

NBER WORKING PAPER SERIES

THE EFFECTS OF TAX-BASED SAVING  
INCENTIVES ON SAVING AND WEALTH

Eric M. Engen  
William G. Gale  
John Karl Scholz

Working Paper 5759

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 1996

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, Michael Fleming, Jon Skinner and Tim Taylor for helpful comments on earlier drafts; Joel Dickson and John Sabelhaus 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. This paper is part of NBER's research program in Public Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1996 by Eric M. Engen, William G. Gale and John Karl Scholz. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

THE EFFECTS OF TAX-BASED SAVING  
INCENTIVES ON SAVING AND WEALTH

**ABSTRACT**

This paper evaluates research examining the effects of tax-based saving incentives on private and national saving. Several factors make this an unusually difficult problem. First, households that participate in, or are eligible for, saving incentive plans have systematically stronger tastes for saving than other households. Second, the data indicate that households with saving incentives have taken on more debt than other households. Third, significant changes in the 1980s in financial markets, pensions, social security, and nonfinancial assets interacted with the expansion of saving incentives. Fourth, saving incentive accounts represent pre-tax balances, whereas conventional taxable accounts represent post-tax balances. Fifth, the fact that employer contributions to saving incentive plans are a part of total employee compensation is typically ignored.

A major theme of this paper is that analyses that ignore these issues overstate the impact of saving incentives on saving. We show that accounting for these factors largely or completely eliminates 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 private or national saving.

Eric M. Engen  
Federal Reserve Board  
Mail Stop 83  
Washington, DC 20551  
and NBER  
eengen@frb.gov

William G. Gale  
The Brookings Institution  
1775 Massachusetts Avenue  
Washington, DC 20036  
wgale@brook.edu

John Karl Scholz  
Department of Economics  
University of Wisconsin-Madison  
Madison, WI 53706  
and NBER  
jkscholz@facstaff.wisc.edu

## (I) Introduction

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 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 pre-retirement 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, a number of studies have examined the effects of saving incentives and reached a variety of conclusions. The crucial issue in this literature is determining what households who have (or are eligible for) saving incentive accounts would have saved in the absence of these incentives. Several factors, however, make this a difficult problem and one 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 the saving behavior of households with and without saving incentives will be biased in favor of 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 indicate that households with saving incentives have taken on more debt than other households. Hence, studies should focus on how saving incentives affect wealth (assets minus debt), not just assets. Because financial assets are small relative to total assets, studies that focus only on the effects of saving incentives on 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 accounts represent pre-tax balances; one cannot consume the entire amount because taxes and

perhaps penalties are due upon withdrawal. In contrast, contributions to other accounts are generally not deductible and one may generally consume the entire balance in a taxable account.<sup>1</sup> Therefore, a given balance in a saving incentive account represents less saving (defined either as reduced previous consumption or increased future consumption) than an equivalent amount in a conventional account.

Fifth, saving incentives that are part of a worker's total compensation can have different effects than other plans. All else equal, a worker that has a 401(k) with employer contributions has higher total compensation than a worker with the same cash wage that does not have a 401(k) and would be expected to have higher saving as well. Ignoring the distinction between total compensation and cash wages 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 efforts in this area—overstate the impact of saving incentives on saving. We show that accounting for these factors largely or completely eliminates 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 private or national saving. Our conclusion does not rule out the possibility that existing saving incentives have raised saving for some individuals, or that saving incentives may eventually raise saving. Nor do we rule out that saving incentives in different forms might raise saving.

The next section sets the stage by describing the main features of existing programs and

---

<sup>1</sup>The major exception applies to the portion of asset value that represents capital gains, which are taxed when the asset is sold. But estimates by Kennickell and Wilcox (1992) imply that unrealized capital gains accounted for only about 10 percent of the value of all household financial assets held outside of retirement accounts in 1989.

examining 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 raise saving with our own results and conclusion that the incentives have not raised saving. The remainder of the paper addresses related issues. These include substitution between 401(k)s and pensions at the firm level; results from simulation models; the effects on public revenues; the effects of saving incentives in Canada; and the application of psychological models to understand saving incentives. The conclusion discusses some of the broader ramifications of our findings. We use two appendices to respond to criticisms by others concerning our previous work.

## (II) Background

Individual Retirement Accounts (IRAs) were established in 1974 for workers without pensions and allowed tax-deductible contributions up to an annual limit. In 1981, eligibility was extended to all workers and the limits were raised to their current levels. For a single worker, the limit is the minimum of earnings or \$2000. The presence of a non-working spouse raises the limit by \$250. The limit in a two-worker household is the minimum of earnings or \$2000 for each worker, for a maximum combined limit of \$4000. The Tax Reform Act of 1986 (TRA86) restricted IRA deductibility. For households with a retirement plan, deductibility is phased out (eliminated) when adjusted gross income exceeds \$40,000 (\$50,000) for joint tax-filers, and \$25,000 (\$35,000) for single tax-filers. TRA86 did not restrict IRA eligibility or the tax-free accrual of interest. Deductible contributions are taxed when withdrawn, while non-deductible contributions are not. In either case, returns on contributions are untaxed until

withdrawn. Withdrawals are subject to income tax and, if the account holder is younger than 59.5, an additional 10 percent penalty.

Although established in 1978, 401(k)s were not generally used until regulations clarifying their tax treatment were issued in 1981. Like IRAs, 401(k)s feature tax-deductible contributions, tax-free accrual of interest, annual contribution limits, and restrictions on withdrawals. But the accounts differ in several respects. 401(k)s are available only to employees of organizations that choose to sponsor the plans.<sup>2</sup> Contributions to 401(k)s are typically made via regular payroll deductions, while IRA contributions can be made on an irregular basis. Employers' 401(k) contributions are tax-deductible and may depend on the level of employee contributions. Most employees offered 401(k)s are also offered employer contributions.<sup>3</sup> The IRS limit on annual employee contributions is \$9,500 in 1996, but many workers face lower limits due to nondiscrimination rules or rules set by the employer. Borrowing against the existing 401(k) balance is allowed in most plans, as are withdrawals due to hardship conditions.

Keogh plans and simplified employee pensions are similar to IRAs but have higher limits and apply to the self-employed and employees of small business.

---

<sup>2</sup>Strictly speaking, section 401(k) of the Internal Revenue Code authorizes plans only for employees of for-profit firms. Sections 403(b) and 457 authorize salary-reduction plans for employees of non-profit institutions and state and local governments, respectively. Federal government employees may participate in thrift saving plans. All of these plans are similar to 401(k) plans, and we refer to them collectively as 401(k) plans.

<sup>3</sup>Using data from the 1993 Current Population Survey, Survey of Employee Benefits, Bassett, Fleming, and Rodrigues (1996) find that among workers offered a 401(k) plan, 72.9 percent were offered a positive employer match, with a mean match rate of 62.3 percent among those with positive match rates.

### (III) Aggregate Trends in Saving Incentives and Personal Saving

Figure 1 compares overall contributions to saving incentives and personal saving in recent years.<sup>4</sup> As measured in the National Income and Product Accounts, personal saving is the difference between personal disposable income and personal spending. 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, but the personal saving rate was flat. Following the Tax Reform Act of 1986, both saving incentive contributions and personal saving fell. The coincident fall from 1986 to 1987 may be suggestive of a causal relation, but personal saving had been declining since 1984, and it fell in 1987 by almost twice as much as saving incentive contributions did. In addition, the amount of saving that occurs in saving incentives includes interest and dividend earnings, plus rollovers, less withdrawals and hence is larger than the contributions alone.<sup>5</sup> 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 accounted for only a small part of the fall in personal saving from 1986 to 1987.

---

<sup>4</sup>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 (EBRI 1995). Keogh contributions rose over the 1980s but have remained below \$9 billion per year (Statistics of Income, various years). Contributions to 401(k) plans grew steadily from \$16 billion in 1984 to \$64 billion in 1992 (Department of Labor, 1996).

<sup>5</sup>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).



Taking the 1980-92 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.<sup>6</sup> None of these patterns provides any evidence that saving incentives influence the level of personal saving.<sup>7</sup>

The aggregate impact of saving incentives would be uninteresting if they were a tiny fraction of total saving. 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 personal saving or at least to expect sharp increases in saving incentive contributions to be reflected in the aggregate statistics. But 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.

#### (IV) Cross-sectional evidence on IRAs

Confronting the heterogeneity in underlying saving behavior between contributors and non-contributors is crucial to understanding IRAs. For example, 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

---

<sup>6</sup> These patterns are similar if one uses alternate measures of personal saving, like the Flow of Funds measure or the Flow of Funds calculated on a NIPA basis.

<sup>7</sup>As noted above, the effect of saving incentives on national saving is of most interest. But if there is no impact on personal saving, the effect on national saving will be worse at least in the short run. National saving has fallen since 1982, due to generally higher federal budget deficits and reduced private saving.

saving hinges on the behavior of households that contribute to IRAs, not of those that do not. And 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 that 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. These transfers provide a tax break for doing what the household would have done and will not raise private saving. Contributors who view IRAs as poor substitutes for other assets will fund contributions by reducing consumption (raising saving). 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 households, IRAs are likely to be good substitutes for other saving. Households older than 59 face no penalty on IRA withdrawals and those with large amounts of other assets can avoid penalties by consuming other assets first. Thus, for most contributors, the effects of IRAs on saving seem likely to be small at best.

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 indicate substantial opportunity at the beginning of the IRA program for contributors to shift existing assets into IRAs and contrast sharply with median

wealth of \$20,448 and total financial assets of \$2,050 for households without IRAs.<sup>8</sup>

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, there is substantial heterogeneity in unobserved tastes or influences on saving (Diamond and Hausman 1984).<sup>9</sup> 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.<sup>10</sup> 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 substitution between IRAs and other saving, limit contributors would need to have less non-IRA saving than non-contributors, controlling for household characteristics.

---

<sup>8</sup>The figures in this paragraph are based on tabulations of the 1983 Survey of Consumer Finances.

<sup>9</sup>For example, households with different discount rates would adjust their saving differently to a one dollar increase in initial wealth (Gale and Scholz 1994).

<sup>10</sup>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 arguments. It is unclear what sort of utility function or preferences would correspond to such a decision function. Also, Venti and Wise restrict the fraction of IRAs that are new saving to be between zero and one, although there is no theoretical justification for doing so. This restriction rules out the possibility that households are target savers and even that income effects of IRAs are larger than the substitution effects. 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 non-contributors, as is now widely understood to be the case (see the discussion in Appendix A). Thus, the Venti and Wise model interprets findings that limit contributors saved more than non-contributors 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. In short, the identifying assumption biases the results toward finding that IRAs raise saving even if they do not.<sup>11</sup>

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. They assume a quadratic utility function, which allows derivation of a specific, estimable saving function.

The model implies that, in the absence of IRAs, households who contributed to the limit in each of three years would have saved the same as other IRA contributors, rather than non-contributors.<sup>12</sup> Thus, the model allows for differences in saving between contributors and

---

<sup>11</sup>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 IRAs. As noted in the text, most contributors and limit contributors may find IRAs to be good substitutes for taxable saving.

<sup>12</sup>This assumption is consistent with the data. See Appendix A.

non-contributors.<sup>13</sup> The model also allows the substitutability between IRAs and other saving to depend on household characteristics.

The model is estimated on households from the 1983-1986 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. 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 even smaller or negative.

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

In Appendix A, we briefly describe the Gale and Scholz (1994) model and address criticisms of the model raised by Poterba, Venti, and Wise (1996) and Bernheim (1996).<sup>14</sup>

---

<sup>13</sup>As noted by Hubbard and Skinner (1995), allowing the saving equations to differ across contributors and non-contributors represents a crucial difference with the Venti-Wise approach and makes the Gale-Scholz model more general than the Venti-Wise model.

<sup>14</sup>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) and Poterba, Venti and Wise (1996).

(V) Cross-sectional evidence on 401(k)s: Is 401(k) Eligibility Exogenous?

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 were 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 while employers ultimately decide on the policy, a relevant issue is whether employers take employee preferences into account. In a survey of a broad range of employers, "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). This should not be surprising; it would be strange if employers created benefits without regard to employee preferences. Moreover, even if firms did provide 401(k)s randomly, we would expect workers with high tastes for saving to seek out firms with 401(k)s or to encourage their firms to provide 401(k)s. These patterns would be consistent with theoretical and empirical models of pensions.<sup>15</sup> But 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

---

<sup>15</sup>Ippolito (1993) explains the growth of 401(k)s precisely in terms of their ability to help firms attract and retain workers with higher tastes for saving. For related empirical work, see Allen, Clark, and McDermed (1993), Curme and Even (1995), Johnson (1993), and Ippolito (1993).

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 non-pension, non-401(k) financial assets as 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. Bernheim (1994a, 1996) shows that evidence in Poterba, Venti, and Wise's own work indicates that differences in median financial assets between eligible and ineligible households are, in several income classes, several times as large as median 401(k) balances for eligible households. Bernheim and Garrett (1995), using cross section data, find that 401(k) eligibility "raises" total wealth by about four times as much as it "raises" retirement wealth. Engen, Gale, and Scholz (1994, Table 8) use a similar sample from the same data set as Poterba, Venti and Wise (1995), a slightly different test format and a longer list of explanatory variables, and find that eligible families have higher levels of non-pension, non-401(k) financial assets, net financial assets and net worth. Unless 401(k) contributions crowd in several times their value in non-401(k) saving, each of these findings suggest that eligibility is positively correlated with tastes for saving.

A second problem is that 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 ineligible households. Moreover, if pension coverage is positively correlated with tastes for saving (see footnote 15), then the difference in coverage is further evidence that 401(k)

eligibility is not exogenous. Even if pension coverage is independent of tastes for saving, the higher pension wealth should show up as lower levels of non-pension wealth for eligibles than ineligibles if any of the pension wealth is offset by reductions in non-pension wealth. In short, the findings on pensions imply that to believe that 401(k) eligibility is exogenous requires additional assumptions that (a) pension coverage is not correlated with tastes for saving, and (b) all pension wealth is new wealth.

The Poterba, Venti, and Wise 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. To determine whether 401(k) eligibility is exogenous requires knowing whether eligible families would have saved more than ineligible families in the absence of 401(k)s. If  $x$  percent of 401(k) wealth would have existed anyway, 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.

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.

Bernheim and Garrett (1995) take a different approach to this problem.<sup>16</sup> They use a cross-section of households and estimate the impact of 401(k) eligibility on self-reported

---

<sup>16</sup>The main purpose of the Bernheim-Garrett study is to examine the effects of financial education on saving, but the regressions also contain estimates of the impact of 401(k) eligibility on saving.



household saving rates, conditioning on initial wealth and other factors.<sup>17</sup> If initial wealth adequately controls for tastes for saving, this procedure may help identify the effects of 401(k)s on saving. Bernheim and Garrett (1995) find that 401(k) eligibility raises self-reported saving rates by 1.5 percentage points. When placed in context, we believe this result supports the finding that only a small portion of 401(k) contributions represent new saving.

The typical 401(k) eligible worker appears to give at least 5 percent of salary to his 401(k).<sup>18</sup> Thus, at most 30 percent (1.5/5) of 401(k) contributions might be new saving. But about 20 percent or more of the contributions represent tax deductions rather than saving, 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 estimate, however, overstates the effect on saving; another problem is that the regressions in Bernheim and Garrett (1995) (as well as those in Poterba, Venti and Wise 1995) overstate the impact of 401(k)s because they control for cash earnings rather than total compensation. Employer contributions are a part of total compensation, but not cash earnings. Thus, if two households have the same cash earnings and one has a 401(k) with employer contributions, the latter has higher total compensation and will have higher overall saving even

---

<sup>17</sup>The specification differs from PVW in that Bernheim and Garrett use a saving rate measure rather than wealth as the dependent variable, and control for wealth as an explanatory variable.

<sup>18</sup>The average employee contribution among participating workers is 7 percent or more of salary. More than 70 percent of those offered 401(k)s also receive employer contributions, with an average matching rate of over 60 percent. Thus, the combined contribution for the average contributor is likely to be in the range of 10-11 percent of salary. Since about two-thirds or more of eligible workers contribute to 401(k)s, a reasonable estimate for the combined contributions of a typical eligible worker would be at least 5 percent of salary. See, for example, EBRI 1994, Table 8; and Bassett, Fleming, and Rodrigues (1996).

if 401(k)s are fully offset by reductions in other saving. If the regression controls for cash earnings, but not for employer 401(k) contributions, the coefficient on 401(k) eligibility will pick up any offset between 401(k)s and other saving—which should produce a zero or negative coefficient—but will also pick up an income or wealth effect associated with the 401(k)-eligible household having higher total compensation and higher saving—which would tend to produce a positive coefficient. Therefore, the coefficient on 401(k) eligibility will overstate the impact of 401(k)s on saving if the regression controls for cash earnings but not total compensation. This bias is described in more detail in Appendix B below and in Gale (1995). Thus, we interpret the Bernheim-Garrett results as showing that less than 10 percent of 401(k) contributions represent net additions to saving.<sup>19</sup>

#### (VI) 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. Median financial assets for families aged 60-64 with saving incentives in 1991 were about \$15,000 higher than for similarly aged families with saving incentives in 1984. Median saving incentive balances were about

---

<sup>19</sup>One caveat to this conclusion is that the dependent variable measures saving out of family income, while the 401(k) contributions variable is measured in our example as a proportion of the worker's earnings, which are less than family income. But 26 percent of the Bernheim-Garrett sample is single and many married couples have only one earner, so it is unclear whether the bias is very large. Assuming that half of the households have one earner, typical contributions would be about 3.75 percent of family income for eligible households (5 percent for single earners, and 2.5 percent for an earner in a dual earner couple), and about 0.75 percent (=20 percent of 3.75 percent) of salary would represent deferred taxes, rather than saving. Hence, the proportion of 401(k) contributions that would represent new saving would be  $(1.5-.75)/3.75 = 20$  percent. Even this would be an overstatement, though, due to use of cash wages rather than total compensation in the regressions.

\$14,000 higher. 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 developments in the 1980s and for data problems in the analysis leaves essentially no room for saving incentives to have raised wealth.<sup>20</sup>

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

For example, 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

---

<sup>20</sup>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.

saving incentive balances grew by less than \$1 trillion.<sup>21</sup> Several major changes occurred during this period, each of which suggests that financial assets would have risen in the 1980s, independent of saving incentives. None of these factors are controlled for in the Venti and Wise analysis, but taken together, they can explain virtually all of the increase in assets shown in table 1.

We consider first the mean results in the top panel. 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 financial assets.<sup>22</sup> In addition, real interest rates were higher between 1984 and 1991 than in the preceding several

---

<sup>21</sup>Board of Governors of the Federal Reserve System, selected years; Employee Benefits Research Institute, 1995.

<sup>22</sup>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 1984 IRA and Keogh balances were \$3,800, of which we estimate (based on EBRI data) about \$950 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 ( $= \$7,150 \times 1.78$ ) in 1991. If the 1991 cohort had instead faced the 5 percent drop in the S&P experienced by the 1984 cohort between 1977 and 1984, the \$7,150 would have fallen 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 documented in table 1. These calculations, however, understate the importance of the stock market boom because the SIPP data used here and by Venti and Wise omit balances held in 401(k)s and thrifts in 1984. Estimates from the 1983 Survey of Consumer Finances show that average thrift balances for 53-57 year olds in 1983 were twice as large as average IRA and Keogh balances. This suggests adding an additional \$1,900 to mean stock and mutual fund holdings of 53-57 year olds in 1984. Making this adjustment and redoing the calculations above with \$9,050 rather than \$7,150 in 1984 holdings for 53-57 year olds implies that differences in stock market performance can account for over \$7,500, or about 92 percent, of the increase in mean financial assets shown in the top panel of table 1.

years.<sup>23</sup> 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 observation in the data.

Changes in four other components of wealth also fueled the increase in financial assets in 1980s. Considering these changes is important because financial assets are only a small part of wealth. In 1991, for example, mean financial assets were only 17 percent of mean net worth for 65-69 year olds and median financial assets were only 5 percent of median net worth (Poterba, Venti, and Wise 1994).

First, the 1980s saw a substantial decline in inflation and marginal tax rates. As a result, investors shifted away from tangible capital (like housing) that was more attractive in the 1970s (Feldstein 1980, Summers 1981 and 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). Third, the value of social security benefits were 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

---

<sup>23</sup>The nominal return on three month Treasury bills less the increase in the consumer price index averaged 2.41 percent between 1977 and 1984 and 3.24 percent between 1984 and 1991. The corresponding figures for AAA-rated corporate bonds are 4.17 percent and 6.11 percent, respectively (Economic Report of the President, 1996).

declines in social security and pensions could have induced an increase in financial assets.<sup>24</sup> A number of data problems also create biases. First, saving incentives contributions are tax-deductible and the withdrawals are taxed; hence, the reported balance is a pre-tax amount. Other saving is treated in the opposite manner, and the balances in such accounts represent post-tax amounts other than the portion of asset value that represents capital gains. Saving can be thought of as either reduced previous consumption or increased future consumption. In either case, the extra \$9,000 held 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  $(1-t) * \$9000$ , where  $t$  is the combined marginal federal and state tax rate. Because the withdrawal is fully taxed, future consumption has increased by only  $(1-t) * \$9000$ . Of course, realized capital gains from taxable accounts are also subject to tax. But Kennickell and Wilcox (1992) estimate that capital gains represented only one-third of the asset value of stocks and mutual funds outside of retirement accounts in 1989. This represents only about 10 percent of the value of total household financial assets outside of retirement accounts in 1989.

Estimating the appropriate " $t$ " requires combining several factors. The average marginal federal tax rate for IRA participants from 1979-88 was 22 percent.<sup>25</sup> Adding in a measure of state tax rates and accounting for the small proportion of taxable financial assets facing capital

---

<sup>24</sup>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.

<sup>25</sup>This figure is based on estimates from 1979-88 using the IRS-Michigan tax panel, as reported in Engen, Gale and Scholz (1994). Characteristics that help determine effective tax rates, such as income, family size, home ownership, are roughly similar between 401(k) participants and IRA participants. Venti and Wise (1986, 1987, 1990, and 1991) estimate a marginal tax rate of about 35 percent for IRA holders in 1982-6.

gains taxes, we conclude that  $t=20$  percent is, if anything, a very conservative estimate of the difference in effective tax rates.<sup>26</sup> If so, then at least \$1,800 of the \$9,000 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. Thrifts were prevalent before 401(k)s became popular, and a substantial portion of 401(k)s appear to have been converted from thrift plans over the course of the 1980s. As these conversions occurred, they appeared in the SIPP as 401(k) balances in 1987 and 1991 even though their equivalent balance as thrift plans did not appear in 1984 or in later years. Thus, omitting after-tax thrift plans in all years and 401(k)s in 1984 understates 1984 wealth and overstates the increase in wealth from 1984 to 1991. Data from the 1983 Survey of Consumer Finances indicates that mean balances in thrifts 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 can be explained by the 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. 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 in the 1983 SCF, including thrift balances raised median total financial assets

---

<sup>26</sup>Hubbard and Skinner (1995) assume that saving incentive balances will be taxed at a 25 percent rate when withdrawn and describe that figure as likely to be an underestimate of the true rate. Using that estimate, if 10 percent of financial assets are subject to capital gains taxation at a rate of 30 percent, for an effective rate of 3 percent on all financial assets, then the appropriate rate at which to adjust saving incentive balances would be 22 (=25-3) percent.

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 non-participants and leads to an overstatement of the increase in relative wealth holdings of participants from 1984 to 1991. Thus, these data issues alone can explain a large proportion of the increase in financial assets for participants relative to non-participants.

In addition, the stock market boom and high real interest rates raised financial assets more in arithmetic terms for participants than for non-participants because participants in 1984 had many times more financial assets than non-participants.<sup>27</sup> Moreover, debt increased more rapidly for 401(k) and IRA participants than for non-participants (see Engen, Gale, and Scholz 1994, Engen and Gale 1995) but debt holdings are omitted from the table.<sup>28</sup>

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, the rise in real interest rates, the role of shifts in non-financial assets, debt,

---

<sup>27</sup>Among households that were 53-57 years old in the 1984 (the cohort that was 60-64 in 1991), median financial assets were \$23,000 among saving incentive participants but only \$1,200 for non-participants. (These figures do not control for household characteristics and so are not directly comparable to the figures in table 1.) If the stock market boom and the rise in real interest rates raised annual rates of return by 3 percentage points between 1984 and 1991 as compared to the period between 1977 and 1984, this would have caused initial balances to rise by over \$5,000 for the median saving incentive participant from 1984 to 1991, but by only \$300 for the median non-participant. Thus, higher rates of return in the 1984-91 period translate mechanically into larger differences in wealth between 60-64 year old participants and nonparticipants in 1991 as compared to 1984.

<sup>28</sup>The shift from non-financial assets to financial assets caused by falling inflation and marginal tax rates likely also caused an increase in the financial assets of saving incentive participants relative to non-participants. 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 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.



pensions and social security, and data problems.

(VII) 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 households 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.<sup>29</sup>

Nevertheless, Poterba, Venti and Wise (1996, p. 27) emphasize that "the critical feature of our 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 asset balances 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 non-financial assets due to falling inflation and tax rates, and the

---

<sup>29</sup>For example, the second panel of table 1 compares 60-64 year old participants in 1984 and 1991, while the test using eligible families described in the text would compare 25-64 year old eligible families in 1984 with 25-64 year old eligible families in 1991.

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.<sup>30</sup> 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 1984-91. Similar results occur for comparisons of eligible and ineligible families conditional on having an IRA. Controlling for IRA status is intended to control for tastes for saving not captured by the other variables in the model. Poterba, Venti and Wise conclude that most 401(k) contributions represent new saving.

As with cohort analysis, however, 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. In addition, mortgage debt increased disproportionately for eligible households relative to ineligible households (Engen and

---

<sup>30</sup>The second and third panels of table 1 compare 60-64 year old saving incentive participants and non-participants over time. Similarly, the successive cross-section test described in the text would compare 25-64 year old eligible and ineligible households over time.

Gale 1995). Both of these factors imply that the financial asset measure used by Poterba, Venti, and Wise 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 thrifts and 1984 401(k) balances are omitted. All of these factors impart biases in favor of finding that reported total 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 thrifts. 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. First, 401(k)s and housing are both illiquid, tax-preferred assets that are often held for long periods, so substitution would not be surprising. Second, there are many ways to substitute between 401(k)s and housing equity. About 14 percent of 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. Thus, a substantial minority of eligible households clearly had direct access to one of these mechanisms.

Substitution can occur in less obvious ways as well, for example, by not accelerating mortgage payments or by not trading up into a bigger house. These channels be difficult to

capture in survey data. 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 downpayments (because of their higher 401(k)s). If they purchase the same size home as the earlier cohort, households in the later cohort would take out a larger loan, and so would have less housing equity, more 401(k) wealth, but the same overall wealth as households in the earlier cohort. Clearly, households in the later cohort substituted 401(k)s for home equity relative to earlier cohorts, even if this substitution was unintentional.<sup>31</sup>

Since most saving incentive participants own their home, 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 and makes saving incentives appear disproportionately important in household portfolios. In addition, because of employer matching of 401(k) contributions, financing a 401(k) with tax-deductible mortgage borrowing can be quite lucrative.<sup>32</sup> Finally, much of the literature on social security and pensions examines the impact of such policies on broad measures of wealth, rather than just financial assets. The literature indicates that looking at broader wealth definitions

---

<sup>31</sup>Engen and Gale (1995) discuss other ways that unintentional substitution can occur, and Stiglitz (1988, page 595) discusses how unintentional substitution could plausibly occur between IRAs and home equity.

<sup>32</sup>Indeed, simple calculations show that with typical employer matching contributions, workers should do everything possible to maximize 401(k) contributions at least up to the match limit. Kusko, Poterba, and Wilcox (1994) report that about 75 percent of 401(k) participants in their sample at one firm contributed at or above the match limit.

generates larger offsets for pension wealth.<sup>33</sup> These results are consistent with Engen and Gale's results.

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 they also show that eligible households did not accumulate more wealth when housing equity is included. For renters, eligibility is estimated to have no significant effect on holdings of financial assets. For homeowners, eligibility raises financial assets, but does not raise the sum of net financial assets and housing equity. 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 summary, 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. In a revised version of the paper, Engen and Gale (1996) also show that the non-401(k) wealth of eligible families fell relative to ineligible families. Finally, we note that the estimated effects on financial assets suffer from some of the same biases as Poterba, Venti, and Wise (1995) and so should not be interpreted as strong evidence that 401(k)s raise financial assets.

In Appendix C, we discuss additional features of the Engen and Gale results and briefly

---

<sup>33</sup>Avery, Elliehausen, and Gustafson (1986), Bernheim (1987), Blinder, Gordon, and Wise (1980), Diamond and Hausman (1984), Dicks-Mireaux and King (1984), and Hubbard (1986) estimate the impact of pensions on broad measures of wealth. 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. Engen, Gale, and Scholz (1994) and Gale (1995) obtain similar results.

address criticisms raised by Bernheim (1996) and Poterba, Venti, and Wise (1996).

In Engen, Gale, and Scholz (1994) we explore an alternative identification strategy based on 401(k) participation as well as eligibility. We use SIPP data to compare 401(k) participants with IRA participants who are ineligible for 401(k) plans. Families in each group have chosen to hold saving incentives, indicating they are "savers," so we make the explicit assumption that in the absence of 401(k)s, after conditioning on observable characteristics, saving of 401(k) contributors would be the same as saving of IRA contributors. Between 1986 and 1991, IRA tax preferences were restricted; moreover, contribution limits were higher for 401(k)s than IRAs. Over this period, 401(k) contributions grew rapidly and IRA contributions plummeted. Thus, if 401(k)s raise private wealth, 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 concluded that 401(k)s did not raise wealth.

This test has been criticized on grounds that average tastes for saving among the 401(k) group fell over time so that the result may be biased toward showing that 401(k)s do not raise wealth (see Bernheim 1994, Hubbard and Skinner 1995, Poterba, Venti and Wise 1996). In response, Engen and Gale (1995) create samples where IRA participation—a plausible indicator of dilution once household characteristics are controlled for—does not change over time.<sup>34</sup> Thus, they compare, for example, the change in wealth for 401(k) participants with IRAs to the change in wealth for non-participants with IRAs. Results mirror those based on eligibility. For homeowners, 401(k)s "raise" financial assets (subject to the caveats made above), but not 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

---

<sup>34</sup>See also the discussion in Engen, Gale, and Scholz (1994).

the view that 401(k)s have raised wealth.

### (VIII) Evidence from Panel Data

In panel data, the same households or individuals are followed over time. This 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.<sup>35</sup> Joines and Manegold 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 their saving prior to IRA eligibility. More specifically, they argue that the increase in saving by new contributors should exceed the change in saving by continuing contributors, who contributed both before and after 1981. This test therefore compares one group of contributors to another, rather than comparing contributors to non-contributors. 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.<sup>36</sup>

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,

---

<sup>35</sup>Additional information on the tax panel can be found in Slemrod (1988, 1990), and Engen, Gale and Scholz (1994, 1996).

<sup>36</sup>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.

show that average non-mortgage debt rose by \$2400 for IRA contributors over this period, but by only \$1350 for non-contributors. Median non-mortgage debt rose by \$2,350 for contributors compared to only about \$100 for non-contributors. Second, the regressions are estimated on a restricted sample that omits taxpayers aged 65 and over. For 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.

An additional concern is that continuing contributors and new contributors may differ in important ways that are not captured by the specification, which can lead to either an understatement or an overstatement of the effects of IRAs.<sup>37</sup> Finally, Poterba, Venti and Wise (1996) note that the Joines and Manegold estimate may understate the effects of IRAs on saving for households that do not contribute the limit amount. This bias is likely to be small, however, because the vast majority of contributions were either at or very close to the limit: 75 percent of contributions in 1983, for example, were limit contributions, and another 13 percent were contributions with one spouse at the limit (Burman, Cordes and Ozanne (1990). The latter are typically households where the combined limit is \$2,250 and the household contributes \$2,000.

Feenberg and Skinner (1989) use the same data set as Joines and Manegold and find that, controlling for initial wealth, IRA holders increased their taxable saving by more than non-contributors from 1982 to 1984. However, since contributors have stronger tastes for saving

---

<sup>37</sup>If continuing contributors have stronger tastes for saving, as Bernheim (1996) claims, and if households with higher assets find IRAs to be better substitutes for saving than other households, as Gale and Scholz (1994) estimate, then it seems likely that the concern with differences between continuing and new contributors will lead to an overstatement of the effects of IRAs on saving in the Joines and Manegold model. See Gale (1996) for further discussion.



than non-contributors, 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 (1995a) use panel data from the SIPP. They examine correlations between changes in IRA contributions and a measure of non-IRA saving during 1984 and 1985 and 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 there is no explicit assumption about saving behavior in the absence of the IRA program. This omission is particularly problematic in this context because the study 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.<sup>38</sup> In addition, 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. Substitution from other assets or current saving to IRAs could occur in numerous ways not captured in this study.

Venti and Wise (1992) examine panel data from the 1983-6 Surveys of Consumer Finances. They examine holdings of financial assets and claim, first, that IRA contributors did not save very much before IRAs became available, and, second, that IRA contributors saved

---

<sup>38</sup>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. However, there was no change in IRA eligibility from 1984 to 1985.

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. Homeownership requires having saved for a downpayment in the past, and paying off the loan over time, 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.<sup>39</sup>

Trends in this broader measure of wealth lead to dramatically different conclusions. In 1983, the median among 1986 IRA contributors was \$51,220.<sup>40</sup> Thus, it is simply incorrect to claim that these families typically did little or no saving prior to 1983. 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 would plausibly have grown by 6 percent per year in nominal terms from 1983 and 1986.<sup>41</sup> This alone would have

---

<sup>39</sup>This is the same wealth definition as used by Engen and Gale (1995).

<sup>40</sup>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 pension wealth) was about \$75,000.

<sup>41</sup>Most classes of financial assets received returns significantly higher than 6 percent. See Gale and Scholz (1994, Appendix A). Constant-quality home prices grew at about 3.6 percent per year over this period, but the actual increase in housing equity for households in the panel

resulted in median 1986 wealth of \$61,000. If the median family had made additional contributions of \$1,200 per year (plus the accumulated interest), it would have attained the actual median 1986 wealth level. If a conservative 20 percent tax rate adjustment is made 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), however, estimate that the marginal tax rate for IRA contributors during this period was about 35 percent. Using this figure, new saving of just \$550 per year would have been sufficient to generate the 1986 wealth level.<sup>42</sup>

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) annually even in the absence of IRAs. This is reinforced by the facts 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 and were \$40,000 in 1986.

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

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

---

would be larger because mortgage debt would be paid down.

<sup>42</sup>A number of features could be added to these estimates. 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.

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 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.<sup>43</sup>

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.<sup>44</sup> 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. This appears to have been a relatively rare response, at least in the early 1980s, (see Kruse 1991), but

---

<sup>43</sup>In the example given, controlling for pension coverage would solve the bias, but the successive cross-section and cohort studies described above 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).

<sup>44</sup>These tabulations were carried out by Joel Dickson at the Federal Reserve Board. A number of studies reach similar conclusions (see Andrews (1992), Papke, Petersen, and Poterba (1993), and Buck Consultants (1989)).

Papke (1995, 1996) finds stronger effects in more recent years. Papke (1996) uses panel data from 1985 to 1992 and estimates that when a plan sponsor started out in 1985 with no 401(k) or other defined contribution plan and then added a 401(k) over time, the number of defined benefit plans offered by the sponsor fell by at least 0.3 more than it otherwise would have over this period. Plan-level estimates imply that if a 401(k) plan is added by a sponsor, the probability that a defined benefit plan is terminated increases by about 18 percentage points, or approximately doubles. These results imply that a sizable minority of 401(k) plans are replacing defined benefit plans.

An additional channel of substitution can occur on the margin—firms could cut back on existing plans in other ways such as restricting or reducing benefit increases. In this case, and in the case when 401(k)s represent converted thrift plans, households with 401(k)s would still have defined benefit pension coverage—as is observed in the data—but 401(k)s would be substituting for other pensions at the margin. Finally, it is also possible that some 401(k) plans are established at firms that would have created another plan had 401(k)s not existed. Little is known about these latter possibilities.

Taken together, these items suggest that a substantial portion of 401(k)s may be replacing other pensions. Since all analyses of the effects of 401(k)s on saving ignore this potential substitution, they are likely to overstate the impact of 401(k)s on saving.

To the extent that 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 the 401(k) can be less than that of the traditional plan it displaces (see Halperin (1987)). If so, 401(k)s could reduce private saving even if they appear to raise saving in surveys that omit pension wealth.

(X) Simulation models

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.<sup>45</sup> 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.<sup>46</sup>

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. Between 22 and 30 percent of IRA contributions represent new saving in the long run. As modelled here, 401(k) plans have stronger effects due to a higher

---

<sup>45</sup>IRAs have tax-deductible contributions, and a limit of either \$2,000 or \$4,000. 401(k) plans have a limit of \$9,000. Employers are assumed to match employee 401(k) contributions at the rate of 50%, up to the first 5% of salary. The early withdrawal penalty is set at 10 percent for both plans.

<sup>46</sup>The simulations 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. The model generates aggregate interest elasticities of saving between 0.15 and 0.35, consistent with empirical estimates, which generally fall between zero and 0.4. 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 real world 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.

contribution limit and the employer match: long-run national saving rises by between 0.5 and 1.0 percentage points.

However, the short-term and transitional effects are very different. As shown in Figure 2, private and national saving fall 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 years of the program by the shifting of already existing taxable assets. The private asset-to-income ratio does not return to its original level until 36 years after IRAs are introduced.

Government saving falls initially due to the shifting of assets. Eventually, government revenues rise as IRA 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). Thus, in the short run, national saving falls by even more than private saving, because of the deferral of taxes on contributions and investment earnings. The national (public plus private) asset-to-income ratio falls from the start and does not regain its initial value until year 49 of the transition.<sup>47</sup>

The simulation results are consistent with our findings in the empirical sections

---

<sup>47</sup>The same intuition and behavioral patterns occur for 401(k) plans. However, since contribution limits are larger for 401(k)s, shifting occurs more rapidly. The private (national) asset-to-income ratio regains its original level by year 22 (35).

described above, because the empirical work covers only the first 10-15 years of the programs, during which very little if any of the contributions to saving incentives should be new saving according to the simulation.

This simulation model, like all such models, is imperfect in a number of ways, and the precise magnitude of the estimates should be regarded with care. Nonetheless, we find the qualitative behavioral patterns generated by the model to be both plausible and robust to a wide variety of sensitivity analyses.

The most important caveat is that, for a number of reasons, we believe the model overstates the long-term effects of saving incentives on private saving. First, general equilibrium effects are omitted. Allowing for these effects would dampen the long-term increase in saving: as the capital stock rose, interest rates would begin to fall, which would diminish the long-run increase in saving. Second, the model overstates the effective 401(k) limit and substantially overstates 401(k) participation among low-income households, for whom saving incentive contributions are largely new saving. Third, the model assumes the saving incentive is financed with an increased public deficit. If instead income tax rates were raised, this would further reduce private saving. Fourth, we have assumed that all wealth other than saving incentives is liquid; this raises the demand for saving incentives.

For a similar set of reasons, the model may also overstate the length of the "transition period" during which the cumulative effect of saving incentives is less than zero (See Hubbard and Skinner 1995). Overstating the length of the transition period, though, is vastly different from claiming there is no transition period.<sup>48</sup> The model suggests that introducing saving

---

<sup>48</sup>The latter claim is much stronger and is implicitly made, for example, by Venti and Wise (1990), who use data from 1982 and find that IRAs raised saving immediately after the expansion of IRA eligibility.



incentives would cause a lengthy transition period, during which contributions on average would not represent increases in saving.

Moreover, making adjustments to reduce the length of the transition period also reduces the estimated long-run impact on saving. For example, introducing illiquid, tax-preferred assets, such as housing, would reduce the extent of initial shifting into saving incentives, and thereby reduce the transition period, but would also reduce long-run demand for saving incentives and their ultimate impact on saving.<sup>49</sup>

Finally, Hubbard and Skinner (1995) characterize the simulations as showing that—in the long run—the ratio of the increase in private capital to the increase in public debt is 16 to 1 for 401(k)s. However, it should also be noted that the ratio is negative for the first 30 to 40 years of the transition. So unless policymakers are certain that the policy will stay in place for a very long time, the model suggests that introducing saving incentives will not raise private and national saving.

#### (XI) Offsetting Revenue Effects?

Although the discussion above focuses primarily on private saving, the impact of tax incentives on national saving depends also on how such incentives affect public saving. Contributions to saving incentives cause an immediate increase in the budget deficit because the contribution is tax-deductible. 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) is eventually withdrawn, raising future tax revenues. And to the extent that saving incentives raise saving,

---

<sup>49</sup>Adding non-liquid wealth to the model has effects similar to raising the penalty for early withdrawal. Engen and Gale (1993, table 7) show that raising the penalty reduces the net effect on long-term national 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.

Feldstein (1995) considers all of the effects above and finds that saving incentives have a substantial, positive impact on public saving. While Feldstein is surely correct in including the impact on corporate taxes, 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.<sup>50</sup> He also may overstate the increase in investment due to an increase in saving. He assumes that 100 percent of increases in domestic saving turns into increased domestic investment. But 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, non-incorporated 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.

---

<sup>50</sup>Even in the simulation models noted above that generated a positive long-run impact of IRAs on saving, only 22 percent of IRA contributions are net additions to saving in the long run when the limit is \$2,000 as Feldstein assumed.

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 effects as well (Ruggeri and Fougere, 1995).<sup>51</sup> In summary, the impact of saving incentives on long-run public saving is a currently unresolved issue and depends critically on their impact on private saving as well as on details on the private economy.

#### (XII) Evidence from Canada

Registered Retirement Saving Programs (RRSPs) are Canadian saving incentives: contributions are tax-deductible, earnings accrue tax-free, and withdrawals are taxed, but there is no early withdrawal penalty. Carroll and Summers (1987) find that Canadian saving rose relative to American saving in the early to mid-1970s, when increased RRSP contribution limits raised RRSP contributions. But Altig (1990) shows that the impact of RRSP contributions vanishes when a borrowing variable—to account for the fact that mortgage interest is not deductible in Canada—is included in the regression.

Venti and Wise (1995b) conclude that RRSPs raise saving, based on a variety of tests using survey data. These tests generally mirror the structure of their tests using American data and hence raise many of the same issues described above.

Sabelhaus (1996) shows that trends in the composition of aggregate saving do not support the view that RRSPs have raised saving. The increase in Canadian relative to American saving in the 1970s occurred because non-retirement saving rose dramatically in

---

<sup>51</sup>Ruggeri and Fougere (1995) note that the assumed real rate of return on private capital, 10 percent, appears to be higher than recent norms and therefore overstates the likely revenue from added investment; and the assumed real return on government bonds, 2 percent, appears to be below recent norms. In addition, they note that when IRAs are introduced, the deficit rises, but in Feldstein's model, there is no mechanism to finance the increase in deficits. Hubbard and Skinner (1995) present alternative estimates of the long-run impact of IRAs.

Canada relative to the United States. Canadian retirement saving (which includes RRSPs) actually fell slightly relative to American retirement saving. The same points apply to longer-term changes comparing the 1960s and 1990s: saving increased in Canada relative to the U.S, but retirement saving did not. The increase is due to large relative changes in non-retirement saving. These trends do not suggest an important role for RRSPs in raising saving.

### (XIII) Saving Incentives and Alternative Models of Saving

Standard economic models suggest that existing saving incentives should have had little effect on saving, except perhaps in the long run. Others, however, argue that saving incentives raise saving and appeal to alternative theories based on psychological approaches to saving (see Bernheim 1996, Thaler 1990). There appears to be a presumption that psychological or behavioral models imply that saving incentives raise saving. In general, however, that presumption is incorrect. There are many ways that saving incentives can fail to raise saving even if psychological motives are present and even if they are dominant.

For example, Bernheim (1996) and Thaler (1994) note that a key element in psychological models of saving is the issue of self-control. If so, psychological models may be most applicable to people who do not save very much. Households with saving incentives tend to save substantial amounts overall. Thus, there could be many households using saving incentives and many households that act mainly according to psychological motives, but the overlap between the two sets may not be very large. If so, the existence of psychological motives would have little impact on existing incentives.<sup>52</sup>

Laibson (1996) demonstrates that illiquid saving accounts can raise saving in a model

---

<sup>52</sup>However, it is not obvious that psychological models are most likely to apply to those that save very little. High-saving households could be those that have solved their self-discipline problems.

where people suffer from self-control problems. But the illiquid accounts in his model require net saving whereas current incentives reward gross contributions. Moreover, as he notes, the positive effect on saving can easily be removed in his model if consumers can take actions in later years to remove the funds prematurely. Since these actions include increasing borrowing, which can be done currently, Laibson's model implies that current saving incentives do not necessarily raise saving even in a model that emphasizes self-control problems.

Second, to the extent that psychological motives do matter, they may have their primary impact on portfolio composition rather than the level of saving.<sup>53</sup> For example, many contributions for a given tax year are made in the following calendar year. It seems to us that people would be much more likely to shift money into an IRA at the last minute if the money represents idle cash than if the money were destined to pay for some large consumption item, say a family vacation. If so, then portfolio composition is affected rather than saving. Advertising provides another example. Most IRA contributors have saved substantial amounts in other forms and so are aware of the opportunities for saving. The information contained in advertising about IRAs may well be the presence of a tax deduction for doing what the household was already doing without a tax deduction. If so, then the primary impact will be on portfolio composition.

Third, target saving—an alternative to the life-cycle model—occurs when households save to reach specific wealth targets—for example, a downpayment on a house, college tuition, or a particular income replacement ratio. Target saving seems to be a plausible way to allow households to apply simple but effective rules to potentially complex long-term saving decisions. In fact, a version of target saving is cited in Doyle and Johnson (1991) as the most

---

<sup>53</sup>Browning and Lusardi (1995) make a similar point.

common retirement planning technique. A distinctive feature of target saving is that higher rates of return reduce the amount of saving needed to reach the target. Thus, to the extent that people are saving to reach specific retirement targets, saving incentive programs will reduce overall private saving, since the saving incentives offer a higher rate of return on a limited amount of contributions.

Finally, what Bernheim (1996) describes as a "scarcity of well-developed organizing principles" means that many possible channels for psychological models need to be explored. For example, Bernheim notes that a tax subsidy for targeted saving may raise saving because it conveys information to people that authorities think saving is important. But people may instead receive the message that authorities think saving is such an unpleasant task that it deserves a tax break. In that case, those without saving incentive opportunities may reduce their saving, and those with such opportunities may reduce their taxable saving by as much or more than their saving incentive contribution.

In summary, we view alternative models of saving as an important new direction for future research, but, to this point at least, we find nothing in the models or interpretations that represents evidence that saving incentives raise saving, and at least one prominent alternative model (target saving) suggests that saving incentives should reduce saving.

Finally, it is often difficult to identify patterns of saving behavior that can distinguish between realistic versions of life-cycle models and psychological models. Researchers sometimes "test" what can only be regarded as straw man versions of the life-cycle model that include many ancillary assumptions (perfect markets, no transaction costs, no uncertainty, perfect information, etc.). Rejecting such models does not prove that more plausible life-cycle

specifications are inappropriate.<sup>54</sup>

But distinguishing between the life-cycle and psychological approaches need not be an "either or" choice, and is unlikely in itself to prove decisive in determining whether saving incentives raise saving. Saving behavior is likely to contain elements of both forward-looking life-cycle behavior as well as psychological motives. The relative importance of these approaches may differ across households, over time, and possibly even over time for a particular household. Ultimately, the best research will develop and test falsifiable hypotheses using insights from both approaches.<sup>55</sup>

#### (XIV) Conclusion

We find that saving incentive plans have had little if any effects on household saving behavior. Other studies have come to the opposite conclusion, and a major goal of the paper has been to reconcile the divergence between our results and the other findings. In particular, we show that each finding that saving incentives raise saving can be traced to various biases that overstate the effects of saving incentives, and that removing the biases also removes the positive effect of the incentives on saving.

For example, cross-sectional analyses of IRAs that assume that contributors and non-contributors have similar tastes for saving are biased in favor of finding that IRAs raise saving; comparing limit contributors to other contributors removes the bias and generates estimates that IRAs had virtually no impact on saving. Cross-sectional studies of 401(k) plans raise

---

<sup>54</sup>For example, the fungibility of assets is sometimes taken to be a primary implication of the life-cycle model by advocates of psychological models (see Thaler 1990). But in the presence of real world considerations such as transactions costs, illiquidity and uncertainty, fungibility is no longer an implication of the life-cycle approach.

<sup>55</sup>For further discussion of psychological models and saving incentives, see Engen, Gale and Scholz (1994) and Gale (1996).

similar concerns: 401(k) eligibility is positively correlated with households' tastes for saving; controlling for initial wealth as an (imperfect) measure of tastes for saving generates estimates that less than 10 percent of 401(k) contributions represent new saving. In cohort studies, there is no role for saving incentives to raise financial assets once allowances are made for the effects of abnormally high equity returns and interest rates during the sample period, the increase in debt, the shift away from non-financial assets caused by reduced inflation and tax rates, the decline of pensions and social security, the difference between pre- and post-tax asset balances, and underlying data problems. Successive cross-section studies suffer from many of the same problems as cohort analysis; measuring the effects of 401(k)s on a broader measure of wealth that includes housing equity indicates virtually no effect of 401(k) plans on saving. In panel studies, examining broader asset measures suggests little impact of IRAs on saving. All of these approaches ignore substitution between 401(k)s and pensions at the firm level, which causes a further overstatement of the impact of saving incentives on saving. Moreover, all of these findings are consistent with the observed trends in the level and composition of aggregate saving and with results from simulation models, which show essentially no impact of saving incentives on saving, at least in the time period considered.

We note, however, that our results are consistent with the possibility that existing saving incentives have raised saving for some individuals, that saving incentives may eventually raise saving, and that saving incentives in different forms might raise saving.

Our findings fit well into the hierarchy of behavioral responses to taxation, as developed by Slemrod (1992) 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 include 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 are consistent with this hierarchy in that we find a strong effect of tax-based saving incentives on the allocation of saving and assets, but little or no effect on the level of saving or wealth accumulation. Saving is painful; it requires a reduction in current living standards in exchange for an increase in future living standards. If people can obtain the same increase in future living standards without the pain of reducing current consumption, then we expect people to choose that route over reducing current 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 to our results is that all of the data sets and experiments used to explore these issues—by us and by others—are faulty in some important way. Thus, it is unclear that there is a completely satisfactory way, given existing data, to address these issues.

Also, 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 1993a). Another possibility is that households have very low intertemporal elasticities of substitution (Hall 1988). If so, it may prove difficult to stimulate saving via any voluntary mechanism. Alternatively, 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 beyond the scope of this paper, but are interesting items for future research.

## APPENDIX A

This Appendix provides more detail on the Gale and Scholz (1994) model and briefly addresses criticisms put forth by Bernheim (1996) and Poterba, Venti and Wise (1996).

The formal model in Gale and Scholz (1994) is given by the following equations:

$$(1) \quad S_I^* = X\beta + u$$

$$(2) \quad S_I = \begin{cases} 0 & \text{if } S_I^* \leq 0 \\ X\beta + u & \text{if } 0 < S_I^* < L \\ L & \text{if } S_I^* \geq L \end{cases}$$

$$(3) \quad S_0 = \begin{cases} X\gamma_1 + \varepsilon_1 & \text{if } S_I^* \leq 0 \\ X\gamma_2 + \varepsilon_2 & \text{if } 0 < S_I^* < L \\ X\gamma_2 + \eta(S_I^* - L) + \varepsilon_2 & \text{if } S_I^* \geq L, \end{cases}$$

Equation (1) shows desired IRA saving,  $S_I^*$ , as a function of observable characteristics  $X$ . Equation (2) shows that the observed IRA contribution,  $S_I$ , must be no smaller than zero and no larger than the annual contribution limit,  $L$ .

Equation (3) describes households' non-IRA saving,  $S_0$ . The effect of a given  $X$  variable (such as age or earnings) is allowed to differ across households that contribute to IRAs ( $\gamma_2$ ) and households that do not ( $\gamma_1$ ). Moreover, the error terms describing unobserved determinants in behavior are also allowed to differ across contributors ( $\varepsilon_2$ ) and noncontributors ( $\varepsilon_1$ ).

The effect on non-IRA saving of a one dollar change in the contribution limit is given by  $\eta$ . The term  $S_I^* - L$  represents the excess of desired IRA saving over the maximum allowed IRA

saving, and  $\eta$  measures how this difference affects non-IRA saving. If  $\eta$  is close to 1, then almost all excess desired IRA saving appears as non-IRA saving. Thus, a change in the IRA contribution limit would have little effect on the overall level saving. If  $\eta$  is close to zero, then almost none of the excess desired IRA saving shows up as non-IRA saving, and increasing the IRA limit would raise total saving.

In the empirical model,  $\eta$  is parameterized as a linear function of the household characteristics. This specification captures the crucial point that substitutability between IRAs and other saving should depend on household characteristics. The elements of  $\eta$  are identified by comparing the non-IRA saving of limit contributors to the non-IRA saving of households that contribute positive amounts but less than the limit, rather than by comparing limit contributors to non-contributors.

Bernheim (1996) notes that if limit contributors have stronger unobservable preferences for saving than interior contributors, the GS model may be biased toward finding that IRAs are not new saving. But he provides no evidence in support of his assumption about unobserved preferences or the quantitative importance of the bias. We believe that even if the assumption is valid, the quantitative importance of the bias is likely to be small. One reason is that GS already control for a large number of observable characteristics to account for heterogeneity in saving behavior across limit and interior contributors. Moreover, we have estimated regressions using the 1983 SCF sample of all limit and interior contributors and controls for all of the covariates in the GS model except those relating to wealth. These regressions show that the coefficient on a dummy variable for limit contributors is negative and insignificant in equations explaining 1983 wealth. If unobserved tastes for saving were higher among limit contributors than interior contributors, this coefficient should be positive and significant. These regressions

are inconsistent with Bernheim's concern that limit contributors have higher tastes for saving than interior contributors.

In contrast, in the Venti-Wise model (1986, 1987, 1990, 1991), limit contributors are compared to (i.e., are required to have the same coefficients as) non-contributors. Regressions using the 1983 SCF, similar to the ones described above but where the sample includes IRA contributors and non-contributors, indicate that the coefficient on the IRA contributor dummy variable is about \$26,000 and is statistically significant. Thus, the regressions show that there are large differences in tastes for saving between IRA contributors and non-contributors, but little if any difference between interior contributors and limit contributors. Therefore, we believe the GS comparison is appropriate and we find no reason to believe that the assumptions underlying Bernheim's bias argument are applicable.

Poterba, Venti and Wise (PVW, 1996) make several claims regarding the Gale and Scholz (1994) model.

The Data and Sample: PVW first criticize the sample choices. In their descriptive tables, GS excluded households under age 25 in 1986 (because the developers of the survey thought the data for these households were unreliable), households that changed marital status (because it is not possible to calculate the applicable IRA limit for such households), and the self-employed (because of difficulties in distinguishing Keogh and IRA contributions).

For their formal estimates, GS also exclude households aged 65 and older in 1983 because of the complications in modelling saving by the elderly. In contrast, GS include these households in the descriptive work because the households are in fact eligible to contribute to IRAs and, for reasons described in the text, the proportion of contributions that come from those older than 59 is a relevant piece of information. The idea that older households are

unlikely to contribute, as Poterba, Venti and Wise assert, is neither supported by evidence nor relevant to whether such households should be in the sample. These households are eligible to contribute, until age 70.5. If they find IRAs to be good substitutes for other saving (because there is no withdrawal penalty for those over 59.5), then excluding these households from the formal estimates will bias the results toward overstating the effects of IRAs on saving.

Finally, for the formal estimates GS also excluded about 180 households (out of 2,822 in the overall sample) that had absolute values of 1983-6 non-IRA saving that exceeded \$100,000. This exclusion is due to difficulties in modelling the saving behavior of the very wealthy. As PVW note, this exclusion has a large effect on sample means of saving. More importantly, as noted by GS and PVW, the estimates of substitution exhibit a good deal of sensitivity to the cutoff points that are chosen. What PVW fail to note, however, is that the exclusion should bias the GS results toward finding less substitution than actually exists, since the exclusion eliminates a disproportionate number of high-income, very high-wealth households that should find IRAs to be very good substitutes for other saving. Notably, both GS (1994) and PVW (1996) in their critique show this to be true: the estimates that impose higher cutoff thresholds generally lead to smaller effects of IRAs on saving than estimates that impose lower cutoffs. The broadest thresholds in the PVW replication, which of course are most closely related to the complete sample of eligible households, show substantial substitution, just as the central GS runs did.

None of these sample restrictions seem implausible or inappropriate.<sup>56</sup>

A Simple Reality Check: PVW (1996, Table 15) show that a measure of the financial

---

<sup>56</sup>Contrary to PVW's assertion (1996, page 65, item (2)), GS did not delete households that they thought were unlikely to contribute to IRAs.

assets of a panel of IRA contributors rose from 1983 to 1986, while financial assets of a panel of noncontributors did not. PVW also claim, based on financial asset data, that IRA contributors were not saving very much before IRAs. PVW perform a "simple reality check" and ask how GS could possibly find, in light of these results, that IRA contributions could be funded by reductions in other wealth or saving.

The "simple reality check" is seriously misleading. First, the GS results are based on comparisons of limit contributors and interior contributors, while the PVW table compares those with and without IRAs. Second, holding observable factors constant, differences in tastes for saving imply that a panel of contributors would be expected to save more than a panel of non-contributors even in the absence of IRAs. That is, comparisons of households with and without IRAs are biased in favor of finding that IRAs raise saving. Indeed, PVW (1996, page 3) note that: "Some people save and others don't, and the savers tend to save more in all forms. For example, families with IRAs also have more conventional savings than families without IRAs." We agree with the quote, and note that it also undermines the logic of the "simple reality check."

The reality check is faulty in other ways as well. The PVW table does not control for any observable household characteristics, while the GS analysis does. Median three-year income among IRA contributors during this period was \$58,000 higher than among non-contributors (see Gale and Scholz 1994). It would not be surprising if an income difference of that size influenced saving. Moreover, PVW examine a narrow measure of financial assets. A broader measure, as in table 2 of our text, shows substantially smaller growth. Because financial assets are only a small portion of wealth, it is simply incorrect to claim that IRA holders did little saving before 1983. Median net worth among IRA holders was over \$75,000

in 1983. This indicates substantial prior saving activity.<sup>57</sup> Finally, many factors besides saving incentives caused financial assets to rise (see section VI on cohort data). Aggregate financial assets rose by \$2.6 trillion between 1983 and 1986, while saving incentive balances rose by less than \$300 billion.

In short, the simple reality check is uninformative. There is nothing inconsistent with it and the results in GS (1994). These issues are explored further in section VIII of this paper, which examines the panel data from the SCF in more detail and interprets the findings as showing little new saving.

Saving Effects of Increasing the Limit: PVW claim that GS severely understate the proportion of contributions or contributors at the limit, since only households that contribute to the limit in all three years are identified as "limit" contributors in GS. This assessment is correct if one is looking at a limit change over a one year period. But over the three-year period covered by the SCF data, it is not accurate. GS (1994, table 1) compare their constructed IRA variable to data from the IRS tax panel. GS show that the constructed variable matches closely the proportion of the population contributing to an IRA at least once and more importantly the percentage of IRA contributors that contributed to the limit in all three years. In the IRS panel, households that contributed to the limit in each of the three years accounted for 68 percent of all of the limit contributions during that period. Therefore, in examining limit increases over a three-year period, the GS approach captures two-thirds of the limit contributions.

---

<sup>57</sup>For example, suppose a household saved \$10,000 a year for five years and then put all the funds into a downpayment on house. After the purchase, an analysis using financial assets would mistakenly conclude that the household had not saved anything in the past, since its financial assets were zero. It is worth noting that most IRA contributors already own their own home.



PVW also claim that using descriptive data from 1986 (as GS do) on financial assets of IRA holders is extremely misleading because it does not show how much wealth was available at the beginning of the program to substitute into IRAs, nor does it show the typical saving behavior of contributors before IRAs. However, since IRA contributions can be funded by shifting pre-existing assets or funds that would have been saved anyway, we do not find 1986 wealth data misleading. As noted above, financial assets rose for many reasons between 1983 and 1986. Looking at the 1986 figures show how much was actually saved in other forms and therefore how much would have been available to shift into IRAs had the limits been larger, which is the question GS address. Moreover, even if one does focus on 1983 data, the typical IRA contributor had already saved substantial amounts by 1983: \$16,500 in NIFA and \$77,299 in net worth.

The Gale Scholz Model: PVW raise a number of issues concerning the GS empirical model. PVW argue that the specification of  $\eta$  should include a marital status term and that the other equations should as well. This variable had been included in earlier versions of the paper (for example, Gale and Scholz (1990)), where its estimated effects were small and statistically insignificant, and where it had little effect on the results. The correlation between family size, which is included in the specification, and marital status may explain why marital status does not affect the results.

PVW also note there is no constant in the specification of  $\eta$ . This is only a problem if the true constant is not zero. But consider what a constant in the specification of  $\eta$  would represent: the substitutability between IRAs and other saving for someone with no assets (or any of the other co-variates such as income or education). Given that substitutability should rise with assets, the appropriate constant is likely to be zero, so that including a constant in  $\eta$

would have little effect on the results. What is particularly puzzling about this criticism is that, in their replication of the GS results, PVW actually show that including a constant in  $\eta$  does not substantially alter the estimated substitutability of IRAs and other saving.<sup>58</sup>

PVW also note—as do Gale and Scholz (1994)—that the mean of  $\eta$  is sensitive to changes in the values of the right hand side variables. We note that this occurs partly because GS did not constrain  $\eta$  to fall within a given range. That the unconstrained value of  $\eta$  estimated in the various models presented in GS fall within the reasonable range of estimates is a strength of the model. In contrast, the VW models (1986, 1987, 1990, 1991) constrain the substitution parameter to be between zero and one, which eliminates even the possibility of households being target savers or having income effects due to IRAs that are stronger than the substitution effects. VW do not report the results of unconstrained estimates, even though actual substitutability could be outside the restricted range.

PVW also undertake a number of curious steps in estimating and examining the GS results: (a) they omit non-IRA financial assets (NIFA) as an explanatory variable; (b) they omit many of the other right hand side variables; (c) they set  $\eta$  equal to a constant across households; and (d) they focus on estimates of non-IRA saving by limit contributors that ignore the effects of the IRA limit. As GS explain in the original paper and as we discuss below, little can be learned from such exercises.

First, consider the specification that omits NIFA. NIFA is a crucial variable, derived from the theoretical model in GS. Families with higher NIFA holdings are more likely to have IRAs and to find IRAs better substitutes for other saving. NIFA also helps control for

---

<sup>58</sup>Their table 20, specification H shows the effects of adding a constant term in  $\eta$  to the GS specification. The estimated mean of  $\eta$  is 1.25, which is slightly higher than the corresponding estimate in specification F without a constant, 1.17.

individual tastes for saving, as it does in the Venti and Wise models.

PVW argue that NIFA is a poor control for individual tastes for saving because the coefficient is negative, implying that those with more assets consume more. As GS note, though, in any cross-sectional study (even in the VW model), existing wealth can influence saving in either of two directions. For a particular person, higher initial wealth should mean higher consumption and therefore less saving, controlling for income. Across people, higher initial wealth will be correlated with tastes for saving and so would indicate more saving. The actual coefficient on NIFA reflects both of these considerations. The fact that the coefficient is negative does not mean the variable is not controlling for individual tastes for saving, it just means that the first effect is larger than the second effect.

PVW also argue that because the GS saving measure is calculated using NIFA in 1983, measurement error is induced and that this measurement error drives the GS results. Measurement error could be present, but we do not think it is of any major quantitative importance for two reasons. First, if measurement error were driving the results, one would also expect the coefficient on NIFA to be negative in the saving equation for non-contributors, but it is not (see GS 1994 Table 4). Second, if measurement error were driving the estimates, then moving to broader measures of saving that do not depend exclusively on NIFA should generate less estimated substitution. In fact, it generates more (see GS 1990). Therefore, we do not believe measurement error is actually a serious problem here. GS (1994) present further evidence on measurement error.

The second curious adjustment is the omission of almost all of the right hand side variables other than income. The purpose of this adjustment is unclear, especially given the tremendous residual variation in asset holdings that GS note in their paper and PVW document

when controlling only for income and age. Our view is that if the problem is trying to control for unobservable differences in tastes for saving, it makes sense first to control for as many observable influences on saving as is practical. That is the strategy GS followed.

The third curious adjustment is setting  $\eta$  equal to a constant. Both theory and common sense (and the empirical patterns) suggest that setting  $\eta$  to a constant is a meaningless exercise. The substitution between IRAs and other saving should vary with individual characteristics. GS included all of the  $X$  variables in  $\eta$  to ensure that they were not identifying the effects of IRAs with arbitrary exclusion restrictions. PVW refer to keeping all of the variables in  $\eta$  as "adding complexity," but it is no different than what VW do in their own work on this topic.

The fourth issue is that PVW remove NIFA from the specification and then examine the effects of doing so on the fitted value of non-IRA saving. These results (PVW 1996, Table 22, second column) show that, when the spillover effect of excess desired IRA saving on non-IRA saving is omitted, adding NIFA to the specification causes the GS model to understate non-IRA saving by limit contributors. The key problem with this exercise is that the impact of the IRA limit is ignored. A crucial feature of the GS model is that the expected non-IRA saving of limit contributors depends on how much IRA saving they are allowed to do. Therefore, it is not interesting to examine the expected non-IRA saving of limit contributors without allowing for impact of the IRA limit. Including the spillover from excess desired IRA saving (PVW 1996, table 22, third column) eliminates the understatement of non-IRA saving from including NIFA. Moreover, the fitted values including NIFA and including the spillover effect from other saving are reasonable.

In summary, it is not surprising that removing key features of the GS model changes the results. The point is that it is not informative. All of these adjustments have direct analogues

in the Venti and Wise (1986, 1987, 1990, 1991) models. If these adjustments were critical or even relevant, it is surprising Venti and Wise do not report the outcome of making these changes in their own model.

## APPENDIX B

This Appendix provides an example of the bias created when cross-section regressions estimate the effects of 401(k) eligibility on wealth, but control for cash earnings rather than total compensation. The particular example is constructed so that the true effect of 401(k) eligibility on wealth is assumed to be zero, but the estimated effect—controlling for cash earnings rather than total compensation—is positive. But the general point holds for any degree of underlying offset between 401(k)s and other saving: the estimated impact will overstate the true impact.

In this example, interest, discount and inflation rates are set to zero, for simplicity. As shown in Appendix Table B-1, worker A has total compensation in period 1 of 110 units, of which 100 units are paid as a cash wage and 10 are paid as employer 401(k) contributions.<sup>59</sup> The worker smooths consumption at 55 units per period, and so holds 55 units of financial assets at the end of period 1. (The manner in which this is divided among 401(k)s and other saving does not matter for the example, but to be concrete, suppose the employee places 10 in the 401(k) along with the employer's 10, so the overall allocation is 20 in the 401(k) and 35 in other saving). Worker B has the same cash wages of 100 in period 1, but no 401(k). The worker smooths consumption at 50 units per period, and holds 50 units of financial assets (all in non-401(k) saving) at the end of period 1.

Now consider the results of estimating the impact of 401(k) eligibility on these two workers at the end of period 1, controlling for cash wages. Let the estimating equation be  $Z =$

---

<sup>59</sup>For purposes of this example, it does not matter whether the contributions are matching or unconditional. Most 401(k) participants receive employer contributions.

$\alpha W + \beta E$ , where  $Z$  is financial assets including 401(k)s,  $W$  is cash wages and  $E$  indicates 401(k) eligibility. The equations would be  $55 = \alpha 100 + \beta$  for worker A and  $50 = \alpha 100$  for worker B. The equations imply estimates of  $\alpha = 0.5$  and  $\beta = 5$ . In short, since worker A has 5 more units of wealth, the estimated coefficient on 401(k) eligibility is 5.

However, the true offset—the reduction in other wealth from deferring a part of compensation, holding total compensation constant—is 100 percent. The true offset can be estimated by regressing wealth on total compensation and 401(k) eligibility. The equations in this case are  $55 = \alpha 110 + \beta$  and  $50 = \alpha 100$ , implying estimates of  $\alpha = 0.5$  and  $\beta = 0$ .

The same biases occur if the goal is to measure the proportion of 401(k) contributions that represent new saving. In the example above, worker A would have 20 in 401(k) balances at the end of year 1. Let the estimating equation be  $Z = \alpha W + \gamma X$ , where  $Z$  and  $W$  are defined as above and  $X$  indicates 401(k) balances. The equations would be  $55 = \alpha 100 + \gamma 20$  for worker A and  $50 = \alpha 100$  for worker B, implying estimates of  $\alpha = 0.5$  and  $\gamma = 0.25$ . That is, the equation would suggest that 25 percent of 401(k) contributions represent new saving even though the true value in this example is zero. As before, controlling for total compensation would provide the correct estimated  $\gamma$  of zero.

For further discussion and generalization of these results, see Bernheim and Scholz (1993b) and Gale (1995).

Appendix Table B-1

Controlling for Total Compensation versus Cash Wages

Worker	Income and Wealth	Period	
		1	2
A	Cash earnings	100	0
	Employer Contribution	10	0
	Consumption	55	55
	Wealth*	55	0
B	Cash earnings	100	0
	Employer Contribution	0	0
	Consumption	50	50
	Wealth*	50	0
—	Estimated impact of 401(k) eligibility using cash earnings	5	—

\*Wealth measures are reported at the end of the period.



## APPENDIX C

This Appendix provides further discussion of the results in Engen and Gale (1995) and responds to criticisms of that paper by Bernheim (1996) and Poterba, Venti, and Wise (1996).

Poterba, Venti, and Wise (1996) raise a variety of issues. First, they claim that increasing mortgage debt may, in the long run, increase saving. This is an odd claim on a number of grounds. Since reducing mortgage debt raises wealth, it is hard to see how increasing mortgage debt also should be interpreted as an increase in future wealth. If households repay the added mortgage by reducing the amount of saving they do in other forms, then net saving will not have increased. Moreover, households can always take out second mortgages or home equity loans in the future; there is nothing permanent about paying off a mortgage. Finally, Manchester and Poterba (1989) show that households with home equity loans have less net worth, controlling for other factors. Thus, the most that can be said is that an increase in mortgage debt is a reduction in current wealth.

PVW then present cohort analyses of saving incentive balances and home equity in 1984-7 and 1987-91. Their basic finding is that "the timing of changes in mortgage debt and net home equity is inconsistent with a causal relationship between personal retirement plan contributions and mortgage debt." They find that from 1984-7, a period of rapid increases in contributions to saving incentives, home equity rose and there was no countervailing increase in mortgage debt. From 1987-91, the increase in personal retirement assets slowed, but mortgage debt rose rapidly.

These results, however, provide no evidence on the validity of the Engen and Gale results. As Engen and Gale note and Poterba, Venti, and Wise (page 48) recognize, market

trends in housing that are unlikely to be induced by saving incentives can have large effects on equity values and mortgage markets; and the tax reform act of 1986 led to shifts in the composition of debt toward mortgages. There is no suggestion in Engen and Gale that saving incentives are driving the changes in the mortgage or housing markets. Obviously, the market for mortgages is very large compared to saving incentive contributions and is affected by many factors. Suppose those factors reduced mortgage debt before 1986 and raised mortgage debt after 1986, as PVW assert. The point made by Engen and Gale is that, after allowing for those factors, and controlling for household characteristics, eligible households ended up with less housing equity relative to ineligible households between 1987 and 1991. Examining the trends across cohorts as PVW do in their cohort analysis, does not provide any information on how eligibles fared relative to ineligibles (or on how participants fared relative to non-participants).<sup>60</sup>

Poterba, Venti and Wise (1996) then undertake wealth comparisons of eligible and ineligible households in 1984 and 1991. These comparisons understate the wealth of eligible households in 1984; as shown above, the degree of understatement is potentially quite large. Thus, for example, the statement that financial asset holdings were similar for eligibles and ineligibles in 1984 is impossible to verify with the SIPP data.

In a revised version of their paper, Engen and Gale (1996) extend their results to cover the 1984 to 1991 period. These results show a substantial reduction in non-401(k) wealth for

---

<sup>60</sup>Poterba, Venti and Wise (1996, page 48) note that "the home equity data may be subject to time effects." If that observation is true, it does not invalidate the Engen and Gale results, because they focus on the difference between eligibles and ineligibles over time, so the time effects are incorporated in the data for both groups. But if the home equity data are subject to time effects, it seems likely that the financial asset data would be, too, which would invalidate the within group cohort analysis of Venti and Wise (1996).

eligible households relative to others. Estimating the impact of 401(k)s on total wealth of course is difficult using the 1984 sample due to missing data on 401(k)s and thrifts.

A second criticism by Poterba, Venti and Wise is that the comparison groups used by Engen and Gale are not similar. This criticism is, at the very least, puzzling. The comparison groups in Engen and Gale are chosen deliberately to mirror those in Poterba, Venti and Wise (1995). The comparison groups are (a) eligible households versus ineligible households, (b) eligibles with IRAs versus ineligibles with IRAs, and (c) eligibles without IRAs versus ineligibles without IRAs. Poterba, Venti and Wise (1995) control for age, education, earnings, and marital status within each comparison, and deem the groups similar. Engen and Gale control for those variables plus pension coverage, race and sex of household head, a two-earner indicator, and family size. Engen and Gale also stratify the groups by homeowner status. It is hard to see how controlling for more variables would make the groups dissimilar.

A third criticism is that Engen and Gale require regression coefficients to be the same for eligible and ineligible households, but Poterba, Venti and Wise claim it would be more appropriate to allow the coefficients to differ across groups because the comparison groups differ significantly.<sup>61</sup> This distinction does not affect the results, however. Engen and Gale (1996) provide estimates of a within-group ("like families") estimator for the 1987-91 period and obtain results that are very similar to Engen and Gale (1995). Engen and Gale (1996) also provide within-group estimators and between-group estimators for the 1984 to 1991 period. These estimates show a substantial reduction in non-401(k) wealth for eligible households relative to ineligible households. Again, however, estimating the impact of 401(k)s, however, is

---

<sup>61</sup>We note, however, that this criticism is only valid if eligible and ineligible households are substantially different from each other, that is, if eligibility is not exogenous.

made more difficult due to the absence of 401(k) balance data in 1984.

Finally, Poterba, Venti and Wise claim that exogenous changes in housing value bias the Engen and Gale results toward finding no effect of 401(k)s on wealth. The claim is that eligibles started with higher housing values and all housing values fell by about the same percentage, resulting in large arithmetic declines for eligibles. These results do not control for household characteristics. We note, however, that the Engen and Gale results demonstrate that controlling for other factors, house value rose for eligibles compared to ineligibles. To address the issue further, Engen and Gale (1996) estimate the effects of 401(k)s on a wealth measure that excludes house value. The remaining wealth variable (net financial assets minus mortgage debt) does not rise for eligible households relative to ineligible households over time, thus showing no effects of 401(k) eligibility on wealth.

Bernheim (1996) raises a variety of additional issues. First, he claims the successive cross-section results in both 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 these claims: first, looking only at eligible families over time; second, 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 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). This increase is unlikely to be due to an increase in employer matching.<sup>62</sup> Our own probit analysis using the SIPP indicates that between 1987 and 1991, controlling for household characteristics (including pension coverage), the 401(k) participation rate of eligible households rose by 8 percentage points, and the increase was statistically significant.<sup>63</sup> Ippolito (1993) provides one explanation of these trends: 401(k)s are more attractive to workers with higher tastes for saving, who value the opportunity to contribute and receive employer-matching contributions. In his model, high savers are attracted to, and remain at, firms that sponsor 401(k)s, while low savers are more likely to exit the firm. Thus, the average tastes for saving plausibly rise over time among eligible workers.

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 ones most likely to become eligible, so there is dilution among ineligible households. The key issue is the relative dilution of the two groups, not the absolute dilution in one group. Probits using the SIPP shows that controlling for household characteristics, IRA participation

---

<sup>62</sup>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).

<sup>63</sup>Bernheim (1996) claims that trends in IRA participation are a "good indication" of dilution. But dilution concerns unobservable characteristics ("tastes for saving"), whereas IRA participation depends on observable and unobservable factors. Controlling for observable factors, IRA participation among eligible households fell by 5.4 percentage points from 1987 to 1991. Moreover, IRAs became less attractive relative to 401(k)s over time due to the phase out of tax deductions after 1986 and the lower contribution limits, so eligible households would naturally have substituted into 401(k)s to some extent (Engen, Gale and Scholz 1994). The likelihood that eligible households participated in any saving incentive-IRAs or 401(k)s-rose by 4.4 percentage points, controlling for other factors.

among eligible households fell by only 1.3 percentage points relative to ineligible households from 1987 to 1991, and saving incentive participation rose for eligibles 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, which raises the average tastes for saving among eligible families and may reduce it among ineligible families.<sup>64</sup> Thus, the bias created could plausibly work in either direction, but more importantly it seems to us that dilution is unlikely to be a major quantitative issue in interpreting the results described above.

Besides dilution, Bernheim offers a series of additional comments. First, he notes that the effects of eligibility on wealth are estimated imprecisely in Engen and Gale (1995). However, almost none of the point estimates in Engen and Gale (1995) would indicate any economically significant effect of 401(k) eligibility on wealth, even if they were statistically significant. In any case, to address this concern, Engen and Gale (1996) estimate the impact of eligibility on non-401(k) wealth. The effects were uniformly large and negative and usually statistically significant, consistent with 401(k)s not raising wealth.

Bernheim also suggests there may be measurement problems with the data on housing wealth. But data problems would have generated the Engen and Gale results for homeowners only if the measurement of housing equity became more negative from 1987 to 1991 for eligible homeowners relative to ineligible homeowners. This seems implausible and Bernheim presents no evidence to back this claim. Moreover, data problems in measuring housing value cannot explain the EG results for renters.

Bernheim asserts that the sample of eligible renters appears to be "more selected" than that of eligible homeowners. However, 401(k) participation rates conditional on income, age

---

<sup>64</sup>See Chang (1996) and Fernandez (1992).

and eligibility are slightly lower for renters than for homeowners. This is not consistent with Bernheim's assertion. Bernheim suggests there could be greater dilution of eligible renters over time. Based on IRA participation, the data do not support this claim.<sup>65</sup>

Bernheim also claims that low wealth among ineligible renters could invalidate the results, because if economic forces were pushing down wealth in both groups of renters, ineligible renters may not have been able to reduce their wealth much, due possibly to liquidity constraints. However, if renters save almost nothing, as Bernheim notes, any increase in 401(k) balances should have shown up very clearly as an increase in wealth. Eligibility did raise financial assets in absolute terms among homeowners, who have higher levels of wealth than eligible renters, so there is little reason to see why, if 401(k)s raise wealth, eligibility should not also have raised renters' financial assets.

Finally, Bernheim questions whether substitution between 401(k)s and housing equity is plausible or likely. We discuss the various ways that substitution could occur in the text, and note here that the Engen and Gale results do not require that all households with 401(k)s substitute between saving incentives and housing equity. The Engen and Gale results refer to changes in mean and median wealth, conditional on explanatory variables, so 401(k)s could actually raise wealth for some people and the EG results could still hold.

Bernheim cites one study that finds that in the past—during a time period with high returns to housing investment—retirees had been reluctant to reduce housing equity. But other studies reach different conclusions (see Hurd 1995). And all of these studies may suffer from an

---

<sup>65</sup>In 1987, controlling for other factors, eligible renters were about 2.2 percentage points more likely to hold an IRA. In 1991, the corresponding figure was also 2.2 percentage points, indicating no symptoms of dilution of the sample of eligible renters compared to the sample of ineligible renters.

important sample selection bias: elderly people who reduce their equity (by selling their house and moving to another, or into a nursing home, or in with relatives) may get dropped from subsequent waves of the survey. Most importantly, whether the elderly wish to consume their housing equity is not directly relevant for interpreting the results in Engen and Gale (1995), who (a) focus on workers, who are accumulating assets, rather than the elderly, who are dissaving, (b) find evidence of reshuffling between mortgage debt and 401(k)s, and (c) generally do not find 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.

Along similar lines, Bernheim notes that many younger households say they view their home equity primarily as a source of financial security. The relevance of his observation is unclear. As Bernheim notes, surveys of people's intentions warrant skepticism. Moreover, Engen and Gale measure people's actions, not their intentions. In addition, survey responses about intentions notwithstanding, there were in fact several home equity lending booms in the 1980s and 1990s indicating that some households were taking equity out of their houses.



## 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.
- Altig, David. 1990. "The Case of the Missing Interest Deductions: Will Tax Reform Increase U.S. Saving Rates?" Federal Reserve Bank of Cleveland Economic Review 26(4):22-34.
- Andrews, Emily S. 1992. "The Growth and Distribution of 401(k) Plans." In John A. Turner and Daniel J. Beller, eds., Trends in Pensions 1992. Washington: U.S. Department of Labor, Pension and Welfare Benefits Administration.
- Attanasio, Orazio, and Thomas De Leire. 1994. "IRAs and Household Saving Revisited: Some New Evidence." Working Paper No. 4900. Cambridge, Mass.: National Bureau of Economic Research (October).
- Auerbach, Alan J. 1996. "Tax Reform, Capital Allocation, Efficiency and Growth." Forthcoming in Henry J. Aaron and William G. Gale, eds., Economic Effects of Fundamental Tax Reform (February).
- Auerbach, Alan J., and Joel Slemrod. 1996. "The Economic Effects of the Tax Reform Act of 1986." Forthcoming, Journal of Economic Literature.
- Avery, Robert B., Gregory E. Elliehausen, and Thomas A. Gustafson. 1986. "Pensions and Social Security in Household Portfolios: Evidence from the 1983 Survey of Consumer Finances." In F. Gerard Adams and Susan M. Wachter, eds., Savings and Capital Formation. Lexington, MA: Lexington Books, 127-60.
- Bassett, William F., Michael J. Fleming, and Anthony P. Rodrigues. 1996. "How Workers Use 401(k) Plans: The Participation, Contribution, and Withdrawal Decisions." Mimeo. Federal Reserve Board of New York (April).
- Bernheim, B. Douglas. 1987. "The Economic Effects of Social Security: Toward a Reconciliation of Theory and Measurement." Journal of Public Economics 33, 273-304.
- Bernheim, B. Douglas. 1994a. "Comment on Chapters 4 and 5." In David Wise, ed., Studies in the Economics of Aging. Chicago: University of Chicago Press and National Bureau of Economic Research.
- Bernheim, B. Douglas. 1994b. "Comment on Engen, Gale and Scholz." Brookings Papers on Economic Activity 1: 152-71.
- Bernheim, B. Douglas. 1996. "Rethinking Saving Incentives." forthcoming, in Alan Auerbach, ed., Fiscal Policy: Lessons from Economic Research, MIT Press, forthcoming.
- Bernheim, B. Douglas and Daniel M. Garrett. 1995. "The Determinants and Consequences of

- Financial Education in the Workplace: Evidence from a Survey of Households." Mimeo. Stanford University (August).
- Bernheim, B. Douglas, and John Karl Scholz. 1993a. "Private Saving and Public Policy." In James M. Poterba, ed., Tax Policy and the Economy, vol. 7. Cambridge, Mass.: MIT Press.
- Bernheim, B. Douglas, and John Karl Scholz. 1993b. "Private Pensions and Household Saving." Mimeo. University of Wisconsin.
- Blinder, Alan S., Roger H. Gordon, and Donald E. Wise. 1980. "Reconsidering the Work Disincentive Effects of Social Security." National Tax Journal 33(4): 431-442 (December).
- Board of Governors of the Federal Reserve System. 1995. Balance Sheets for the U.S. Economy 1945-1994.
- Browning, Martin and Annamaria Lusardi. 1995. "Household Saving: Micro Theories and Micro Facts," Mimeo. April.
- Buck Consultants, Inc. 1989. "Current 401(k) Plan Practices: A Survey Report." Secaucus, N.J.: Buck Consultants, Inc.. (September).
- Burman, Leonard Burman, Joseph Cordes, and Larry Ozanne. 1990. "IRAs and National Savings," National Tax Journal, September. XLIII:3, 259-85.
- Carroll, Chris and Lawrence H. Summers. 1987. "Why Have Private Savings Rates in the United States and Canada Diverged?" Journal of Monetary Economics 20: 249-279.
- Chang, Angela E. 1996. "Tax Policy, Lump Sum Pension Distributions, and Household Saving." National Tax Journal. 49(2): 235-252.
- Curme, Michael A. William E. Even. (1995). "Pension Coverage and Borrowing Constraints." Journal of Human Resources. 30(4): 701-712.
- Diamond, P.A., and J.A. Hausman. 1984. "Individual Retirement and Savings Behavior." Journal of Public Economics 23: 81-114.
- Dicks-Mireaux, Louis, and Mervyn King. 1984. "Pension Wealth and Household Savings: Tests of Robustness." Journal of Public Economics 23: 81-114.
- Doyle, Robert J., Jr., and Eric T. Johnson. 1991. "Readings in Wealth Accumulation Planning, fourth edition, Bryn Mawr, Pennsylvania: The American College.
- Economic Report of the President. 1996. U.S. Government Printing Office, Washington (February).
- Employee Benefits Research Institute. 1994. Issue Brief (November).

Employee Benefits Research Institute. 1995. Databook on Employee Benefits, Third Edition. Washington, D.C.

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

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

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

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

Feenberg, Daniel, and Jonathan Skinner. 1989. "Sources of IRA Saving." In Lawrence H. Summers, ed., Tax Policy and the Economy, vol. 3. Cambridge, Mass.: MIT Press.

Feldstein, Martin. 1980. "Inflation, Portfolio Choice, and the Prices of Land and Corporate Stock." American Journal of Agricultural Economics 62(5): 910-916 (December).

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

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

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

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

Gale, William G. 1996. "Comments on Bernheim's "Rethinking Saving Incentives." in Alan Auerbach, ed., Fiscal Policy: Lessons from Economic Research, MIT Press, forthcoming.

Gale, William G. and John Karl Scholz. 1990. "IRAs and Household Saving." Mimeo. May.

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

Gravelle, Jane. 1991. "Do Individual Retirement Accounts Increase Saving?" Journal of Economic Perspectives 5: 133-49 (Spring).

- Hall, Robert E. 1988. "Intertemporal Substitution in Consumption." Journal of Political Economy 96(2): 339-57 (April).
- Halperin, Daniel I. 1987. "Tax Policy and Retirement Income: A Rational Model for the 21st Century." In Jack L. Vanderhei, ed., Search for a National Retirement Income Policy. Homewood, Illinois: Richard D. Irwin, Inc., 159-195.
- Hubbard, R. Glenn. 1984. "Do IRAs and Keoghs Increase Saving?" National Tax Journal 37: 43-54.
- Hubbard, R. Glenn. 1986. "Pension Wealth and Individual Saving." Journal of Money, Credit, and Banking 18(2): 167-178 (May).
- Hubbard, R. Glenn, and Jonathan S. Skinner. 1995. "The Effectiveness of Saving Incentives: A Review of the Evidence." Mimeo. (June).
- Hurd, Michael. 1995. "Mortality Risk and Consumption by Couples," Presentation at the Conference on the Microeconomics of Saving and Consumption Growth, Institute for Fiscal Studies and Bank of Portugal, Estoril Portugal. November.
- Internal Revenue Service. 1994. Statistics of Income Bulletin. (Spring) Washington D. C.
- Ippolito, Richard A. 1993. "Selecting and Retaining High-Quality Workers: A Theory of 401(k) Pensions." Unpublished paper. Pension Benefit Guaranty Corporation, Washington (April).
- Johnson, Richard W. 1993. "The Impact of Worker Preferences on Pension Coverage in the HRS." Health and Retirement Study Working Paper Series Paper No. 94-018 (December).
- Joines, Douglas H., and James G. Manegold. 1995. "IRAs and Saving: Evidence from a Panel of Taxpayers." Mimeo. USC.
- Kennickell, Arthur B. and David W. Wilcox. 1995. "The Value and Distribution of Unrealized Capital Gains: Evidence from the 1989 Survey of Consumer Finances." Mimeo. Federal Reserve Board (November).
- Kruse, Douglas S. 1991. "Pension Substitution in the 1980's: Why the Shift toward Defined Contribution Pension Plans?" Working Paper 2882. Cambridge, Mass.: National Bureau of Economic Research (October).
- 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).
- Laibson, David. 1996. "Hyperbolic Discount Functions, Undersaving, and Savings Policy," NBER Working Paper No. 5635, June.

Long, James A. 1990. "Marginal Tax Rates and IRA Contributions." National Tax Journal 43: 143-154 (June).

Papke, Leslie E. 1995. "Does 401(k) Introduction Affect Defined Benefit Plans?" National Tax Association Proceedings of the Eighty-Sixth Annual Conference, 1994, 122-3.

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

Papke, Leslie E., Mitchell Petersen, and James M. Poterba. 1993. "Did 401(k) Plans Replace Other Pensions?" Working Paper 4501. Cambridge, Mass.: National Bureau of Economic Research (October).

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

Poterba, James M., Steven F. Venti, and David A. Wise. 1994. "Targeted Retirement Saving and the Net Worth of Elderly Americans." American Economic Review 84(2): 180-185 (May).

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

Poterba, James M., Steven F. Venti, and David A. Wise. 1996. "Personal Retirement Saving Programs and Asset Accumulation: Reconciling the Evidence," NBER Working Paper No. 5599, May.

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

Sabelhaus, John. 1996. "Public Policy and Saving Behavior in the U.S. and Canada." Mimeo. (February).

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

Skinner, Jonathan, and Daniel Feenberg. 1990. "The Impact of the 1986 Tax Reform on Personal Saving." In Joel Slemrod, ed., Do Taxes Matter? The Impact of the Tax Reform Act of 1986. Cambridge, Mass. and London: MIT Press, 50-79.

Slemrod, Joel. 1988. "The 1979-84 Linked Panel of Tax Return Data: Sampling and Linking Methodology." Mimeo. University of Michigan (November).

Slemrod, Joel. 1990. "Notes on the 1979-86 Panel of Individual Income Tax Returns." Mimeo. University of Michigan (July).

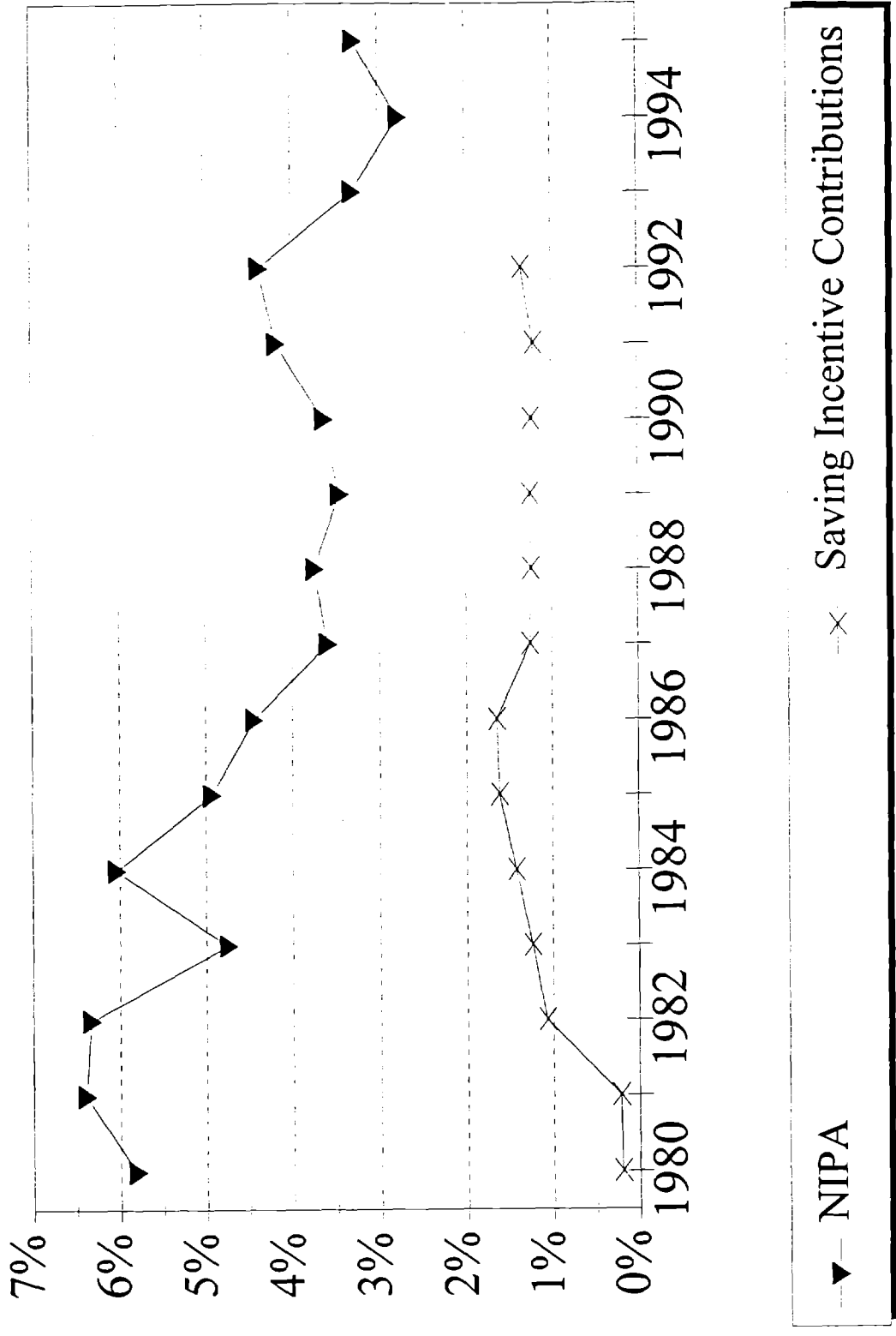
Slemrod, Joel. 1992. "The Economic Impact of the Tax Reform Act of 1986." In Joel Slemrod,

- ed., Do Taxes Matter? The Impact of the Tax Reform Act of 1986. MIT Press.
- Slemrod, Joel. 1994. "Comments and Discussion on 'Do Saving Incentives Work?'" Brookings Papers on Economic Activity 1: 166-171.
- Stiglitz, Joseph E. 1988. Economics of the Public Sector, second edition, New York: W.W. Norton and Company.
- Summers, Lawrence H. 1981. "Inflation, the Stock Market, and Owner-Occupied Housing." American Economic Review 71(2): 429-434 (May).
- Thaler, Richard. 1990. "Anomalies: Saving, Fungibility, and Mental Accounts." Journal of Economic Perspectives 4(1): 193-205 (Winter).
- Thaler, Richard. 1994. "Mental Accounts and Household Saving." American Economic Review (May).
- United States Department of Labor, Pension and Welfare Benefits Administration. 1996. Private Pension Plan Bulletin, Abstract of 1992 Form 5500 Annual Reports, Number 5 (Winter).
- Venti, Steven F., and David A. Wise. 1986. "Tax-Deferred Accounts, Constrained Choice, and Estimation of Individual Saving." Review of Economic Studies 53: 579-601.
- Venti, Steven F., and David A. Wise. 1987. "IRAs and Saving." In Martin Feldstein, ed., The Effects of Taxation on Capital Accumulation. Chicago: University of Chicago Press.
- Venti, Steven F., and David A. Wise. 1990. "Have IRAs Increased U.S. Saving?: Evidence from Consumer Expenditure Surveys." Quarterly Journal of Economics 105: 661-698 (August).
- Venti, Steven F., and David A. Wise. 1991. "The Saving Effect of Tax-Deferred Retirement Accounts: Evidence from SIPP." In B. Douglas Bernheim and John B. Shoven, eds., National Saving and Economic Performance. University of Chicago Press and NBER.
- Venti, Steven F., and David A. Wise. 1992. "Government Policy and Personal Retirement Saving." In James M. Poterba, ed., Tax Policy and the Economy, vol. 6. Cambridge, Mass.: MIT Press.
- Venti, Steven F., and David A. Wise. 1995a. "Individual Response to Retirement Saving Programs: Results From U.S. Panel Data." Recherche Economique 49:235-254.
- Venti, Steven F., and David A. Wise. 1995b. "RRSPs and Saving in Canada." Mimeo. (December).
- Venti, Steven F., and David A. Wise. 1996. "The Wealth of Cohorts: Retirement Saving and the Changing Assets of Older Americans." NBER working paper No. 5609.

Figure 1

# Personal Saving as a Percentage of GDP

## Selected Measures, 1980-1995



**Table 1**  
**Cohort Effects at Age 60-64<sup>1</sup>**

<i>All Households Aged 60-64</i>	1984	1991	Difference
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
 <i>Contributors</i>			
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
 <i>Non-Contributors</i>			
Median Financial Assets	2,687	2,134	-553

Source: Poterba, Venti, and Wise (1996).



**Table 2**  
**IRA Holders in the 1986 SCF<sup>2</sup>**

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 <sup>3</sup> (Median)	10,200	21,500
IRA plus Keogh	250	6,500
Other Financial Assets	7,699	12,000
Net Financial Assets <sup>4</sup> plus Housing Equity (median)	51,220	64,897
Sample Size	618	

Source: Authors' calculations.

1. 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.

2. All dollar figures are in nominal dollars.

3. Financial assets include stocks and mutual funds, bonds, checking and savings accounts, IRA and Keogh accounts, money market accounts and certificates of deposit.

4. Net financial assets equals financial assets minus consumer debt.