NBER WORKING PAPER SERIES

THE LABOR SUPPLY EFFECTS OF THE SOCIAL SECURITY EARNINGS TEST

Leora Friedberg

Working Paper 7200
http://www.nber.org/papers/w7200

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1999

I would like to thank Josh Angrist, George Borjas, Peter Diamond, Gary Engelhardt, Jon Gruber, Steve Pischke, Jon Skinner, Aaron Yelowitz, two anonymous referees; participants of seminars at MIT, UCSD, the NBER Summer Institute, and several other universities; and especially Jerry Hausman and Jim Poterba for very thoughtful comments and discussions. I gratefully acknowledge financial support from the National Institute on Aging and the National Bureau of Economic Research Fellowship in Aging and Health Economics. All opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

The Labor Supply Effects of the
Social Security Earnings Test

Leora Friedberg

NBER Working Paper No. 7200
June 1999
JEL No. J22, J14, H55

ABSTRACT

The Social Security earnings test reduces benefits at a 33-50% rate once earnings pass a threshold amount - among the highest marginal tax rates in the economy. Previous research dismissed the importance of the earnings test but failed to take advantage of three changes in the earnings test rules in order to identify its impact. Each change applied to some age groups and not others - which make them useful for identifying the effect of tax rules on the labor supply of working beneficiaries. Beneficiaries in the Current Population Survey satisfy the strongest prediction: many keep their earnings just below the exempt amount, and this bunching shifts with the earnings test rule changes. The rule changes are then incorporated into an econometric model of labor supply to identify income and substitution elasticities. The resulting elasticity estimates suggest considerable deadweight loss suffered by working beneficiaries. Simulations predict a substantial boost to labor supply from eliminating the earnings test, and at a minimal fiscal cost. However, a slight decrease in labor supply is predicted from the recently legislated increase in the exempt amount.

Leora Friedberg
Department of Economics, 0508
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
lfriedbe@weber.ucsd.edu

THE LABOR SUPPLY EFFECTS OF THE SOCIAL SECURITY EARNINGS TEST

Leora Friedberg

University of California, San Diego and NBER

June 1999

ABSTRACT

The Social Security earnings test reduces benefits at a 33-50% rate once earnings pass a threshold amount – among the highest marginal tax rates in the economy. Previous research dismissed the importance of the earnings test but failed to take advantage of three changes in the earnings test rules in order to identify its impact. Each change applied to some age groups and not others – which make them useful for identifying the effect of tax rules on the labor supply of working beneficiaries. Beneficiaries in the Current Population Survey satisfy the strongest prediction: many keep their earnings just below the exempt amount, and this bunching shifts with the earnings test rule changes. The rule changes are then incorporated into an econometric model of labor supply to identify income and substitution elasticities. The resulting elasticity estimates suggest considerable deadweight loss suffered by working beneficiaries. Simulations predict a substantial boost to labor supply from eliminating the earnings test, and at a minimal fiscal cost. However, a slight decrease in labor supply is predicted from the recently legislated increase in the exempt amount.

JEL Codes: H55, J14, J22

I would like to thank Josh Angrist, George Borjas, Peter Diamond, Gary Engelhardt, Jon Gruber, Steve Pischke, Jon Skinner, Aaron Yelowitz, two anonymous referees; participants of seminars at MIT, UCSD, the NBER Summer Institute, and several other universities; and especially Jerry Hausman and Jim Poterba for very thoughtful comments and discussions. I gratefully acknowledge financial support from the National Institute on Aging and the National Bureau of Economic Research Fellowship in Aging and Health Economics.

The Social Security earnings test generates some of the highest marginal tax rates in the economy today. The earnings test reduces someone's Social Security benefits once earnings pass a threshold amount. When the earnings test was introduced in 1939, the intent was to push older workers out of the labor force, so beneficiaries lost an entire month's benefits when monthly earnings exceeded $15. Since the 1950s, the earnings test has been gradually relaxed as the emphasis has shifted towards encouraging work and saving. In 1998, the earnings test took away $1 in benefits for every $2 in earnings above $9,120 for a beneficiary under the age of 65 – a 50% tax on wages. A beneficiary aged 65-69 with earnings above $14,500 faced a 33% earnings test tax rate.

Altogether, a 64-year old beneficiary in 1995 faced a combined marginal tax rate of 83% on average, and possibly up to 114%, for earnings above the exempt amount.[1] In consequence, the earnings test receives substantial popular attention, typified by the title of the 1989 study *Paying People Not to Work* (Robbins and Robbins). The 1999 *Economic Report of the President* reviewed the impact of the earnings test.[2] Concern over distortions to labor supply led to a 1996 law gradually raising the earnings exempt amount for older beneficiaries to $30,000 by 2002. On the other hand, the earnings test will grow more restrictive for some beneficiaries beginning in 2000, as the tighter rules for 62-64 year olds are extended while the normal retirement age slowly rises from 65 to 67.

In spite of substantial popular attention to the earnings test, the existing literature concludes it has a small impact on the labor supply of the elderly.[3] Earlier papers used the Retirement History Survey, which ended in 1979 and has become dated. More importantly, there were virtually no changes in the earnings test during the time period covered by the Retirement History Survey, so estimates of the earnings test's effect are identified primarily from potentially endogenous cross-sectional variation in wages and non-labor income.

---

[1] 83% reflects the earnings test, payroll tax, and federal and state income tax, according to the National Center for Policy Analysis (1995). The tax rate climbs higher when Social Security benefits becomes taxable at higher income levels.

[2] Pages 142-145.

[3] Burtless and Moffitt (1985), Gustman and Steinmeier (1986).

In this paper I investigate three changes in the earnings test rules instituted between 1978 and 1990. These changes, which remain mostly unexplored, involved different combinations of shifts in the earnings test tax rate and the earnings exempt amount. Also, each applied to beneficiaries of some ages and not others, so the behavior of unaffected age groups can be used to control for other changes in labor supply.

This sets the stage for a "natural experiment" analysis of the earnings test of the sort used to study many social programs and tax changes.[4] In this vein, I find that the working elderly bunch in substantial numbers at and just below the earnings exempt amount. The bunching reacts directly to the changes in the earnings test rules. The evidence of bunching casts doubt on earlier findings that the earnings test has little impact on behavior and contrasts with other circumstances where people fail to bunch as predicted at smaller convex kinks in the budget constraint.[5] The bunching at the kink also demonstrates that econometric analysis needs to allow for these responses.

This paper extends the natural experiment approach by exploiting the changes in the earnings test rules to estimate labor supply elasticities. The estimation incorporates the reactions to the earnings test along the entire piecewise linear budget constraint and decomposes those reactions into income and substitution elasticities. This type of estimation is typically faulted for imposing assumptions on the specification that drive the estimates and for overlooking identification. However, the earnings test budget constraint is well-defined and easy to measure, avoiding the pitfalls that sometimes arise with this estimation method. And, to get the most out of the changes in the earnings test, I combine data before and after the major rule changes. The substantial variation in the budget constraint both cross-sectionally and over time drives the resulting econometric estimates, as will be demonstrated.

The estimation yields relatively large income and substitution elasticities that imply substantial deadweight loss from the earnings test. Eliminating the earnings test is predicted to raise average hours worked by 5.3% for those currently at or above the exempt amount. In contrast, the increase in the exempt amount to $30,000 is predicted to *lower* aggregate hours slightly because of the negative substitution and income effects imposed on high earners. It is important to note that the fiscal cost of

4 These methods and their applications are described in Meyer (1995).

5 Heckman (1983), MaCurdy (1992).

easing the earnings test over the medium term is small and declining, because of the rules governing future benefits. These rules ultimately erase most of the fiscal gains from the earnings test, but they are complicated and do not appear to reduce the labor supply distortions.

The elasticities estimated here are also of broader interest in understanding the elderly who continue to work, in contrast to the general focus in the literature on retirement. It will be important to keep in mind this apparent sensitivity of older workers to tax and transfer rules conditional on working, which will affect policies that aim to induce later retirement. The results suggest a potentially severe negative effect on labor supply if benefits are means tested – a proposal gaining attention recently. Another important consideration involves 62-64 year olds. They continue to face a much more restrictive earnings test, unchanged since the mid-1970s, which will gradually be extended to older beneficiaries as the normal retirement age begins to rise from 65 to 67 in 2000. As the median retirement age continues to drop – from 65 in the 1970s to 62 today – the earnings test is growing more binding for the younger group over time.

The rest of this paper is divided into five sections. Section I describes the earnings test, its impact on beneficiaries, and the earlier literature. The actual response of workers to the earnings test is demonstrated in Section II. Section III incorporates the piecewise linear budget constraint generated by the earnings test into an econometric model of labor supply. Section IV describes the resulting maximum likelihood estimates, and Section V concludes.

I. BACKGROUND

This section describes how the earnings test works at present and in the past. It follows with evidence on the number of beneficiaries affected by the earnings test and then discusses earlier papers on the earnings test.

A. The Earnings Test Rules

Once a Social Security beneficiary earns more than a certain amount, his or her benefits are reduced at a rate proportional to additional earnings. The benefit reduction rate is equivalent to a tax applied to wages until benefits are gone. In 1998 a beneficiary aged 62-64 could earn up to $9,120 –

the earnings *exempt amount* – with no reduction in benefits. When he works more, he loses $1 in benefits for every $2 he earns – a 50% *tax rate.* For 65-69 year olds, the earnings test is less restrictive with a higher exempt amount and a 33% tax rate, as reported in Table 1. The earnings test no longer applies once a person turns 70.[6]

As shown in Figure 1, the earnings test generates a piecewise linear budget constraint with one convex kink corresponding to the exempt amount and one nonconvex kink corresponding to the exhaustion of benefits. The point where each segment of the budget constraint intercepts the zero-hours-of-work axis measures virtual income, which is the analog on that segment to non-labor income and is due to nonlinear taxation. When a change in the earnings test shifts the budget constraint, the shift in the virtual income intercept measures the income effect and the shift in the slope measures the substitution effect.[7]

At a convex kink a person faces distinct marginal wages when he considers working slightly more or less. Choosing hours just on the kink is compatible with a range of indifference curves. There is no other reason for people to mass at a particular point, so observing a cluster at the kink will be the strongest evidence that people react to the earnings test.[8]

The earnings test has been altered several times in recent years, as highlighted in Table 1.[9] In 1990 the tax rate for 65-69 year olds was reduced from 50% to 33%, which rotates upwards the middle segment of the budget constraint in Figure 1. Intuitively, one expects labor supply to rise when the penalty for working is lowered. People at the kink do enjoy a higher marginal wage and an incentive to work more. However, people on the middle segment already losing benefits become better off when their income rises and may work more or less. The income effect grows along the middle segment, and people just above the nonconvex kink experience a negative substitution effect as well and will lower their hours. It is more difficult to search for this range of predicted responses in the raw data than to

---

[6] Nonlabor income is not subject to the earnings test. The exempt amount for 62-64 year olds is raised yearly according to average earnings. 1996 legislation raises the exempt amount for older beneficiaries $1,000 a year through 1999 and to $17,000 in 2000, $25,000 in 2001, and $30,000 in 2002. The normal retirement age will be raised from age 65 to 67 by two months per year from 2000 to 2005 and from 2017 to 2022.

[7] Figure 1 abstracts from income taxes for simplicity, but taxes are incorporated in the econometric analysis.

[8] Another prediction is that people should avoid the nonconvex kink, depending on the convexity of their indifference curves. It is difficult to detect in the data because the location of the nonconvex kink differs across people and is unknown for those on the upper segment.

look for a reaction at the kink, especially lacking longitudinal data. Thus, understanding how all beneficiaries react demands a more structured framework for analyzing the data.

The other changes in the earnings test rules also shifted the budget constraint. In 1983 the earnings test was eliminated entirely for 70-71 year olds. This is the most substantial change, and it creates two natural comparison groups: younger people who face the same budget constraint as 70-71 year olds before 1983, and older people who face the same budget constraint after. In 1978 the exempt amount was raised about 25% for 65-71 year olds. A comprehensive picture emerges from the set of rules changes, involving different combinations of shifting the convex kink and the slope of the budget constraint. And, since each change affected a particular age group, the earnings test response is further isolated by comparing the change in behavior to the unaffected group as a way to control for other trends in labor supply over time.

As illustrated, Figure 1 presumes that workers can freely choose their hours on the entire budget constraint. Research on prime-age workers, though, shows that hours choices are constrained by fixed costs, minimum hours requirements, and tied wage-hours contracts.[10] The responsiveness of older workers to the earnings test suggests that they enjoy greater hours flexibility. Ruhm (1990) documented the importance of flexible "bridge jobs" taken by many people who continue working after leaving their career jobs. Also, older workers are much more likely to work part-time. 44.7% of 65-69 year old men working in 1990 worked part-time hours, compared to 4.2% of 40-49 year old men.[11]

It is important to mention another set of rules governing the relationship between the earnings test and future benefits. These rules substantially reduce the long-run fiscal gains from the earnings test and should reduce its impact on labor supply. Just as people are rewarded with higher future benefits if they delay claiming, beneficiaries who work are compensated for current benefits lost to the earnings test. For beneficiaries aged 62-64 the actuarial adjustment raises future benefits by about 7% for each year's worth of benefits forgone. For beneficiaries aged 65-69 the Delayed Retirement Credit was introduced in 1973 at 1%, raised in 1982 to 3%, and since 1990 is being raised gradually to 8%. A 7-

---

[9] The history of earnings test changes is detailed in Friedberg (1998).

[10] For example, Card (1990), Altonji and Paxson (1988), Cogan (1981), and Hausman (1980).

[11] Data on usual hours of work in the previous year from the March 1991 CPS, reported in Friedberg (1999). Part-time is defined as less than 35 hours per week.

8% credit is meant to be actuarially fair on average, so someone with average life expectancy and no borrowing constraints should be indifferent between either receiving benefits today or forgoing them and later receiving higher benefits over the rest of the expected lifetime.[12]

However, the credits do not appear to affect behavior and may not be well understood. The earnings test's impact is confirmed by the clustering at the convex kink demonstrated in the next section. Recent articles in *Money* and the *Los Angeles Times* describing the earnings test failed to note that future benefits are raised if current benefits are reduced.[13] Similarly, Reimers and Honig (1993, 1996) found no evidence of increased labor force re-entry when the Delayed Retirement Credit was raised. In the estimation described in Sections III and IV I tried allowing the effective earnings test tax rate to be reduced below the statutory rate by the Delayed Retirement Credit, but no such effect emerged. In the rest of this paper I take the view that the credits adjusting future benefits do not influence the response to the earnings test.

B. How Many Are Affected

Figure 1 made it clear that the earnings test will change the incentives of people who want to work a significant amount. Data from the *Social Security Bulletin* summarized in Table 2 suggests that a large fraction of older workers are affected by the earnings test.

About 926,000 out of 9.8 million retired-worker beneficiaries aged 62-69 in 1989 suffered a reduction in benefits due to the earnings test, according to Bondar (1993). Leonesio (1990) also noted the impact of the earnings test on an estimated 173,700 beneficiaries aged 65-69 in the immediate vicinity of the kink and 582,000 eligibles with earnings who did not claim benefits and should be located on the upper segment of the budget constraint in Figure 1. Those and other articles in the *Social Security Bulletin* argued that a small percentage of the elderly are affected by the earnings test.[14] However, with over one-third of working 65-69 year olds losing benefits to the earnings test and

---

[12] Burkhauser (1980) and Blinder, Gordon, and Wise (1980) noted this potential impact of the actuarial adjustment, although the Delayed Retirement Credit is only slowly becoming actuarially fair. Beneficiaries are compensated for each full month's worth of benefits lost.

[13] Simon (1996), Kristof (1997). Coile, et al, (1999) found that many beneficiaries for whom the adjustment is actuarially fair or better claim benefits early at age 62, more evidence that the credits are not fully taken into account. Burtless and Moffitt (1985) did not incorporate the credits, as people do not appear to react to them.

[14] Packard (1990), Lingg (1986).

hundreds of thousands more affected as well, it appears that the earnings test could have an important impact.[15]

C. Previous Literature

In spite of the substantial popular attention, the previous literature on the earnings test generally dismissed its importance. The most thorough papers to date are by Burtless and Moffitt (1985), who noted the reaction of beneficiaries at the exempt amount, and Gustman and Steinmeier (1986).[16] They incorporated the effect of the earnings test into joint models of retirement and "post-retirement" labor supply. The post-retirement hours choice was modeled using the piecewise linear budget constraint approach, which is described in Section III.

Each paper used the Retirement History Survey, a rich longitudinal survey that lasted from 1969 to 1979, but with a major weakness for studying the earnings test. During that time beneficiaries of all ages faced virtually unchanged earnings test parameters.[17] What remains to identify the estimates of post-retirement hours elasticities is other variation in the budget constraint across individuals. However, cross-sectional variation due to wages and non-labor income is likely to be correlated with omitted individual characteristics, biasing the estimates. Moreover, time series variation in post-retirement labor supply may be confounded by a strong trend in retirement during those years.[18] Krueger and Pischke (1992) pointed out the potential for spurious correlation, since real Social Security benefits rose sharply at the same time.

Both papers estimated significant elasticities, but both concluded that the earnings test was unimportant. Burtless and Moffitt, like the articles in the *Social Security Bulletin*, judged that the percentage affected by the earnings test was minor, although the previous subsection showed that over

---

[15] These statistics do not account for the possible impact on retirement, arising if there are restrictions on hours choices, as discussed earlier. The focus is on the cleanest predictions, relating to hours of work conditional on working.

[16] Vroman (1985) also documented the bunching of beneficiaries at the exempt amount in the 1970s.

[17] In 1972 a higher earnings test "tax bracket" with a 100% tax rate was eliminated. However, the 1972 change is less useful for identifying the reaction to the earnings test. It applied to beneficiaries of all ages, yielding no natural comparison group, an important factor with the sharp contemporaneous increase in retirement. Also, there was no apparent reaction to the higher kink or to its elimination. While that bounds the underlying responsiveness to the earnings test, it is still consistent with the strong reaction to the initial kink demonstrated here. The propensity to locate at the higher kink is reduced because of its proximity to the nonconvex kink, which is avoided.

[18] The proportion of non-working men aged 64 in the CPS rose from 44.0% in 1969 to 54.1% in 1977.

one-third of working beneficiaries lose benefits to the earnings test. Gustman and Steinmeier's conclusions were influenced by their treatment of the actuarial adjustment and Delayed Retirement Credit, which should reduce the impact of the earnings test for many beneficiaries.[19] However, as I noted previously, beneficiaries do not appear to be taken into account in the reaction to the earnings test.

II. EARNINGS DISTRIBUTIONS AND THE EARNINGS TEST

The conclusion in earlier research, therefore, is that the earnings test has a minor impact on behavior. On the other hand, beneficiaries appear to be well informed about the earnings test, and the popular view is that it deters people from working.[20] How can one determine its influence on labor supply? Empirical strategies vary in the degree to which they formalize individual behavior: the more structured the approach, the more closely the conclusions resemble theoretical concepts of interest; but also the more the accuracy of the conclusions depends on the formalization itself being a good description of reality.

In the rest of this paper, I combine two approaches. In order to focus on the strongest predictions, this section analyzes raw data on earnings relative to the earnings exempt amount. Comparing earnings before and after the different rule changes decomposes how people respond to the exempt amount and to the tax rate.[21] Further, comparing these shifts to other unaffected groups over time controls for aggregate trends which also move the earnings distribution. However, capturing the responses of people along other parts of the budget set, and quantifying those responses in terms of income and substitution elasticities, will require a more structured framework. Section III will incorporate the entire piecewise linear budget constraint illustrated in Figure 1 in a model of labor supply.

---

[19] Simulations involving the earnings test appear in Gustman and Steinmeier (1985, 1991).

[20] Leonesio (1990) reported that 73% of retirees under age 72 in the 1982 New Beneficiary Survey knew of the earnings test. Simon (1996) in *Money* and Kristof (1997) in *The Los Angeles Times* are recent examples of articles in the popular press describing the earnings test.

[21] The earnings test affects narrow age ranges, so the response to rule changes generally occurs among new cohorts entering these age ranges and facing new earnings test rules, rather than among the same people facing new rules and adjusting their work hours.

Figure 2 begins by showing the earnings distribution *relative to the exempt amount* before and after the exempt amount was raised for 65-71 year olds in 1978. Using data from March Current Population Surveys (CPS), the graphs compare the previous year earnings of 67-69 year old men and 63-64 year old men, who did not experience any change in the exempt amount. Figure 2-A shows, before 1978, how many of the older and younger groups had earnings in each $1000 interval above and below the convex kink defined by the exempt amount, as a proportion of the total number of people in the age group.[22,23]

Figure 2-A demonstrates a strong response to the earnings kink. People in both age groups bunch just below the kink. Roughly the same number of people appear in each increment for several steps below the kink, followed by a big drop – of over 2% of the sample for the 63-64 year olds and 4% for the 67-69 year olds (about 8% of working 67-69 year olds) – in the step from just below to just above the kink. The visible reaction to the earnings test is somewhat in contrast to earlier conclusions that the earnings test has little impact.

After 1978, the cluster of 67-69 year olds moves up to the new kink. Figure 2-B shows earnings of both age groups in relation to the unchanged exempt amount of the younger group. The 63-64 year olds cluster just below it as before, but the earnings of 67-69 year olds have clearly shifted upwards. Figure 2-C shows them bunching at their new higher kink.

Quantifying the visual evidence allows conclusions about statistical significance. The first panel of Table 3 reports the same data as Figure 2. Each column reports the proportion of individuals with earnings within a few intervals around the kink. The last row computes the difference between the percentage of people just below and above the kink and demonstrates that the bunching in every case is both significant and significantly different from the behavior across any other interval. The first and fourth columns report the distribution of people in both age groups before 1978. The other columns

---

[22] Figures 2-4 show earnings from wages and salaries and from self-employment. Age groups were narrowed because the CPS reports age in March following the working year, instead of exact birthdays. The sample is restricted to men because spousal benefits are complex. A spouse chooses whether to receive benefits as a dependent or a retiree, and the benefit type is not reported in the CPS. Dependent spouses lose benefits from both the dependent's and the retiree's earnings, while retiree spouses lose benefits only from the retiree's earnings.

[23] While theory predicts a cluster exactly at the kink, measurement error or restrictions in hours choices will plausibly spread out the cluster in the neighborhood of or just below the kink. The interval width of $1000 was chosen because respondents sometimes round off reported earnings to the nearest thousand, so a different interval length confounds the measurement of

compare the patterns of earnings at the new separate kinks and demonstrate that each group now responds strongly to its own kink. Together, the visual and statistical evidence confirms that a significant number of people react to the earnings test by holding down their labor supply.

Figure 3 makes the same comparisons of earnings around the kink before and after 1983, when the earnings test was eliminated for 70-71 year olds. Figures 3-A and 3-B illustrate earnings patterns before 1983.[24] Figure 3-A compares the earnings of the affected group to that of a younger group that faces the earnings test both before and after 1983. The clustering at the kink by those of both ages is, again, substantial. Figure 3-B gives a sense of the counterfactual by comparing the affected group to an older group of 73-75 year olds who do not face the earnings test and whose earnings decline smoothly over the same range.

Figures 3-C and 3-D make the same comparisons after 1983. Now, the earnings of the affected 71-72 year olds resembles the older group, declining smoothly over the range of the earnings kink. The younger group continues to bunch at the kink. The evidence is quantified in Table 3, comparing the affected 71-72 year olds (in the middle columns) with the younger and older groups. In the first three columns, among those facing the earnings test, the percentage located just below the kink is large: 4.1% of the younger group before 1983, 2.2% of the younger group after, and 2.3% of the 71-72 year olds before; versus 0.8-1.3% for those with no earnings test, including the 71-72 year olds after 1983. In this case, the usefulness of both comparison groups is clear. Comparing the 71-72 year olds to the younger group could be ambiguous because the younger group were bunching less as well.[25] On the other hand, after 1983 the 71-72 year olds look more like the older group than the younger as they did before. The reaction to the 1983 change shows, similarly to 1978, how the kink governs people's behavior.

The analysis concludes with Figure 4, focusing on the 1990 decline in the tax rate from 50% to 33% for 65-69 year olds. Figure 4-A shows the earnings of 63-64 and 67-69 year olds before 1990

---

bunching.

[24] Figure 3 shows 71-72 year olds, since they were 70-71 when the earnings were earned. 1982 is omitted because the change had been scheduled for 1982 but was postponed in 1981 by one year.

[25] 67-69 year olds might bunch less because the Delayed Retirement Credit was raised from 1% to 3% in 1982. However, earnings did not rise, and labor force participation declined instead of rising, which is not consistent with a response to the DRC. Reimers and Honig (1993, 1996) found no evidence of increased labor market re-entry by the affected age group.

and Figure 4-B shows earnings after 1990. In both, we see the familiar piling up at the kink. Yet, comparing the graphs, it is difficult to detect a change in the degree of bunching by the older group relative to the younger. Table 3 also shows the lack of a significant reaction. While this bounds the underlying responsiveness of labor supply, it does not contradict the evidence of strong reactions to the other changes. The 17-percentage point decline in the tax rate is smaller than earlier changes when the tax rate at the kink went from 50% to zero, and a small aggregate response is predicted from the elasticities estimated later.[26]

In sum, the visual and statistical analysis of earnings patterns demonstrates substantial bunching at the convex kink generated by the earnings test, along with considerable responsiveness to the shifts in the kink when the earnings test rules change. The next section will expand the analysis to consider the entire budget constraint.

III. PIECEWISE LINEAR BUDGET SET MODELLING

This section describes the econometric model of labor supply that arises out of utility maximization subject to a kinked budget constraint. The estimation will capture the bunching, together with reactions along other parts of the budget constraint which are difficult to detect in the raw data. This will permit broader conclusions about the earnings test while providing estimates of labor supply elasticities that are of more general interest.

A. Methodology

In estimating the determinants of labor supply, the problem of distinguishing the effect of the budget constraint on labor supply from the effect of preferences on the budget constraint come to the fore. Past labor supply may shape the budget constraint through the current wage, Social Security and other income, and may also be correlated with current labor supply. The resulting estimates of the impact of the current wage and non-labor income will be inconsistent. The substantial variation in the net wage and virtual income across age groups and over time caused by changes in the earnings test rules are key for identifying the estimated elasticities.

---

[26] Also, the smaller rule change may have been noted less widely by beneficiaries.

Most tax and transfer programs generate nonlinearities in the budget constraint and discontinuities in labor supply. Ignoring them mis-specifies labor supply, but using the net instead of the gross wage as the regressor makes the net wage endogenous because it is correlated with hours through the nonlinear tax schedule. Taking account of taxes in the estimation is not only important but also useful because policy changes are a source of variation that serve to identify the estimates. Accounting fully for the nonlinearities generally demands maximum likelihood techniques, introduced by Burtless and Hausman (1978) and outlined in Moffitt (1986).[27] This involves selecting a labor supply function, specifying the source of stochastic variation, and forming the likelihood function. The likelihood function takes into account the choice of hours over the entire exogenous tax schedule – in this way removing the endogeneity of jointly choosing hours and a tax rate on a particular segment.

However, the econometric application of the piecewise linear budget constraint method has been called into question by the work of MaCurdy, Green, and Paarsch (1990). They, and Pencavel (1986) earlier, showed that the probability of locating at a convex interior kink is positive, and the log likelihood is defined, only if the estimated coefficients yield a positive compensated substitution effect. When this condition was not satisfied, researchers imposed it by constraining the income coefficient to be negative. MaCurdy, Green, and Paarsch suggested further that the piecewise linear budget constraint method automatically imposes a positive compensated effect. Blomquist (1995) explained that this conclusion is not warranted. The compensated effect may be estimated to be positive without the researcher imposing it, and MaCurdy, Green, and Paarsch obtained estimates with a negative compensated effect, so it is not automatically imposed. The real trouble is with the estimate of a negative compensated effect, which has occurred in many econometric settings. This outcome casts doubt on either the theoretical underpinnings or the econometric implementation of the problem.[28]

The earnings test estimation described in the next section does not encounter such trouble. The compensated substitution effect is positive for all observations without imposing restrictions on the

---

27 Instrumental variables methods have been employed in some cases, but they do not deal with the kink observations, who face two different marginal tax rates and virtual incomes.

28 Two papers have focused on functional form. Blundell, Duncan, and Meghir (1998) excluded observations around the convex kink and added a selection term to account for it. That approach is unappealing here because the kink observations provide so much information. Blomquist and Newey (1996) estimated a nonparametric labor supply function subject to a parametric budget constraint. That approach requires major simplifications to be feasible. It omits covariates, only incorporates convex kinks, and

coefficients. The literature on kinked budget constraints suggests that the problems arise in the case of income taxes. We do not observe a person's precise tax schedule, which depends on other family income, household structure, and unobserved filing status and deductions.[29] This is exacerbated when there are many small kinks, as in the tax code before 1986 and 1981. In contrast, the earnings test creates a simple budget constraint with a large kink that has the same location for everyone and that appears well understood, which explains why we observe much more bunching. Finally, there is considerably more variation in elderly labor supply, which aids in identification.

These reasons explain why the earnings test estimation escapes the problems described by MaCurdy, Green, and Paarsch. The outcome here suggests the circumstances when the piecewise linear budget constraint method is most useful.

B. The Labor Supply Model

A beneficiary's choice of hours H is assumed to be determined by the linear function,

$$(1) \quad H(w, Y, X, \alpha) = \kappa + X\beta + \gamma w(1-\tau) + \delta Y_V + \alpha = Z\theta + \alpha$$

where w is the gross wage, w(1-τ) is the net wage, Y_V is virtual income, X are demographic characteristics that influence labor supply, and α is a random variable which represents unobserved heterogeneity in preferences.

The net wage and virtual income terms are defined in terms of the piecewise linear budget constraint generated by the earnings test. Let E denote the earnings exempt amount and Y_{SS} denote Social Security benefits. Then τ and Y_V are defined as follows:

(2) $\tau = \tau_1 \equiv 0$ if earnings < E;
$Y_V = Y_{V1} \equiv$ non-labor income

$\tau = \tau_2 \equiv \tau_{ET}$ if E < earnings < E + Y_{SS}/τ_{ET};
$Y_V = Y_{V2} \equiv Y_{V1} + E^*\tau_2$

$\tau = \tau_3 \equiv 0$ if E + Y_{SS}/τ_{ET} < earnings.
$Y_V = Y_{V3} \equiv Y_{V1} - Y_{SS}$

does not account for the participation decision, an important consideration for older workers.

29 Heckman (1983) pointed out that measurement error in the location of the kinks gets incorporated into the log likelihood and causes inconsistent estimates. Using Monte Carlo methods, Blomquist (1992) found that imputing income tax deductions causes

The structure in (2) makes apparent the correlation of the net wage and virtual income terms with hours as hours and earnings increase along the budget constraint. It is through (2) that the changes in the earnings test rules will be incorporated. While there may be concern that α is correlated with the gross wage and non-labor income, the identifying assumption is that the substantial shifts in the net wage and virtual income terms generated by the policy changes will dominate the potential endogeneity of α. The extent to which the policy variables drive the estimates will be demonstrated after the estimates are presented.

Assuming α in equation (1) is distributed over the population as $N(0,\sigma_\alpha^2)$ yields an expanded tobit. The log likelihood function in (3) expresses the probability of α occurring such that the person wishes to locate on the segment or kink where observed:

$$(3)\quad \log L(H_i) = S_{1i} * \log\left[\frac{1}{\sigma_\alpha}\,\phi\left(\frac{\alpha_i = H_i - Z_{1i}\theta}{\sigma_\alpha}\right)\right] \qquad \textit{(lower segment)}$$

$$+\ K_i * \log\left[\int_{H_i - Z_{1i}\theta}^{H_i - Z_{2i}\theta} \frac{1}{\sigma_\alpha}\,\phi\left(\frac{\alpha_i}{\sigma_\alpha}\right) d\alpha_i\right] \qquad \textit{(kink)}$$

$$+\ S_{2i} * \log\left[\frac{1}{\sigma_\alpha}\,\phi\left(\frac{\alpha_i = H_i - Z_{2i}\theta}{\sigma_\alpha}\right) * \Phi\left(\frac{\alpha'}{\sigma_\alpha}\right)\right] \qquad \textit{(middle segment)}$$

$$+\ S_{3i} * \log\left[\frac{1}{\sigma_\alpha}\,\phi\left(\frac{\alpha_i = H_i - Z_{3i}\theta}{\sigma_\alpha}\right) * \left\{1 - \Phi\left(\frac{\alpha'}{\sigma_\alpha}\right)\right\}\right] \qquad \textit{(upper segment)}$$

$$-\ \log\left[1 - \int_{-\infty}^{-Z_{1i}\theta} \frac{1}{\sigma_\alpha}\,\phi\left(\frac{\alpha_i}{\sigma_\alpha}\right) d\alpha_i\right] \qquad \textit{(truncation at zero hours)}$$

$$H_i - Z_{ji}\theta = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_j) - \delta Y_{V\,ji} \quad j = 1, 2, 3$$

ϕ (.) is the standard normal probability density function. Each element of (3), multiplied by an indicator for the budget segment where the person is located, expresses the likelihood that he chooses that segment conditional on his observed characteristics Z and on the budget parameters denoted by the subscript on Z. Desired hours are observed along the budget segments but not at the convex kink. The

severe bias in the piecewise linear method.

term for a person at the kink expresses the condition that desired hours are below the kink when he faces the budget parameters of the middle segment and above the kink when he faces the budget parameters of the lower segment, neither of which is feasible.[30]

The last term conditions the hours choice on the decision to work positive hours. This specification entails truncating nonworkers from the sample. Truncating focuses the estimation on the hours choices of those who work, instead of trying to explain the work decision too, which is far from trivial with so much of the sample retired. The strategy places stress on the useful policy variation, occurring well away from the zero hours point. Truncating is also practical because observed instead of imputed wages can be used.[31] Truncating might raise concern if people of different ages exhibit different trends in retirement, but there is little evidence of this.[32]

A few more details will make (3) operational. The major features of the tax code are incorporated. The estimation includes the payroll tax and the kink generated when people first enter the income tax system and face a tax rate of 11-15%. It is the only one of a comparable magnitude to the earnings test during this time period. Modeling all the additional small tax kinks would increase the demands of the type noted by MaCurdy, Green, and Paarsch, subjecting it to considerably more measurement error without adding meaningful information. Bunching at the initial tax kink is minor at best and bunching at any other kink is not discernible.[33]

A final issue involves assigning people to the kink. As Section II demonstrated, people are massed near the earnings test kink, but only some are located exactly on the kink, with many more in a small range below. One way that other researchers have handled this is by assigning observations from a small band on both sides of the kink to the kink itself. The earnings distributions showed no sign of

---

30 The terms involving α' account for the nonconvex kink that people avoid. α' is the value of α corresponding to the indifference curve tangent to both the middle and upper segments. This is discussed in more detail in the Appendix.

31 Heckman and MaCurdy (1981) pointed out that the usual sample selection procedure to impute the wage yields inconsistent estimates in the piecewise linear setting. The way to handle missing wages is to integrate over the stochastic variation in the imperfectly observed wage as well as over the other sources of randomness.

32 Labor force participation in the sample fell 2.9 percentage points before and after the 1983 earnings test change. Though it fell more for the affected 71-72 year olds than for the older and younger workers, the trends were similar.

33 The Appendix discusses the details. The nonconvex kink from the payroll tax ceiling on taxable earnings is not included, a minor detail in this sample with few high earners. Income taxation of benefits began in 1984 but is ignored here, as it affects few beneficiaries and only those with earnings well above the exempt amount, and it requires information about tax-exempt interest income which is unavailable. Taxes did not change differentially across ages, so a simplified treatment should not confound the identification strategy.

bunching just above the kink, however, so I assign observations occurring within $1000 below the kink to the kink itself.[34]

C. The Data

The model is estimated on older working men from the March Current Population Surveys. I focus on 66-75 year olds from the three years before and after the 1983 elimination of the earnings test for 70-71 year olds. The sample includes the affected group along with two comparison groups – younger men who face the same budget constraint before 1983, and older men who face the same budget constraint after. I also estimate the model on the three years before and after the 1978 increase in the exempt amount for 65-71 year olds, to check robustness. The reaction of the affected 65-71 year olds was strong relative to unaffected 63-64 year olds; however, only one natural comparison group is available.[35]

Statistics for the sample are reported in Table 4, and the variables are defined in the Appendix. 80% of the 1983 sample does not work. Within the subsample that works, 57.2% located below the earnings test kink, 8.9% on the kink, 16.6% on the middle segment above the kink, and 17.3% on the upper segment.[36] Their annuals hours will be explained as a function of their net wage, virtual income, and other personal characteristics.

---

[34] As discussed earlier, the interval length of $1000 is appropriate because some people round off their earnings to the nearest thousand. Burtless and Moffitt similarly assigned people within 100 annual hours on either side of the earnings test kink to the kink. Other papers without major bunching, such as those on income taxes, formalize measurement or optimization error to explain why individuals locate near, but not on, the kink. I tried estimating such a model, with observed hours equal to desired hours plus white noise, but it did not converge. Moffitt noted that the measurement error variance is identified by dispersion in bunching at the kink, which does not occur here.

[35] The availability of two comparison groups serves to identify the two elasticities of interest. Longitudinal data would be more useful, but the Retirement History Survey ended in 1979, the Survey of Income and Program Participation did not begin until 1984, and the Panel Study of Income Dynamics is too small. A weakness with annual cross-sectional data is that people claim benefits through the year and may switch to a different part of the budget constraint, but annual earnings and benefits determine where they are assigned. Previous research on retirement and the earnings test has generally used annual or biannual data as well. The timing of claiming is not well understood, and with this data it is impossible to avoid mis-assigning some people. However, this will not be systematically related to the identifying earnings test variation and is unlikely to involve the key observations on the kink.

[36] The samples exclude those with self-employment income, negative non-labor non-Social Security income, or a real wage not between $1 and $100, and those who earn less than the exempt amount but receive no benefits, since they do not claim benefits for a reason unrelated to the earnings test. Though this might introduce selection bias, the alternative of modeling benefit claiming

IV. RESULTS OF THE STRUCTURAL ESTIMATION

This section discusses the estimates of the model in equation (3). It follows by considering a number of specification checks that demonstrate the robustness of the results and the closeness of fit. It concludes with policy simulations based on the estimates.

A. Estimates of the Labor Supply Model

The estimates are reported in Table 5. The coefficients have the predicted signs, and the wage and income terms are estimated precisely. The estimated elasticities are large compared to many in the literature, especially for prime age males. Still, they follow from the strong reaction to the earnings test demonstrated earlier and also from the greater variation in hours among older workers. The 1983 estimates indicate that a $1 increase in the real wage would lead to an increase of 48 hours worked per year, implying an uncompensated wage elasticity of 0.316 at the sample means. The income coefficient indicates that a $1000 increase in non-labor income would lower hours worked by 17, yielding an income elasticity of -0.332 at the sample means.

The 1978 estimates are generally slightly smaller but similar in magnitude, a sign that the results are robust. In the rest of the discussion, I will focus on the 1983 results because they are based on more substantial identifying variation and because they relate directly to the proposed policy changes today. These involve changes in the earnings test for ages 65-69, and in particular eliminating the earnings test entirely, as occurred in 1983 for ages 70-71.

The large elasticity estimates imply significant deadweight loss imposed by the earnings test. Exact deadweight loss can be computed using the unique indirect and direct utility functions corresponding to the linear labor supply function, as in Hausman (1981). The 1983 estimates indicate average deadweight loss for people located on the kink of $1923 – without the earnings test, we would have to take away $1923 to make them as unhappy as they are with the earnings test. The average deadweight loss for people on the middle segment, facing a 50% tax rate, is $892. While their total utility loss is higher, worth $4603 on average, their deadweight loss is smaller because of the benefits

---

would be a stretch with the available data.

they forfeit to the earnings test.[37] However, it is important to recall that most of the benefits are returned over the long-run because of the actuarial adjustment and Delayed Retirement Credit, leaving only the deadweight loss. Thus, the earnings test imposes high efficiency costs, relative to shrinking fiscal gains.

The estimation controls for education, race, marital status, and age. People with at least a high school education and nonwhites work more. In 1978 married people work slightly more, and in 1983 they work slightly less.[38,39] Age has the usual negative effect on labor supply. To verify that the wage and income coefficients are not picking up nonlinear age effects on labor supply – which might occur because the policy variation is a function of age – I tried adding a quadratic in age. Including the quadratic does not change the wage and income coefficients. It makes both age terms individually insignificant and negative but jointly significant; and as the quadratic does not add appreciably to the explanatory power of the estimation, I left it out. I also tried including a linear time trend or a dummy for the years following the rule changes, to allow for a time effect on labor supply that might be confounded with the earnings test variation; the coefficients were insignificant in both samples.

The changes in the earnings test rules over time and across age groups are presumed to be the source of identifying variation. Our confidence in the estimates depends on the degree to which they are determined by the earnings test variation, as opposed to the potentially endogenous variation in wages and non-labor income. A series of exercises, summarized in Table 6, confirm the importance of the earnings test variation in generating the estimates.

First, estimating labor supply using OLS with the gross wage and other income – instead of the net wage and virtual income – yields insignificant coefficients with the wrong signs, so accounting for the shape of the budget constraint is crucial for obtaining sensible estimates. A similar point is demonstrated by including income and payroll taxes but not the earnings test. The estimation does not converge, so the remaining variation is insufficient to yield results. Estimating the model using non-labor income instead of virtual income in effect turns off the income variation from the shifts in the kinked budget

---

[37] People on the upper segment of the budget constraint experience no deadweight loss because they face only an income effect from losing all their benefits, but no substitution effect.

[38] The effects of education, race and marital status are not statistically different if freed up by age group or time.

[39] As discussed earlier, the earnings test treats spousal earnings differently if the spouse receives dependent or retiree benefits, which is not reported in the CPS. This might make the results sensitive to marital status, so I tried stratifying the sample. The estimates were almost identical, with substitution and income elasticities in the 1983 sample of 0.318 (0.016) and -0.328 (0.026)

constraint. The result is that the earnings test variation gets channeled into a much larger wage coefficient. The coefficient on non-labor income becomes less significant, so the variation due to the earnings test is key.

I also tried estimating the model for various subsets of the sample, eliminating some of the earnings test variation to gauge its importance. In the pre-1983 sample the estimation did not converge. In the other cases the estimates were less precise and somewhat, though not statistically, different. The comparison with 66-70 year olds has the strongest influence. Since they face the earnings test throughout, it suggests again its importance in the estimation. These results suggest that both the cross-age and time-series earnings test variation is important.

I tried another specification check allowing the Delayed Retirement Credit to mitigate the impact of the earnings test. I undertook this by freeing up the earnings test tax rate from both the net wage and virtual income terms and including an additional coefficient affecting the tax terms jointly. The coefficient can be viewed as a measure of tax sensitivity, which should fall below one if the Delayed Retirement Credit reduces the impact of the earnings test. The estimated coefficient did not drop below one, so the Delayed Retirement Credit does not appear to affect the response to the earnings test.

Lastly, to check the model's fit, Table 7 compares actual location on the budget constraint for 66-70 year olds to predicted location.[40] The model correctly predicted segment location for 85% of the group, a very close fit overall. The model also predicted hours relatively closely for the entire sample, and for those at or above the earnings test kink and being directly affected by the earnings test. This suggests that the estimates will reliably predict the response to other possible changes in the earnings test rules.

B. Policy Simulations

The labor supply estimates in Table 5 allow us to simulate the impact of other changes in the earnings test rules. The simulations also give a better sense about the estimates, since elasticity calculations pertain only to marginal changes away from the kinks. Table 8 reports simulations removing the earnings test for 65-69 year olds or raising their exempt amount to $30,000, which it will reach in

---

for married and 0.332 (0.040) and -0.331 (0.053) for single people.

2002. The first simulation simply repeats the 1983 change for another age group, while the second simulation amounts to doing so for many, so the estimates should be informative. Since the earnings test has been altered since 1983, the simulations are best compared to those identified as the Current Benchmark predictions, where the earnings test parameters have been updated to resemble the current rules. The Benchmark predicts mean hours of 1251 for all workers and 1782 for those located at or above the kink.[41]

Removing the earnings test would raise mean hours for people initially at or above the kink from 1782 to 1876, a 5.3% increase. The positive substitution effect from eliminating the earnings test dominates in the aggregate, therefore. The lower panel makes this more explicit, reporting the change in average earnings by initial location on the budget constraint. The substitution effect for those at the kink causes a 50% increase in average earnings, from $8,758 under the benchmark to $13,145 predicted without the earnings test. While the negative income effect grows along the middle segment, average earnings are still predicted to rise 18% from $18,600 to $21,983. The income effect for those initially on the upper segment leads them to work less and reduce their earnings 4%, from $43,892 to $42,128. Nevertheless, the overall effect on hours and earnings of eliminating the earnings test is strongly positive.

While the high initial cost is one deterrent to eliminating the earnings test, the medium run cost is small and declining.[42] The cost would gradually be offset because the Delayed Retirement Credit would not be granted to beneficiaries who otherwise get higher benefits later when they lose benefits to the earnings test today, as Honig and Reimers (1989) pointed out. At present, the cumulative fiscal cost of eliminating the earnings test is diminishing as the Delayed Retirement Credit is increased every other year. Once the Delayed Retirement Credit becomes fully actuarially fair for the average beneficiary, then the cost of eliminating the earnings test today will be virtually canceled out within several years.[43]

---

[40] Tables 7 and 8 focus on 66-70 year olds, who are the 65-69 year olds facing the earnings test throughout.

[41] In 1995 65-69 year olds faced a tax rate of 33% instead of 50% and an exempt amount about 20% higher in real terms. Separately lowering the tax rate or raising the exempt amount are each predicted to reduce average hours, but by less than one percent. Lowering the tax rate is predicted to induce a small increase in earnings (about 3%) even for those located at the kink, which is consistent with the lack of reaction in 1990 observed in Section II.

[42] Leonesio (1993) reported Social Security Administration forecasts that eliminating the earnings test for ages 65-69 would cost $4.3 billion in the first year. Increased tax payments would offset 14.8% of that. The forecasts are based on Hanoch and Honig's (1983) estimates of 0.17 for the uncompensated wage elasticity and virtually zero for the income elasticity. The elasticities estimated here imply a greater boost to labor supply and tax collections.

[43] The medium-run offset might not be complete due to adverse selection: the lower a person's life expectancy, the less

Lastly, Table 8 reports the results of simulating the increase in the exempt amount to $30,000. Interestingly, aggregate hours for those at or above the kink would be virtually unchanged. The breakdown of earnings by initial location gives more insight. Average earnings for those initially on the middle segment barely rise, compared to the substantial increase from eliminating the earnings test. Average earnings for those on the upper segment fall by 10%. They experience not only the negative income effect, as they would if the earnings test were eliminated, but also a negative substitution effect from facing the earnings test tax rate.[44] In sum, raising the earnings exempt amount makes the earnings test bind for a new group of people with higher earnings, who would lower their labor supply considerably.

VI. CONCLUSIONS

The earnings test has been the subject of substantial popular attention, but less academic interest in recent years. This paper revisits the evidence on the earnings test using more recent data and a new identification strategy. Several changes in the earnings test rules altered the budget constraint for beneficiaries of certain ages and not other ages. Comparing the reactions of beneficiaries before and after the changes, using the unaffected groups to control for other changes in labor supply, isolates the reaction to the earnings test. This strategy is implemented here using a combination of methods.

The first approach focuses on the strongest implication of the theory – that we should observe bunching at the convex kink induced by the earnings test. The data show that people respond to the earnings test by clustering at the kink in substantial numbers. The clustering moves when the kink moves and disappears when the earnings test is eliminated. The clustering is evidence that the earnings test leads some beneficiaries to hold down their labor supply.

The behavior around the exempt amount is the most noticeable but is not a complete picture of the earnings test's impact. Therefore, I also estimate a model of labor supply that characterizes

---

constrained they will be to postpone filing for benefits without the earnings test.

[44] This result in particular depends on the assumption that everyone can adjust their labor supply flexibly. However, while those near the kink can (and do) adjust their hours, others who work a lot may have less flexibility. If labor supply of those on the upper segment were completely inflexible, then the static response to raising the exempt amount would be less negative – and similarly, the response to eliminating the earnings test would be more strongly positive. On the other hand, if the incentives become great enough, full-time workers might choose to leave their jobs for part-time work – a stronger negative response.

behavior along the entire budget constraint. A typical criticism of this type of estimation, however, is that it overlooks identification of the parameter estimates. This is where the variation in the earnings test rules, causing major shifts in the budget constraint, is most important. The policy changes induce substantial variation in the right-hand side variables to identify the model's estimates. The resulting income and substitution elasticities are relatively large, implying considerable deadweight loss from taxation of older workers.

The estimates from a formal model also yield predictions about the effect of various policies to liberalize the earnings test. Thus, it was interesting to find that the model predicts a slight decline in aggregate labor supply among 65-69 year olds resulting from raising the exempt amount to $30,000, which will occur by 2002. The positive effect on hours worked by low earners gets offset by a negative effect for high earners. On the other hand, the model predicts that eliminating the earnings test would raise aggregate hours by 5.3%, with a major increase for people at and near the kink. The fiscal cost of such a measure would shrink over the medium run and is approaching zero, because increased benefits from the Delayed Retirement Credit would no longer be triggered as beneficiaries no longer lose benefits. Further savings might arise because workers age 65 and over are covered first by their employer-provided health insurance and only secondarily by Medicare. Finally, those who continue to work into their relative old age even with only moderate earnings will be able to save more for their eventual retirement. These results demonstrate the usefulness of the substantial changes in the earnings test rules in exposing underlying labor supply preferences.

The estimated elasticities also tell us about the labor supply of the working elderly more generally, a distinct focus compared to most previous research on retirement. The estimates imply that rules affecting hours choices conditional on working will influence the success of other policies that aim to encourage later retirement. The results also suggest a potentially severe negative effect on labor supply if benefits start to be means tested – a proposal that has gained attention in Congress. Another important consideration will involve 62-64 year olds. Compared to 65-69 year olds they face a much more restrictive earnings test, with a 50% tax rate and an exempt amount virtually unchanged in real terms since the mid-1970s. As the median retirement age continues to drop, the earnings test is likely to grow more binding on the labor supply of this age group over time. Furthermore, these tighter rules will gradually be extended to 65 and 66 year olds as the normal retirement age is raised beginning in 2000. In sum, the earnings test is an important consideration in understanding the decisions of the elderly who continue to work.

REFERENCES

Altonji, Joseph G., and Christina H. Paxson, "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-offs," *Journal of Labor Economics* 6 (1988), 254-276.

Berger, Jacqueline, "The Social Security Earnings Test and Labor Supply of the Elderly," unpublished Ph.D. thesis, Princeton University (1996).

Blinder, Alan S., Roger H. Gordon, and Donald E. Wise, "Reconsidering the Work Disincentive Effects of Social Security," *National Tax Journal* 33 (1980), 431-442.

Blomquist, Sören, "Estimation Methods for Male Labor Supply Functions: How to Take Account of Taxes," Working Paper 7, Department of Economics, Uppsala University (1992).

Blomquist, Sören, "Restrictions in Labor Supply Estimation: Is the MaCurdy Critique Correct?" *Economics Letters* 47 (1995), 229-235.

Blomquist, Sören, and Whitney Newey, "Nonparametric Estimation of Labor Supply Functions Generated by Piecewise Linear Budget Constraints," Working Paper 1997:24, Department of Economics, Uppsala University, Sweden (1997).

Blundell, Richard, Alan Duncan and Costas Meghir, "Estimating Labour Supply Responses Using Tax Reforms," *Econometrica* 66 (1998), 827-862.

Bondar, Joseph, "Beneficiaries Affected by the Annual Earnings Test, 1989," *Social Security Bulletin* 56 (1993), 20-28.

Burkhauser, Richard V., "The Early Acceptance of Social Security: An Asset Maximization Approach," *Industrial and Labor Relations Review* 33 (1980), 484-492.

Burtless, Gary, and Jerry Hausman, "The Effect of Taxes on Labor Supply," *Journal of Political Economy* 86 (1978), 1103-1130.

Burtless, Gary, and Robert A. Moffitt, "The Joint Choice of Retirement Age and Post-Retirement Hours of Work," *Journal of Labor Economics* 3 (1985), 209-236.

Card, David, "Labor Supply with a Minimum Hours Threshold," *Carnegie-Rochester Conference Series on Public Policy* 33 (1990), 137-168.

Cogan, John F., "Fixed Costs and Labor Supply," *Econometrica* 49 (1981), 945-963.

Coile, Courtney, Peter Diamond, Jonathan Gruber, and Alain Jousten, "Delays in Claiming Social Security Benefits," unpublished manuscript, Massachusetts Institute of Technology (1999).

Economic Report of the President (Washington, DC: United States Government Printing Office, 1999).

Friedberg, Leora, "The Social Security Earnings Test and Labor Supply of Older Men," in James Poterba (ed.), *Tax Policy and the Economy*, Volume 12 (Cambridge, MA: MIT Press, 1998).

Friedberg, Leora, "The Trend Towards Part-Time Work among Older Workers," unpublished manuscript, University of California, San Diego (1999).

Gustman, Alan L., and Thomas L. Steinmeier, "The 1983 Social Security Reforms and Labor Supply Adjustments of Older Individuals in the Long Run," *Journal of Labor Economics* 3 (1985), 237-253.

Gustman, Alan L., and Thomas L. Steinmeier, "A Structural Retirement Model," *Econometrica* 54 (1986), 555-584.

Gustman, Alan L., and Thomas L. Steinmeier, "Changing the Social Security Rules for Work After 65," *Industrial and Labor Relations Review* 44 (1991), 733-745.

Hanoch, Giora, and Marjorie Honig, "Retirement, Wages, and Labor Supply of the Elderly," *Journal of Labor Economics* 1 (1983), 131-151.

Hausman, Jerry A., "The Effect of Wages, Taxes, and Fixed Costs on Women's Labor Force Participation," *Journal of Public Economics* 14 (1980), 161-194.

Hausman, Jerry A., "Exact Consumer's Surplus and Deadweight Loss," *American Economic Review* 71 (1981), 662-676.

Heckman, James J., "Comment," in Martin Feldstein (ed.), *Behavioral Simulation Methods in Tax Policy Analysis* (Chicago: The University of Chicago Press, 1983).

Heckman, James J., and Thomas E. MaCurdy, "New Methods for Estimating Labor Supply Function: A Survey," in Ronald Ehrenberg (ed.), *Research in Labor Economics*, Volume 4 (Greenwich, CT: JAI Press Inc., 1981).

Honig, Marjorie, and Cordelia Reimers, "Is It Worth Eliminating the Earnings Test?" *American Economic Association Papers and Proceedings* 79 (1989), 103-107.

Kristof, Kathy, "Personal Finance," *Los Angeles Times* (September 5, 1997), D5.

Krueger, Alan B., and Jörn-Steffen Pischke, "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation," *Journal of Labor Economics* 10 (1992), 412-437.

Leonesio, Michael V., "Effects of the Social Security Earnings Test on the Labor Market Activity of Older Americans: A Review of the Evidence," *Social Security Bulletin* 53 (1990), 2-21.

Lingg, Barbara A., "Beneficiaries Affected by the Annual Earnings Test in 1982," *Social Security Bulletin* 49 (1986), 25-32.

MaCurdy, Thomas, "Work Disincentive Effects of Taxes: A Reexamination of Some Evidence," *American Economic Association Papers and Proceedings* 82 (1992), 243-249.

MaCurdy, Thomas, David Green, and Harry Paarsch, "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *Journal of Human Resources* 25 (1990), 415-490.

Meyer, Bruce, "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics* 13 (1995), 151-162.

Moffitt, Robert, "The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method," *Journal of Business and Economic Statistics* 4 (1986), 317-328.

National Center for Policy Analysis, "Tax Fairness for the Elderly: Eliminating the Social Security Earnings Penalty," Policy Backgrounder No. 137, Dallas, TX, (1995).

Packard, Michael D., "The Earnings Test and the Short-Run Work Response to Its Elimination," *Social Security Bulletin* 53 (1990), 2-16.

Pencavel, John, "Labor Supply of Men," in Orley Ashenfelter and Richard Layard (eds.), *Handbook of Labor Economics*, Volume 1 (Amsterdam: North-Holland, 1986).

Reimers, Cordelia, and Marjorie Honig, "The Perceived Budget Constraint under Social Security: Evidence from Reentry Behavior," *Journal of Labor Economics* 11(1993), 184-204.

Reimers, Cordelia, and Marjorie Honig, "Responses to Social Security by Men and Women: Myopic and Far-Sighted Behavior," *Journal of Human Resources* 31 (1996), 359-382.

Robbins, Aldona, and Gary Robbins, *Paying People Not to Work: The Economic Costs of the Social Security Retirement Earnings Limit*, National Center for Policy Analysis Policy Report No.142 (1989).

Ruhm, Christopher, "Bridge Jobs and Partial Retirement." *Journal of Labor Economics* 8 (1990), 482-501.

Simon, Ruth, "How To Be Sure You Never Go Broke," *Money* 25 (October 1996), 100-114.

Vroman, Wayne, "Some Economic Effects of the Retirement Test," in Ronald Ehrenberg (ed.), *Research in Labor Economics*, Volume 7 (Greenwich, CT: JAI Press Inc., 1985).

</ref_section>

APPENDIX

Data

The data are extracted from March Current Population Surveys. The variables and tax parameters are defined as follows:

Variable	*Definition*
earnings	wage and salary income plus self-employment income earned last year
hours	usual hours of work per week last year times weeks worked last year
completed high school	dummy variable, completed 12 years of education or more
married	dummy variable, married or separated
nonwhite	dummy variable, race is black or other
wage	real gross hourly wage: earnings divided by hours, deflated by the CPI to 1987 dollars
other income	real non-labor non-Social Security income: total family income minus individual earnings and Social Security benefits, deflated by the CPI to 1987 dollars
virtual income	intercept of each budget constraint segment: defined in equation (2) as a function of other income, Social Security benefits, and the earnings test
initial tax kink	maximum of zero or the standard deduction plus the personal exemption (times two if married, times two if over age 65) minus other income
initial tax rate	marginal federal income tax rate applying to the first tax bracket
payroll tax rate	employee's share of the OASHDI payroll tax

The Log Likelihood Function

The nonconvex kink. The log likelihood function in (3) includes terms involving α'. α' refers to the value of α, the heterogeneity error term defined in (1), that sets equal the indirect utility functions on the middle and upper segments. The log likelihood specifies that someone locates on the middle segment if their value of α is smaller than α' and locates on the upper segment if their value of α is greater.

The indirect utility function. The explicit indirect utility function is required for the estimation and for simulations. From Hausman (1981), the indirect utility function that corresponds to the linear labor supply function is the following:

$$v(w, y) = e^{\delta w}\left(y + \frac{\gamma}{\delta}w - \frac{\gamma}{\delta^2} + \frac{\kappa + X\beta + \alpha}{\delta}\right)$$

Taxes. A simplified version of the income tax system is incorporated into the estimation, although it is not specified in (3). Payroll taxes are assumed to apply to all earnings. The lowest federal income tax rate is also incorporated, along with the kink corresponding to that tax bracket for people with low income, who do not face income taxes for the first hour of work. In some cases the location of the initial tax kink and the earnings test kink coincide within $1000 (8% of the sample) or are reversed (10% of the sample).

This could raise the concern that the location of the tax kink might interfere with the analysis of bunching at the earnings test kink. However, since the income tax treatment is the same across age groups, comparisons will difference out the effect of the tax kink. Furthermore, the data suggest that the income tax kink is not very important. There does not appear to be significant bunching at the initial tax kink: the number of people within $1000 below the initial tax kink is less than one-third of the number within $1000 below the earnings test kink and causes no visible or statistically significant spike in the earnings distribution. Moreover, the kinks coincide for only a small proportion of the sample. Finally, 71-72 year olds in particular exhibit no noticeable bunching at the initial tax kink after 1983, which serves as a check on their pre-1983 behavior when this might be of concern. Therefore, the reaction to the tax kink does not appear to interfere with the earnings test analysis.

Those who do not face the earnings test. The log likelihood for older people who are exempted from the earnings test omits the terms specific to the earnings test. Those with high non-labor income face a single budget segment and pay income and FICA taxes, and those with low non-labor income face the income tax kink.

TABLE 1

The Earnings Test Rules in 1998

	62-64 year olds	65-69 year olds
Exempt amount	$9,120	$14,500
Tax rate	50%	33%

Changes in the Earnings Test Rules

Year	What changed	Ages affected	Ages not affected
1978	Raised exempt amount about 25%	65-71	62-64
1983	Eliminated the earnings test	70-71	62-69
1990	Lowered tax rate to 33%	65-69	62-64
1996	Raised exempt amount to $30,000 by 2002	65-69	62-64
2000+	Rules for 62-64 year olds will be extended as the normal retirement age rises from 65 to 67.	65-66	62-64, 67-69

TABLE 2

Individuals Affected by the Earnings Test, 1989

Aged 65-69	
Number of retired worker beneficiaries [b]	7,229,512
who did not work [a]	5,253,500
who worked and	
had benefits withheld [b]	757,560
had earnings within 90-110% of exempt amount [a]	173,700
Number who had not claimed benefits and still worked [a]	582,000
Aged 62-64	
Number of retired worker beneficiaries [b]	2,549,084
who worked and had benefits withheld [b]	168,782
Total, retired worker beneficiaries who worked and had benefits withheld	926,342
Total, dependents and survivors who worked and had benefits withheld [b]	314,938

[a] Leonesio (1990).

[b] Bondar (1993).

TABLE 3

Percentage with Earnings in Each Earnings Intervals

1978 Change	63-64 year olds			67-69 year olds		
	1975-77	1978-81		1975-77	1979-81	
Earnings intervals		62-64 kink	65-71 kink		62-64 kink	65-71 kink
-$3000 to -$2000	0.035	0.022	0.027	0.043	0.027	0.029
-$2000 to -$1000	0.036	0.029	0.034	0.038	0.029	0.034
-$1000 to kink	0.045	0.036	0.018	0.058	0.033	0.045
kink to +$1000	0.021	0.019	0.015	0.018	0.037	0.014
+$1000 to +$2000	0.017	0.015	0.012	0.007	0.028	0.008
+$2000 to +$3000	0.016	0.013	0.014	0.007	0.010	0.010
Difference at kink	0.024	0.016	0.003	0.041	-0.004	0.030
	(0.006)	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)

1983 Change	67-69 year olds		71-72 year olds		73-75 year olds	
	1980-81	1984-86	1980-81	1984-86	1980-81	1984-86
-$3000 to -$2000	0.027	0.016	0.028	0.011	0.022	0.010
-$2000 to -$1000	0.031	0.021	0.019	0.009	0.018	0.010
-$1000 to kink	0.041	0.022	0.023	0.012	0.013	0.008
kink to +$1000	0.013	0.008	0.009	0.005	0.008	0.007
+$1000 to +$2000	0.007	0.006	0.005	0.004	0.006	0.003
+$2000 to +$3000	0.012	0.006	0.010	0.005	0.010	0.004
Difference at kink	0.028	0.016	0.014	0.006	0.005	0.001
	(0.006)	(0.003)	(0.006)	(0.004)	(0.004)	(0.003)

1990 Change	63-64 year olds			67-69 year olds		
	1988-89	1991-92		1988-89	1991-92	
-$3000 to -$2000	0.020	0.016		0.014	0.015	
-$2000 to -$1000	0.023	0.014		0.016	0.010	
-$1000 to kink	0.027	0.024		0.022	0.018	
kink to +$1000	0.012	0.015		0.009	0.006	
+$1000 to +$2000	0.011	0.009		0.010	0.007	
+$2000 to +$3000	0.007	0.012		0.004	0.008	
Difference at kink	0.015	0.008		0.012	0.012	
	(0.006)	(0.006)		(0.005)	(0.004)	

Cells show the percentage of men who have earnings in each $1000 interval, defined relative to the earnings test kink. Data from the March Current Population Surveys, details in the Appendix. Standard errors are in parentheses or, where omitted, they are 0.001 to 0.004.

TABLE 4

Summary Statistics

	1978 rule change		1983 rule change	
	all	working	all	working
Number of observations	20,647	6,663	23,889	4,876
Number aged: 63-64 *	6,144	3,331	-	-
66-70 *	10,246	2,565	13,453	3,435
71-72 *	4,257	767	4,544	676
73-75 *	-	-	5,892	765
Completed high school	0.425	0.518	0.471	0.579
Nonwhite	0.102	0.094	0.097	0.093
Married	0.816	0.864	0.811	0.842
Annual hours	478 (850)	1473 (873)	237 (604)	1162 (844)
Gross hourly wage	-	11.20 (9.54)	-	10.68 (10.34)
Net hourly wage	-	7.58 (7.03)	-	7.69 (8.04)
Non-labor income	17,658 (15,503)	15,546 (15,796)	21,052 (17,773)	20,802 (19,131)
Social Security income	4,595 (3,083)	2,884 (3,192)	6,080 (2,692)	5,201 (3,339)
Location on the budget set: zero hours	0.675	-	0.796	-
below earnings test kink	0.076	0.234	0.117	0.572
earnings test kink	0.045	0.137	0.018	0.089
above earnings test kink, benefits>0	0.053	0.163	0.034	0.166
above earnings test kink, benefits=0	0.152	0.466	0.035	0.173

Men, March Current Population Surveys of 1976-78 and 1979-81 for 1978, 1981-83 and 1984-87 for 1983. Income data are retrospective for the previous year and are deflated using the CPI to 1987 dollars. The sample excludes those with self-employment income or negative non-labor income non-Social Security income; earnings less than the exempt amount but without Social Security income; or a real wage less than 1 or more than 100. Standard deviations are in parentheses. Variable definitions are reported in the Appendix.
* Age is reported in March following the working year. For example, 70-71 year olds (for whom the earnings test was eliminated in 1983) are approximately 71-72 in the CPS.

TABLE 5

Maximum Likelihood Estimates

Dependent Variable: Annual Hours	1978 rule change		1983 rule change	
Completed high school	259		577	
	(29)		(67)	
Nonwhite	177		362	
	(47)		(91)	
Married	119		-79	
	(39)		(40)	
Age-65	-168		-156	
	(49)		(12)	
Net wage	44		48	
	(2)		(3)	
Virtual income, $1000	-25		-17	
	(1)		(2)	
Constant	1603		963	
	(40)		(92)	
Standard deviation, σ_α	1276		1356	
	(10)		(29)	
Uncompensated wage elasticity *	0.225 (0.013)		0.316 (0.021)	
Income elasticity *	-0.297 (0.013)		-0.332 (0.044)	
Average annual hours	1473		1161	
Log likelihood	-26369		-17726	
Number of observations	6663		4876	
Age of sample	63-64, 67-72		66-75	
	Deadweight loss **	vs. reduction in benefits	Deadweight loss **	vs. reduction in benefits
Average for individuals located…				
at the kink	1610	0	1923	0
on middle segment (losing benefits)	234	3814	892	3711

Maximum likelihood estimates of the choice of hours conditional on working, as modeled in equation (3). Asymptotic standard errors are in parentheses. Sample: working men from the March CPSs three years before and after each rule change. Sample statistics and selection are reported in Table 4. * Calculated at the sample means of 1162 hours, 7.69 net wage, 22.449 virtual income for 1983; and 1473, 7.58, 17.534 for 1978. ** The indirect and direct utility functions from Hausman (1981) are used to compute each person's equivalent variation, the amount of income he would have to forfeit to keep utility unchanged if the earnings test were removed. Equivalent variation minus benefits lost to the earnings test equals deadweight loss.

TABLE 6

The Importance of the Change in the Earnings Test Rules

1983 rule change	Coefficient on wage	Coefficient on income
Model estimates (Table 5)	48	-17
	(3)	(2)
What if the model is estimated...		
using OLS, no earnings test or tax parameters	insignificant, opposite sign	
with income taxes, no earnings test	estimation does not converge	
with earnings test, but using non-labor income instead of virtual income	77	-14
	(3)	(7)
pre-1983 sample only	estimation does not converge	
post-1983 sample only	37	-13
	(10)	(9)
affected 71-72 year olds only	54	-23
	(12)	(11)
affected 71-72 year olds and 73-75 year olds	64	-29
	(10)	(10)
affected 71-72 year olds and 66-70 year olds	45	-15
	(3)	(4)

Variations on the maximum likelihood estimates reported in Table 5.

TABLE 7

The Within-Sample Fit of the Model

*66-70 year olds (based on 1983 rule change) **	Actual	Predicted
Number of individuals located on:		
budget constraint below earnings test kink**	1521	1635
kink	389	293
middle segment	717	659
upper segment	808	848
Average hours		
all 66-70 year olds	1224	1254
66-70 at or above kink	1738	1786

Predictions based on Table 5 estimates. α was calculated for those not at the kinks and drawn from a truncated normal distribution for those at the kinks. This was used to compute desired hours, earnings and utility on each segment. The hours choice is the feasible hours giving the highest utility. Hours were then averaged over all observations.

* Recall that those who are 66-70 in the sample were 65-69 when their earnings occurred. ** For those with relatively little non-labor income, this comprises two segments and the initial income tax kink.

TABLE 8

Policy Predictions

66-70 year olds (based on 1983 rule change)	Average hours All	 At or above kink
Actual	1224	1738
Predicted	1254	1786
Policy simulations		
Current Benchmark: exempt amount up 20%, tax rate 33%	1251	1782
Earnings test eliminated	1300	1876
Change from benchmark		*5.3% increase*
Exempt amount raised to $30,000	1249	1778
Change from benchmark		*0.2% decrease*

	Predicted average earnings (according to initial segment of the budget constraint)			
	Lower	Kink	Middle	Upper
Current Benchmark	$3,757	$8,758	$18,600	$43,892
Earnings test eliminated	$3,757	$13,145	$21,983	$42,128
Exempt amount raised to $30,000	$3,757	$11,743	$19,801	$39,457

Predictions based on the estimates in Table 5 and calculations described in the notes to Table 7.

FIGURE 1: The Earnings Test

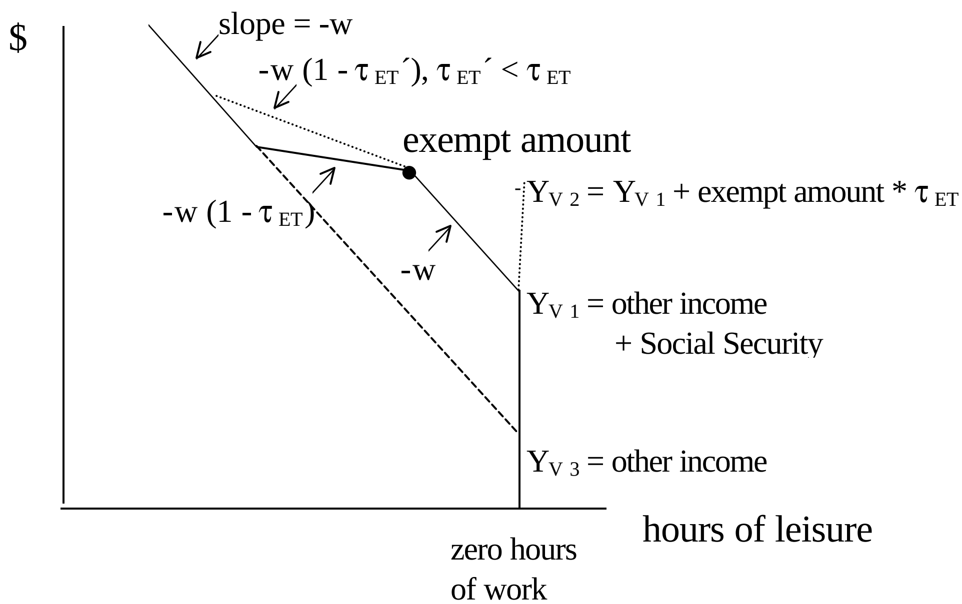

FIGURE 2-A
1975-77

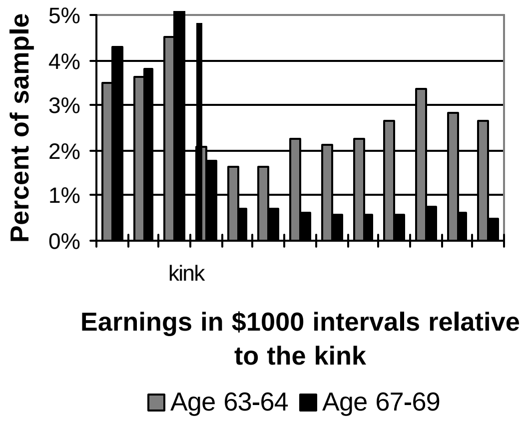

FIGURE 2-B
1979-81

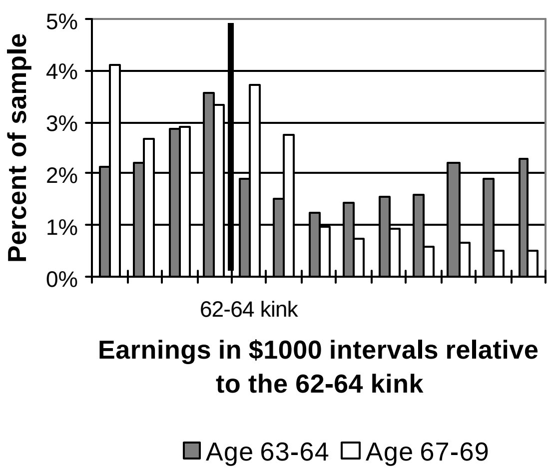

FIGURE 2-C
1979-81

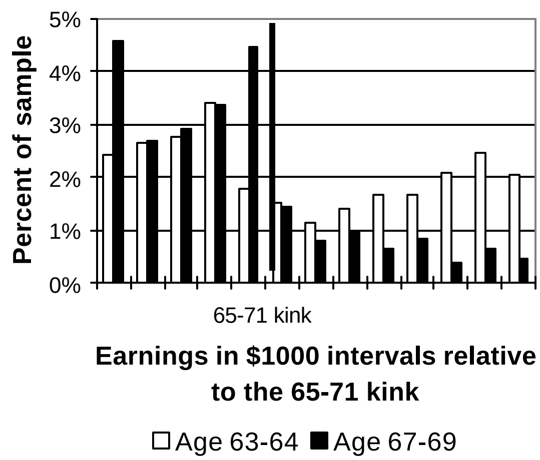

FIGURE 3-A

1980-81

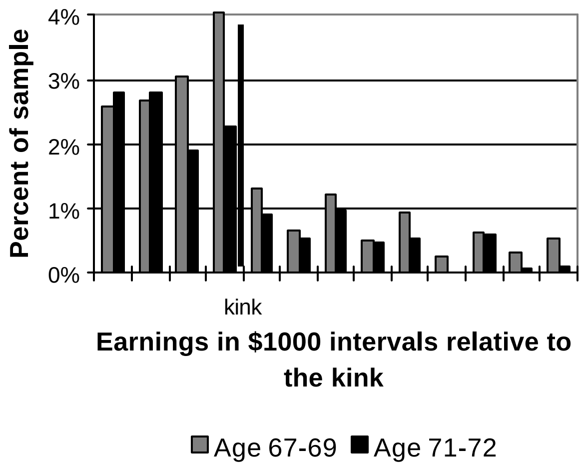

FIGURE 3-B

1980-81

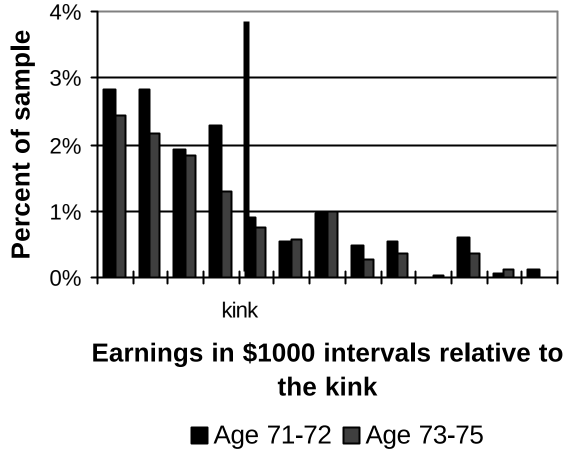

FIGURE 3-C

1984-86

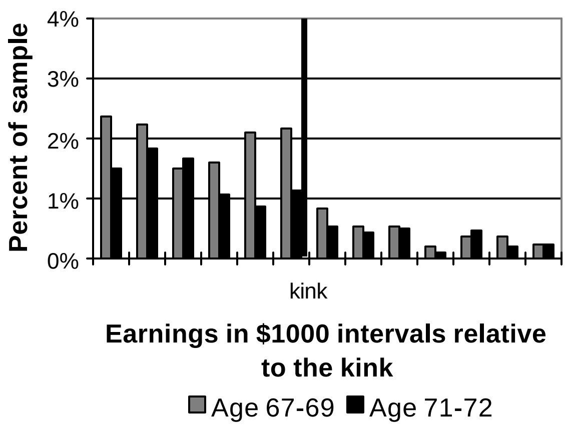

FIGURE 3-D

1984-86

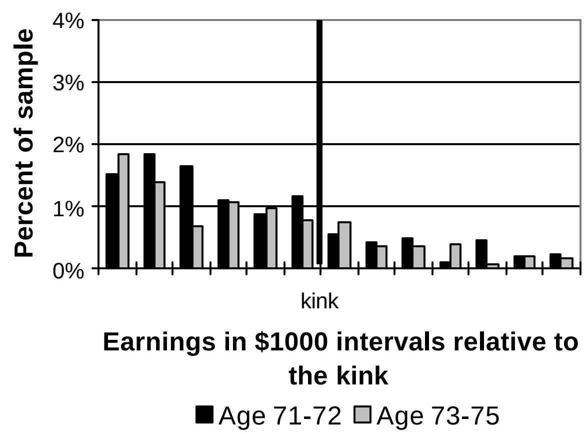

FIGURE 4-A

1988-89

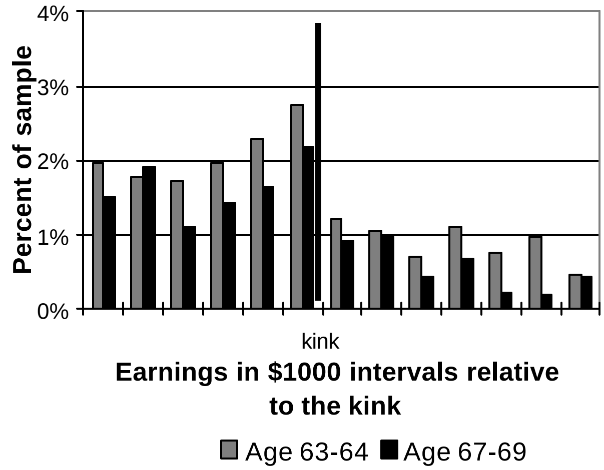

FIGURE 4-B

1991-92