Journal of Economic Perspectives—Volume 10, Number 4—Fall 1996—Pages 113–138

The Illusory Effects of Saving Incentives on Saving

Eric M. Engen, William G. Gale, and John Karl Scholz

American saving rates fell dramatically in the 1980s and have remained low since then. The decline in saving has raised concerns that the economy may be unable to finance investment and sustain growth and that households may not be saving adequately for retirement. One response to these concerns has been the development of special saving accounts, such as Individual Retirement Accounts, 401(k) plans and Keogh plans. These voluntary accounts, which we refer to as "saving incentives," feature preferential tax treatment of contributions and investment earnings, annual contribution limits and penalties for early withdrawals.

The question addressed in this paper is the extent to which saving incentives have raised private and national (public plus private) saving. Contributions and investment earnings are typically tax deferred, thus reducing public saving (increasing the budget deficit) in the short run. The long-run impact on public saving is less obvious; if the incentives increase private saving, they may also increase income and tax revenue.

Saving incentives raise private saving when households finance contributions with reductions in consumption or increases in labor supply. Private saving rises even if the contributions are financed by the associated tax cut; this reinforces the importance of examining the net impact on public and private saving. Saving incentives do not raise private saving when households finance contributions with reductions in existing assets, with saving that would have been done even in the absence of the incentive, or with increases in debt. It is also possible that the higher after-tax return on saving incentives

■ *Eric M. Engen is Senior Economist, Federal Reserve Board, Washington, D.C. William G. Gale is Senior Fellow, The Brookings Institution, Washington, D.C. John Karl Scholz is Associate Professor of Economics and Public Policy, University of Wisconsin, Madison, Wisconsin. Their e-mail addresses are eengen@frb.gov, wgale@brook.edu, and jkscholz@facstaff.wisc.edu, respectively.*

could *reduce* private saving. For example, "target" saving occurs when people follow the common financial planning technique of saving enough to replace a fixed percentage of their preretirement income in retirement (Doyle and Johnson, 1991). Higher returns reduce the amount of saving needed to reach a given target. Even for life-cycle savers, the higher after-tax return on saving incentives yields ambiguous effects on saving due to opposing income and substitution effects.

In recent years, several studies have examined these issues. The crucial issue in this literature is determining what households who have saving incentive accounts, or are eligible for one, would have saved in the absence of the incentives. This problem is difficult and subject to a series of biases that generally overstate the impact of saving incentives on saving.

First, saving behavior varies significantly across households. Households that participate in, or are eligible for, saving incentive plans have systematically stronger tastes for saving than other households. Thus, a simple comparison of saving done by households with and without saving incentives will be biased toward showing that the incentives raise saving.

Second, saving and wealth are net concepts: if a household borrows $1000 and puts the money in a saving incentive account, net private saving is zero. The data show that households with saving incentives have taken on more debt than other households. Hence, studies should focus on how the incentives affect wealth—assets minus debt—not just assets. Because financial assets are small relative to total wealth, studies that focus on how the incentives affect financial assets may have particularly limited significance.

Third, since the expansion of IRAs and 401(k)s in the early 1980s, financial markets, pensions and Social Security have undergone major changes. Omitting interactions among these changes and saving incentives can overstate the effects of the incentives on saving.

Fourth, saving incentive contributions are generally tax-deductible, and saving incentive balances represent pretax balances; one cannot consume the entire amount because taxes (and perhaps penalties) are due upon withdrawal. In contrast, contributions to other accounts are not deductible, and one may generally consume the entire balance upon withdrawal. A given balance in a saving incentive account therefore represents less saving, defined either as reduced previous consumption or deferred consumption, than the same balance in a conventional account. The common practice in the literature of comparing saving incentive balances and taxable assets, without adjusting for this difference, overstates the impact of the incentives on saving.

Fifth, incentives that are a part of a worker's total compensation can have different effects from other plans. Consider two workers with equal cash wages, but one of whom also has a 401(k) with employer contributions. The worker with the 401(k) has higher income and would be expected to have higher saving. Ignoring the distinction between cash wages and total compensation can lead to a systematic overstatement of the impact of 401(k)s on saving.

A major theme of this paper is that analyses that ignore these issues—including most previous research in this area—overstate the impact of saving incentives on

saving. We show that efforts to account for these factors largely or completely eliminate the estimated positive impact of saving incentives on saving found in the literature. Thus, we conclude that little, if any, of the overall contributions to existing saving incentives have raised saving. Our conclusion does not rule out the possibility that existing incentives have raised saving for some people, or that saving incentives may eventually raise saving.

The next section sets the stage by describing aggregate trends in U.S. saving and saving incentives. We then organize our comments around the various types of data and methods that have been used to examine these issues, including cross-sectional studies of IRAs and 401(k)s, studies using cohort analysis, successive cross-sections and panel data. In each case, our goal is to use the five considerations listed above to reconcile estimates suggesting that saving incentives have raised saving with our own results and conclusion that the incentives have not raised saving. The remainder of the paper addresses related issues: substitution between 401(k)s and pensions at the firm level, results from simulation models and the effects on public revenues. The conclusion discusses some of the broader ramifications of our findings.

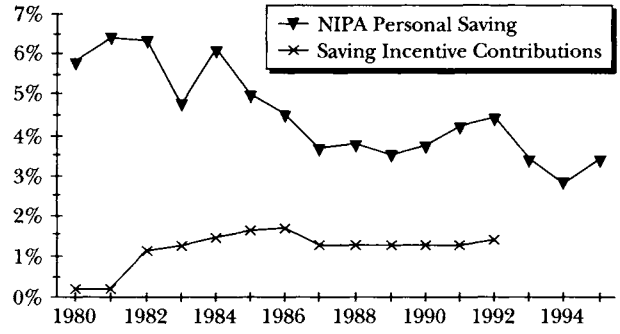
Aggregate Trends in Saving Incentives and Personal Saving

Figure 1 compares overall contributions to saving incentives and personal saving in recent years. As measured in the National Income and Product Accounts, personal saving is the difference between personal disposable income and personal spending.[1] Major changes in the ratio of saving incentive contributions to GDP occurred only twice. From 1981 to 1982, with the expansion of IRAs and the issuance of clarifying regulations on 401(k)s, saving incentive contributions rose by almost a full percentage point of GDP, while the personal saving rate was flat. Following the Tax Reform Act of 1986, both saving incentive contributions and personal saving fell. This coincidence may suggest a causal relation, but personal saving had been declining since 1984. In addition, the amount of saving that occurs in saving incentives is larger than the contributions alone, because it includes interest and dividend earnings, plus rollovers, less withdrawals.[2] Accounting for these factors, saving in saving incentives declined by only $5 billion from 1986 to 1987, while personal saving fell by $28 billion. Thus, saving incentives could have accounted for no more than a fifth of the fall in personal saving from 1986 to 1987.

[1] The composition of saving incentive contributions has changed over time. IRA contributions totalled $5 billion in 1981, rising to $34 billion annually from 1982 to 1986, following the rule changes in 1981. After the 1986 Tax Act restricted IRA deductions, contributions fell to $20 billion in 1987 and $16 billion in 1990 (Employee Benefits Research Institute, 1995). Keogh contributions rose over the 1980s but have remained below $9 billion per year (Internal Revenue Service, various years). Contributions to 401(k) plans grew steadily from $16 billion in 1984 to $64 billion in 1992 (U.S. Department of Labor, 1996).

[2] This saving measure is the same as the change in overall balances in saving incentive accounts, less capital gains. Capital gains are excluded because they are also excluded from personal saving and national income. The calculations required to make these estimates allocate IRA and 401(k) balances to stocks and other assets using the percentages described in Hubbard and Skinner (1995).

Figure 1
Personal Saving and Saving Incentive Contributions as a Percentage of GDP, 1980–1995



Sources: National Income and Product Accounts and the references in footnote 1.

Taking the 1980–1992 period as a whole, contributions to saving incentives rose by 1.1 percent of GDP, and the broader measure of saving in saving incentive accounts rose by almost 2 percent of GDP, but personal saving declined. None of these patterns provides any evidence that saving incentives influence the level of personal saving.

The aggregate impact of the incentives would be uninteresting if saving incentives were tiny. But from 1982 to 1994, the incentives accounted for one-third or more of personal saving. If these plans represent mostly new saving, it would not be unreasonable to expect saving of this magnitude to raise aggregate personal saving to some extent. That is not what the data show. It remains possible, however, that saving would have fallen even further if the incentives had not been enacted. For this and other reasons, research has turned to microeconomic data.

Cross-Sectional Evidence on IRAs

A commonly made claim in favor of the idea that IRAs raise saving is that the typical household holds few financial assets, and thus IRA contributions must be largely new saving rather than a reallocation of existing assets (Feldstein and Feenberg, 1983; Venti and Wise, 1991). This claim is misleading, however, since the typical household does not have an IRA. The extent to which IRAs raise private saving hinges on the behavior of households that contribute to IRAs, not on those that do not. If the question is the effects of raising the annual contribution limit, as it is in the formal econometric models described below, analysis should focus on households that contribute to the limit.

Contributors who view IRAs as good substitutes for taxable assets are more

likely to transfer into IRAs existing assets or saving that would have been done anyway. In effect, such transfers provide the household with a tax break for doing what it would have done anyway, and they will not raise private saving. Contributors who view IRAs as poor substitutes for other assets will fund contributions by reducing consumption (raising saving). Households where IRAs are likely to be good substitutes for other saving include those where the head is older than 59, since they face no penalty on IRA withdrawals, and those with large amounts of other assets, who can avoid penalties for early withdrawal by consuming other assets first. From 1983 to 1985, almost 70 percent of IRA contributions were made by households with heads older than 59 or with 1986 non-IRA financial assets in excess of $20,000. About 78 percent of households that contributed to the limit in each of the three years were in these groups (Gale and Scholz, 1994). For these contributors, the effects of IRAs on saving seem likely to be small at best.

At the beginning of the IRA program, there was substantial opportunity for contributors to shift existing assets into IRAs. In 1983, the median IRA contributor held $77,000 in net worth, excluding pensions, and $16,000 in non-IRA financial assets, while the median household that contributed to the limit for each year from 1983 to 1985 held $101,000 in net worth and $30,000 in non-IRA financial assets. These figures, based on tabulations from the 1983 Survey of Consumer Finances, contrast sharply with median wealth of $20,448 and total financial assets of $2,050 for households without IRAs.

Some of this heterogeneity in saving is due to observable factors such as age, earnings, or family size. But even after controlling for such factors, substantial heterogeneity remains in unobserved tastes or influences on saving (Diamond and Hausman, 1984; Gale and Scholz, 1994). Allowing for heterogeneity while estimating the effects of IRAs has proven difficult. For example, Hubbard (1984) estimates that households with IRAs have higher ratios of net worth to income than do other households, but his estimates do not allow for differences in tastes for saving across the two groups, so the results are difficult to interpret.

Venti and Wise (1986, 1987, 1990, 1991) estimate that raising the annual contribution limit would raise IRA saving and that 45 to 66 percent of the increased contributions would come from reductions in consumption.[3] However, these estimates are based on the identifying assumption that in the absence of IRAs, households that contributed the limit amount would have saved the same amount as households that did not contribute to IRAs. Thus, to generate a finding of substitution between IRAs and other saving in the Venti-Wise framework, limit contributors would need to have less non-IRA saving than noncontributors.

[3] Between 3 and 20 percent of the increased contributions would be financed by reductions in other saving, and about 35 percent would come from reduced tax payments. Some features of the Venti-Wise framework are worth noting. For example, consumers are assumed to maximize a function that has IRA saving and other saving as specific separate arguments. It is unclear what sort of utility function or preferences would correspond to such a decision function. Also, Venti and Wise constrain the fraction of IRAs that are new saving to be between zero and one, although there is no theoretical justification for doing so. For details and other discussion, see Gale and Scholz (1994).

The identifying assumption, of course, is not valid if contributors have stronger tastes for saving than noncontributors, as is now widely understood to be the case. Thus, the Venti-Wise model interprets findings that limit contributors saved more than noncontributors as evidence that IRAs raise saving, even though that finding is perfectly consistent with the view that limit contributors have higher tastes for saving and IRAs do not raise saving. The identifying assumption biases the results toward finding that IRAs raise saving even if they do not.[4]

Gale and Scholz (1994) develop and estimate a model that addresses these and other concerns. They assume that households maximize a utility function that depends on current and expected future consumption. Individuals like the higher return on IRAs relative to other saving, but face uncertain income and are hesitant to lock up funds in an IRA. Thus, IRAs are imperfect substitutes for other saving, but as assets and age rise, IRAs become better substitutes because the withdrawal penalty becomes less important. Gale and Scholz assume a quadratic utility function, which allows derivation of a specific, estimable saving function. The model implies that, in the absence of IRAs, households that contributed to the limit in each of three years would have saved the same as other IRA contributors, rather than noncontributors.[5] Thus, the model allows for differences in saving between contributors and noncontributors.[6] The model also allows the substitutability between IRAs and other saving to depend on household characteristics.

The model is estimated on households from the 1983–86 Surveys of Consumer Finances (SCF). The estimates generate reasonable saving functions and show that people with higher assets find IRAs and other saving to be better substitutes. Hence, for these people, IRA contributions are less likely to represent new saving. The estimates suggest that increases in the IRA contribution limit in 1983–86 would have generated little, if any, new saving.[7] The central estimate is that 2 percent of the additional contributions would have represented net additions to *national* saving if the reduction in tax payments due to increased deductions had been completely saved. If some of the tax cut had been spent, the effect on national saving would have been smaller or negative. Gale and Scholz (1994) also formally test and reject

[4] The bias can also be seen by noting that for IRAs to be a perfect substitute for other saving in the Venti-Wise model, every household that saved any amount would have to hold an IRA. Since only about 20 percent of the sample held IRAs, it is not surprising that Venti and Wise reject the hypothesis of perfect substitutability for the whole sample. However, what matters is whether IRA contributors—not every household that saves—find IRAs to be good substitutes for other saving.

[5] This assumption is consistent with the data. For example, Engen, Gale and Scholz (1996) report regression results indicating that the average three-year limit IRA contributor from 1983 to 1986 did not have a statistically significant difference in wealth in 1983 compared to the average of other contributors, after controlling for other household characteristics.

[6] As noted by Hubbard and Skinner (1995), allowing the saving equations to differ across contributors and noncontributors represents a crucial difference with the Venti-Wise approach and makes the Gale-Scholz model more general than the Venti-Wise model.

[7] Limit contributors in the model are those who contributed to the limit in all three years. Gale and Scholz (1994) show, using IRS data, that households that contributed to the limit in all three years accounted for 68 percent of all limit contributions during this period.

the idea that contributors and noncontributors have similar non-IRA saving equations. Thus, it is not valid to impose the same non-IRA saving equation for contributors and noncontributors or to assume that in the absence of IRAs, contributors and noncontributors would have saved the same amount.

Naturally, the effects of IRAs are subject to some uncertainty.[8] First, Gale and Scholz (1994) note that the estimated substitutability is a function of household characteristics and can be sensitive to changes in the value of those characteristics. This occurs, however, partly because Gale and Scholz did not constrain the proportion of IRAs that are new saving to be between zero and one. That the unconstrained estimates fall within a reasonable range is a strength of the model. Second, the central estimates in Gale and Scholz use a sample of households whose saving was less than $100,000 in absolute value over the period. Gale and Scholz also show that changing the saving threshold changes the estimated impact of IRAs on saving. This should not be surprising; as the threshold is reduced, high savers, for whom IRAs should be very good substitutes for other saving, are removed. When the largest thresholds are used, very high levels of substitution are obtained.[9]

Cross-Sectional Evidence on 401(k)s: Is 401(k) Eligibility Exogenous?

If 401(k) eligibility is distributed independently of underlying propensities to save, the effects of 401(k)s could be measured from simple comparisons of the saving or wealth of eligible and ineligible households. However, if eligibility is positively correlated with underlying tastes for saving, then cross-sectional comparisons of eligibles and ineligibles that do not control for tastes for saving will systematically overstate the effects of 401(k)s on saving.

At first glance, the idea that 401(k) eligibility is exogenous may seem plausible; as Poterba, Venti and Wise (1995, p. 10) note, "eligibility is determined by employers." But employers often take employee preferences into account; indeed, it would be strange if employers created benefits without regard to employee preferences. In a broad survey of employers, "perceived employee interest" was the

[8] Poterba, Venti and Wise (1996) claim that the Gale and Scholz results are driven by measurement error due to the way the saving variable is constructed. However, as Engen, Gale and Scholz (1996) note, if measurement error were quantitatively important in this regard, the estimated effects of IRAs on saving should get larger as broader measures of the dependent variable—which are constructed differently—are used, but they became smaller instead. Similarly, measurement error would be expected to cause a negative relationship between wealth and saving for households not contributing to IRAs, but the results do not show this pattern. Gale and Scholz (1994) discuss other possible sources of measurement error.

[9] Recent papers by Poterba, Venti and Wise (1996) and Bernheim (1996) criticize the Gale-Scholz model. Engen, Gale and Scholz (1996) provide responses and clarifications to the criticisms. Gravelle (1991) and Skinner (1992) provide surveys of the literature on IRAs. Attanasio and De Leire (1994) present additional evidence on IRAs, which they interpret as being consistent with the Gale and Scholz (1994) results. This interpretation has been challenged by Hubbard and Skinner (1995).

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). Moreover, even if firms did provide 401(k)s randomly, workers with high tastes for saving should seek out firms with 401(k)s (Allen, Clark and McDermed, 1993; Curme and Even, 1995; Johnson, 1993; Ippolito, 1993). If employers do consider employee preferences, or if some employees prefer firms that offer 401(k)s, then eligibility is likely to be positively correlated with tastes for saving.

Ultimately, whether 401(k) eligibility is exogenous is an empirical issue. Poterba, Venti and Wise (1995) present regressions showing that eligible households have about the same level of nonpension, non-401(k) financial assets as do ineligible households, controlling for income and other factors. They interpret these results as evidence that 401(k) eligibility is exogenous with respect to tastes for saving.

But the evidence and interpretation are at best fragile. Engen, Gale and Scholz (1994, Table 8) use a similar sample from the same data set, a slightly different test format and a longer list of explanatory variables, and find that eligible families have higher levels of nonpension, non-401(k) financial assets, net financial assets and net worth. Moreover, Poterba, Venti and Wise (1995) omit pensions. Families eligible for 401(k)s are between 24 and 33 percentage points more likely to be covered by a defined benefit pension plan than other families, controlling for other factors (Engen, Gale and Scholz, 1994). Again, this implies that eligible households have higher non-401(k) wealth than do ineligible households.[10]

The Poterba, Venti and Wise (1995) test has another problem that creates a potentially large bias in favor of finding that eligibility is exogenous: the test ignores all 401(k) wealth and thereby *assumes* that all 401(k) saving is new saving. However, if x percent of 401(k) wealth would have existed even in the absence of 401(k)s, an appropriate test of exogeneity compares the non-401(k) assets of ineligible families to the sum of non-401(k) assets plus x percent of the 401(k) wealth of eligible families. Clearly, assuming that all 401(k) saving is new saving ($x = 0$)—as in the Poterba, Venti and Wise test—creates a bias in favor of finding that eligibility is exogenous.

An additional bias arises because the regressions used by Poterba, Venti and Wise (1995) control for cash earnings rather than total compensation. If two households have the same cash earnings, one without and one with a 401(k) with employer contributions, the latter has higher total compensation and would be expected to have higher total saving even if 401(k)s are not new saving. However, a regression that controls only for cash earnings will treat these as two households with the same income. Then, the coefficient on eligibility will overstate the impact of 401(k)s on saving because it will pick up not only substitution between 401(k)s and other saving (which would generally give a negative coefficient), but also the

[10] Along similar lines, Bernheim and Garrett (1995) find that 401(k) eligibility "raises" total wealth by about four times as much as it "raises" retirement wealth. They note that unless one is willing to believe that 401(k) contributions crowd in several times their own value in non-401(k) saving, these findings suggest that eligibility is positively correlated with tastes for saving.

Table 1
Cohort Effects at Age 60–64

	1984	*1991*	*Difference*
All Households Aged 60–64			
Mean Financial Assets	42,250	50,419	8,169
Mean Saving Incentives	5,118	14,156	9,038
Mean Other Financial Assets	37,132	36,263	−869
Contributors			
Median Financial Assets	34,975	50,182	15,207
Median Saving Incentives	8,171	22,148	13,977
Median Other Financial Assets	22,983	21,528	−1,455
Noncontributors			
Median Financial Assets	2,687	2,134	−553

Notes: All figures are in 1991 dollars. The figures represent conditional means and medians, controlling for age, income, education and marital status, based on data from the Surveys of Income and Program Participation. The data omit balances in 401(k)s in 1984 and in after-tax thrift plans in all years.
Source: Poterba, Venti and Wise (1996).

effect of having higher total compensation on saving (which would give a positive coefficient) (Gale, 1995; Engen, Gale and Scholz, 1996, Appendix B).

For all of these reasons, we conclude that eligibility is positively correlated with tastes for saving and that comparisons of eligible and ineligible households that do not control for tastes for saving are biased toward showing that 401(k)s raise saving.

Some related evidence from Bernheim and Garrett (1995) supports our belief that 401(k)s contribute little to new saving. In a cross-sectional study of how financial education affects saving, they control for initial wealth as a proxy for tastes for saving and find that 401(k) eligibility raises self-reported saving rates by 1.5 percentage points. The typical eligible worker probably contributes about 5 percent or more of salary in employer and employee contributions, so about 30 percent of contributions might be new saving (1.5/5). But at least 20 percent of the total contributions represent tax deductions rather than reductions in consumption, so the impact on saving rates should be reduced by about 1 percentage point (20 percent of 5 percent), leaving an increase in saving of 0.5 percentage points, suggesting that only 10 percent of 401(k) contributions represent new saving. Even this figure is biased upward because the regressions control for cash earnings rather than total compensation. Engen, Gale and Scholz (1996) discuss related issues and details of these calculations.

Cohort Analysis

A cohort is a group of people born within a given time interval. Table 1 reports results from cohort analyses in Poterba, Venti and Wise (1996) and Venti and Wise

(1996). Mean financial assets for families aged 60–64 in 1991 were about $8,200 higher than for families aged 60–64 in 1984. Mean saving incentive balances were $9,000 higher. For families aged 60–64 with saving incentives, median financial assets in 1991 were about $15,000 higher than for similarly aged families with saving incentives in 1984, with saving incentive balances about $14,000 higher for the 1991 group. For families without saving incentives, median financial assets fell slightly.

Venti and Wise (1996, p. 11–12) interpret the results as showing that saving incentives have raised private saving: "The basic assumption is that younger cohorts—that reached a given age in later calendar years—had a longer period in which to contribute to personal retirement accounts. But that in other respects the cohorts are similar (after correcting for earnings. . .). Thus differences in asset accumulation can be attributed to the differential availability of these programs." Our view, however, is that accounting for other changes in the 1980s and data problems in the analysis leaves essentially no room for saving incentives to have raised wealth.[11]

A fundamental problem with cohort analysis is that it is impossible to identify separate age, time and cohort effects without making strong assumptions. Cohort effects apply to groups born in a common period. Time effects affect all groups at a point in time. Age effects refer to behavior at different points of the life cycle. Problems arise because age equals time minus cohort, so the three variables provide only two pieces of information. Thus, "cohort analyses" reflect an unknown combination of age, time and cohort effects.

To illustrate this point, consider the claim by Venti and Wise (1996) that, other than increased exposure to 401(k)s and IRAs, there were no systematic differences between the cohort aged 60–64 in 1984 and the one aged 60–64 in 1991 that would have affected accumulation of financial assets. It is implausible to attribute all or even most of the growth in financial assets to saving incentives: between 1984 and 1991, aggregate real financial assets grew by $4 trillion, while saving incentive balances grew by less than $1 trillion (Employee Benefit Research Institute, 1995). In fact, several major changes during this period, none controlled for in the Venti and Wise analysis, can explain virtually all of the increase in assets shown in Table 1.

The most obvious reason for the growth in financial assets was the stock market boom. From 1984 to 1991, the S&P 500 Index rose by 78 percent in real terms. In contrast, from 1977 to 1984, the real index fell by 5 percent. This difference alone can explain most of the difference in mean financial assets shown in the top panel.[12]

[11] Venti and Wise (1996) present findings similar to those in Table 1 for several other cohorts. These results are subject to criticisms of the same nature as those noted in the text.

[12] Calculations using the SIPP show that the cohort aged 60–64 in 1991 had mean stock and mutual fund holdings outside of retirement accounts in 1984 of $6,200 (1991 dollars). Mean IRA and Keogh balances were $3,800, of which we estimate (based on EBRI data) about 25 percent was held in stocks and mutual funds, so that total mean stock and mutual fund holdings were $7,150. A passive investment strategy that held the S&P 500 Index would have raised these initial holdings to $12,727 ($7,150 × 1.78) in 1991. If the 1991 cohort had instead faced the 5 percent fall in stocks experienced by the earlier cohort, the $7,150 would have dropped to $6,800. Thus, differences in stock market returns can account for a wealth difference of $5,927 (=$12,727–$6,800), or about 72 percent of the entire $8,169 increase

Real interest rates were also higher between 1984 and 1991 than in the preceding seven years. Thus, the 1991 cohort of 60–64 year-olds experienced higher returns on their pre-existing financial assets than the 1984 cohort did in the seven years prior to being observed in the data.

Changes in four other types of wealth also fueled the increase in financial assets in the 1980s. First, when inflation and marginal tax rates declined in the 1980s, investors shifted away from tangible capital (like housing) that had been more attractive in the 1970s (Feldstein, 1980; Summers, 1981; Poterba, 1984). Hence, part of the increase in financial assets was just a shift in the composition of assets. Second, mortgages and overall household debt rose relative to income or assets. At least a quarter of the rise in financial assets from 1984 to 1991 was matched by an increase in debt (Board of Governors of the Federal Reserve System, 1995). Third, the value of Social Security benefits was reduced in the 1983 reforms. Poterba, Venti and Wise (1994) estimate that for households aged 65–69 in 1991, real mean and median Social Security wealth was about $6,000 lower than for households aged 65–69 in 1984. Fourth, pension coverage other than 401(k)s also fell in the 1980s. Calculations using the SIPP indicate that 55.4 percent of working households aged 60–64 were covered by a defined benefit or non-401(k) defined contribution plan in 1984 compared to 46.9 percent of 60–64 year olds in 1991. These declines in Social Security and pensions could have induced an increase in financial assets.[13]

A number of data problems also create biases. First, saving incentive balances are pretax amounts, whereas balances in other accounts (other than unrealized capital gains) are generally post-tax amounts.[14] Thus, the $9,000 increase in saving incentive balances by the 1991 cohort does not represent $9,000 of increased saving. Since the contribution is deductible, previous consumption would have fallen by at most only $9000 minus the federal and state taxes that would have been owed. The relevant marginal federal and states tax rates for saving incentive participants is conservatively estimated at 20 percent (Engen, Gale and Scholz, 1996), implying that at least $1,800 of the $9000 increase in tax-preferred assets does not represent saving.

A second data problem is that the SIPP data contain no information on balances in 401(k)s in 1984 or in after-tax thrift plans in any year. Thrift plans were generally supplemental pensions with tax-deductible employer contributions, but after-tax employee contributions, and were prevalent before 401(k)s became popular. A substantial portion of 401(k)s appear to have been converted from thrift plans over the course of the 1980s. As the conversions occurred, they appeared in the SIPP as 401(k) balances

documented in Table 1. These calculations understate the importance of the stock market boom because they omit stocks and mutual funds held in 401(k)s and thrifts in 1984. Adjusting for these holdings implies that stock market performance can account for over 90 percent of the increase in mean financial assets shown in the top panel of Table 1.

[13] See Bernheim (1987) and citations therein on the effects of Social Security on saving. Gale (1995) reviews the literature on how pensions affect saving and provides new estimates.

[14] Unrealized capital gains represented only one-third of the value of stocks and mutual funds outside of retirement plans in 1989 (Kennickell and Wilcox, 1992), or 10 percent of total household financial assets outside of retirement plans.

in 1987 and 1991 even though their equivalent balance as thrift plans did not appear in 1984 or in later years. Data from the 1983 Survey of Consumer Finances indicates that mean balances in such plans for the entire cohort aged 60–64 in 1983 was $5,860 (1991 dollars). Although caution is warranted in comparing data from different sources, almost all of the $8,200 increase in mean financial assets in the top panel of Table 1 can be explained by the two data problems alone.

The same issues arise in understanding why median financial assets of saving incentive participants rose relative to nonparticipants over this period, as shown in the bottom panel of Table 1. First consider the data problems. The $14,000 rise in median saving incentive balances shown in the bottom panel of Table 1 is overstated by $2,800 if a tax rate of 20 percent is assumed. The omission of thrift plans appears especially important. Among households aged 60–64 with an IRA or thrift plan in the 1983 SCF, including thrift balances raised median total financial assets by $8,519 in 1991 dollars. The omission of 401(k)s from the 1984 data further reduces measured 1984 wealth of participants relative to nonparticipants and leads to an overstatement of the increase in relative wealth holdings of participants from 1984 to 1991. Again, data issues alone can explain a large proportion of the increase in financial assets for participants relative to nonparticipants.

In addition, the stock market boom and higher real interest rates raised the initial financial assets of participants more in arithmetic terms than for nonparticipants, because participants in 1984 had many times more financial assets than did nonparticipants. Moreover, debt increased more rapidly for 401(k) and IRA participants than for nonparticipants (Engen, Gale and Scholz, 1994; Engen and Gale, 1995), but debt holdings are omitted from the table.[15]

Thus, we find little evidence in "cohort" analyses that saving incentives raise wealth. It is plausible to attribute virtually all of the reported increase in financial assets to the stock market boom, high real interest rates, shifts in nonfinancial assets, debt, pensions and Social Security, the bias caused by comparing pre- and post-tax balances and omitted 1984 data on thrift and 401(k) balances.

Evidence from Successive Cross-Sectional Studies

Successive cross-sectional studies use random cross-sections of households from two or more years. Households eligible for 401(k)s in later years had access to 401(k)s for longer, on average, than eligible households in earlier years. Other things equal, if 401(k)s raise wealth, the wealth of cross-sections of eligible house-

[15] The shift from nonfinancial assets to financial assets caused by falling inflation and marginal tax rates also likely caused an increase in the financial assets of saving incentive participants relative to nonparticipants. If participants and nonparticipants shifted the same proportion of their portfolio to financial assets, this would generate a larger arithmetic increase in financial assets for participants, because their initial total wealth level was much larger. Moreover, participants may have shifted a greater percentage to financial assets because they had greater access to financial assets that are tax preferred.

holds should increase over time. This comparison of eligible or "like" families over time is very similar to cohort analysis and suffers from the same problems.[16]

Nevertheless, Poterba, Venti and Wise (1996, p. 27) emphasize that "the critical feature of . . . [their] approach to controlling for heterogeneity is comparison of the *within group change* in non-IRA-401(k) assets as IRA and or 401(k) assets grow . . ." (emphasis in the original). While much of the research reported in Poterba, Venti and Wise (1996) adopts this approach, it is difficult to learn anything from within-group comparisons of eligible (or "like") families over time. Within-group comparisons never explicitly confront the critical issue in this literature: namely, establishing what households would have done in the absence of saving incentives. The implicit assumption is that other financial assets would have stayed the same, so papers typically show that other financial assets do not fall over time as saving incentive balances rise and conclude that saving incentives are largely new saving. But the implicit assumption that other financial assets for these households would not have grown in the absence of saving incentives is extreme, because of the stock market boom, high real interest rates, the rise in debt, the shift away from nonfinancial assets due to falling inflation and tax rates, and the decline in pensions and Social Security over the period being studied.

A cleaner test relies on the intuition that if 401(k)s increase wealth, wealth should rise more over time for cross-sections of eligible households than for ineligible households. Even this comparison across groups and time, however, shares many of the same features and problems of cohort analysis.[17] A key assumption is that, over time, changes in the level or composition of wealth for eligible households relative to ineligible households were due only to lengthier access to saving incentive plans.

A recent example of successive cross-section analysis is Poterba, Venti and Wise (1995), who find that from 1987 to 1991, median financial assets rose by about $2,900 (in 1987 dollars) for eligible households relative to others, controlling for age, income, education and marital status. They also show that non-IRA, non-401(k) financial assets did not fall for eligible households relative to others over 1984–1991. They conclude that most 401(k) contributions represent new saving.

As with cohort analysis, interpreting the results must be tempered by consideration of changes during the 1980s (other than saving incentives) and data problems. For example, eligible households had greater access to *tax-preferred* financial assets and so would have found it more attractive to reallocate wealth to financial assets and away from nonfinancial assets after the reduction in inflation and tax rates. The rise in mortgage debt occurred disproportionately for eligible house-

[16] For example, the second panel of Table 1 compares 60 to 64 year-old participants in 1984 and 1991, while the test using eligible families described in the text would compare 25 to 64 year-old eligible families in 1984 with 25 to 64 year-old eligible families in 1991.

[17] The second and third panels of Table 1 compare 60 to 64 year-old saving incentive participants and nonparticipants over time. Similarly, the successive cross-section test described in the text would compare 25 to 64 year-old eligible and ineligible households over time.

holds relative to ineligible households (Engen and Gale, 1995). Both of these factors imply that the financial asset measure used by Poterba, Venti and Wise (1995) may not be a consistent measure of wealth trends over time or across groups. Data problems are an additional concern. The SIPP data used in the study overstate the saving done via saving incentives over time because saving incentive contributions are tax-deductible and because thrift plans and 1984 401(k) balances are omitted. All of these factors impart biases in favor of finding that reported financial assets rose for eligible households relative to ineligible households, even if 401(k)s did not raise wealth.

Engen and Gale (1995) address several of these issues. They use the same comparison groups as Poterba, Venti and Wise and a similar sample from the SIPP. They avoid using data from 1984, due to the missing data on 401(k)s and thrift plans. The major innovation of their study is to measure the impact of 401(k) eligibility on broader measures of wealth, such as the sum of net financial assets and home equity (called "wealth" below), as well as on narrow wealth measures, such as financial assets.

Examining substitution between 401(k)s and housing wealth is motivated by several considerations. About 14 percent of 401(k)-eligible households had home equity loans in 1991; about 19 percent bought new homes between 1987 and 1991; and Bernheim (1996) estimates that about 12 percent extracted equity from their home via a refinancing between 1987 and 1991. Substitution can occur in less obvious ways as well: for example, by not accelerating mortgage payments or not trading up into a bigger house. Or consider new homeowners in different years who are identical, except those in later years have longer access to 401(k)s and higher 401(k) balances. Suppose also that the later cohort has fewer liquid assets available for down payments, because of their higher 401(k)s. Therefore, they will either buy smaller homes or take out more debt to purchase a home of the same size as the earlier cohort. In either case, households in the later cohort clearly substituted 401(k)s for home equity relative to earlier cohorts, even if this substitution was unintentional. Engen and Gale (1995) and Stiglitz (1988, p. 595) discuss other ways that unintentional substitution can occur between housing equity and saving incentives.

Since most saving incentive participants own their own homes, and housing represents a substantial portion of most households' wealth, omitting saving that occurs in owner-occupied housing systematically understates how much saving these households have done. In addition, because of employer matching of 401(k) contributions, effectively financing a 401(k) with tax-deductible mortgage borrowing can be quite lucrative. A considerable literature on Social Security and pensions examines the impact of such policies on broad measures of wealth, rather than just financial assets, and finds that broader wealth definitions generate larger offsets for pension wealth; for example, Avery, Elliehausen and Gustafson (1986) find that only 11 percent of pension wealth shows up as a

reduction in liquid assets, but that 66 percent of pension wealth shows up as a reduction in other net worth.[18]

Engen and Gale (1995) find that between 1987 and 1991, controlling for a standard list of covariates, 401(k)-eligible households accumulated more financial assets than other households. But Engen and Gale also show that eligible households did *not* accumulate more wealth when housing equity is included. House value rose for eligibles relative to others, but mortgage debt grew even faster, so that housing equity fell for eligible homeowners relative to other homeowners. Similar results occur when the sample is stratified by IRA status. In fact, no group of eligible families raised their wealth (including housing equity) from 1987 to 1991 relative to the similar group of ineligible families. The results suggest that any impact of 401(k)s on financial assets is limited to homeowners, and even for them, 401(k)s did not raise broader wealth measures. Moreover, the estimated effects for financial assets should not be taken at face value since they suffer from some of the same biases as Poterba, Venti and Wise (1995).[19]

Bernheim (1996) claims the results in Poterba, Venti and Wise (1995) and Engen and Gale (1995) understate the impact of 401(k)s due to "dilution," which he describes as the problem that occurs if the average "taste for saving" among eligible households falls over time. The logic of the claim is that the most dedicated savers were most likely to become eligible for 401(k)s early on. As less dedicated savers became eligible, average tastes for saving fell among eligible households over time.

There are two ways to address this concern: looking only at eligible families over time, and comparing eligible and ineligible families over time. It is not at all clear that the group of eligible families became more diluted. If participation in a saving incentive plan is taken as an indicator of strong tastes for saving and dilution of the eligible sample were empirically important, then—other things equal—it would be reasonable to expect that the proportion of 401(k)-eligible workers making contributions would have fallen over time. Instead, data from the Current Population Survey show that it rose from 57 percent in 1988 to 65 percent in 1993 (Bassett, Fleming and Rodrigues, 1996).[20] Our own analysis using the SIPP indicates that between 1987 and 1991, controlling for household characteristics (including pension coverage), the rise in the 401(k) participation rate of eligible households was 8 percentage points and was statistically significant, and the likelihood of participating in any saving incentive plan rose by 4.4 percentage points and was significant. Ippolito (1993) provides one

[18] See also Bernheim (1987), Blinder, Gordon and Wise (1980), Diamond and Hausman (1984), Dicks-Mireaux and King (1984) and Hubbard (1986). Engen, Gale and Scholz (1994) and Gale (1995) obtain similar results.

[19] A separate literature on whether the elderly wish to consume their housing equity has reached mixed conclusions, but is not directly relevant for interpreting Engen and Gale (1995), who focus on workers, who are accumulating assets, rather than the elderly, who are dissaving. Also, Engen and Gale find evidence of reshuffling between mortgage debt and 401(k)s, but generally do not find an offset between house value and 401(k)s. Thus, even if people do not want to trade off the size of their house for other consumption (elderly) or other assets (workers), they may still reshuffle housing equity with 401ks through debt.

[20] This increase is unlikely to be due to an increase in employer matching. In 1993, 60 percent of eligible workers that did not receive a match contributed. This is larger than the overall 1988 average probability of contributing (Bassett, Fleming and Rodrigues, 1996).

explanation of these trends: because of employers' matching contributions, 401(k)s are more attractive to workers with high tastes for saving who are more likely to contribute. Apparently, high savers are attracted to, and remain at, firms that sponsor 401(k)s, while low savers are more likely to leave such firms. Thus, the average tastes for saving could plausibly rise over time among eligible workers rather than fall.

The direction of the net bias caused by dilution is also unclear when comparing eligible and ineligible households. Over time, the most dedicated savers among ineligible households are the most likely to become eligible, so there is also dilution among ineligible households. Thus, the key issue is the relative dilution of the two groups, not just dilution in one group. Analysis using the SIPP shows that controlling for household characteristics, IRA participation among 401(k)-eligible households fell by only 1.3 percentage points relative to ineligible households from 1987 to 1991, but overall saving incentive participation rose for 401(k)-eligible relative to ineligible households by 15 percentage points. Moreover, workers with low tastes for saving can and frequently do liquidate their 401(k) upon leaving a firm (Chang, 1996; Fernandez, 1992), which raises the average tastes for saving among eligible families and may reduce it among ineligible families. Thus, the bias created could plausibly work in either direction. More importantly, however, it seems to us that dilution is unlikely to be a major quantitative issue in interpreting the results described above.

A number of other criticisms and sensitivity analyses of Engen and Gale (1995) merit attention, and while these issues are discussed in Engen, Gale and Scholz (1996, Appendix C), some of them also deserve mention here. Poterba, Venti and Wise (1996) note that Engen and Gale require regression coefficients to be the same for eligible and ineligible households, and they claim it would be more appropriate to allow them to differ across groups because the comparison groups differ significantly. They also claim that exogenous changes in housing value bias the Engen and Gale (1995) results toward finding no effect of 401(k)s on wealth. However, these factors do not in fact affect the results, as Engen and Gale (1996) show in a set of revised estimates, which also includes a response to other criticisms.[21]

In Engen, Gale and Scholz (1994), we create a test based on 401(k) participation as well as eligibility. We compare 401(k) participants with IRA participants who are ineligible for 401(k) plans, using SIPP data. Families in each group have chosen to hold saving incentives, indicating they have strong tastes for saving. Between 1986 and 1991, 401(k)s had higher limits than IRAs, and IRA tax preferences were restricted. Over this period, 401(k) contributions grew rapidly and IRA contributions plummeted. If 401(k)s raise private wealth, controlling for other factors, wealth should have risen for 401(k) participants relative to IRA participants not eligible for 401(k) plans. We found that this did not occur, and we concluded that 401(k)s did not raise wealth. This test has been criticized on grounds that the 401(k)

[21] Engen and Gale (1996) also provide estimates for the 1984 to 1991 period, which show a substantial reduction in non-401(k) wealth for eligible households relative to others. Estimating the impact of 401(k)s on total wealth is difficult using the 1984 sample, as the reader will remember from the earlier discussion, due to missing data on 401(k)s and thrift plans.

group became diluted (Bernheim, 1994; Poterba, Venti and Wise, 1996). In response, Engen and Gale (1995) create samples where IRA participation—a plausible indicator of tastes for saving once household characteristics are controlled for—does not change over time. Results mirror those based on eligibility. For homeowners, 401(k)s "raise" financial assets (subject to the caveats made above), but do not raise broader wealth measures that include home equity. For renters, 401(k)s have no economically or statistically significant impact on net financial assets. These results provide no support for the view that 401(k)s have raised wealth.

Evidence from Panel Data

In panel data, the same households or individuals are followed over time, which controls for certain types of household-specific heterogeneity in saving behavior. In practice, however, the advantage has proven difficult to exploit in a satisfactory way.

Joines and Manegold (1995) use data from the IRS-Michigan Tax Panel, which provides a random sample of tax returns in 1979 and follows the same taxpayers in subsequent years.[22] They note that if IRAs stimulate saving, those taxpayers made newly eligible for IRAs after the rule changes in 1981 should increase their saving relative to the change in saving by continuing contributors, who contributed both before and after 1981. This test therefore compares one group of contributors with another, rather than comparing contributors with noncontributors. Controlling for taxpayer-specific effects, period effects and taxpayer characteristics, they conclude that between 19 and 26 percent of IRA contributions were increases in national saving.

This estimate may be overstated, however. First, the effects of IRAs are measured on a narrow measure of assets, which does not take into account the fact that IRA contributions can be financed by increases in consumer debt, investment debt, or mortgage debt, a reduction in tax-exempt bond holdings, a shift from tangible assets, or a reduced contribution to a thrift, pension or a 401(k) plan. For example, Engen, Gale and Scholz (1994), using a similar sample, show that average nonmortgage debt rose by $2,400 for IRA contributors over this period, but by only $1,350 for noncontributors. Median nonmortgage debt rose by $2,350 for contributors but by only $100 for noncontributors. Second, the regressions are estimated on a restricted sample that omits taxpayers aged 65 and over. For these elderly households, IRAs are likely to be good substitutes for taxable saving, since there is no withdrawal penalty. The elderly represent about 12 percent of IRA contributions in the unrestricted sample.[23]

[22] Additional information on the tax panel can be found in Slemrod (1988, 1990a) and Engen, Gale and Scholz (1994, 1996).

[23] Engen, Gale and Scholz (1994) use a different extract of the tax panel to estimate similar regressions and find somewhat less favorable results for IRAs. Feenberg and Skinner (1989) use the tax panel and find that, controlling for initial wealth, IRA holders increased their taxable saving by more than noncontributors from 1982 to 1984. However, since contributors have stronger tastes for saving than noncontributors, even after controlling for observable characteristics, the Feenberg and Skinner finding is consistent with any view of the effects of IRAs on saving.

Venti and Wise (1995) use panel data from the SIPP to examine changes in IRA contributions and a measure of non-IRA saving during 1984 and 1985. They find that the fall in non-IRA saving when people begin to contribute to an IRA is at most only a small fraction of the typical IRA contribution. They conclude that IRAs raise overall saving and may even crowd in additional non-IRA saving. The main problem with this test is that it does not control for any individual characteristics or period-specific individual or aggregate shocks to saving. Thus, any change that moves IRA saving and non-IRA saving in the same direction—for example, receipt of an inheritance, birth of a child, changes in income, or a decision to start saving for retirement—is interpreted as evidence that the two forms of saving are not substitutes.[24] Their sample omits persons over age 60, even though such households account for a large portion of IRA contributions and are likely to find IRAs and other saving to be good substitutes (Gale and Scholz, 1994). Also, the measure of non-IRA saving used is extremely narrow, including only bank accounts, bonds and stocks. In short, substitution from other assets or current saving to IRAs could occur in numerous ways not captured in their study.

Venti and Wise (1992) use panel data on financial assets from the 1983–86 Surveys of Consumer Finances and claim, first, that IRA contributors did not save very much before IRAs became available, and, second, that IRA contributors saved substantial amounts after IRAs became available. They conclude (p. 24) that "the 1986 contributors, prior to 1983, had not been accustomed to saving nearly as much as they saved over the next three years." Hence, they conclude that IRAs represent new saving.

But financial assets are only one component of wealth. Table 2 shows that most IRA contributors held substantial amounts of housing equity, which is also a form of saving. Moreover, IRAs can substitute for other forms of wealth besides financial assets, and the reduced inflation and tax rates over this period encouraged this shift. Therefore, it seems reasonable to examine broader wealth measures such as net financial wealth plus housing equity.

Trends in this broader measure of wealth lead to dramatically different conclusions. In 1983, the median among 1986 IRA contributors was $51,220.[25] Thus, it is simply incorrect to claim that these families typically did little or no saving prior to 1983. What is true is that financial assets were only a small portion of wealth for this group. Over the next three years, all of these families contributed to IRAs and their median wealth grew to $64,897.

Does that increase represent more than those households would have saved without IRAs? We construct one estimate as follows. The initial wealth level

[24] This would be less of a problem had there been a change in IRA eligibility rules over the period examined, as in the period studied by Joines and Manegold (1995). However, there was no change in IRA eligibility from 1984 to 1985.

[25] This figure is in 1983 dollars and omits all pension wealth (thrift, defined benefit and defined contribution plans), vehicle equity, second homes, business equity and other items. The median of net worth (omitting only pension wealth) was about $75,000.

Table 2
IRA Holders in the 1986 SCF

Variable	*1983*	*1986*
Age of Head (Median)	42	45
Household Earnings (Median)	32,464	40,000
Home Ownership Rate	0.772	0.791
Housing Equity (Median)	39,398	45,168
Financial Assets[a] (Median)	10,200	21,500
IRA plus Keogh	250	6,500
Other Financial Assets	7,699	12,000
Net Financial Assets[b] plus Housing Equity (median)	51,220	64,897

[a] Financial assets include stocks and mutual funds, bonds, checking and savings accounts, IRA and Keogh accounts, money market accounts and certificates of deposit.
[b] Net financial assets equals financial assets minus consumer debt.
Source: Authors' calculations. All dollar figures are in nominal dollars. Sample Size = 618.

would plausibly have grown by at least 6 percent per year in nominal terms from 1983 and 1986 (Engen, Gale and Scholz, 1996). This alone would have resulted in median 1986 wealth of at least $61,000. If the median family had made an additional contribution of $1,200 per year (plus the interest), it would have attained the actual median 1986 wealth level. If a conservative 20 percent tax rate adjustment is made to account for the fact that $6,000 of the 1986 wealth was accumulated with tax-deductible IRAs, new saving of less than $900 per year would have been sufficient to generate the 1986 wealth level. Venti and Wise (1986, 1987, 1990, 1991) estimate that the marginal tax rate for IRA contributors was about 35 percent. Using this figure, new saving of just $550 per year would have been sufficient to generate the 1986 wealth level.[26] Since the median 1986 contributor was age 42 in 1983 and had net financial assets plus housing equity over $51,000, we find it quite plausible to believe that the typical contributor would have saved at least $900 (or $550) even in the absence of IRAs. This is reinforced by noting that saving rates typically rise over the life cycle, and median earnings for the households in the sample rose by about 12 percent in real terms from 1983 to 1986, reaching $40,000 in 1986.

[26] These estimates could be modified in a number of ways. For example, some of the housing wealth in 1983 is due to capital gains on housing in the 1970s. To adjust for this, we examined the sum of net financial assets plus half of housing equity. These calculations implied that households would have needed to save $1,300 per year ($1,000 with the 20 percent tax adjustment or $750 with the 35 percent adjustment) to reach the 1986 wealth median.

Substitution Between Pensions and 401(k)s at the Firm Level

All of the research described above ignores interactions at the firm level between 401(k)s and other pensions. This can create important biases in favor of finding that 401(k)s raise saving. Suppose a worker in year 1 has $10,000 in assets and $5,000 in a pension (or after-tax thrift) plan. In year 2, an otherwise identical worker has $10,000 in assets, but the firm has converted the pension to a 401(k), and the worker has $5,000 in the 401(k). The cohort and successive cross-section analyses described above omit data on pension and thrift balances, and so they would "show" that the worker in year 2 was eligible for a 401(k) and had $5,000 more in assets than the worker in year 1, who was not eligible for a 401(k). The conclusion that 401(k) eligibility raised saving would be inappropriate, however; the 401(k) balance just represents what otherwise would have been in a pension.[27]

Many 401(k)s appear to have been converted directly from previously existing pension or thrift plans. Because 401(k)s were not popular until the IRS issued clarifying regulations in 1981, most plans created before 1982 are thought to be conversions. In 1985, these plans accounted for 85 percent of 401(k) balances, 39 percent of the plans, 65 percent of participants with nonzero balances and two-thirds of contributions.[28] Even in 1991, the latest year evaluated in successive cross-sections by Poterba, Venti and Wise (1995) and in cohort analysis by Venti and Wise (1996), the majority of assets, 42 percent of participants and 47 percent of contributions were in plans created before 1982. This suggests the possibility that much of the 1991 401(k) wealth would have existed even in the absence of 401(k) plans.

Adopting a 401(k) could also cause the outright termination of other plans. Papke (1996) finds considerable substitution between 401(k)s and other pension plans between 1985 and 1992. In addition, firms could cut back on existing plans in other ways such as restricting or reducing benefit increases. In this case, and when 401(k)s represent converted thrift plans, households with 401(k)s would still have traditional pension coverage—as observed in the data—but 401(k)s would be substituting for other pensions at the margin. Another possibility is that some 401(k)s are established at firms that would have created another plan had 401(k)s not existed.

When 401(k)s displace other plans completely or at the margin, an additional issue comes into play. All workers covered by traditional plans participate, but workers may well opt out of a 401(k). Thus, the saving impact of a 401(k) can be less than that of the pension it displaces (Halperin, 1987). If so, 401(k)s could *reduce* private saving even if they appear to raise saving in surveys that omit pension wealth.

[27] In the example given, controlling for pension *coverage* would solve the bias, but the successive cross-section and cohort studies that claim saving incentives raise saving do not control for pension coverage. Moreover, controlling for pension coverage, as opposed to pension balances, would not help in more realistic situations where the worker in period 2 still had a pension, but the pension balance was smaller than it otherwise would have been, with the remainder of the balance made up by a 401(k).

[28] These tabulations were carried out by Joel Dickson using the Form 5500 reports. A number of studies reach similar conclusions (Andrews, 1992; Papke, Petersen and Poterba, 1993; Buck Consultants, 1989).

Since all analyses of saving to date ignore the substitution between 401(k)s and other pensions, they overstate the impact of 401(k)s on saving.

Simulation Analysis and the Long Term

Engen and Gale (1993) and Engen, Gale and Scholz (1994) examine saving incentives in a stochastic life-cycle simulation model that uses estimates of behavioral and economic characteristics, like preference parameters and stochastic earnings patterns, to develop quantitative predictions for saving behavior and its response to government policies. In the model, households maximize expected lifetime utility by choosing consumption and allocating saving to a tax-preferred saving incentive or a conventional, fully taxed asset. Decisions are made subject to a lifetime budget constraint, net asset constraints, uncertain future earnings and uncertain mortality. For present purposes, the most interesting implications of the model involve the plausible impact of saving incentives and the ability to distinguish short-term and long-term effects on saving.[29]

The model indicates that in the long run, both IRAs and 401(k) plans can stimulate national saving. After 70 years, IRAs raise the steady state national saving rate from 5.9 percent to 6.2 percent; 401(k)s have somewhat stronger effects due to a higher contribution limit and the employer match.

However, the short-term and transitional effects are very different. Private and national saving *decline* following the implementation of IRAs (or other saving incentives) and do not recover their original level for decades. The intuition is that before IRAs are introduced, all assets are held in taxable forms. When IRAs become available, people start shifting out of taxable assets into IRAs. This process takes time because of the annual contribution limit. Eventually, households run out of funds to shift or reach optimal portfolios, and new households (with no initial taxable wealth to shift) enter the economy, so the extent of shifting falls and more of the contributions represent new saving. The interest elasticity of saving in the model is positive, but its effect is swamped in the first decades of the program by the shifting of already existing taxable assets.

Government saving falls initially because asset shifting creates tax deductions. Thus, in the short run, national saving falls by even more than private saving. Eventually, government revenues rise as saving incentive balances are withdrawn, but even in the long-term, government revenues are below what they would have been in the absence of IRAs (the national debt is positive). In plausible specifications,

[29] The simulations also demonstrate that many observed empirical patterns concerning IRAs, 401(k)s and saving can be generated by a well-specified model of utility maximization. At the aggregate level, model results are broadly consistent with U.S. historical data. At the individual level, the model generates hump-shaped consumption-age profiles and age-wealth profiles that are consistent with microdata for much of the population. Like the data, the model shows saving incentive participation rising with age and income. Model results also match well with data on the proportion of households with a saving incentive, the percentage of contributors that contribute to the limit, and the persistence of contributions over time.

the national (public plus private) asset-to-income ratio falls from the start and does not regain its initial value until 35 to 50 years later.

The simulation results are consistent with our findings described above, because the empirical work covers only the first 10 to 15 years of the programs, during which very little if any of the contributions to saving incentives should be new saving according to the simulation.

The simulations, like all models, are imperfect in a number of ways, and the magnitude of the estimates should be regarded with care. Nonetheless, the qualitative behavioral patterns generated by the model seem both plausible and robust. The most important caveat is that for a number of reasons, we believe the model overstates the long-run impact of saving incentives on private saving and may also overstate the transition period—the length of time after implementation during which saving incentives reduce national saving. But even after making adjustments for these factors, the model would imply a lengthy transition period, which casts doubt on findings that IRAs or 401(k)s raise saving immediately after being implemented.

Offsetting Revenue Effects?

Although the discussion above focuses primarily on private saving, the impact of tax incentives on national saving depends also on how public saving is affected. Because they are tax-deductible, saving incentive contributions increase the budget deficit immediately. To the extent that taxable private saving is shifted into a saving incentive, current and future tax revenues from that saving and the investment earnings are lost. In the long term, the saving incentive contribution and associated earnings are withdrawn, raising future tax revenues. If saving incentives raise saving, they may also raise domestic investment, which may produce additional corporate tax revenues.

In simulations in Engen and Gale (1993) and Engen, Gale and Scholz (1994), public saving is negative for many years after saving incentives are introduced. Annual public saving eventually turns positive as the funds are cashed in, but the public debt is immediately and permanently higher with the incentives than without. Corporate tax payments are not considered, but if they had been, revenues would have fallen further in the short and medium term as national saving fell, before rebounding in the long term.

In contrast, Feldstein (1995) finds that after including the impact on corporate taxes, saving incentives have a substantial, positive impact on public saving. However, his estimates overstate and may even produce the wrong sign for the revenue effects. First, he assumes that in the long run at least half or more of IRA contributions are new saving. For all of the reasons noted in this survey, this estimate appears high. He also assumes that 100 percent of increases in domestic saving turns into increased domestic investment, although the correlation between domestic saving and domestic investment is less than one and appears to be falling over time. Feldstein also overstates the effective tax rate on new investment. He assumes that new investment faces a business-level tax of 17 to 34 percent as well as individual-level taxes of 15 to 25 percent.

But investments in housing, state and local governments, unincorporated businesses, or foreign countries can face lower effective tax rates. Auerbach (1996) estimates a combined business- and individual-level average marginal tax rate on new investment of 16 percent. Addressing these and other issues not only dramatically reduces the favorable impacts of IRAs on public revenues, but under plausible conditions reverses the sign of the effect as well (Ruggeri and Fougere, 1995).

Conclusion

Our findings fit well into the hierarchy of behavioral responses to taxation, as developed by Slemrod (1990b) and Auerbach and Slemrod (1996) in their careful reviews of studies of the Tax Reform Act of 1986 on economic behavior. They find that decisions concerning the timing of economic transactions are the most clearly responsive to tax considerations. The next tier of responses includes financial and accounting choices, such as allocating a given amount of saving to tax-preferred versus other assets. The least responsive category of behavior applies to agents' real decisions, such as the level of saving. Our findings reflect this hierarchy in that we find a strong effect of saving incentives on the allocation of saving and wealth, but little or no effect on the level. If people can use a saving incentive account to raise their future living standards without the pain of reducing current consumption, we expect people to choose that route over reducing consumption. Our reading of the evidence is that people have by and large taken the less painful route to financing saving incentives, or at least have taken that route first. The plausibility of this notion and the fact that similar patterns have been discovered in a wider range of economic activities gives us more confidence in its applicability to saving issues as well.

One caveat is the data sets used to explore these issues—by us and by others—are all faulty in some important way. It is unclear if there is a completely satisfactory way, given existing data, to address these issues.

We cannot distinguish among various reasons why saving incentives have not appeared to work well to date (Slemrod, 1994). One possibility is that existing incentives are poorly designed (Bernheim and Scholz, 1993). Another is that households have low intertemporal elasticities of substitution (Hall, 1988), so that it will prove difficult to stimulate saving via any voluntary mechanism. People may be uninformed about the need and opportunities for saving, in which case educational programs may help. Or saving incentives may eventually raise saving, even if they have not done so yet, in which the case the policy implication is to be patient.

Finally, we do not address whether saving incentives are a desirable feature of tax policy. If saving incentives do raise private saving, there is a question as to whether they do so at an acceptable cost. Even if they do not raise saving at all, there may be equity reasons to provide access to saving incentives to certain groups, such as households that do not have pension coverage. Issues relating to the equity and efficiency of tax-based saving incentives are interesting items for future research.

■ *A longer and more detailed version of this paper is available on request from the authors. Our views on these issues have been shaped by discussions with many people; in particular, we thank Doug Bernheim and Jon Skinner for comments over several years; Alan Auerbach, Glenn Hubbard, Jon Skinner and especially Timothy Taylor for helpful comments on earlier drafts; Joel Dickson for generously providing some of the data for this paper; Joe Milano and Jasper Hoek for outstanding research assistance; and the National Institute on Aging and the National Science Foundation for research support. The views presented are our own and should not be taken to represent the views of any of the institutions with which we are affiliated.*

References

Allen, Steven G., Robert L. Clark, and Ann A. McDermed, "Pensions, Bonding, and Lifetime Jobs," *Journal of Human Resources,* 1993, *28*:3, 463–81.

Andrews, Emily S., "The Growth and Distribution of 401(k) Plans." In Turner, John A., and Daniel J. Beller, eds., *Trends in Pensions 1992.* Washington, D.C.: U.S. Department of Labor, Pension and Welfare Benefits Administration, 1992, pp. 149–76.

Attanasio, Orazio, and Thomas De Leire, "IRAs and Household Saving Revisited: Some New Evidence." Working Paper No. 4900, National Bureau of Economic Research, Cambridge, Mass., October 1994.

Auerbach, Alan J., "Tax Reform, Capital Allocation, Efficiency and Growth." In Aaron, Henry J., and William G. Gale, eds., *Economic Effects of Fundamental Tax Reform.* Washington, D.C.: Brookings Institution, 1996, forthcoming.

Auerbach, Alan J., and Joel Slemrod, "The Economic Effects of the Tax Reform Act of 1986," *Journal of Economic Literature,* forthcoming 1996.

Avery, Robert B., Gregory E. Elliehausen, and Thomas A. Gustafson, "Pensions and Social Security in Household Portfolios: Evidence from the 1983 Survey of Consumer Finances." In Adams, F. Gerard, and Susan M. Wachter, eds., *Savings and Capital Formation.* Lexington, Mass.: Lexington Books, 1986, pp. 127–60.

Bassett, William F., Michael J. Fleming, and Anthony P. Rodrigues, "How Workers Use 401(k) Plans: The Participation, Contribution, and Withdrawal Decisions," mimeo, Federal Reserve Board of New York, April 1996.

Bernheim, B. Douglas, "The Economic Effects of Social Security: Toward a Reconciliation of Theory and Measurement," *Journal of Public Economics,* August 1987, *33,* 273–304.

Bernheim, B. Douglas, "Comment on Engen, Gale and Scholz," *Brookings Papers on Economic Activity,* 1994, *1,* 152–71.

Bernheim, B. Douglas, "Rethinking Saving Incentives," mimeo, March 1996.

Bernheim, B. Douglas, and Daniel M. Garrett, "The Determinants and Consequences of Financial Education in the Workplace: Evidence from a Survey of Households," mimeo, Stanford University, August 1995.

Bernheim, B. Douglas, and John Karl Scholz, "Private Saving and Public Policy." In Poterba, James M., ed., *Tax Policy and the Economy.* Vol. 7., Cambridge, Mass.: Massachusetts Institute of Technology Press, 1993, pp. 73–110.

Blinder, Alan S., Roger H. Gordon, and Donald E. Wise, "Reconsidering the Work Disincentive Effects of Social Security," *National Tax Journal,* December 1980, *33*:4, 431–42.

Board of Governors of the Federal Reserve System, *Balance Sheets for the U.S. Economy: Selected Years.* Washington, D.C.: U.S. Government Printing Office, 1995.

Buck Consultants, Inc., *Current 401(k) Plan Practices: A Survey Report.* Secaucus, N.J.: Buck Consultants, Inc., September 1989.

Chang, Angela E., "Tax Policy, Lump-Sum Pension Distributions, and Household Saving," *National Tax Journal,* 1996, *49*:2, 235–52.

Curme, Michael A. and William E. Even, "Pension Coverage and Borrowing Constraints," *Journal of Human Resources,* 1995, *30*:4, 701–12.

Diamond, Peter A., and Jerry A. Hausman, "Individual Retirement and Savings Behavior," *Journal of Public Economics,* February/March 1984, *23,* 81–114.

Dicks-Mireaux, Louis, and Mervyn King, "Pension Wealth and Household Savings: Tests of Robustness," *Journal of Public Economics,* February/March 1984, *23,* 115–39.

Doyle, Robert J., Jr., and Eric T. Johnson, *Readings in Wealth Accumulation Planning.* 4th ed., Bryn Mawr, Penn.: American College, 1991.

Employee Benefits Research Institute, *EBRI Databook on Employee Benefits.* Washington, D.C.: Employee Benefits Research Institute, 1995.

Engen, Eric M., and William G. Gale, "IRAs and Saving in a Stochastic Life-Cycle Model," mimeo, UCLA and Brookings Institution, April 1993.

Engen, Eric M., and William G. Gale, "Debt, Taxes and the Effects of 401(k) Plans on Household Wealth Accumulation," mimeo, October 1995.

Engen, Eric M., and William G. Gale, "Debt, Taxes and the Effects of 401(k) Plans on Household Wealth Accumulation," revised mimeo, 1996.

Engen, Eric M., William G. Gale, and John Karl Scholz, "Do Saving Incentives Work?," *Brookings Papers on Economic Activity,* 1994, *1,* 85–180.

Engen, Eric M., William G. Gale, and John Karl Scholz, "Effects of Tax-Based Saving Incentives on Saving and Wealth," mimeo, Brookings Institution, August 1996.

Feenberg, Daniel, and Jonathan Skinner, "Sources of IRA Saving." In Summers, Lawrence H., ed., *Tax Policy and the Economy.* Vol. 3., Cambridge, Mass.: Massachusetts Institute of Technology Press, 1989, pp. 25–46.

Feldstein, Martin, "Inflation, Portfolio Choice, and the Prices of Land and Corporate Stock," *American Journal of Agricultural Economics,* December 1980, *62:*5, 910–16.

Feldstein, Martin, "The Effects of Tax-Based Saving Incentives on Government Revenue and National Saving," *Quarterly Journal of Economics,* May 1995, *110:*2, 475–94.

Feldstein, Martin, and Daniel R. Feenberg, "Alternative Tax Rules and Personal Saving Incentives: Microeconomic Data and Behavioral Simulations." In Feldstein, Martin, ed., *Behavioral Simulation Models in Tax Policy Analysis.* Chicago: University of Chicago Press and NBER, 1983, pp. 173–209.

Fernandez, Phyllis A., "Preretirement Lump-Sum Distributions." In Turner, John A., and Daniel J. Beller, eds., *Trends in Pensions 1992.* Washington, D.C.: U.S. Government Printing Office, 1992, pp. 285–317.

Gale, William G., "The Effects of Pensions on Wealth: A Re-Evaluation of Theory and Evidence," mimeo, June 1995.

Gale, William G., and John Karl Scholz, "IRAs and Household Saving," *American Economic Review,* December 1994, *84:*5, 1233–60.

Gravelle, Jane, "Do Individual Retirement Accounts Increase Saving?," *Journal of Economic Perspectives,* Spring 1991, *5,* 133–49.

Hall, Robert E., "Intertemporal Substitution in Consumption," *Journal of Political Economy,* April 1988, *96:*2, 339–57.

Halperin, Daniel I., "Tax Policy and Retirement Income: A Rational Model for the 21st Century." In Vanderhei, Jack L., ed., *Search for a National Retirement Income Policy.* Homewood, Ill.: Richard D. Irwin, Inc., 1987, pp. 159–95.

Hubbard, R. Glenn, "Do IRAs and Keoghs Increase Saving?," *National Tax Journal,* March 1984, *37,* 43–54.

Hubbard, R. Glenn, "Pension Wealth and Individual Saving," *Journal of Money, Credit, and Banking,* May 1986, *18:*2, 167–78.

Hubbard, R. Glenn, and Jonathan S. Skinner, "The Effectiveness of Saving Incentives: A Review of the Evidence," mimeo, June 1995.

Internal Revenue Service, *Statistics of Income Bulletin.* Washington, D.C.: U.S. Government Printing Office, various years.

Ippolito, Richard A., "Selecting and Retaining High-Quality Workers: A Theory of 401(k) Pensions," unpublished paper, Pension Benefit Guaranty Corporation, Washington, April 1993.

Johnson, Richard W., "The Impact of Worker Preferences on Pension Coverage in the HRS." Health and Retirement Study Working Paper Series Paper No. 94-018, December 1993.

Joines, Douglas H., and James G. Manegold, "IRAs and Saving: Evidence from a Panel of Taxpayers," mimeo, USC, 1995.

Papke, Leslie E., "Are 401(k) Plans Replacing Other Employer-Provided Pensions? Evidence from Panel Data," mimeo, Michigan State University, August 1996.

Papke, Leslie E., Mitchell Petersen, and James M. Poterba, "Did 401(k) Plans Replace Other Pensions?" Working Paper No. 4501, National Bureau of Economic Research, Cambridge, Mass., October 1993.

Poterba, James M., "Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach," *Quarterly Journal of Economics,* November 1984, *99:*4, 729–52.

Poterba, James M., Steven F. Venti, and David A. Wise, "Targeted Retirement Saving and the

Net Worth of Elderly Americans," *American Economic Review,* May 1994, *84*:2, 180–85.

Poterba, James M., Steven F. Venti, and David A. Wise, "Do 401(k) Contributions Crowd Out Other Personal Saving?," *Journal of Public Economics,* September 1995, *58,* 1–32.

Poterba, James M., Steven F. Venti, and David A. Wise, "Do Retirement Saving Programs Increase Saving? Reconciling the Evidence." NBER Working Paper No. 5599, 1996.

Ruggeri, G. C., and M. Fougere, "The Effects of Tax-Based Saving Incentives on Government Revenue and National Saving: Comment," mimeo, 1995.

Skinner, Jonathan, "Do IRAs Promote Saving? A Review of the Evidence," *Tax Notes,* January 13, 1992.

Slemrod, Joel, "The 1979–84 Linked Panel of Tax Return Data: Sampling and Linking Methodology," mimeo, University of Michigan, November 1988.

Slemrod, Joel, "Notes on the 1979–86 Panel of Individual Income Tax Returns," mimeo, University of Michigan, July 1990a.

Slemrod, Joel, "The Economic Impact of the Tax Reform Act of 1986." In Slemrod, Joel, ed., *Do Taxes Matter? The Impact of the Tax Reform Act of 1986.* Cambridge, Mass.: Massachusetts Institute of Technology Press, 1990b, pp. 1–12.

Slemrod, Joel, "Comments and Discussion on 'Do Saving Incentives Work?,'" *Brookings Papers on Economic Activity,* 1994, *1,* 166–71.

Stiglitz, Joseph E., *Economics of the Public Sector.* 2nd ed., New York: W.W. Norton and Company, 1988.

Summers, Lawrence H., "Inflation, the Stock Market, and Owner-Occupied Housing," *American Economic Review,* May 1981, *71*:2, 429–34.

United States Department of Labor, Pension and Welfare Benefits Administration, "Private Pension Plan Bulletin, Abstract of 1992 Form 5500 Annual Reports, Number 5," Winter 1996.

Venti, Steven F., and David A. Wise, "Tax-Deferred Accounts, Constrained Choice, and Estimation of Individual Saving," *Review of Economic Studies,* August 1986, *53,* 579–601.

Venti, Steven F., and David A. Wise, "IRAs and Saving." In Feldstein, Martin, ed., *The Effects of Taxation on Capital Accumulation.* Chicago: University of Chicago Press, 1987, pp. 7–48.

Venti, Steven F., and David A. Wise, "Have IRAs Increased U.S. Saving?: Evidence from Consumer Expenditure Surveys," *Quarterly Journal of Economics,* August 1990, *105,* 661–98.

Venti, Steven F., and David A. Wise, "The Saving Effect of Tax-Deferred Retirement Accounts: Evidence from SIPP." In Bernheim, B. Douglas, and John B. Shoven, eds., *National Saving and Economic Performance.* Chicago: University of Chicago Press and NBER, 1991, pp. 103–30.

Venti, Steven F., and David A. Wise, "Government Policy and Personal Retirement Saving." In Poterba, James M., ed., *Tax Policy and the Economy.* Vol. 6, Cambridge, Mass.: Massachusetts Institute of Technology Press, 1992, pp. 1–41.

Venti, Steven F., and David A. Wise, "Individual Response to Retirement Saving Programs: Results From U.S. Panel Data," *Ricerche Economiche,* 1995, *49,* 235–54.

Venti, Steven F., and David A. Wise, "The Wealth of Cohorts: Retirement Saving and the Changing Assets of Older Americans." NBER Working Paper No. 5609, 1996.

This article has been cited by:

1. Raj Chetty. 2015. Behavioral Economics and Public Policy: A Pragmatic Perspective. *American Economic Review* **105**:5, 1-33. [Abstract] [View PDF article] [PDF with links]
2. Michal Grinstein-Weiss,, Michael Sherraden,, William G. Gale,, William M. Rohe,, Mark Schreiner,, Clinton Key. 2013. Long-Term Impacts of Individual Development Accounts on Homeownership among Baseline Renters: Follow-Up Evidence from a Randomized Experiment. *American Economic Journal: Economic Policy* **5**:1, 122-145. [Abstract] [View PDF article] [PDF with links]
3. Alexander M. Gelber. 2011. How do 401(k)s Affect Saving? Evidence from Changes in 401(k) Eligibility. *American Economic Journal: Economic Policy* **3**:4, 103-122. [Abstract] [View PDF article] [PDF with links]
4. Andrew A. Samwick, Jonathan Skinner. 2004. How Will 401(k) Pension Plans Affect Retirement Income?. *American Economic Review* **94**:1, 329-343. [Citation] [View PDF article] [PDF with links]
5. Orazio P. Attanasio, Susann Rohwedder. 2003. Pension Wealth and Household Saving: Evidence from Pension Reforms in the United Kingdom. *American Economic Review* **93**:5, 1499-1521. [Abstract] [View PDF article] [PDF with links]
6. B. Douglas Bernheim,, Jonathan Skinner,, Steven Weinberg. 2001. What Accounts for the Variation in Retirement Wealth Among U.S. Households?. *American Economic Review* **91**:4, 832-857. [Abstract] [View PDF article] [PDF with links]