect saving.⁸⁹ To mitigate this problem, we estimate models that eliminate time-invariant household characteristics. Third, there is nonrandom attrition from the balanced panel.⁹⁰ We have not examined the role that attrition plays in our analysis.

Table 10 provides information on selected characteristics of taxpayers from 1979 to 1986. The first row shows that mean and median real

86. Feenberg and Skinner (1989) and Joines and Manegold (1991).

87. Scholz (1992, p. 281) finds that no systematic biases are induced when dividend yields are capitalized to infer equity holdings in his study of dividend clienteles using the 1983 Survey of Consumer Finances.

88. Capitalizing taxable interest and dividend payments will understate wealth for very high-income taxpayers, as they hold a disproportionate share of their wealth in tax-exempt bonds (Feenberg and Poterba, 1991). The direction of bias created in the regressions is uncertain.

89. Age exemptions on the tax return indicate only whether the taxpayer or spouse is older than 65.

90. Christian and Frischmann (1989).

Table 10. Selected Characteristics of IRA Contributors and Other Taxpayers, 1979–88
1986 dollars, except as noted

<i>Characteristic</i>	<i>IRA contributors</i>	<i>Non-contributors</i>	<i>All taxpayers</i>
<i>Average income, 1979–88^a</i>			
Mean	46,230	24,164	31,682
Median	38,487	20,635	26,132
<i>Average taxable wealth, 1979–81^b</i>			
Mean	31,016	22,729	25,552
Median	5,789	261	1,110
<i>Average non-IRA saving, 1979–81^c</i>			
Mean	2,224	541	1,114
Median	107	0	0
<i>Average non-IRA saving, 1982–86^c</i>			
Mean	5,340	3,430	4,081
Median	854	34	125
<i>Average years of IRA contributions, 1982–86</i>			
Mean	3.2	0	1.1
Median	3	0	0
<i>Average IRA contributions, 1982–86</i>			
Mean	1,680	0	572
Median	1,418	0	0
<i>Average nonmortgage liabilities, 1979–81^d</i>			
Mean	6,294	3,221	4,268
Median	1,368	0	0
<i>Average nonmortgage liabilities, 1982–86^d</i>			
Mean	8,690	4,569	5,973
Median	3,725	126	1,395
Number of taxpayers	1,811	3,504	5,315

Source: Authors' calculations using data from the IRS-Michigan Tax Panel data set, and the *Economic Report of the President 1993*, tables B-69 and B-91, for Aaa bond yield and average dividend yield, respectively.

a. Average income is adjusted gross income plus IRA contributions averaged from 1979 to 1988.

b. Average taxable wealth is the average value of capitalized interest and dividends from 1979 to 1981. Taxable wealth is constructed by capitalizing reported dividends by the average dividend yield and interest receipts by the Aaa bond yield.

c. Average non-IRA (gross) saving from year x to year y is defined as taxable wealth in year $y + 1$ minus taxable wealth in year x divided by $(y + 1 - x)$.

d. Average nonmortgage liabilities are calculated by capitalizing personal interest and investment interest deductions.

average income, or, following Slemrod, time-exposure income, is higher for IRA contributors than it is for taxpayers who did not contribute to IRAs.⁹¹ Contributors also had greater wealth before 1982 and, as discussed above, save more than noncontributors in all forms. However, contributors' debt also increased more rapidly, which is consistent with

91. Slemrod (1992b).

Table 11. Comparison of Saving by Continuing and New IRA Contributors^a
1986 dollars

	<i>Average saving, 1979–81</i>	<i>Average saving, 1982–86</i>	<i>Difference over time</i>
Continuing IRA contributor ^b	2,071 (1,461)	8,877 (3,649)	6,806 (2,188)
New IRA contributor ^c	2,532 (85)	6,560 (2,185)	4,028 (2,100)
Group difference	–461 (1,376)	2,317 (1,464)	. . .
Difference in difference	–2,778 (–88)
Never contributed ^d	513 (0)	3,390 (33)	2,877 (33)

Source: Authors' calculations from the IRS-Michigan Tax Panel data set.

a. Medians are shown in parentheses. Gross saving in year t is defined as the difference in taxable wealth in year $t + 1$ and year t plus IRA contributions in year t . Continuing IRA contributors are those that contributed to IRAs prior to 1982. New IRA contributors are those that did not contribute prior to the period of universal eligibility but contributed at least one year between 1982 and 1986. The remaining group is those taxpayers who never contributed to an IRA.

b. Sample size is 348.

c. Sample size is 1,515.

d. Sample size is 3,452.

some IRA contributions being financed by borrowing. Large differences in means and medians in each row reflect a skewed wealth distribution. Accordingly, we present both mean and median regressions below.

Did IRAs Increase Household Saving?

Our first test of the effects of IRAs on household saving relies on intuition discussed in Joines and Manegold, who note that if IRAs stimulate saving, then those who were made newly eligible for IRAs in 1982 should have increased their saving relative to their saving prior to IRA eligibility. Because the change in saving before and after 1981 was presumably also affected by changes in interest rates, tax rates, and other factors, Joines and Manegold suggest a “difference in difference” approach: if IRAs stimulate saving, the difference in saving by new contributors before and after the 1981 tax reform should exceed the difference in saving by those that were always eligible.

Table 11 shows that the mean and median saving of continuing contributors were \$461 lower and \$1,376 higher, respectively, than that of new contributors for 1979–81. When eligibility was made universal, new

contributors raised their mean saving by \$4,028 and their median saving by \$2,100. These increases, however, were smaller than the \$6,806 and \$2,188 increases for continuing contributors. Mean and median saving, therefore, rose \$2,778 and \$88 less for new contributors than for continuing contributors. These results provide no support for the proposition that universal IRA eligibility stimulated saving among new contributors.

There are some caveats, however. First, the standard errors in the regression analog to table 11 are large.⁹² Second, the “control group” for the analysis—continuing contributors—was not covered by an employer-provided pension plan from 1979 to 1981. Because pension coverage may be correlated with saving behavior, the comparison of continuing and new contributors may not yield accurate estimates of the effects of IRAs on saving. In short, heterogeneity complicates the analysis of IRAs and saving.⁹³

The tax panel allows us to account for heterogeneity in a more flexible way than in previous work (other than Joines and Manegold).⁹⁴ Because the data follow the same taxpayers over time, we estimate models that eliminate the effects of time-invariant characteristics that are not observed in the data, such as education or taste for saving. In addition, the data span three periods of statutory variation in the limits on deductible IRA contributions, which helps us to identify the effects of the limit changes on household saving. Before 1982, only workers not covered by employer-provided pensions were eligible for IRAs. From 1982 to 1986, all workers could make deductible IRA contributions. Following the 1986 tax reform, married (single) workers with incomes above \$40,000

92. A regression of saving in 1979–81 and 1982–86 on a constant, a dummy variable for the 1982–86 period, a dummy variable for being a new contributor, and the interaction of the two dummy variables yields the “difference in difference” estimate of the saving effects of IRAs. The standard error of the OLS regression is 2,272, the standard error of the median regression is 161. Difference-in-difference estimates that incorporate a number of additional household characteristics yield nearly identical results on the key interaction term.

93. Gale and Scholz (1994) and Venti and Wise (1988), among others, show that taxpayers who save tend to save in all forms. For example, in table 11, new contributors increased mean saving by \$1,151 more than noncontributors, but this comparison reveals little about the effects of IRAs because, as shown in table 10, noncontributors had lower incomes, had fewer assets and liabilities, and saved less in all forms than contributors. Thus, drawing inferences about IRAs by comparing the saving of new IRA contributors and noncontributors would be inappropriate.

94. Joines and Manegold (1991).

(\$25,000) had IRA deductibility curtailed if they were covered by an employer-provided pension. We aggregate the data into three periods corresponding to these periods of statutory IRA variation.

Our baseline model is a standard fixed-effects regression:

$$S_{it} = \alpha_i + \beta'Z_{it} + \epsilon_{it},$$

where S_{it} is gross saving and α_i is the individual-specific effect that is taken to be constant across all periods, t , and is specific to the cross-sectional unit i . Z_{it} is a vector of time-varying household characteristics and ϵ_{it} is an independently and normally distributed random error term.

Gross saving in period t is defined as taxable wealth (the capitalized sum of interest and dividend receipts) in period $t + 1$ minus taxable wealth in period t plus IRA contributions in period t . A number of variables are included in Z . Deductible personal and investment interest payments are capitalized to impute a measure of liabilities for households that itemize.⁹⁵ We expect that higher levels of personal and investment interest expense will be positively correlated with gross saving. We also include dummy variables that show whether the taxpayer used the home mortgage interest deduction. To the extent that homeowners are older than renters, the expected sign of the variable is positive (recall we do not observe age in the tax panel). Gary Englehardt suggests, however, that households may save rapidly to meet the downpayment constraint on a house, and, upon purchase, consumption increases sharply.⁹⁶ This suggests that for some taxpayers gross saving might be negatively correlated with owning a home.

Adjusted gross income (AGI) and AGI-squared are also included. The level of saving should be positively correlated with AGI. In cross-sectional work, it is difficult to independently identify the effects of taxes and income because tax rates are determined by income and demographic characteristics. The tax panel, however, spans a period of large statutory changes in marginal tax rates, which helps us distinguish tax from income effects. The marginal tax rate, which reduces the net

95. We use the Baa corporate bond rate from table B-69 in the *Economic Report of the President 1993* to calculate liabilities. In 1987, 65 percent of personal interest was deductible; in 1988, the number was 40 percent. Mortgage debt could also be used for investment purposes, but because we have no way of including housing wealth in the definition of saving, we exclude mortgage interest from our measure of liabilities.

96. Englehardt (1991).

return to saving, should be negatively correlated with saving. Interpretation is still problematic, however, as tax rates are endogenous to saving decisions and, in some circumstances, high marginal tax rates may increase the incentive to save.⁹⁷

We include the number of child exemptions claimed on the tax return and dummy variables for those filing a single return or claiming an age exemption. Conditioning on other factors, we normally expect families with more child exemptions to save less, but, because we do not condition on age, the number of children may pick up life-cycle patterns of saving and be positively correlated with saving. All else being equal, we expect single taxpayers to save more than married taxpayers, while elderly taxpayers presumably save less.

We assume the IRA contribution limit was \$1,500 for single taxpayers who contributed to an IRA before 1982, \$3,000 for a married couple who contributed,⁹⁸ and \$0 for everyone else. From 1982 to 1986, single taxpayers could contribute \$2,000, couples with one earner could contribute \$2,250, and a two-worker household could contribute the lesser of actual earnings or \$2,000 for each worker. The 1986 tax reform limited deductible contributions based on the taxpayer's income, so the limits on deductible contributions are assumed to be unchanged for married (single) taxpayers with incomes below \$40,000 (\$25,000), are phased out over the next \$10,000 interval, and are \$0 for all other taxpayers. We restrict the sample to taxpayers who contributed to an IRA at least once during the period spanned by the data, because it is difficult to interpret the limit variable for taxpayers who do not contribute to IRAs.⁹⁹ Appendix table A2 gives sample statistics for these variables.

We estimate the model in first-differenced form.¹⁰⁰ The first panel of

97. Slemrod (1992a, p. 254).

98. Both spouses could make the maximum contribution (\$1,500) if each had income above \$1,500. We do not know the earnings of the lower-earning spouse before 1982 and after 1986, so we assume that couples are eligible to make the maximum contribution.

99. Restricting the sample to IRA contributors raises sample selection issues. Treating selection in a panel model is complicated, however (see Hsiao, 1986, pp. 198–202). We do not address this complication here.

100. Differencing eliminates the fixed effect and makes it computationally simpler to estimate quantiles. We deviate from the strict representation of the model by including a constant term in the estimated specification and, in the two-observation model (the difference of three periods), a dummy variable for the second difference.

Table 12. Mean and Median Regression of Changes in Gross Saving on Changes in IRA Limits, 1979–87

<i>Regression type and independent variable^a</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>t-statistic</i>
<i>OLS estimates^b</i>			
Liabilities	0.04	0.15	0.239
Home owner	-1,955.56	7,643.39	-0.256
Adjusted gross income (AGI)	0.63	0.08	7.705
AGI-squared ^c	-16.10	1.97	-8.190
Number of children	-1,255.81	3,854.19	-0.326
Single	15,417.26	11,369.36	1.356
Age exemption	18,901.28	11,486.91	1.645
IRA limit	0.20	1.45	0.140
Marginal tax rate	-742.99	332.39	-2.235
Second period dummy	-12,194.29	7,233.53	-1.686
Constant	832.81	5,308.89	0.157
<i>Median regression estimates^d</i>			
Liabilities	0.01	0.01	1.039
Home owner	-629.26	223.50	-2.815
Adjusted gross income (AGI)	0.01	0.03	0.289
AGI-squared ^c	-2.26	29.40	-0.077
Number of children	-271.00	127.24	-2.130
Single	778.37	475.56	1.637
Age exemption	-1,744.25	1,127.62	-1.547
IRA limit	0.31	0.07	4.246
Marginal tax rate	31.39	31.83	0.986
Second period dummy	-1,842.74	209.40	-8.800
Constant	1,088.05	286.78	3.794

Source: Authors' regressions using data from the IRS-Michigan Tax Panel data set.

a. All variables are first differenced.

b. The estimates are from a first-differenced fixed-effects model. The sample size for the regression is 3,880, and the regression has an R^2 of 0.0189.

c. Adjusted gross income-squared is divided by 10^8 .

d. The estimates are from a first-differenced fixed-effects model. The sample size for the regression is 3,880, and the pseudo- R^2 is 0.0178. (The pseudo- R^2 is defined as the sum of the weighted deviations about the estimated median divided by the sum of the weighted deviations about the raw median, all subtracted from one.) Standard errors are estimated with bootstrap methods.

table 12 presents the mean (OLS) regression estimates of the fixed-effects model. The coefficient estimates generally have the expected signs. Saving is positively related to AGI, though at a decreasing rate, and marginal tax rates are negatively related to saving. None of the other coefficients is precisely estimated. The coefficient on the primary vari-

able of interest, the limit variable, is 0.20, which suggests that a \$1 increase in the IRA limit increases gross saving by 20 cents. This effect, however, is estimated very imprecisely.

The second panel shows median regression estimates. Most of the signs of the coefficients are similar to the OLS specification, though the age exemption variable becomes negative, as we originally expected, and the marginal tax variable becomes positive, which is unexpected. The coefficient on the limit variable increases to 0.31 and is precisely estimated.¹⁰¹ This suggests that at the medians of the observed distributions, 31 percent of IRA contributions represent increases in gross private saving. To gauge the effect on national saving requires an adjustment for the loss in tax revenue owing to the deductibility of IRA contributions. Accurate long-run calculations of the revenue loss of IRAs are complicated, as they require knowledge of current and future tax rates and information on the source and tax treatment of the funds used to finance the IRA. Short-run calculations are more straightforward. The mean and median marginal tax rate for IRA contributors was 27 percent during 1979–87. This suggests that about 4 percent of IRA contributions represented increases in national saving during 1982–86.

There may be problems with treating the 1986 tax changes in IRA deductibility in a symmetric fashion with the extension of universal IRA eligibility. First, in 1986 only high-income taxpayers with pensions had deductibility restricted; all taxpayers remained eligible to make contributions. Second, one of the advantages of collapsing the data into periods corresponding to the tax treatment of IRAs is the presumption that averaging measures of taxable wealth reduces the measurement error associated with the wealth imputations. The last period in the sample, however, is composed of only one year. For these reasons, we estimate the model in difference form using data from the 1979–81 and 1982–86 periods in table 13.¹⁰²

101. We bootstrap the standard errors for the median regression because we are concerned about heteroscedasticity.

102. This specification differs from Joines and Manegold (1991) in several respects. We use two additional years of data (1985 and 1986). Their specification includes time-exposure income and average taxable wealth in 1979–81. Both these period-invariant characteristics difference out in a fixed-effects regression, so we include only time-varying characteristics. Last, we estimate the model with both mean and median regressions.

Table 13. Mean and Median Regression of Changes in Gross Saving on Changes in IRA Limits, 1979–86

<i>Regression type and independent variable^a</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>t-statistic</i>
<i>OLS estimates^b</i>			
Liabilities	0.13	0.07	1.885
Home owner	-202.02	3,173.88	-0.064
Adjusted gross income (AGI)	0.15	0.06	2.709
AGI-squared ^c	18.70	1.92	9.743
Number of children	-2,007.96	1,450.15	-1.385
Single	4,467.63	4,132.56	1.081
Age exemption	-9,879.60	5,136.18	-1.924
IRA limit	-0.71	0.72	-0.993
Marginal tax rate	-448.57	158.05	-2.838
Constant	4,260.58	2,221.61	1.918
<i>Median regression estimates^d</i>			
Liabilities	-0.04	0.02	-2.109
Home owner	-1,464.03	413.10	-3.544
Adjusted gross income (AGI)	0.13	0.05	2.396
AGI-squared ^c	20.40	7.86	2.602
Number of children	-626.17	118.28	-5.294
Single	1,533.62	464.88	3.299
Age exemption	-2,283.86	2,258.78	-1.011
IRA limit	0.09	0.13	0.682
Marginal tax rate	-186.66	64.76	-2.883
Constant	1,158.70	344.32	3.365

Source: Authors' regressions using data from the IRS-Michigan Tax Panel data set.

a. All variables are first differenced.

b. The estimates are from a first-differenced fixed-effects model. The sample size for the regression is 1,940, and the regression has an R^2 of 0.0565.

c. Adjusted gross income-squared is divided by 10^7 .

d. The estimates are from a first-differenced fixed-effects model. The sample size for the regression is 1,940, and the pseudo- R^2 is 0.026. (The pseudo- R^2 is defined as the sum of the weighted deviations about the estimated median divided by the sum of the weighted deviations about the raw median, all subtracted from one.) Standard errors are estimated with bootstrap methods.

Table 13 shows that dropping the last period raises the explanatory power of the ordinary least squares regression. The coefficient on the changes in IRA limits is -0.71 but is estimated imprecisely. In the median regressions, the coefficient on IRA limit changes is 0.086 and is also imprecisely estimated.¹⁰³

103. The dependent variable in the regression is calculated by capitalizing interest and dividend income. Interest and dividends are also included in the measure of AGI. If we use AGI less interest and dividend income instead of AGI, the effects of IRAs on saving are smaller in each of the specifications in tables 12 and 13.

Additional Issues with IRAs and Saving

The regression results above suggest that only a small portion of IRA contributions, if any, represented increases in national saving during 1982–86. This result is consistent with the full-sample findings of Joines and Manegold, using a different extract from the same data source, and with those of Gale and Scholz, who use the 1983–86 Survey of Consumer Finances and a different empirical approach.¹⁰⁴ Venti and Wise, however, find that IRA contributions are largely new saving and that increases in limits would significantly increase household saving.¹⁰⁵ Venti and Wise and Richard Thaler suggest that standard economic theory cannot explain the effects of IRAs on household saving and propose interpretations of the Venti and Wise results that emphasize psychological factors and the promotion of IRAs.¹⁰⁶ In this subsection, we discuss three aspects of these issues.

The first was raised by Feenberg and Skinner, who report that 39 percent of all IRA contributors in 1985 were “falsely constrained,” which they define as a taxpayer who contributed exactly \$2,000 but was eligible to contribute more.¹⁰⁷ This result has been used as evidence of a widespread misunderstanding of the contribution limits that taxpayers face. Venti and Wise, for example, acknowledge that while transaction costs associated with opening a spousal account may explain a portion of this behavior, “it is likely that the promotion, in which the amount \$2,000 figured prominently, played a key role.”¹⁰⁸

After examining data from the tax panel, however, it is clear that nowhere near 40 percent of all IRA contributors are “falsely constrained.” For example, using the 1985 tax panel cross-section, we find that 19 percent of all IRA contributors contributed exactly \$2,000 when they could have contributed more.¹⁰⁹

In table 14, we examine patterns of “false constraints” during the pe-

104. Joines and Manegold (1991) and Gale and Scholz (1994).

105. Venti and Wise (1986, 1987, 1990, and 1991).

106. Venti and Wise (1992) and Thaler (1994).

107. Feenberg and Skinner (1989).

108. Venti and Wise (1992, p. 33).

109. In private correspondence, Jonathan Skinner reports that programming errors were found in the original calculations.

Table 14. Distribution of IRA Contribution Amounts, 1982–86
Number of contributors

Contribution	Number of taxpayers				Total
	Taxpayers with \$2,000 limit	Taxpayers with \$2,250 limit	Taxpayers with \$2,251–\$3,999 limit	Taxpayers with \$4,000 limit	
\$2,000 (not at limit)	. . .	410	64	401	875
Other nonlimit contribution	296	250	79	647	1,272
Limit contribution	964	1,210	161	1,368	3,703
Total	1,260	1,870	304	2,416	5,850

Source: Authors' calculations using Michigan-IRS Tax Panel data set.

riod of universal IRA eligibility.¹¹⁰ The unit of observation is annual IRA contributions. That is, if a taxpayer contributed to an IRA three times between 1982 and 1986, he or she appears as three observations in the table. Table 14 shows that in our sample there were 875 IRA contributions of exactly \$2,000 that could have been more. This represents 15 percent of all contributors and 19 percent of the 4,590 contributors who faced limits above \$2,000. To place bounds on the possible importance of false constraints, we made two additional calculations.

First, we found that 156 of the 875 contributions were preceded by a contribution larger than \$2,000 by the same taxpayer. These taxpayers obviously were aware that larger contributions were allowed. Thus, at the very most, 12.3 percent (719/5,850) of all contributors and 15.7 percent (719/4,590) of married taxpayers (all of whom have limits above \$2,000) were falsely constrained.

Second, a plausible assumption is that married taxpayers who were eligible for less than a \$4,000 limit and contributed \$2,000 did so purposefully.¹¹¹ Transaction costs presumably play a large role in explaining why couples do not open a spousal IRA of less than \$2,000.

110. Taxpayers that are (i) single, (ii) married but filing separately, (iii) unmarried heads of household, or (iv) surviving spouses are given a \$2,000 contribution limit. Taxpayers filing joint returns where the second-earner credit exceeded \$100 in 1982 or \$200 in 1983–86 are given a contribution limit of \$4,000. (The two-earner deduction increased from 5 percent in 1982 to 10 percent in 1983–86.) All other joint filers are given a limit of the larger of either \$2,250 or \$2,000 plus the earnings of the lower-earning spouse.

111. Of these taxpayers, 86 percent (410/474) were eligible for an additional contribution of only \$250.

There are costs to opening and keeping separate accounts, closing such accounts, and understanding the rules governing spousal IRA contributions.¹¹² If we assume that the only taxpayers who are falsely constrained are those that (i) contribute \$2,000, (ii) have a \$4,000 limit, and (iii) did not contribute more than \$2,000 to an IRA in any previous year, then 5.7 percent of all IRA contributors and 7.2 percent of married contributors were falsely constrained. These findings cast significant doubt on the quantitative importance of “false constraints” in IRA decisionmaking.

The second claim is that a bunching of IRA contributions occurs immediately preceding April 15, the last day on which deductible IRA contributions can be made for the previous tax year.¹¹³ But the forgone benefits of waiting until the end of the tax year are small, and, because IRA balances are illiquid, it may be perfectly reasonable for households to postpone locking up funds until they have enough information to make an appropriate decision.

The third issue focuses on investor behavior following the Tax Reform Act of 1986. James Long suggests that IRA contributions fell by more than would have been predicted given the changes in deductibility rules.¹¹⁴ Venti and Wise write, “This ‘overreaction’ is at least in part attributable to widespread misunderstanding of the legislation (often reported at the time to have eliminated IRAs) and to the marked decline in the promotion of IRAs.”¹¹⁵ If misinformation caused the overreaction following the 1986 tax reform, it is surprising that IRA contributions have not rebounded. Brokerage firms and other marketers of IRAs have every incentive to reduce taxpayer confusion, yet contributions to IRAs have *fallen* in each year following the act (see table 1).

There are several other possibilities for the decline in IRA contributions that have received less attention in the literature. First, lower marginal tax rates after 1986 reduced the benefits of tax-free interest accrual. Second, restricted deductibility and the compression of marginal tax rates reduced the tax benefits further. Both effects should reduce the

112. Burman, Cordes, and Ozanne (1990) provide an excellent discussion of the false constraints issue. They note the fact that couples are composed of two decisionmakers whose objectives are interrelated but not necessarily identical, which may lead to married couples purposefully choosing to contribute exactly \$2,000.

113. Summers (1986).

114. Long (1990).

115. Venti and Wise (1992, p. 34).

number of contributors. Third, the increased availability of 401(k) plans and other tax-preferred saving options may have lessened the desire to contribute to IRAs.¹¹⁶ Fourth, to the extent that IRAs are funded through transfers of existing assets or new saving that would have occurred anyway, contributions may have declined even in the absence of the 1986 tax reform. We leave the difficult project of disentangling the role of these and other explanations to future research. Finally, we note that none of the suggested psychological factors—false constraints, bunching of contributions, or reaction to the 1986 tax reform—has any specific implications for whether IRAs raise saving.

Saving Incentives in a Stochastic Life-Cycle Simulation Model

In this section, we examine saving incentives in the context of a stochastic life-cycle simulation model. The simulation model presented here utilizes estimates of behavioral and economic characteristics to develop quantitative and testable predictions for saving behavior and its response to government policies. These predictions can then be compared with data to check the model. The model can also help determine whether certain interpretations of the data are appropriate. The simulations demonstrate that many observed empirical patterns concerning IRAs, 401(k) plans, and saving can be generated by a well-specified model of utility maximization.

Simulation analysis is particularly helpful for analyzing policies when data are limited. For example, 401(k)s and IRAs have been implemented for only relatively short periods of time. Simulations can distinguish short-term and long-term effects of saving incentives, while empirical analysis is currently limited by data capturing only short-term effects.

Model Description

Conventional life-cycle models assume certainty regarding future income and other factors, so saving is primarily for retirement. Some re-

116. Contributions to variable annuities appear to have grown rapidly since 1986, but reliable data are difficult to find.

cent research has stressed “buffer stock” saving where people save to offset uncertain shocks to income.¹¹⁷ The model in this paper incorporates both motives for saving by extending the standard life-cycle framework to include stochastic earnings and uncertain lifespan. Individuals save for retirement and as a precaution against downturns in future earnings and the possibility of outliving one’s assets once retired.¹¹⁸ The model consists of five parts: (i) consumers’ preferences, (ii) consumers’ budget constraints, (iii) the stochastic economic environment, (iv) government finance, and (v) the overlapping generations framework.¹¹⁹

People enter the model at age 21 and face an age-varying probability of dying. The maximum lifespan is 90 years. In each period, individuals maximize expected lifetime utility by making consumption and portfolio choices.¹²⁰ There are two portfolio options: a tax-preferred saving incentive asset or a conventional, fully taxed asset. The following assumptions make the model more tractable while still capturing important elements of household saving decisions. Labor supply and retirement decisions are assumed to be exogenous. Utility is assumed to be separable over time, and, within a time period, separable among consumption, leisure, and an exogenously provided government good. The utility function for consumption is assumed to exhibit constant relative risk aversion (constant intertemporal elasticity of substitution), which implies that uncertain future income leads to precautionary saving. Using household data from the Panel Study of Income Dynamics (PSID) and the Consumer Expenditure Survey, Engen estimates an intertemporal elasticity of substitution of 0.33 and a time preference rate of 0.04, which are used here to specify household-level preference parameters.¹²¹

Consumption and portfolio choices are subject to a lifetime budget

117. See Carroll (1992) and Deaton (1991).

118. Hubbard, Skinner, and Zeldes (1993a, 1988b) also develop a model with precautionary saving against uncertain lifespan and uncertain income and includes retirement saving. Their model adds uncertain health expenditures and a government-provided minimum consumption level.

119. The basic structure of the model is described in Engen (1993b, 1993c) and Engen and Gale (1993). A technical appendix that describes this model in detail is available from the authors upon request.

120. There is no intentional bequest motive in the model. All bequests are accidental and arise from assets held at the unexpected time of death. In each period, bequests are distributed equally to each member of the generation that is 45 years old.

121. Engen (1993a).

constraint, net asset constraints,¹²² uncertain future earnings, and uncertain lifespan.¹²³ Consumption can be financed by after-tax labor earnings before retirement; by partially indexed, earnings-based annuity income from social security and private pensions after retirement; and by assets in saving incentive accounts or conventional forms. Both of the latter types of assets earn the same nonstochastic pretax rate of return. The return on conventional assets is fully taxed. Contributions to saving incentive accounts are tax deductible and constrained by annual limits. Contributions and investment earnings are not taxed until withdrawn. A penalty is imposed on funds withdrawn before the account holder reaches age 60.¹²⁴

Household earnings are modeled as the sum of a certain and an uncertain component. The certain component is a mean age-earnings path based on estimates by Engen of a log earnings regression using panel data on earnings and other characteristics for employed heads of household between the ages of 21 and 65 from the PSID. Estimates were made separately for three groups with different levels of educational achievement. Age-earnings profiles for college graduates rise more steeply and peak later than for other groups. The stochastic process for earnings innovations is modeled as an AR(1) process using data from the IRS-Michigan Tax Panel.¹²⁵

The government runs a balanced budget in the initial steady-state equilibrium. All revenues go toward purchasing a government-provided good that is allocated equally to all individuals and is constant over time. There is a progressive income tax structure, similar to the U.S. system in 1989, with tax brackets of 15, 28, and 33 percent. Individuals are allowed a personal exemption and the standard deduction, as well as de-

122. In the model, saving incentive accounts and conventional asset balances are constrained to be nonnegative. This is a tighter constraint than requiring net worth to be nonnegative, but Engen and Gale (1993) show that these constraints are realistic for IRAs. The constraints rule out uncollateralized loans and financing saving incentive accounts with a negative position in conventional assets, but they do not restrict collateralizing debt, reshuffling of existing assets into a saving incentive account, or borrowing against existing assets and placing the funds in a saving incentive account.

123. There is assumed to be no private market for purchasing insurance against risky income or private annuities (except pensions) to insure against uncertain lifespans.

124. Some features of actual saving incentive plans, such as the loan provisions of some 401(k) plans, are not modeled here.

125. Engen (1993b).

ductions for IRA or 401(k) contributions. There is a 20 percent estate tax on accidental bequests in the model.

Because the model does not have a closed-form solution, and the analytical solution would be intractable, we use a numerical method to solve individuals' consumption-saving problem. Earnings innovations for each individual over the life cycle are simulated with a random number generator. Individuals each receive different earnings shocks and thus end up with different realized earnings, consumption, and saving. Calculating means or medians for each age gives representative profiles for consumption, saving, wealth, and earnings. The model sets 25 percent of the population in each age cohort as high school dropouts, 50 percent as high school graduates, and 25 percent as college graduates, which approximates the observed distribution of educational attainment. Cohorts of different ages are incorporated into an overlapping generations framework that accounts for mortality and annual population growth of 1 percent in order to calculate aggregate saving and assets. The model is not general equilibrium because gross wages and gross returns to capital are held constant when the capital stock changes in response to government policy.

Comparing the Model's Implications with Empirical Patterns

To help assess the usefulness of the model, we consider the extent to which the model can match empirical patterns of consumption, saving, and wealth.¹²⁶ At the aggregate level, the model allocates approximately 75 percent of gross income to labor and 25 percent to capital. Aggregate tax revenue (and government spending) is approximately one-third of output. The aggregate saving rate is almost 6 percent. Asset-income ratios are approximately 3.6. These figures are broadly consistent with U.S. historical experience.

At the individual level, the model generates hump-shaped consumption-age profiles. Consumption tracks income more closely in this model than in certainty models, and simulated consumption is more sensitive to income shocks in this model than in a certainty equivalence model. All of these are well-documented features of actual consumption data.

126. Engen (1993b) and Engen and Gale (1993) document the comparisons mentioned below. Engen (1993b) also shows that the comparisons analyzed in that paper are not sensitive to reasonable changes in parameter values.

The model generates life-cycle wealth profiles that are consistent with estimated profiles from microdata for much of the population. However, the model does not reflect the fact that some households never accumulate significant amounts of assets.¹²⁷ For individuals aged 25–65, the model generates mean holdings of about \$119,000 and median asset holdings of \$77,000. In the 1986 SCF, mean assets are \$111,000 and median assets are \$41,000, for the 25–65 age group. Saving rates generally increase with age before leveling off before retirement.

An important parameter for interpreting the effects of taxes on saving is the after-tax interest rate elasticity of saving.¹²⁸ The implied (uncompensated) saving elasticity in the model is between 0.15 and 0.35, consistent with empirical estimates, which generally fall between zero and 0.4.

In a previous study of IRAs using a related model, Engen and Gale report the following comparisons between simulated and empirical patterns for IRA saving.¹²⁹ First, like the data, the model shows IRA participation rising with age and income. Second, the model overstates the overall probability of holding an IRA, relative to 1986 SCF data, by about 10 percentage points. But IRAs are the only retirement tax shelter in the model, and the proportion of households that hold an IRA, Keogh, or 401(k) in 1986 SCF data is virtually the same as the proportion holding an IRA in the model. Third, the percentage of IRA contributors who contribute at the limit in the model is very similar to that observed in the data. Fourth, the probability of a taxpayer making an IRA contribution in one year conditional on having contributed in the previous year is approximately the same as that calculated from the data.

Long-Run Effects of Saving Incentives

We now turn to calculating the long-term effects of saving incentives on private and national saving. IRAs are modeled as having tax-deductible contributions, a limit of either \$2,000 or \$4,000, and an early withdrawal penalty of 10 percent. 401(k) plans are modeled as having a limit of \$9,000. The withdrawal penalty is set either at 10 percent or 100 per-

127. See Hubbard, Skinner, and Zeldes (1993a, 1993b).

128. Engen (1993c).

129. Engen and Gale (1993).

cent.¹³⁰ Employers are assumed to match employee contributions at the rate of 50 percent up to the first 5 percent of salary and to provide no unconditional contributions. Two alternative assumptions about wage offsets resulting from employer matching contributions are employed. In the first scenario, wages are offset on an individual basis; each worker's wages are reduced by the amount of the employer match for that employee. This essentially removes any income or substitution effects of the employer match. In the second and probably more realistic scenario, wages are offset on an average basis; each worker's wages are offset by the average employer match. Thus, a worker making the average 401(k) contribution ends up with no inframarginal (or income) effect from the matching contribution, yet a marginal substitution effect is still operative. Workers making above-average 401(k) contributions will not have their individual wages fully offset and, in fact, are subsidized by workers making below-average contributions, whose wages are reduced by more than the employer match that they receive.

In the results reported below, the government is not required to run a balanced budget after the implementation of a saving incentive plan. In other words, spending on the government-provided good is held constant so that the revenue shortfall from the tax deductibility of contributions to saving incentives creates a budget deficit. However, we also discuss the effects of simultaneously introducing a saving incentive and either reducing government spending or raising other tax revenues so as to keep the budget in balance.

Table 15 reports the long-run effects of these saving incentive plans on various measures of saving and wealth. The long run is defined as the length of time necessary for everyone in the model to have had access to saving incentives for their entire life. It takes 70 years in this model to reach the long-run equilibrium. The first two rows show that IRAs have a relatively small effect on long-run saving and wealth. The level of national saving increases by about 3 to 5 percent. With an initial saving rate of 5.9 percent, changes of this magnitude lead to a new saving rate (not shown in table 15) of about 6.2 percent, at most. This is also reflected in the asset-to-income ratio, which is 3.6 in the steady state without saving

130. As described above, early withdrawals from 401(k)s are subject to a 10 percent penalty if allowed, but are only allowed in certain situations, so the effective penalty is between 10 and 100 percent.

Table 15. Long-Run Effects of Saving Incentives

Percent, except as noted

<i>Saving incentive</i>	<i>Asset-income (ratio)^a</i>	<i>Participation rate^b</i>	<i>Contributions that are net saving</i>	<i>Increase in national saving</i>
<i>IRA^c</i>				
\$2,000 annual limit	3.69	56.6	22.0	3.3
\$4,000 annual limit	3.76	56.6	30.0	5.4
<i>401(k)^d</i>				
<i>Wage offset per individual</i>				
<i>10 percent penalty</i>				
on early withdrawal	4.11	56.6	50.1	15.0
<i>100 percent penalty on</i>				
early withdrawal	3.80	45.1	86.3	7.9
<i>Wage offset on average</i>				
<i>10 percent penalty on</i>				
early withdrawal	4.19	63.4	51.0	17.2
<i>100 percent penalty on</i>				
early withdrawal	3.89	52.2	87.4	9.1

Source: Authors' simulations as described in the text.

a. Initial (before saving incentives) asset-to-income ratio is 3.61, and saving rate is 0.059.

b. Participation of population aged 21–65.

c. Penalty for early withdrawal is 10 percent of the interest and principal.

d. Annual contribution limit is \$9,000.

incentives and 3.7 or 3.8 in the steady state with IRAs, depending on whether the contribution limit is \$2,000 or \$4,000, respectively. In the long-run equilibrium, over half of households headed by individuals aged 21–65 choose to hold IRAs, and between 22 and 30 percent of funds contributed in the 70th year following the introduction of IRAs and in future years represent new national saving.

The bottom two panels show that 401(k) plans, as modeled here, have stronger positive effects than IRAs on long-run national saving. This occurs for two reasons. First, the higher contribution limit means that more savers face increased rates of return on saving at the margin. Second, when employer matching contributions are assumed to reduce wages on an average basis, the income effect of the employer matching contributions is removed but the substitution effect remains, providing a positive stimulus to saving.

With 401(k)s, the two lower panels show that long-term national saving rises by between 8 and 17 percent, so the saving rate increases from

5.9 percent to a range of 6.4 to 6.9 percent (not shown in table 15). The asset-to-income ratio rises from 3.6 to between 3.8 and 4.2. When early withdrawals from a 401(k) account are prohibited, effectively raising the penalty to 100 percent, the proportion of 401(k) contributions that represent new saving increases substantially. Nevertheless, far fewer contributions are made,¹³¹ so the effect on national saving with no early withdrawals is about half as big as the effect with a 10 percent penalty on early withdrawals.

When employers reduce wages on an average basis, the effects of 401(k) plans are slightly larger than when employers reduce wages on an individual basis. The former removes only the income effect of the employer match, while the latter removes both the income effect and the substitution effect. Since the substitution effect raises saving, removing it via the individual wage offset reduces the effects of 401(k)s on saving.

In the long run, both IRAs and 401(k) plans can stimulate national saving. These results are robust to reasonable changes in model parameters.¹³² However, the increase in the saving rate is small relative to recent declines in national saving.

Table 15 shows the results of saving incentives 70 years after they have been implemented. However, as Alan Auerbach and Laurence Kotlikoff, Kenneth Judd, and others have shown, the transitional effects of changes in the taxation of all capital income can be in a different direction than the long-term effects.¹³³ Engen and Gale note that this phenomenon is even more likely to occur for targeted saving incentives because of the possibility of asset substitution to finance saving incentive contributions.¹³⁴

Short-Run and Transitional Effects of Saving Incentives

To model the short-run and transitional effects of saving incentives, we assume that the implementation of such plans is completely unanti-

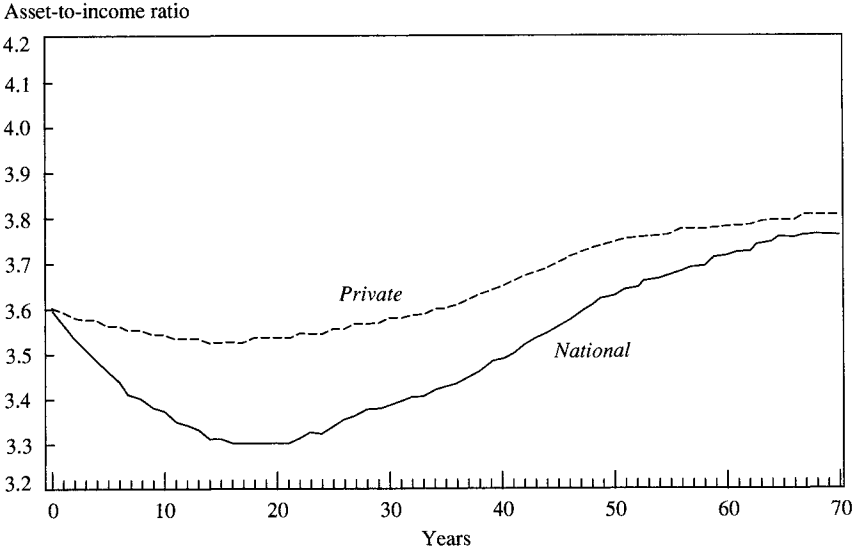
131. In the new long-run equilibrium, 401(k) plans with 10 percent penalties would account for about 30–35 percent of saving, while 401(k) plans with 100 percent penalties would account for only about 10 percent of saving. By comparison, IRAs would be about 15 to 18 percent of saving.

132. For further sensitivity analysis of the long-run effects of IRAs and other tax policies, see Engen and Gale (1993) and Engen (1993c).

133. Auerbach and Kotlikoff (1987) and Judd (1985).

134. Engen and Gale (1993).

Figure 4. Aggregate Asset-to-Income Ratio with an IRA Program Introduced at Time Zero



Source: Authors' simulations as described in the text. The annual contribution limit is \$4,000 and the penalty for early withdrawal assessed on both earnings and principal is 10 percent.

pated but permanent. The transitions show the effects of saving incentives after each year of their implementation.

Figure 4 shows the transitional effects for IRAs with a \$4,000 contribution limit. The figure plots the ratios of national and private wealth to income, with the difference between them being the ratio of government debt to income (when private wealth exceeds national wealth). Following the implementation of the IRA program, private saving falls to a small degree. The private asset-to-income ratio falls slowly but steadily from 3.6 in year 0 to 3.5 in years 14 to 18. It does not return to its pre-IRA value until year 36. In the short run, of course, national saving falls by more than private saving, because of the deductibility of IRA contributions. The government asset-to-income ratio starts at zero in year 0, falls for the next 20 years, and then rises steadily until almost returning to zero by year 70. Hence, the national asset-to-income ratio falls from the start and does not regain its initial value of 3.6 until year 49.

The intuition behind these numbers is straightforward. Government revenue falls as the initial shifting of assets to saving incentive accounts reduces the tax base; this occurs because contributions reduce taxable

income and because asset income generated in saving incentive accounts is no longer taxed. Initially there are many more tax-deductible contributions than taxable withdrawals. Over time, those who reach retirement age begin to withdraw funds, which raises government revenues in later years.

The intuition behind the initial drop and subsequent recovery in private saving is perhaps more interesting. Before IRAs are established, households hold all of their assets in taxable forms. The opportunity to shift some of those funds into tax-deferred accounts is valuable, even if it means tying up the funds in an illiquid account. Hence, substantial amounts of funds are shifted in the years immediately following the introduction of a saving incentive. Over time, however, as households shift some of their portfolios, they may run out of funds to shift¹³⁵ or become less likely to give up additional liquidity in exchange for tax preferences. Hence, the extent of shifting falls, and more of the contributions represent new saving. Recall that the model generates a positive, but small, interest elasticity of saving. This is a persistent impetus in favor of IRAs raising private saving, but its effect is swamped in the first decades of the program by the shifting of already existing assets into tax-preferred accounts. Over time, the backlog of initial assets to be shifted falls, more households find that saving incentives affect the marginal decision to save, and the results of the small but positive saving elasticity can be seen more clearly.¹³⁶

Figure 5 shows the transitional effects when 401(k)s are instituted, with the assumption that employer matching contributions offset wages on an individual basis. The results are qualitatively the same as those for IRAs.¹³⁷ The higher limit causes a larger short-run drop in personal saving because more asset shifting occurs earlier. Even so, the transition takes a long time. It takes 22 years for the private asset-to-income ratio to return to its original level and 35 years for the national asset-to-income ratio to return to its original level.

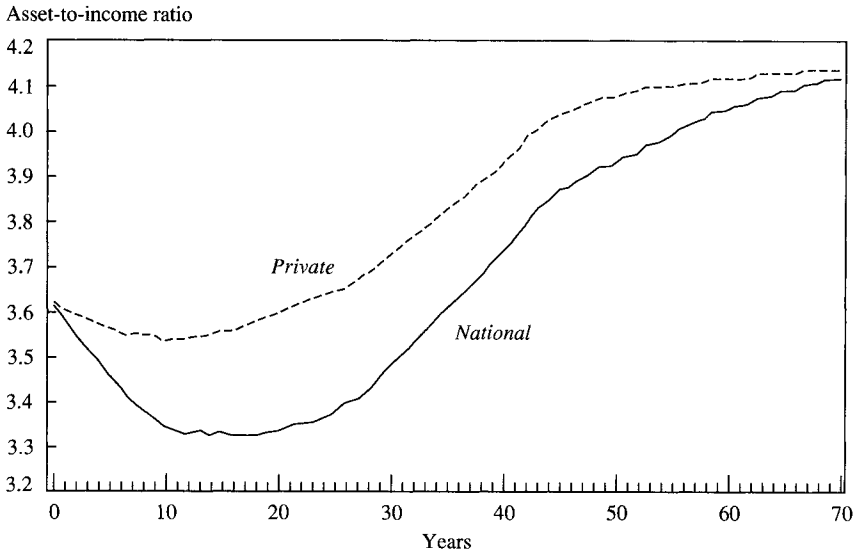
In the transitional results presented in figures 4 and 5, the government is allowed to run a deficit when revenues fall as a consequence of insti-

135. This scenario has been put forth by Feldstein and Feenberg (1983).

136. Results with a limit of \$2,000 are qualitatively similar, except that, since the limit is lower, the initial asset-shifting phase takes longer, and the positive effects on national saving are ultimately smaller (table 15).

137. Results using the individual offsets and a 100 percent penalty, and using the average offsets and either the 10 percent or 100 percent penalty, are qualitatively similar.

Figure 5. Aggregate Asset-to-Income Ratio with a 401(k) Program Introduced at Time Zero



Source: Authors' simulations as described in the text. The annual contribution limit is \$9,000, and the penalty for early withdrawal assessed on both earnings and principal is 10 percent. Wages are offset for employer contributions on an individual basis.

tuting a saving incentive program. These results can provide information on the transitional effects on saving of the government running a balanced budget instead. If budget balance were achieved by reducing spending on the government-provided good, there would be no effect on private consumption and saving decisions since the government good is assumed to be separable in individual utility functions. Therefore, private asset accumulation would not change, but the government deficit would disappear. As a consequence, national saving would simply track the time path for private saving in figures 4 and 5. Both national and private saving would fall in the early years of the saving incentive program and would increase only in the long run.

Alternatively, if budget balance were achieved by raising all income tax rates, the deficit would again disappear. However, raising all income taxes would have the likely effect of reducing the level of private saving.¹³⁸ Thus, the level of private asset accumulation shown in the figures

138. Note the positive elasticity of saving with respect to the after-tax return embedded in the model. Also, see Engen (1993c).

provides an upper bound for the level of national and private saving that would occur.

Allowing for general equilibrium effects would dampen both the initial decline and the subsequent increase in the asset-to-income ratio. As the capital stock fell in the initial phase of the transition, interest rates would start to rise, which would offset some of the decline in saving. As the positive effects of saving incentive programs became evident in the longer term, interest rates would start to fall as the capital stock rises, which would diminish the long-run increase in saving. Thus, the partial equilibrium results tend to overstate both the short-term losses and the long-run gains from saving incentive programs. However, the partial equilibrium changes in the capital stock are small, especially in the initial phase of the transition, so that the differences between general equilibrium and partial equilibrium results would be meager.

Additional Points

The results in figures 4 and 5 are consistent with the empirical results presented earlier. Available data on saving incentives are limited to the initial phase of the transition. Universal eligibility for IRAs was established in 1982, the same year that the number of 401(k) plans began to grow rapidly. Hence, the data sources used above represent saving information from only the first nine years of these programs. At that stage of the transitions in the simulations, the effects of saving incentives on private saving are essentially zero and the effects on national saving are negative.

One concern with using this model to study saving behavior is that it is clear that not all economic agents optimize along the lines presented above. This should come as no surprise. All economic models are necessarily false in that they involve extracting the most important elements of a situation and omitting others. Nonetheless, the model reproduces accurately many aspects of real world economic behavior. Moreover, economic models are used to ensure that statements are internally consistent. In the case of saving incentives, intuition suggests that when a saving incentive program is first introduced, people will shift taxable assets into tax-preferred assets.¹³⁹ After a time, this effect will diminish and a greater percentage of the contributions will represent net saving.

139. Feldstein and Feenberg (1983).

The extent of this shifting and the time it takes depend critically on the parameters of households' preferences and opportunities (for example, time preference rates, earnings uncertainty, mortality risk, and tax rates). The model allows us to quantify these effects in a way that is consistent with what is known about these parameters.

It is clear that the model does not accurately capture the saving behavior of all households, and most notably misspecifies important aspects of how low-income, low-wealth households make saving decisions. For some issues, this would be a critical flaw.¹⁴⁰ For the analysis of voluntary saving incentives, it is probably not. Participants in saving incentives typically have many times the financial wealth of nonparticipants (as shown in table 3), and low-wealth households that do participate contribute less than other participants. As documented above, the model is able to capture many features of IRA contribution behavior. Moreover, the model implies that the contributions of low-income households are predominantly new saving.¹⁴¹ Thus, to the extent that the model overstates contributions by this group, it overstates the positive effects of saving incentives on saving.

Conclusion

Over the past 20 years, the United States has experimented with a variety of tax-deferred saving incentive plans. Although the plans are intended to raise the saving rate, they are designed so that a contributor does not need to have positive saving to exploit the tax advantages. Because they are capped by annual limits, the subsidies do not provide incentives to save at the margin for some households. Moreover, the tax deductibility of contributions reduces short-run public saving.

The aggregate data show no correlation between contributions to saving incentive plans and a variety of measures of personal saving. Using survey data, we generate a test of the effects of saving incentives by using two groups that can each be described as "savers" and by exploiting the simultaneous expansion of 401(k) opportunities and the contraction

140. For example, Hubbard, Skinner, and Zeldes (1993a, 1993b) show that the behavior of low-income people can be modeled much more accurately if one imposes a government-provided, means-tested consumption floor.

141. Engen and Gale (1993).

of IRA incentives that occurred after 1986. If 401(k) contributions are, at least in part, new saving, and our comparison group is a good one, the wealth of 401(k) participants should have risen relative to that of IRA participants not eligible for 401(k) plans from 1987 to 1991. A variety of specifications showed no economically or statistically significant increase in wealth for 401(k) participants relative to IRA participants not eligible for 401(k) plans. Further, the results show that non-401(k) wealth fell for 401(k) participants relative to the wealth of IRA participants not eligible for 401(k)s. We also show that these two groups are similar in important respects. Finally, we show that the results hold when other groups are compared and when different asset measures are used as well. Thus, the regression results are consistent with the view that 401(k) plans have not raised private saving. Clearly, if they have not raised private saving, 401(k) plans have not raised national saving.

Tests using the tax panel revealed little evidence that IRAs substantially raised private saving. After accounting for the decline in tax revenues, the estimates imply little, if any, positive effect on national saving. We also presented new evidence that false constraints are substantially less important than previous evidence suggests and that there appears to be substitution between IRAs and 401(k) plans and between pensions and 401(k) plans.

These findings are consistent with an analysis of saving incentives in a stochastic, life-cycle simulation model. Model results indicate that saving incentive plans reduce private and national saving in the short run (0–20 years) because of substantial shifting of assets or saving that would have occurred anyway from taxable into tax-preferred accounts. Over much longer periods, the simulation model indicates that saving incentives are capable of raising national saving and wealth.

APPENDIX A

Sample Statistics for Saving Regressions

THIS APPENDIX presents more detailed summary statistics for several of the regressions discussed in the main text. Table A1 provides background data for table 5 in the main text. Table A2 does the same for tables 12 and 13 in the main text.

Table A1. Sample Statistics for the Regression in Table 5^a

<i>Variable</i>	<i>Mean</i>	<i>Standard deviation</i>	<i>Median</i>	<i>Minimum</i>	<i>Maximum</i>
Net financial assets ^b	35,710	96,950	14,454	-1,855,480	2,356,381
Age (years)	44	10.6	43	25	64
Age-squared	0.203	0.095	0.185	0.063	0.410
Income ^b	53,672	30,032	47,637	-4,695	294,510
Income-squared ^b	0.378	0.482	0.227	0	8.674
Age-income interaction ^b	0.239	0.156	0.202	-0.027	1.708
12 years of education	0.314	0.464	...	0	1
13-15 years of education	0.217	0.412	...	0	1
16 years or more of education	0.384	0.486	...	0	1
Defined benefit pension	0.540	0.498	...	0	1
Married	0.707	0.455	...	0	1
Male	0.759	0.427	...	0	1
White	0.923	0.267	...	0	1
Two earners	0.436	0.495	...	0	1
Family size	2.75	1.37	3	1	11
Family member with 401(k) (<i>PART</i>)	0.553	0.497	...	0	1
1991 sample dummy (<i>IN91</i>)	0.514	0.499	...	0	1
Interaction variable [<i>(IN91) (PART)</i>]	0.343	0.475	...	0	1

Source: Authors' calculations using data from the SIPP.

a. The sample pools 1987 and 1991 families that either (i) have a 401(k) plan or (ii) have an IRA but are not eligible for a 401(k) plan. The sample size is 7,341.

b. 1991 dollars.

Table A2. Sample Statistics for the Regressions in Tables 12 and 13^a

<i>Table and variable</i>	<i>Mean</i>	<i>Standard deviation</i>	<i>Median</i>	<i>Minimum</i>	<i>Maximum</i>
<i>Table 12^b</i>					
Saving ^c	576	144,934	450	-2,042,087	6,392,784
Liabilities ^c	2,447	15,626	0	-159,662	249,346
Average gross income (AGI) ^c	4,700	78,576	1,763	-485,835	4,446,515
AGI-squared ^c	0.619 ^d	31.8 ^d	0.00473 ^d	6	1,980 ^d
Number of children	-0.132	0.622	0	-3.8	4
Single	-0.019	0.217	0	-1	1
Age exemption	0.053	0.206	0	-1	1
IRA limit ^c	713	2,552	1,470	-4,248	4,248
Marginal tax rate	-3.82	7.67	-3.53	-45.2	30.8
<i>Table 13^b</i>					
Saving ^c	4,528	40,447	1,844	-548,577	735,448
Liabilities ^c	2,372	12,773	227	-113,210	212,146
Average gross income (AGI) ^c	1,181	22,505	812	-431,783	258,629
AGI-squared ^c	0.0508 ^d	0.0496 ^d	0.00397 ^d	6	18.6 ^d
Number of children	-0.129	0.650	0	-3	4
Single	-0.039	0.237	0	-1	1
Age exemption	0.053	0.184	0	-0.333	1
IRA limit ^c	2,667	1,298	2,389	-1,865	4,248
Marginal tax rate	-2.62	7.49	-2.40	-41.8	30.8

Source: Authors' calculations using data from the IRS-Michigan Tax Panel data set.

a. Table entries refer to the first-differenced values of the variables.

b. Sample size for table 12 is 3,880; sample size for table 13 is 1,940.

c. 1986 dollars.

d. Divided by 10¹⁰.

Comments and Discussion

B. Douglas Bernheim: The paper by Eric Engen, William Gale, and Karl Scholz launches a frontal assault on the notion that tax incentives stimulate significant new saving. The authors deserve credit for devising some novel and clever analytical innovations that shed new light on an old question. I congratulate them on a thorough and careful analysis of the available data.

I confess to having my own reservations about previous studies of saving incentives, including (but not limited to) those that purported to establish the existence of large beneficial effects. I have always regarded the existing evidence, such as that by James Poterba, Steven Venti, and David Wise, as interesting and suggestive but not conclusive, and certainly open to other interpretations.¹ Unfortunately, reading this paper, I feel much the same way, despite the authors' laudable efforts. As far as I am concerned, the jury on saving incentives is still out. This is not, however, because I am entirely agnostic about our ability to answer this question through empirical analysis. Rather, I question whether the type of data that have been analyzed to date can ever provide a reliable answer to the central policy question.

The paper by Engen, Gale, and Scholz is divided into three sections. The first section is an analysis of 401(k)s based on the SIPP data. The second section examines the effect of IRAs using the IRS-Michigan Tax Panel. The third section investigates the effects of tax incentives in a life-cycle simulation model. I will discuss each in turn.

Evidence from the SIPP on 401(k)s

Using data from the SIPP, the authors attempt to determine whether participation in a 401(k) affects total saving. They are correct in noting

1. Poterba, Venti, and Wise (1992, 1994).

that one cannot infer the effect of 401(k)s by comparing the behavior of participants and nonparticipants, since those with strong preferences for saving will choose both to participate and to save more. This point is widely acknowledged in the literature on saving incentives. The central questions with respect to sample selection are whether it is possible to control for sample selection in a way that yields convincing results and whether Engen, Gale, and Scholz accomplish this objective.

Like Poterba, Venti, and Wise, the authors examine data on assets obtained from a series of cross-sectional data sets. The surveys in question were conducted in 1984, 1987, and 1991 (unfortunately, data on 401(k)s are not available for 1984). By comparing cross-sectional asset patterns across years, they hope to identify the effect of 401(k)s. In particular, since 401(k)s were “newer” in 1987 than in 1991, they presumably could not have had as much of an effect on the level of the cross-sectional asset profile. Thus, under the hypothesis that 401(k)s increase saving, Engen, Gale, and Scholz argue that one would expect to see an upward migration of the asset profile—relative to some appropriate control group—between 1987 and 1991.

Table 3 of the paper demonstrates that various measures of median wealth (financial assets, net financial assets, and net worth) declined for 401(k) participants and that this decline exceeded the reduction in median wealth for nonparticipants. Superficially, this would appear to contradict the hypothesis that 401(k)s stimulated saving. Yet the authors have already admonished us about the differences between participants and nonparticipants, and they are therefore justifiably reluctant to make too much of this result. It is, for example, difficult to imagine a significant decline in the median net financial assets of nonparticipants, given that they start out (in 1987) with nothing. Clearly, participants and nonparticipants have different underlying attitudes toward saving.

Engen, Gale, and Scholz note, however, that there may be identifiable subgroups of nonparticipants whose attitudes toward saving are similar to those of participants. In particular, they point out that IRA participants also tend to be high savers. Moreover, eligibility for IRAs was significantly curtailed in the Tax Reform Act of 1986. Consequently, they argue that it is appropriate to use IRA participants who are not eligible for 401(k)s as a control group when evaluating the shift in the estimated asset profile between 1987 and 1991.

A direct comparison of the subsample medians in tables 3 and 4

(which contains summary statistics for the control group) would not, of course, control for differences in age, income, education, and other characteristics. Consequently, the authors estimate a median regression, which describes net financial assets as a function of household characteristics (table 5). They also allow participation in a 401(k) plan to affect both the level of the age-wealth profile (the coefficient on *PART*) and the shift in the age-wealth profile between 1987 and 1991 (the coefficient on *(IN91)(PART)*). Their central finding is that the shift in the age-wealth profile for 401(k) participants is not larger, and indeed may be smaller, than the shift in the age-wealth profile for the control group.

The central problem with this approach is that, just as there is heterogeneity across various identifiable subgroups (such as participants and nonparticipants), there is also heterogeneity within these subgroups, and, moreover, the composition of these subgroups may change through time. For this reason, the authors' methodology continues to be flawed by the same sample selection issues that have plagued this entire literature.

In principle, there are excellent reasons to believe that the composition of 401(k) participants changed significantly over the relevant period. It stands to reason that the most motivated "serious" savers would start participating as soon as these accounts became available at their firms. Less motivated, "occasional" savers are much more likely to pass on their first opportunities to open tax-favored saving accounts. However, it also seems likely that, with the passage of time, more occasional savers would sign up. 401(k)s may have become especially common for less serious savers after 1986, because of the more demanding nondiscrimination requirements that were established for private pension plans. Consequently, between 1987 and 1991, serious savers may represent a declining fraction of 401(k) participants. In that case, the authors' result would simply reflect dilution of the 401(k) participant group with less serious savers.

Likewise, it is also possible that the composition of the Engen, Gale, and Scholz control group changed through time. The curtailment of IRAs significantly reduced new account formation and may have increased account termination. If termination is particularly likely among occasional savers, then serious savers may represent an increasing fraction of IRA participants between 1987 and 1991 ("reverse dilution").

In the final revision of their paper, Engen, Gale, and Scholz acknowledge these criticisms and attempt to deflect them. Their analysis of the issue is unconvincing, however. In part, their response is based on the suggestion that, in theory, there may be considerations that work against dilution. Despite these theoretical considerations, it strikes me as unlikely that serious savers represent a constant or increasing fraction of 401(k) participants through time.² In any case, I do not believe that this issue can or should be resolved on purely theoretical grounds.

Engen, Gale, and Scholz present evidence that, they claim, fails to support the hypothesis that either the composition of 401(k) participants or the composition of their control group changed significantly between 1987 and 1991. In effect, they estimate wealth profiles separately for the two subsamples, allowing in each case for the possibility that the profiles may have shifted between 1987 and 1991. With respect to 401(k)s, they find that the financial asset profile shifted upward between 1987 and 1991, while the net financial asset profile changed little, if at all. They assert that these findings do not support the view that the 401(k) participant group became diluted with more occasional savers between 1987 and 1991. With respect to the control group, they find that neither the financial asset profile nor the net financial asset profile shifted significantly between 1987 and 1991. They argue that these patterns do not support the hypothesis that the control group experienced reverse dilution over the relevant time period.

I am puzzled by this analysis. The authors' methodology is predicated on the assumption that the preferences of 401(k) participants, and the preferences of the control group, did not change between 1987 and 1991. This is an identifying restriction. It appears to me that Engen, Gale, and Scholz are, in effect, attempting to test their identifying restriction without the addition of new information. If they have in mind some formal model that allows for group-specific effects and time-specific effects, as well as for the possibility that the composition of various population subgroups changes through time, and that nevertheless per-

2. For example, they argue that while some occasional savers were opening 401(k) accounts, others were cashing out their accounts. But unless initiation and termination probabilities changed through time in some particularly serendipitous fashion, one would expect to observe monotone convergence of participation rates from an initial value of zero to some positive steady-state value.

mits one to infer the effects of 401(k) plans on saving using only the data discussed in this section, then they should write this model down explicitly so that the skeptics among us can verify its validity.

For myself, I very much doubt that such a model exists. Indeed, the patterns noted by Engen, Gale, and Scholz are easily reconciled with the existence of the sample dilution effects described above. Consider their evidence on shifting preferences (dilution) for 401(k) participants. Their analysis is explicitly based on the premise that the wealth profile for 401(k) participants should have shifted upward (relative to an appropriate control group) through time. The dilution effect works in the opposite direction. Hence, if the net effect is positive, this does not imply that dilution is absent. Rather, it only proves that the behavioral effect is larger than the dilution effect. If the net effect is zero, one can only conclude that the two effects are offsetting; both may be very large, or both may be very small. One would expect to observe a downward shift in the wealth profile only if the dilution effect is larger than the behavioral effect.

Next consider the authors' evidence on shifting preferences (reverse dilution) for the control group. It is easily conceivable that, absent reverse dilution, the wealth profile for this group might have shifted downward between 1987 and 1991, possibly in response to changing macroeconomic conditions or because of differences in the accuracy of the surveys. (Indeed, Engen, Gale, and Scholz find that, overall, wealth is lower in 1991 than in 1987.) But in that case, the absence of a downward shift in the wealth profile is entirely consistent with reverse dilution.

Although the authors' evidence sheds little light on the importance of dilution, other evidence is suggestive. According to table 2 of the paper, 401(k) participation nearly doubled, from 13.7 percent of families in 1987 to 25.4 percent of families in 1991. The spread of 401(k)s was attributable to two factors. First, participation rates within plans rose from 62.7 percent to 69.8 percent. This is consistent with the notion that, through time, 401(k)s achieved greater penetration to less dedicated savers. Second, the fraction of families eligible for 401(k)s rose from 21.9 percent to 36.4 percent. Engen, Gale, and Scholz themselves make the case that eligibility is endogenous and strongly related to preferences for saving. This argument naturally implies that 401(k)s were first offered at companies where they were in the greatest demand.

Additional evidence on the importance of compositional effects can be gleaned from a more detailed analysis of 401(k) contributors. Suppose for the moment that it is possible to observe some variable that Engen, Gale, and Scholz did not control for in their empirical specification (table 5) and that is positively correlated with residual unobserved preferences toward saving. Then, under the dilution hypothesis, the average value of this variable among 401(k) contributors should have declined between 1987 and 1991.

It is possible to implement the test described in the preceding paragraph using data on IRA participation. IRA participation among 401(k) contributors is certainly observable; Engen, Gale, and Scholz explicitly argue that it is correlated with underlying preferences;³ and the authors did not control for IRA participation among 401(k) contributors in table 5. According to their paper, 46 percent of 401(k) participants also held IRA accounts in 1987. By 1991, this fraction had declined to 35 percent. This observation is consistent with the view that the composition of the 401(k) participant sample shifted toward less serious savers between 1987 and 1991 and that this mechanically reduced the upward movement in the estimated asset profile for this group. It does not, however, constitute clear proof, since the decline in IRA participation among 401(k) participants could, in principle, be explainable by changes in characteristics that the authors did control for. At my suggestion, Gale has investigated this issue. His preliminary results show that changes in observable characteristics explain roughly 70 percent of the decline in IRA participation among 401(k) participants, relative to a control group. The existence of a nontrivial unexplained residual is consistent with the dilution hypothesis. However, the results of this preliminary analysis are open to mixed interpretations, owing to the imprecision of the estimates.

In their work, Poterba, Venti, and Wise noted this shift away from IRAs among 401(k) participants. Their solution was to drop households with both IRAs and 401(k)s, comparing 401(k) participants without IRAs to nonparticipants. I do not endorse this procedure. The decline in IRA participation among 401(k) participants between 1987 and 1991 is merely a symptom of dilution. IRA participation is a highly imperfect

3. This appears to be true even among 401(k) participants, since those who participate in both 401(k)s and IRAs have significantly more wealth than those who participate in 401(k)s alone.

proxy for unobservable propensities to save. Removing IRA participants from the analysis does not homogenize either the 401(k) sample or the control group and therefore cannot preclude either dilution or reverse dilution.⁴

After presenting their preferred estimates, Engen, Gale, and Scholz also explore the Poterba, Venti, and Wise approach described in the previous paragraph. Surprisingly, their results differ from those of the previous authors. The source of the discrepancy is difficult to discern. On the one hand, Engen, Gale, and Scholz control for a larger set of household characteristics than do Poterba, Venti, and Wise. On the other hand, the latter authors allow the coefficients of the independent variables to differ for 401(k) participants and nonparticipants, whereas Engen, Gale, and Scholz do not. Regardless of the true explanation for the discrepancy, I would reiterate my conclusion that the exclusion of IRA participants treats a symptom of the dilution problem, without solving the problem itself. With respect to the Engen, Gale, and Scholz estimates, I would also add two observations. First, when the authors use this approach, their point estimates are generally more favorable to the hypothesis that 401(k)s stimulate saving than the point estimates based on their original approach. The sensitivity of their results is at least suggestive of an underlying sample selection problem. Second, the imprecision of the authors' estimates renders this portion of their analysis inconclusive. Oddly, Engen, Gale, and Scholz note that the estimates do not differ significantly from zero, but they fail to add that one is also unable to reject the hypothesis that all 401(k) contributions represent new saving.

Finally, Engen, Gale, and Scholz themselves present evidence suggesting that the dilution problem contaminates estimates based on their preferred approach. As can be seen in tables 3 and 4 in the paper, net worth includes both financial and nonfinancial assets. As long as both kinds of assets are related to some of the same unobservable characteristics that create preferences for saving, the dilution effect should be

4. The direction of the bias resulting from this procedure is not obvious. On the one hand, it may be the case that, among 401(k) participants without IRAs, unobserved preferences for saving are correlated with earlier participation. On the other hand, in light of the statutory limitations imposed on IRAs in 1986, the failure to have an IRA may be more indicative of a low propensity to save in 1987 than in 1991.

larger for net worth than for financial wealth. If dilution roughly offsets the behavioral effect of 401(k)s when attention is confined to financial wealth (as the authors' results suggest), then dilution should more than offset the behavioral effect of 401(k)s when both financial and nonfinancial assets are considered. This is precisely what one observes. Note that the net worth of 401(k) participants (including those who held IRAs) fell by \$22,111—more than 27 percent—between 1987 and 1991. In comparison, the net worth of IRA participants remained roughly constant.⁵ It seems to me that the relative decline in the net worth of 401(k) participants is difficult to explain in the absence of relative dilution. After all, it is very unlikely that the availability of 401(k)s depresses total saving.

In summary, I am not convinced that Engen, Gale, and Scholz have dealt effectively with the sample selection problems that have plagued previous attempts to measure the behavioral effects of 401(k)s and IRAs. Indeed, there is some evidence that the dilution effect is important and that it biases the approach preferred by the authors against the finding that 401(k)s stimulate saving. Certainly, other considerations may bias their results in the opposite direction. But I see no reason to believe that the aggregate effect of multiple biases would be offsetting.

Engen, Gale, and Scholz also use the SIPP data to explore the exogeneity of 401(k) eligibility. They conclude that eligibility is strongly related to underlying preferences for saving. This issue is important, since Poterba, Venti, and Wise have attempted to identify the effects of 401(k)s by using eligibility as an instrument. On this point, I see eye-to-eye with Engen, Gale, and Scholz. In fact, I have, in the past, criticized Poterba, Venti, and Wise's analysis on precisely these grounds.⁶ This does not, however, support the inference that 401(k)s are ineffective. Rather, I am simply left feeling very pessimistic about our ability to draw reliable inferences concerning the effects of 401(k)s from currently available data.

5. These comparisons do not control for changes in household characteristics between 1987 and 1991. Preliminary analysis by Gale suggests that roughly 60 percent of the difference-in-differences noted in my comparison is attributable to changes in household characteristics. The residual, which amounts to more than \$8,500, is unexplained and consistent with significant dilution.

6. Bernheim (1994a).

Evidence on IRAs and Household Saving

Using data from the IRS-Michigan Tax Panel, Engen, Gale, and Scholz attempt to determine whether IRA eligibility affects total saving. The availability of a true panel permits the authors to reduce the potentially adverse effect of sample selection problems by allowing for fixed effects. Even so, a number of serious problems remain.

First, the data contain no direct measure of asset balances. The authors infer wealth by capitalizing reported taxable interest and dividend receipts. As they acknowledge, this measure of wealth is somewhat suspect and certainly very noisy. Engen, Gale, and Scholz difference wealth to obtain a measure of saving. They then difference saving to eliminate fixed effects. Thus, the key dependent variable in their analysis is the second difference of a noisy construct. One must seriously question how much “news” is left over after these operations. Since the dependent variable is significantly related to certain independent variables, it presumably contains either some news or systematic noise. Although I am unable to identify an obvious source of systematic noise, I remain skeptical about the hypothesis that the second difference of capitalized interest and dividend income could contain much news. I am therefore not surprised that results on the effects of changes in IRA contribution limits are imprecise and, on the whole, inconclusive.

Second, the empirical specification is problematic. Although the authors include a variety of explanatory variables, they omit wealth, despite the fact that wealth varies with time. Both theory and existing evidence suggest that wealth is an important determinant of consumption, and therefore of saving. The inclusion of wealth as an independent variable would, however, raise more complicated econometric issues, since wealth is also used to construct the dependent variable. The authors also omit age (because of data limitations). One can think of the constant in the differenced specification as the coefficient of age (since the time difference of age is a constant). However, the existing evidence on age-wealth profiles suggests that rates of saving accelerate as households age. To capture this pattern, one would have to allow age to affect the first difference of saving. When age is omitted (as it is, of necessity, in the authors’ specifications), spurious relationships may appear between the second difference of wealth and any variable that is correlated with age.

The omission of variables like wealth and age is important because the authors' measure of the change in the IRA contribution limit is correlated with household characteristics. Prior to 1982, eligibility depended on private pension coverage, which is certainly related to wealth and age. After 1986, eligibility depended on adjusted gross income (AGI). Although the authors control for AGI, it is not clear that this control is adequate. Since they take differences, the independent variables include the differenced IRA limit, differenced AGI, and differenced AGI-squared. But the differenced IRA limit is related to the *level* of AGI in 1987, which does not appear as an explanatory variable in the differenced equation. It is easily conceivable that the level of AGI affects the acceleration of saving (for example, AGI is related to age). Consequently, the IRA limit variable may measure spurious effects.

None of the issues raised above suggests that the results should be biased systematically in any particular direction. I can think of stories that go both ways. Nevertheless, I am left with little confidence that the coefficient on the IRA contribution limit measures what it is intended to measure.

Engen, Gale, and Scholz also challenge the view that standard economic theory cannot explain the effects of IRAs on household saving. I regard this as a central issue and also discuss it below. If IRAs affected saving through nontraditional mechanisms, then the authors' analysis of the IRS-Michigan Tax Panel data would not necessarily pick this up. Suppose, for example, that the expansion of the IRA program in 1982 encouraged financial institutions to promote retirement saving and that this served an important educational function. Then all households may have increased their saving in response to the expansion of eligibility, regardless of whether they were previously eligible. This would invalidate inferences based on the comparisons in table 11 or the estimates in tables 12 and 13.

Engen, Gale, and Scholz discuss three different types of evidence that have been offered as support for the view that the response to IRAs was not grounded in intertemporal rationality. The first concerns the "false contribution limit." I myself have cited Feenberg and Skinner on this point many times, and I was disturbed to hear that their result is in error.⁷ Nevertheless, it seems to me that some intriguing patterns re-

7. Feenberg and Skinner (1989).

main. I am struck, for example, by the fact that, among those who could have contributed more than \$2,000 but who contributed less than the limit, 47 percent (875 out of 1,851) contributed exactly \$2,000. Even among those with a \$4,000 limit, 38 percent (401 out of 1,048) of those contributing less than the limit contributed exactly \$2,000.⁸ I think that Engen, Gale, and Scholz have correctly inferred from the data that few individuals were falsely constrained by misinformation. Yet the data do invite an interpretation that I have always favored: that the well-publicized, “officially endorsed” \$2,000 figure created a focal target for saving, and that the very existence of this target may have influenced the behavior of many less serious savers (such as those contributing less than the limit).

The second type of evidence concerns the bunching of IRA contributions at the end of the tax year. The authors suggest that households may delay contributions as long as possible to preserve liquidity. I am skeptical. As noted in table 4, in 1984, the typical IRA contributor had roughly \$13,000 in financial assets other than IRAs, and nearly \$93,000 of net wealth other than IRAs. It is hard to imagine that such individuals would benefit significantly from the additional liquidity associated with a one-year delay in an IRA contribution. Even if households are concerned about liquidity, it is hard to rationalize the fact that more of them do not make a series of smaller contributions to their IRAs as the year progresses. This would be particularly natural for individuals who pay estimated taxes on a quarterly basis.

The third type of evidence concerns the “overreaction” of IRA contributions to the Tax Reform Act of 1986. According to other analysts,

8. Incidentally, I do not agree that “(T)ransactions costs presumably play a large role in explaining why couples do not open a spousal IRA of less than \$2,000.” The costs of opening and maintaining an IRA are fairly trivial. Moreover, these one-time costs must be weighed not against the benefits of a \$250 contribution but against the benefits of a \$250 contribution that recurs for 10, 20, or 30 years. I am also very skeptical about the proposition that one could rationalize the \$2,000 contribution phenomenon by arguing that spouses have conflicting objectives. It is hard to imagine how this would emerge in a formal model of household bargaining, without the introduction of significant transaction costs. Moreover, it should be noted that, if saving results from interspousal bargaining, then behavioral theories of saving (such as the notion that decisions reflect the resolution of conflict between a “planner” and a “doer,” as in Shefrin and Thaler, 1988) may be more descriptive of actual choices, and of the effects of tax incentives, than the standard life-cycle theory.

the sharp decline in contributions was attributable to misinformation. Engen, Gale, and Scholz argue that, under the misinformation hypothesis, IRAs should have rebounded. Engen, Gale, and Scholz also suggest four alternative explanations for the sharp decline. The first two concern other features of the 1986 tax reform that reduced the tax incentives for saving through IRAs. I agree with Engen, Gale, and Scholz that, in principle, IRAs may have declined sharply because of these other changes and that this issue is not yet settled. However, their third and fourth explanations (the growth of 401(k)s and the depletion of non-IRA financial assets) concern gradual phenomena and cannot explain the suddenness of the decline in contributions. These phenomena can, however, account for the absence of a rebound under the hypothesis that misinformation was gradually corrected.

Life-Cycle Simulations

With respect to the final section of the paper, the basic question that I would like to raise is whether one should have much faith in policy simulations generated from highly sophisticated life-cycle models. I do not believe that these types of models are descriptive of either decision processes or behavior, and I am particularly suspicious of relying on these models to predict behavioral responses to changes in environmental parameters. Many of my views on this issue are laid out in a recent paper, but it is worth summarizing the central points here.⁹

Given the state of economic and financial literacy among the general population, one ought to be skeptical of the view that personal saving behavior is the consequence of highly sophisticated deliberation. Indeed, the evidence suggests that much of the population is ill-equipped to make even the most basic economic calculations. For example, only 20 percent of adults can determine correct change using prices from a menu,¹⁰ and many have trouble determining whether a mortgage at 8.6 percent is better than a mortgage at 8 3/4 percent.¹¹ My own analysis of

9. Bernheim (1994b).

10. Mary Jordan, "Literacy of 90 Million is Deficient: U.S. Survey Sounds Alarm over Reading, Arithmetic," *The Washington Post*, September 9, 1993, p. A1.

11. Albert B. Crenshaw, "For Too Many, Managing Money Isn't Child's Play," *Washington Post*, October 3, 1993, p. H1.

a recent survey sponsored by Merrill Lynch reveals that two-thirds of adults will not even hazard a guess as to the level of the Dow Jones average, despite the fact that this number is widely, prominently, and frequently reported by newspapers, television, and radio. Financial concepts, such as compound interest, are widely misunderstood,¹² despite the fact that they are at the very heart of the intertemporal maximization problem. According to the Merrill Lynch survey, roughly one-third of adults believe that \$1,000, left in the bank for 30 years at 8 percent, will accumulate to less than \$5,000 (whereas the correct answer is well over \$10,000). The number one source of information and advice on financial matters among baby boomers is parents. Advice from employers and the government, collectively, ranks behind prayer. Clearly, households are not solving the problem described by Engen, Gale, and Scholz's simulation model.

Of course, the authors' analysis only requires households to act "as if" they are solving the problem described by the simulation model. In general, I am sympathetic to "as if" arguments. However, these arguments are not universally applicable. They tend to be plausible in the context of frequently repeated decisions, where individuals can learn from trial and error. But most important activities related to saving are either not repeated or repeated very few times. Consider, in particular, the problem of saving for retirement. Virtually all individuals expect to retire only once. They do not have the opportunity to practice the life-cycle process. They do not have the luxury of learning appropriate rates of saving through trial and error. Rather, if they have overprovided or underprovided for retirement, they must live with their mistakes. There are no second chances.

In principle, individuals might also learn through vicarious experience, from observing others. But there are several problems with this. First, the experiences of others are difficult to observe. Financial information is generally regarded as private; a person's income and wealth are usually not thought of as appropriate subjects for general discussion. Second, vicarious information is necessarily either inconclusive or stale. On the one hand, if one learns about the decisions of those who have not yet retired, or who are not far enough along into the retirement period, one cannot observe the ultimate outcome of their financial

12. Ng (1992).

choices. On the other hand, if one acquires information about the very old, then one might be able to observe both decisions and consequences. However, in that case, the observations would be of little relevance since the economic environment changes dramatically over the span of decades.

It is perhaps conceivable that, despite the obstacles listed above, individuals do learn basic rules of thumb that produce approximately optimal decisions under appropriate economic conditions. For that reason, it is not too surprising that one can fit the data reasonably well by simulating a life-cycle model, as long as the model allows for sufficient parametric flexibility. However, there is an enormous difference between fitting observed behavior and fitting behavior out-of-sample after significant environmental changes. Even if individuals use an optimal rule of thumb in one economic environment, this does not mean that they will modify their rule of thumb optimally if the environment changes. Indeed, unless they understand the implications of the environmental change, neither personal experience nor vicarious experience will lead them to alter their retirement savings behavior.

One possible defense of the authors' simulations is that 401(k) and IRA participants tend to be more financially sophisticated than the rest of the population; consequently, the model may describe their behavior better than the behavior of a typical household. My preliminary investigations of this issue do reveal that levels of financial literacy are higher among those who save more. However, prevailing levels of financial literacy are quite low even among high savers. I encourage devotees of the life-cycle hypothesis to examine carefully the methods and practices of expert financial planners, who are almost certainly more sophisticated than the average 401(k) or IRA participant. The dominant methods of financial planning are astonishingly simplistic and certainly bear no resemblance whatsoever to the simulation model in this paper. It is easily demonstrated that these methods yield behavioral patterns that differ both qualitatively and quantitatively from the predictions of life-cycle models.

I emphasize that this is not a blanket criticism of all models that depict households as sophisticated decisionmakers. On the contrary, I heartily endorse the use of these models when decisions are frequently repeated, when the experiences of others are easily observed, or when there is a well-functioning market for "expert" guidance. However, none of these

conditions are satisfied in the context of retirement planning. Different analytical tools are useful in different situations. It is important for us to recognize when particular tools are applicable and when they are not.

Joel Slemrod: The low and declining U.S. national saving rate has been blamed by many economists for nearly all the ills that have befallen the U.S. in the past decade or two. The low saving rate has been blamed by many on a tax system that drives a large wedge between the pretax and after-tax rate of return. Partly in response to this argument, over the past 15 years a series of plans have been proposed to give tax-related incentives to saving.

In this paper Engen, Gale, and Scholz do not evaluate whether these programs are appropriate or misguided. Nor do they discuss whether it makes sense to use any kind of fiscal instrument to increase saving. Instead, in this paper they investigate a necessary, though not sufficient, condition for rendering a favorable judgment on such programs—that they increase national saving. (Note that even if they do increase saving, there may be better ways to do so.) The authors conclude that saving incentives have not appreciably increased, and may have reduced, national saving.

I tend to agree with this conclusion. I agreed with it before I read this paper. Reading this paper has succeeded in making me more suspicious of the arguments against this view. I am, however, less persuaded that the authors have sealed the case that existing saving incentive plans are ineffective in raising saving.

The tactics used in this paper for approaching this problem are somewhat unusual. The authors first perform some analysis of microeconomic data to address the relationship between participation in saving incentive plans and changes in wealth, concluding that no evidence links the two. They *then* present a stochastic life-cycle simulation model that not only reproduces many of the stylized facts about saving and participation in saving incentive plans, but is also consistent with their principal empirical finding. Interestingly, the simulation model predicts significant saving effects in the long run, although these effects do not appear for two generations. I will follow the plan of the paper by first discussing the empirical work and then the simulation model.

The aim of this work is to establish the effect on saving of the availability of a saving plan. To do this, one must first estimate the effect on

private saving and then adjust for the effect on national saving of a decrease in tax receipts.

The first problem one encounters is measuring saving. No data set provides a very accurate measure of it. The data sets used here provide poor measures. The method employed is to compute a measure of wealth at the beginning of a period and at the end of a period and call the difference saving. With the SIPP data for 401(k) plans one compares self-reported measures of wealth at the beginning and end of the period. Ignoring the problems of accuracy of self-reported wealth, this measure of saving will pick up not only planned net additions to wealth but also capital gains. Capital gains arguably belong in a measure of saving for some purposes, but not, I think, for the purpose of measuring whether participation in a 401(k) increased saving. This is especially true for the period January 1987–May 1991, a period of large movements in stock and housing prices. Could this volatility explain their finding that the real wealth of 401(k) plan participants actually fell over this period, while that of nonparticipants rose? This conclusion may not be valid holding constant the portfolios of wealthholders, which I expect differ substantially between participants and nonparticipants.

When tax data are used to study IRAs, a different problem arises. The measure of saving is again the change in measured wealth, but this time wealth is imputed by capitalizing reported interest and dividend receipts. This will to some extent avoid the capital gains problem I just mentioned but surely introduces a large amount of error in the measure of saving, since, for one thing, it excludes a large part of wealth. Differences in portfolios within asset categories (such as bank accounts versus long-term bonds) can be important and may be systematically related to saving plan participation. I do not have solutions for these problems. At a minimum, I urge the authors to openly address the issue of errors in the measurement of saving and discuss what biases might arise from their procedures.

I have a more fundamental concern about the data analysis. The authors make much of the notion that household heterogeneity is important to account for, because it implies that families with high propensities to save will make up a disproportionate share of those who participate in voluntary plans like IRAs and 401(k)s. They argue that the presence of heterogeneity invalidates the making of inferences about the effects of saving incentives from the observation that, in a cross-section, participants have higher *wealth* than nonparticipants.

They then go on to compare *changes* in wealth between participants and nonparticipants. I simply do not see how, if heterogeneity of saving propensities invalidates inferences about wealth differences between participants and nonparticipants, it does not invalidate similar inferences about changes in wealth or saving.

The authors try to avoid this problem in their 401(k) analysis by not comparing savers to nonsavers but by comparing one group of savers (401(k) participants) to another group of savers (IRA participants who are not eligible for 401(k) plans). But, as the authors admit, 401(k) eligibility is not randomly assigned, so it remains an open question whether 401(k) participants are *bigger* savers than non-401(k) eligible households. If so, one cannot sort out differences in the propensity to save from the effects of plan availability. This is a difficult endeavor, requiring the investigator to observe variables that influence the probability of being eligible for a plan but are not correlated with saving propensity, and vice versa. The authors pursue a few analyses that suggest that their results are not biased by self-selection, but I do not find these analyses completely convincing.

This problem aside, they find that 401(k) participants' net financial assets did not rise relative to the wealth of IRA holders not eligible for 401(k). Given what I think is the direction of heterogeneity bias (that 401(k) participants have a *higher* unobserved propensity to save than others), this is a striking finding, suggesting that these plans do not raise private saving.

I find more convincing their test of the substitutability of 401(k)s and IRAs, because it relies on an exogenous event—the fact that the 1986 tax reform restricted the attractiveness of IRAs for some households but not others. The authors find that those households that had their IRA deductibility removed increased their 401(k) balances, other things equal, implying that they are substitutes.

This is more convincing, but there is still a problem, albeit a different one. The criteria for removal of eligibility were based on two factors—income and the presence of another pension plan. Thus, relying on this estimate means relying on how precisely the effects of income and the presence of other pension plans is estimated.

The same set of issues applies to the IRA analysis. The 1981 tax act made many households newly eligible for IRAs. The authors compare the subsequent saving of this newly eligible group to the group whose eligibility was unchanged. This *is* an exogenous change, except that the

criterion for prior eligibility was coverage by a pension plan, which, as the authors note, is arguably correlated with (unobservable) saving propensities. So heterogeneity rears its ugly head again and must temper the interpretation of the results.

Dealing with heterogeneity is tricky. The authors discuss the possibility of dealing with it in a formal econometric model but reject this approach. In the absence of a formal model, the potential biases introduced by heterogeneity blur the interpretations of the estimated coefficients, and suggestive tests of the biases remain inconclusive.

I will not discuss, but find very convincing, their debunking of some earlier studies that purport to find large positive saving effects of savings incentives. Thus, the paper makes the “little or no increase” case at least as plausible as the “large saving increase” case but does not decisively establish the former.

Having offered some evidence that saving incentive plans do not increase saving, the authors never say exactly why this is so. There are a couple of possibilities. One is that the intertemporal elasticity of substitution is close to zero, so that a compensated increase in the after-tax return to saving, whether achieved via saving incentives, a reduction in the tax wedge on saving, or an increase in pretax real interest rates, will fail to increase saving.

A second possibility is that the saving plans studied here and used in the United States have fatal design flaws. One flaw is that they all have caps on the amount of saving that qualifies for the tax break, both to limit the revenue loss and to prevent too much of the tax break from going to wealthy high savers. But this means that much of the tax break is infamarginal and will fail to budge saving. Perhaps the most crippling design flaw is that the tax break is not tied to saving but rather to contributions to an account. Such contributions can be made, in the absence of saving, either by transferring assets into the account or by borrowing.

Which of these two, or which combination of the two, causes the ineffectiveness of savings plans matters for judging whether there exists *some* fiscal scheme that will increase saving. If the intertemporal elasticity of substitution is near zero, even dumping the income tax for a consumption tax will not work; transferring resources from the poor to the rich might still work, although this strategy did not seem to help much in the 1980s.

If design flaws are the problem, they will be difficult to overcome. Measuring saving right is difficult for investigators; it would be just as

difficult for the IRS. A personal expenditure tax would be extremely complicated. The feasible saving incentive plan is to replace the income tax with a value-added tax. This is feasible operationally but is probably not politically feasible, for it would be a wrenching and regressive change in the tax system and would eliminate all the personalization of the tax burden. For all this cost, it would increase saving only to the extent that the intertemporal elasticity of saving is above zero.

The authors do not resolve the important issue of why saving plans have been ineffective. Instead they present a stochastic, life-cycle simulation model that generates the short-run ineffectiveness of saving incentive plans and also replicates fairly closely the basic features of saving, savings plan participation, and so on.

This model is impressive, yields many insights, and is a great advance over the theoretical models underlying previous empirical studies of saving incentive plans. It generates the important result that the introduction of an IRA plan will *depress*, not increase, national saving for at least two generations but after that, and forever, national saving will be increased; the first result is consistent with the empirical results from the first part of the paper. The second result implies that, if the country could only stick with the plans until asset shifting is exhausted, they would work.

Before this conclusion is accepted, a few cautions are in order. First of all, this is presumably not the only parameterization of a stochastic life-cycle model that can generate short-run saving plan ineffectiveness and also be consistent with other stylized facts. In particular, this one features what, to my mind, is a high intertemporal elasticity of substitution of 0.33. The authors should report whether a similar model with an elasticity of, say, 0.10 could also be consistent with short-term savings plan ineffectiveness and other stylized facts; my guess is that the answer is yes. On this line I would like to see the authors use the paper's model and parameterization to simulate the effect of a tax cut like that of 1981, in order to see whether the predicted saving response in the model is a plausible one.

This same point can be made on a large number of dimensions. Many models can generate ineffectiveness of saving incentive plans. This is a very impressive one that represents a lot of hard work and careful thinking. But one should guard against the fallacy that because the model follows the empirical sections it must be the appropriate one.

One final point. I believe an exclusive focus on the average or median saving rate is incomplete and may be misleading. For example, a large fraction of private saving is done by the top 10 percent of savers, but little is known about the saving behavior of the rich—what is their true marginal rate on the income from savings, and what kind of portfolio responses can they make in response to changes in the taxation of capital income?

A related point is that, for a policy evaluation of saving incentive plans, it is important to refer back to the reason for wanting to increase saving. Is it that some particular people do not save enough? If so, who? Do saving incentive plans influence these people? If the real problem is the inadequate saving of low-income people, tax-based measures are unlikely to be very effective. This issue is especially troublesome because the model developed here does not capture well the empirical fact that many people do not save at all.

The key policy question raised by this paper is, given the apparent ineffectiveness of these plans, (i) should they be dropped with the revenue thus gained used to lower tax rates? or (ii) should the country institute new and improved saving plans, ones that address the design flaws of the current plans? One of the authors of this paper, together with one of the discussants, has advocated such a plan. I am less sanguine about the workability of such plans and suggest that more radical tax reform would be required to raise saving, given that the intertemporal elasticity of saving is high enough to make this effective. That leaves the question of whether the change from an income tax to a value-added tax is worthwhile, the answer to which depends on many more considerations than the expected change in national saving.

General Discussion

Many members of the panel focused on how heterogeneity and sample selection bias might complicate the interpretation of the empirical results for 401(k) plans and IRAs. Chris Sims and Gary Burtless pointed out that—because table 3 is based on separate cross-sections rather than on a panel—the change in assets between periods may not represent saving. In particular, if the characteristics of individuals in each cross-section differ significantly, then the change over time could represent

differences in the composition of the samples rather than actual saving behavior. For example, Sims pointed out that the median net worth of 401(k) participants in table 3 fell by about \$22,000 between 1987 and 1991, suggesting that the 1991 sample differed significantly from the 1987 sample. Gale responded that all of the reduction in financial assets and net financial assets for 401(k) participants, and most of the reduction in net worth, could be explained by changes in observable characteristics, which were also controlled for in the regression-based tests.

Brigitte Madrian called attention to the large increase in the participation in 401(k) plans between 1987 and 1991. Because these plans pulled in a wider spectrum of individuals in 1991, she suggested that the 1991 sample of 401(k) participants likely includes more “nonserious” savers than the 1987 sample. Further, because IRAs became less attractive after 1987, the 1991 IRA sample likely includes more “serious” savers than the 1987 sample. Gale responded that tests revealed little evidence of an increased proportion of “nonserious” savers among 401(k) participants or an increased proportion of “serious” savers among IRA participants not eligible for 401(k) plans, once controls for observable characteristics were included. Moreover, the authors obtained similar results for the effects of 401(k)s using other comparison groups. Barry Bosworth emphasized the importance of heterogeneity in understanding saving behavior. In particular, there are many zero savers in the United States whose behavior may not be captured well by the medians used in the paper. He reported that Japan has a smaller fraction of zero savers and that this difference is an important factor in explaining Japan’s higher saving rate.

Robert Moffitt offered a way to exploit possible heterogeneity in the sample. He suggested that many segments of the population would not be expected, a priori, to respond to IRAs. For example, low-income taxpayers face low or zero marginal tax rates and so may not respond to tax incentives; high-income taxpayers are likely saving more than the maximum IRA deduction so that the tax incentives in an IRA will not matter at the margin; and young people can be ruled out because they will not want to lock up funds in an IRA. Given these likely differences across segments of the population, Moffitt suggested examining whether the effects of IRAs are different among these groups.

Panelists also focused on the interpretation of the simulation model. Burtless agreed with Douglas Bernheim’s formal comment that the simulation model should not be taken literally. Nonetheless, he found the

model useful as part of the policy debate because it rests on assumptions that are widely used by other economists: what would happen if every agent correctly solved his or her dynamic optimization problem. By showing that, even under this strained assumption, tax preferences were not effective in boosting saving, the simulation added force to the other empirical results in the paper. William Brainard noted that it would be useful to know how sensitive the simulation results are to the specification and parameter values used in the model. He conjectured that, in reality, many individuals are not saving on a margin that is affected by the IRAs. Some may be saving zero and others may be contributing up to or beyond the allowable IRA limit. These individuals' marginal rates of return are unaffected by tax deductibility. The more individuals there are in this position, the less sensitive aggregate short-run saving behavior will be to alternative assumptions about elasticities. Indeed, for individuals who save beyond the IRA limits, there is no substitution effect to raise saving at the margin. Instead, the tax reduction boosts income, leading to higher consumption and lower saving.

George Perry warned that the wording of survey questions about asset balances was critically important. He noted that the surveys used in some earlier work on IRAs were ambiguous about whether respondents did or did not include IRA assets in their reported holding of assets by type—such as stocks, certificates of deposit, and mutual funds. Estimates of the effect of IRAs on total saving would be drastically different if it were assumed that respondents did not include IRA assets in their listing of assets by type (the usual assumption). For example, if, upon the introduction of IRAs, reported assets by type rose by \$100, as they had the year before, while reported IRA assets rose by \$100, the first assumption would show that all IRA contributions were a net increment to saving; the second assumption would show IRAs added nothing to saving.

Benjamin Friedman noted that part of the increased participation in 401(k) plans has occurred because companies have been dropping defined benefit plans, thus ridding themselves of certain future obligations. He added that there was some indication that defined contribution plans would be increasing for state and local government employees as well, and for the same reason. Friedman also suggested that marginal analysis of 401(k) plans may miss a key feature of how employees view them. Although economists typically treat 401(k) plans as a great opportunity for employees because of the tax break they provide, many employers be-

lieve that employer-matching is necessary to induce employees to contribute to these plans. If so, something very different is going on than marginal responses to marginal changes in the rate of return. Madrian expanded on this, suggesting that many employers offer matching to get enough low-wage employees involved to satisfy nondiscrimination rules.

Some panelists discussed the reasons that policy is concerned about private saving and whether current tax incentives address the concern. Greg Mankiw noted that there are two reasons policymakers might care about saving. First, the level of saving affects the capital available to future generations. Second, people may not save enough for retirement. If policymakers are mostly concerned about future generations, then they should care about bequest saving as well as retirement saving. But, he noted, current saving incentives fail to encourage bequests because IRA balances are taxed at death. Bosworth suggested that the main reason for saving incentives is to ensure that individuals save enough for retirement. To this end, they are not successful because they leave dissavers out. For example, defined benefit plans put retirement savings in place for everyone, but as firms shift to 401(k) plans some individuals can opt out. Further, many workers cash in their 401(k) holdings when they change jobs. For these individuals, the plans simply provide tax benefits for postponed consumption rather than preparation for retirement. Bernheim proposed a novel saving incentive, motivated by the idea that the short-run impact of IRAs was muted because they did not represent net new saving for many older savers who just shifted assets around. He suggested making IRAs available only to people born after 1970 because IRA contributions are more likely to represent net new saving for these younger individuals.

Dan Sichel reiterated that the key question is not whether saving incentives increase private saving, but whether they cause enough of an increase to offset the revenue loss. Thus, even granting a wide band of uncertainty around the results for private saving in the paper (reflecting heterogeneity or sample selection problems), the evidence in the paper strongly suggests that saving incentives decrease national saving, except perhaps in the very long run. Mankiw noted that saving incentives could increase national saving if they were accompanied by other policies that ensured a balanced-budget rule. They effectively shift the tax burden from savers to nonsavers.

References

- Allen, Steven G., Robert L. Clark, and Ann A. McDermed. 1993. "Pensions, Bonding, and Lifetime Jobs." *Journal of Human Resources* 28(3): 463–81.
- Andrews, Emily S. 1992. "The Growth and Distribution of 401(k) Plans." In *Trends in Pensions 1992*, edited by John A. Turner and Daniel J. Beller. Washington: U.S. Department of Labor, Pension and Welfare Benefits Administration.
- Andrews, William D., and David F. Bradford. 1988. "Savings Incentives in a Hybrid Income Tax System." In *Uneasy Compromise: Problems of a Hybrid Income-Consumption Tax*, edited by Henry J. Aaron, Harvey Galper, and Joseph A. Pechman. Washington: Brookings.
- Auerbach, Alan J., and Laurence J. Kotlikoff. 1987. *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press.
- Bernheim, B. Douglas. 1992. "Is The Baby Boom Generation Preparing Adequately for Retirement?" Technical report prepared for Merrill Lynch & Co., Inc. (September).
- . 1994a. "Comment on Chapters 4 and 5." In *Studies in the Economics of Aging*, edited by David Wise. Chicago: University of Chicago Press and National Bureau of Economic Research.
- . 1994b. "Personal Saving, Information, and Economic Literacy: New Directions for Public Policy." In *Tax Policy for Economic Growth in the 1990s*. Washington: American Council for Capital Formation.
- , and John Karl Scholz. 1993. "Private Saving and Public Policy." In *Tax Policy and the Economy*, vol. 7, edited by James M. Poterba. Cambridge, Mass.: MIT Press.
- Bosworth, Barry, Gary Burtless and John Sabelhaus. 1991. "The Decline in Saving: Evidence from Household Surveys." *BPEA*, 1:1991, 183–241.
- Bound, John, David A. Jaeger, and Regina Baker. 1993. "The Cure Can be Worse Than the Disease: A Cautionary Tale Regarding Instrumental Variables." Technical Paper 137. Cambridge, Mass.: National Bureau of Economic Research (June).
- Buck Consultants, Inc. 1989. "Current 401(k) Plan Practices: A Survey Report." Secaucus, N.J.: Buck Consultants, Inc. (September).
- Burman, Leonard, Joseph J. Cordes, and Larry Ozanne. 1990. "IRAs and National Saving." *National Tax Journal* 43(3): 259–83.
- Carroll, Christopher D. 1992. "The Buffer Stock Theory of Saving: Some Macroeconomic Evidence." *BPEA*, 2:1992, 61–135.
- Chang, Angela E. 1993. "Tax Policy, Lump Sum Pension Distributions, and Household Saving." Unpublished paper. Federal Reserve Bank of New York (November).
- Christian, Charles W., and Peter J. Frischmann. 1989. "Attrition in the Statistics of Income Panel of Individual Returns." *National Tax Journal* 42(4): 495–501.

- Collins, Julie H., and James H. Wyckoff. 1988. "Estimates of Tax-Deferred Retirement Savings Behavior." *National Tax Journal* 41(4): 561–72.
- Congressional Budget Office. 1993a. *Assessing the Decline in the National Saving Rate*. Washington: Government Printing Office.
- Congressional Budget Office. 1993b. *Baby Boomers in Retirement: An Early Perspective*. Washington: Government Printing Office.
- Deaton, Angus. 1991. "Saving and Liquidity Constraints." *Econometrica* 59(5): 1221–48.
- Eisner, James. 1993. "The Effect of Tax Policy on 401(k) Plans." Unpublished paper. University of Wisconsin, Madison (November).
- Engelhardt, Gary V. 1991. "Consumption Distortions from the Tax Treatment of Owner-Occupied Housing." Unpublished paper. Massachusetts Institute of Technology (December).
- Engen, Eric M. 1993a. "A Stochastic Life-Cycle Model with Mortality Risk: Estimation with Panel Data." Unpublished paper. University of California, Los Angeles (February).
- . 1993b. "Consumption and Saving in a Life-Cycle Model with Stochastic Earnings and Mortality Risk." Unpublished paper. University of California, Los Angeles (March).
- . 1993c. "Precautionary Saving and the Structure of Taxation." Unpublished paper. University of California, Los Angeles (June).
- , and William G. Gale. 1993. "IRAs and Saving in a Stochastic Life-Cycle Model." Unpublished paper. University of California, Los Angeles, and Brookings (April).
- Feenberg, Daniel R., and Jonathan Skinner. 1989. "Sources of IRA Saving." In *Tax Policy and the Economy*, vol. 3, edited by Lawrence H. Summers. Cambridge, Mass.: MIT Press for the National Bureau of Economic Research.
- Feenberg, Daniel R., and James M. Poterba. 1991. "Which Households Own Municipal Bonds? Evidence From Tax Returns." *National Tax Journal* 44(4): 93–103.
- . 1993. "Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns." In *Tax Policy and the Economy*, vol. 7, edited by James M. Poterba. Cambridge, Mass.: MIT Press for the National Bureau of Economic Research.
- Feldstein, Martin S. 1976. "Personal Taxation and Portfolio Composition: An Econometric Analysis." *Econometrica* 44(4): 631–50.
- . 1992. "The Effects of Tax-Based Saving Incentives on Government Revenue and National Saving." Working Paper 4021. Cambridge, Mass.: National Bureau of Economic Research (March).
- . 1993. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." Working Paper 4496. Cambridge, Mass.: National Bureau of Economic Research.
- , and Daniel R. Feenberg. 1983. "Alternative Tax Rules and Personal Saving Incentives: Microeconomic Data and Behavioral Simulations." In *Be-*

- havioral Simulation Methods in Tax Policy Analysis*, edited by Martin Feldstein. Chicago: University of Chicago Press and the National Bureau of Economic Research.
- Fernandez, Phyllis A. 1992. "Preretirement Lump Sum Distributions." In *Trends in Pensions 1992*, edited by John A. Turner and Daniel J. Beller. Washington: U.S. Department of Labor, Pension and Welfare Benefits Administration.
- Gale, William G., and John Karl Scholz. 1994. "IRAs and Household Saving." *American Economic Review* (forthcoming).
- Galper, Harvey, and Charles Byce. 1986. "Individual Retirement Accounts: Facts and Issues." *Tax Notes* 31(9): 917–21.
- Gravelle, Jane G. 1991. "Do Individual Retirement Accounts Increase Savings?" *Journal of Economic Perspectives* 5(2): 133–48.
- Gruber, Jonathan. 1994. "The Efficiency of a Group-Specific Mandated Benefit: Evidence from Health Insurance Benefits for Maternity." *American Economic Review* (forthcoming).
- , and Alan B. Krueger. 1991. "The Incidence of Mandated Employer-Provided Insurance: Lessons from Workers' Compensation Insurance." In *Tax Policy and the Economy*, vol. 5, edited by David Bradford. Cambridge, Mass.: MIT Press for the National Bureau of Economic Research.
- Hausman, Jerry A., and James M. Poterba. 1987. "Household Behavior and the Tax Reform Act of 1986." *Journal of Economic Perspectives* 1(1): 101–19.
- Hewitt Associates. 1993. *401(k) Plan Hot Topics*. Lincolnshire, Ill.: Hewitt Associates.
- Holloway, Thomas M. 1989. "Present NIPA Saving Measures: Their Characteristics and Limitations." In *The Measurement of Saving, Investment, and Wealth*, edited by Robert E. Lipsey and Helen Stone Tice. Chicago: University of Chicago Press.
- Hsiao, Cheng. 1986. *Analysis of Panel Data*. Cambridge: Cambridge University Press.
- Hubbard, R. Glenn. 1984. "Do IRAs and Keoghs Increase Saving?" *National Tax Journal* 37(1): 43–54.
- , Jonathan Skinner, and Stephen P. Zeldes. 1993a. "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving." Unpublished paper. Columbia University, University of Virginia, and University of Pennsylvania (September).
- . 1993b. "Precautionary Saving and Social Insurance." Unpublished paper. Columbia University, University of Virginia, and University of Pennsylvania (October).
- Huizinga, Harry. 1991. "Individual Retirement Accounts in an Optimal Tax Framework." Unpublished paper. Stanford University (June).
- Ippolito, Richard A. 1993. "Selecting and Retaining High-Quality Workers: A Theory of 401(k) Pensions." Unpublished paper. Pension Benefit Guaranty Corporation, Washington (April).

- Joines, Douglas H., and James G. Manegold. 1991. "IRAs and Saving: Evidence from a Panel of Taxpayers." Unpublished paper. University of Southern California (September).
- Judd, Kenneth L. 1985. "Short-Run Analysis of Fiscal Policy in a Perfect Foresight Model." *Journal of Political Economy* 93(2): 298–319.
- Kotlikoff, Laurence J. 1990. "The Crisis in U.S. Saving and Proposals to Address the Crisis." *National Tax Journal* 43(3): 233–46.
- Kusko, Andrea, James M. Poterba, and David W. Wilcox. 1994. "Employee Decisions With Respect to 401(k) Plans: Evidence from Individual Level Data." Working Paper 4635. Cambridge, Mass.: National Bureau of Economic Research (February).
- Long, James E. 1990. "Marginal Tax Rates and IRA Contributions." *National Tax Journal* 43(2): 143–53.
- Manchester, Joyce M., and James M. Poterba. 1989. "Second Mortgages and Household Saving." *Regional Science and Urban Economics* 19(2): 325–46.
- Montgomery, Edward, Kathryn Shaw, and Mary-Ellen Benedict. 1992. "Pensions and Wages: An Hedonic Price Theory Approach." *International Economic Review* 33(1): 111–28.
- Nelson, Charles R., and Richard Startz. 1990. "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator." *Econometrica* 58(4): 967–76.
- Ng, Yew Swang. 1992. "Do Individuals Optimize in Intertemporal Consumption/Savings Decisions? A Liberal Method to Encourage Savings." *Journal of Economic Behavior and Organization* 17(1): 101–14.
- O'Neil, Cherie J., and G. Rodney Thompson. 1987. "Participation in Individual Retirement Accounts: An Empirical Investigation." *National Tax Journal* 40(4): 617–24.
- Papke, Leslie E. 1992. "Participation in and Contributions to 401(k) Pension Plans: Evidence from Plan Data." Working Paper 4199. Cambridge, Mass.: National Bureau of Economic Research (October).
- , Mitchell Petersen, and James M. Poterba. 1993. "Did 401(k) Plans Replace Other Employer-Provided Pensions?" Working Paper 4501. Cambridge, Mass.: National Bureau of Economic Research (October).
- Poterba, James M., Steven F. Venti, and David A. Wise. 1992. "401(k) Plans and Tax-Deferred Saving." Working Paper 4181. Cambridge, Mass.: National Bureau of Economic Research (October).
- . 1993. "Do 401(k) Contributions Crowd Out Other Personal Saving?" Working Paper 4391. Cambridge, Mass.: National Bureau of Economic Research (June).
- . 1994. "Do 401(k) Contributions Crowd Out Other Personal Saving?" Unpublished paper. Massachusetts Institute of Technology, Dartmouth College, and Harvard University (April).
- Scholz, John Karl. 1992. "A Direct Examination of the Dividend Clientele Hypothesis." *Journal of Public Economics* 49(3): 261–85.
- . 1994. "Portfolio Choice and Tax Progressivity: Evidence from the Sur-

- veys of Consumer Finances." In *Tax Progressivity and Income Inequality*, edited by Joel Slemrod. New York: Cambridge University Press.
- Shefrin, Hersch M., and Richard H. Thaler. 1988. "The Behavioral Life-Cycle Hypothesis." *Economic Inquiry* 26(4): 609–43.
- Skinner, Jonathan. 1992. "Individual Retirement Accounts: A Review of the Evidence." *Tax Notes* 54(2): 201–12.
- . 1994. "Comment on '401(k) Plans and Tax-Deferred Saving'." In *Studies in the Economics of Aging*, edited by David Wise. Chicago: University of Chicago Press (forthcoming).
- , and Daniel Feenberg. 1990. "The Impact of the 1986 Tax Reform on Personal Saving." In *Do Taxes Matter? The Impact of the Tax Reform Act of 1986*, edited by Joel Slemrod. Cambridge, Mass.: MIT Press.
- Slemrod, Joel B. 1988. "The 1979–84 Linked Panel of Tax Return Data: Sampling and Linking Methodology." Unpublished paper. University of Michigan (November).
- . 1990. "Notes on the 1979–86 Panel of Individual Income Tax Returns." Unpublished paper. University of Michigan (July).
- . 1992a. "Do Taxes Matter? Lessons From the 1980s." *American Economic Review* 82(2): 250–56.
- . 1992b. "Taxation and Inequality: A Time-Exposure Perspective." In *Tax Policy and the Economy*, vol. 6, edited by James M. Poterba. Cambridge, Mass.: MIT Press for the National Bureau of Economic Research.
- Staiger, Douglas, and James H. Stock. 1993. "Asymptotics for Instrumental Variables Regressions with Weakly Correlated Instruments." Unpublished paper. Harvard University (July).
- Summers, Lawrence H. 1986. "Summers Replies to Galper and Byce on IRAs." *Tax Notes* 31(10): 1014–16.
- Thaler, Richard H. 1994. "Psychology and Savings Policies." *American Economic Review* 84(2): 186–92.
- U.S. General Accounting Office. 1988a. *401(k) Plans: Incidence, Provisions, and Benefits*. Washington: General Accounting Office.
- . 1988b. *401(k) Plans: Participation and Deferral Rates by Plan Features and Other Information*. Washington: General Accounting Office.
- U.S. Internal Revenue Service. Various issues. *Statistics of Income Bulletin*. Washington: U.S. Department of the Treasury.
- U.S. Pension and Welfare Benefits Administration. 1993. *Private Pension Plan Bulletin: Abstract of 1990 Form 5500 Annual Reports*. Washington: U.S. Department of Labor.
- Venti, Steven F., and David A. Wise. 1986. "Tax-Deferred Accounts, Constrained Choice, and Estimation of Individual Saving." *Review of Economic Studies* 53(4): 579–601.
- . 1987. "IRAs and Saving." In *The Effects of Taxation on Capital Accumulation*, edited by Martin Feldstein. Chicago: University of Chicago Press.
- . 1988. "The Determinants of IRA Contributions and the Effect of Limit Changes." In *Pensions in the U.S. Economy*, edited by Zvi Bodie, John B. Shoven, and David A. Wise. Chicago: University of Chicago Press.

- . 1990. "Have IRAs Increased U.S. Saving?: Evidence from Consumer Expenditure Surveys." *Quarterly Journal of Economics* 105(3): 661–98.
- . 1991. "The Saving Effect of Tax-Deferred Retirement Accounts: Evidence from SIPP." In *National Saving and Economic Performance*, edited by B. Douglas Bernheim and John B. Shoven. Chicago: University of Chicago Press and the National Bureau of Economic Research.
- . 1992. "Government Policy and Personal Retirement Saving." In *Tax Policy and the Economy*, vol. 6, edited by James M. Poterba. Cambridge, Mass.: MIT Press for the National Bureau of Economic Research.
- Wilson, John F., and others. 1989. "Measuring Household Saving: Recent Experience from the Flow-of-Funds Perspective." In *The Measurement of Saving, Investment, and Wealth*, edited by Robert E. Lipsey and Helen Stone Tice. Chicago: University of Chicago Press.
- Yakoboski, Paul J., and Sarah Boyce. 1993. "Pension Coverage and Participation Growth: A New Look at Primary and Supplemental Plans." *Employee Benefit Research Institute Issue Brief* 144 (December).