The Effect of Government Programs on the Labor Supply of the Elderly

by

Leora Friedberg

B.A. in Economics, The Johns Hopkins University, 1988

Submitted to the Department of Economics
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 1996

Signature of Author . . . .
Department of Economics
May 17, 1996

Certified by . . . .
James Poterba
Professor of Economics
Thesis Supervisor

Certified by . . . .
Jerry Hausman
John and Jennie S. MacDonald Professor of Economics
Thesis Supervisor

Accepted by . . . .
Richard S. Eckaus
Chair, Departmental Committee on Graduate Students

The Effect of Government Programs on the Labor Supply of the Elderly

by

Leora Friedberg

Submitted to the Department of Economics on May 17, 1996
in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Economics

ABSTRACT

Two chapters of this dissertation focus on elderly labor supply and how it has been influenced by means or earnings tested government transfers. The third chapter studies how divorce behavior responded to changes in state divorce laws. All three chapters use the strategy of studying the response to changes in rules as a way of identifying how individuals make decisions.

Chapter 1 examines the Social Security earnings test, which reduces benefits when a beneficiary works. The marginal tax rates of 33-50% for earnings above an exempt amount are among the economy's highest. I examine three changes in the earnings test rules to identify its impact. First, I show that many individuals bunch their earnings just below the exempt amount, evidence of reduced labor supply. Then, I model the earnings test formally and estimate sizable elasticities. The elasticities imply substantial deadweight loss and suggest that eliminating the earnings test would raise labor supply by 14%, with minimal fiscal cost.

Chapter 2 analyzes the impact of the first transfer program for the elderly on retirement. Economists have extensively studied rising retirement rates without reaching a consensus, focusing mostly on the role of Social Security since the 1960s. Old Age Assistance was introduced in 1935 and dwarfed Social Security at least until 1950. Retirement rates began to increase during that same period. Moreover, states determined benefit levels, inducing cross-sectional variation that is lacking for Social Security to identify labor supply estimates. Using 1940 and 1950 Census data, I find that Old Age Assistance contributed significantly to the early increase in retirement.

Chapter 3 revisits evidence on the impact of unilateral divorce laws on divorce. Most states switched from requiring mutual consent to allowing unilateral or no-fault divorce between 1970 and 1985, while the divorce rate more than doubled after 1965. According to the Coase theorem, the legal shift should not affect divorce rates. I employ a panel of state divorce rates, which controls in a very flexible way for unobservables that influence divorce and vary across states and over time. This approach reveals a strong effect of divorce laws -- switching to unilateral divorce explained 16% of the increased divorce rate.

Thesis Supervisor: James Poterba
Title: Professor of Economics

Thesis Supervisor: Jerry Hausman
Title: Professor of Economics

ACKNOWLEDGEMENTS

Choosing MIT was one of the better decisions I have made. I have had a wonderful time here. The work was very hard, and I experienced the usual moments of self-doubt. However, more resources were available here at MIT to combat those moments than anywhere else I can imagine -- professors who were accessible, encouraging, and honest, and classmates who were the same, and who were not afraid to express their doubts either.

This thesis would not be what it is, and I would not have gotten such a marvelous job, without the crucial support of my thesis advisors. In particular, Jim Poterba is a model teacher and researcher, and advisor too. Jerry Hausman and Jon Gruber lent their expertise when I most needed it, reading my papers and thinking about my ideas with enthusiasm and honesty. Comments from Josh Angrist, Dora Costa, Peter Diamond, Whitney Newey and Steve Pischke are also appreciated.

Advisors are useful, but my classmates made MIT fun. Vandy Howell, Mary Hallward, Todd Sinai and I formed a study group that got me through the first two years of classes. In addition, Jason Abrevaya, Matt Barmack, Patrizia Canziani, Harrison Hong, Jeff Kubik, Peter Sørensen, Karl Whelan, Catherine Wolfram, Jennifer Wu and especially Robert Shimer made the hours spent in the computer room, office and pub much more fun. Aaron Yelowitz showed me with his example and his advice how to do interesting research. And special mention to Gary King, who knows everything about everything at MIT and shares it with the utmost courtesy. Last but not least, thanks to João Ejarque, who has been my best friend and who has believed in me always.

None of the past four years would have been possible without financial support from the Jacob K. Javits Fellowship, the National Institute on Aging, and MIT.

Many individuals helped me on the road to MIT. Carl Christ, Bruce Hamilton and Lou Maccini introduced me to economics at Johns Hopkins with excellent results. Later, John Page gave me a great job at the World Bank and told me when it was time to go to grad school.

Most importantly, my parents and sister have given me constant love, support and encouragement. Thanks.

TABLE OF CONTENTS

INTRODUCTION

Retirement rates have been on the rise for several decades. According to Ransom and Sutch (1986), the labor force participation rate of men aged 60 and over fell from 64.5% in 1930 to 32.2% in 1980. Combined with declining savings rates, this has placed a severe burden on our public resources, even as the elderly now become the fastest growing segment of our population. Economists have been trying many years to understand the role of means and earnings tested government programs transfer programs in contributing to the decline in labor supply, but with mixed success. A major problem arises because transfer levels are determined as a function of current or past earnings, which are most likely correlated with preferences for work and result in inconsistent estimates of the impact of benefits on current labor supply. My strategy in this dissertation is to understand individuals' behavior by analyzing how they respond to changes in the rules and laws which govern benefits. This approach yields evidence of considerable responsiveness to means and earnings tested transfers.

Chapter 1 analyzes the impact on labor supply of the Social Security earnings test, which reduces benefits when a beneficiary works. In 1995, a beneficiary under the age of 65 who earned more than $8,160 lost $1 in benefits for every $2 in additional earnings -- which functions as a 50% tax on wages. A 65-69 year old lost benefits at a 33% rate for earnings above $11,280. The marginal tax rates are among the highest in the economy. Concern over the distortions imposed by the earnings test led Congress to raise the exempt amount in 1996 gradually to $30,000 by 2002. Nevertheless, the existing literature argues that the earnings test has a small impact on the labor supply of the elderly. The principal papers in this literature by Burtless and Moffitt (1985) and by Gustman and Steinmeier (1986) used the Retirement History Survey, which ended in 1979 and has become somewhat dated. More importantly, there were virtually no changes in the earnings test during the time period the Retirement History Survey covered, so estimates of the earnings test's effect are identified primarily from potentially endogenous cross-sectional variation in wages.

My approach is to take advantage of three changes in the earnings test rules over the last twenty years in order to identify its impact. Each change involved a different combination of changing the exempt amount and the tax rate, and each applied to some age groups and not others -- which make the changes extremely useful for identifying the effect of tax rules on the labor supply of working beneficiaries.

The empirical analysis begins by examining the most straightforward prediction: beneficiaries should bunch just below the exempt amount, although bunching is not observed in many applications of kinked budget sets. Using the Current Population Survey, I find substantial bunching of earnings just below the exempt amount, evidence that some individuals are reducing their labor supply. The bunching moves when the exempt amount moves and smooths out when the earnings test was eliminated for one age group. Then, I model formally how the earnings test affects labor supply, which decomposes the reactions to the earnings test along the entire budget set into income and substitution elasticities. This approach, again, relies on the substantial variation in the budget set induced by the changes in the earnings test rules to identify the elasticities. Estimating the model using maximum likelihood methods yields significant elasticities that suggest substantial deadweight loss suffered by many working beneficiaries. The estimates imply that eliminating the earnings test would raise labor supply by 14% for 65-69 year olds currently affected by the earnings test, and at a minimal fiscal cost. Thus, eliminating the earnings test will raise the amount of resources available to those who continue to work during relative old age for their post-work years.

Chapter 2 studies the influence of the first means-tested transfer program for the elderly on retirement. Most of the attention in the literature on declining participation rates has gone to analyzing retirement since the 1960s and to explaining the role of Social Security, with mixed findings, and, more recently, the role of private pensions. However, while Social Security and pensions are the main source of income for the elderly today, another public transfer program dwarfed Social Security before 1950, as Parsons (1991) pointed out. Old Age Assistance (OAA) was a means-tested program for those 65 and older established by the same legislation that created Social Security in 1935. By 1940, 22% of the aged population was receiving OAA. The size of annual Social Security transfers only began to exceed that of OAA during the 1950s. Thus, the heyday of OAA coincides with the early stages of the modern decline in the elderly's labor force participation and may have played a significant role.

Another distinctive feature of OAA was that benefit levels were determined by the states. This is an advantage for studying OAA compared to Social Security because differences in benefit generosity -- for instance, changes in benefits within a state over time -- can be associated with the corresponding changes in labor supply. In constrast, because Social Security is a national program, all individuals face the same rules at a point in time. Variation across individuals in Social Security benefits arises because of past earnings histories, which may well be correlated with present labor supply and thus bias estimates of the influence of Social Security. Studying OAA may be one of the best opportunities for identifying how income flows for the elderly that are tied to low earnings or to retirement affect labor supply, in addition to giving us insight into early retirement trends.

Earlier work in this area by Parsons employed state-level aggregate data. In this chapter, I estimate the relationship between state benefit levels and the labor force participation of the elderly using individual data from the Censuses of 1940 and 1950. In this way, I can include individual as well as state level characteristics to explain the participation decision, and I can allow the effect of OAA to vary across individuals according to those characteristics. I find that allowing the impact of OAA to vary, especially by education level and by household structure, reveals important aspects about how OAA affected participation. Additional results point to a stronger effect of OAA when it is identified from changes within a state over time in OAA generosity, and when state per capita income is included. With state income and OAA benefits positively correlated, the measured effect of OAA had been picking up part of the positive effect of state income on participation when income was omitted. The estimates demonstrate that OAA accounted for as much as 60% of the increase in retirement rates between 1940 and 1950.

Chapter 3 focuses on divorce, a very different subject, but employs the same strategy as the first two chapters by studying the response of individual behavior to variation in divorce laws. Most states switched from requiring mutual consent to allowing unilateral or no-fault divorce between 1970 and 1985, while the national divorce rate more than doubled after 1965. Researchers have found conflicting theoretical and empirical evidence on whether the legal changes raised divorce rates. The question has arisen again in 1996 as several states consider whether to tighten divorce laws, while some opponents argue that the law does not affect the likelihood of divorce. In Chapter 3 I find that divorce rates rose significantly in response to the earlier wave of liberalization.

In the previous literature, Peters (1986) and Becker (1981) argued that the legal shift should have had no effect on divorce rates as an application of the Coase theorem. Using the same cross-sectional data, Peters (1986, 1992) found no effect and Allen (1992) estimated a significant effect. Their differences centered on controls for state-level heterogeneity in divorce behavior.

In this chapter I use a panel of state-level divorce rates, which includes virtually every divorce in the U.S. over the entire period of the law changes and also permits thorough controls for differences across states in divorce behavior which may be correlated with the divorce law. This turns out to be crucial to the results. Employing an extremely flexible parameterization for the unobservables that govern divorce behavior and that vary across states and over time reveals a strong effect of the divorce law. The results indicate that switching to unilateral divorce raised a state's divorce rate by about 7%, accounting for about 16% of the increase in the U.S. divorce rate between 1970 and 1985.

The papers in this dissertation highlight the usefulness of changes in rules and laws for identifying estimates of individual behavioral responses. This approach yields new evidence about policy issues that have been debated for years. The first two chapters found considerable responsiveness of older workers to tax and transfer rules, which has important policy implications as we try to structure public policy to deal with an aging population. The third chapter concluded that switching from mutual to unilateral divorce raised divorce rates, as several states debate whether to repeal unilateral and no-fault divorce laws.

REFERENCES

Allen, Douglas, W. 1992. "Marriage and Divorce: Comment." *American Economic Review* 82: 679-85.

Becker, Gary S. 1981. *A Treatise on the Family*, Cambridge: Harvard University Press.

Burtless, Gary, and Robert A. Moffitt. 1985. "The Joint Choice of Retirement Age and Post-Retirement Hours of Work." *Journal of Labor Economics* 3: 209-36.

Gustman, Alan L., and Thomas L. Steinmeier. 1986. "A Structural Retirement Model." *Econometrica* 54: 555-84.

Parsons, Donald O. 1991. "Male Retirement Behavior in the United States, 1930-1950." *Journal of Economic History* 51: 657-74.

Peters, H. Elizabeth. 1986. "Marriage and Divorce: Informational Constraints and Private Contracting." *American Economic Review* 76: 437-54.

Peters, H. Elizabeth. 1992. "Marriage and Divorce: Reply." *American Economic Review* 82: 686-93.

Ransom, Roger L., and Richard Sutch. 1986. "The Labor of Older Americans: Retirement of Men On and Off the Job." *Journal of Economic History* 46: 1-30.

CHAPTER 1

The Labor Supply Effects of the Social Security Earnings Test

With the life expectancy of those who reach old age having risen substantially and with the elderly continuing to grow as a percentage of the population, their economic well-being has been a major policy focus in recent decades. Government support programs such as Social Security have contributed to increasing the income of the elderly both in absolute terms and relative to other groups of the population from the late 1960s on. More recently, the policy focus has shifted to encouraging the aged to become more reliant on their own resources.

The Social Security earnings test has been one focus of such efforts. The earnings test reduces a recipient's Social Security benefits at a proportional rate as he or she earns more than an earnings floor. When the earnings test was introduced in 1939, the intent was to push older workers out of the labor force: beneficiaries lost an entire month's benefits when monthly earnings exceeded $15. Since the 1950s, the earnings test has been gradually relaxed. In 1995, a beneficiary under the age of 65 who earned more than $8,160 lost $1 in benefits for every $2 in additional earnings -- which functions as a 50% tax on wages. A 65-69 year old lost benefits at a 33% rate for earnings above $11,280.

Even following the liberalizations of the earnings test, these tax rates remain among the economy's highest. A 1995 press release by the National Center for Policy Analysis noted that a 64-year old beneficiary faces, on average, a combined marginal tax rate of 83%

for earnings above $8,160, and possibly up to a 114% marginal tax rate.[1] In consequence, the earnings test receives substantial popular attention, as exemplified by the 1989 book *Paying People Not to Work: The Economic Costs of the Social Security Earnings Limit* (Robbins and Robbins), which was widely cited in the media. Moreover, retirees appear to be well-informed about the earnings test from popular magazines like *Money* and AARP's *Modern Maturity*.[2] Concern over the distortions induced by the earnings test led Congress in early 1996 to raise the earnings exempt amount gradually to $30,000 by 2002.

In spite of the substantial popular attention to the earnings test, the existing literature concludes it has a small impact on the labor supply of the elderly. The principal papers in this literature by Burtless and Moffitt (1985) and Gustman and Steinmeier (1986) used the Retirement History Survey, which lasted from 1969 to 1979 and has become somewhat dated. More importantly, there were virtually no changes in the earnings test during the time period the Retirement History Survey covered, so estimates of the earnings test's effect are identified primarily from potentially endogenous cross-sectional variation in wages.

I propose to examine three changes in the earnings test rules that were instituted between 1978 and 1990. Those changes, which remain mostly unexplored, offer different combinations of shifts in the earnings test tax rate and the level of the earnings exempt amount. Also, they apply only to certain beneficiaries differentiated by age, so the behavior of unaffected age groups may be used to control for unrelated changes in labor supply. This sets the stage for a "natural experiment" analysis of the earnings test of the sort recently used to study other social insurance programs.[3] The substantial changes in the earnings test also provide strong exogenous variation to identify estimates of a structural labor supply model.

The rest of this paper is divided into six sections. The first section sets the stage for

---

[1] 83% is the sum of the earnings test tax rate, federal income and payroll taxes and a 4% state income tax. The tax rate can climb higher because benefits become taxable when income plus benefits plus tax-free interest income exceeds $25,000 for a single filer or $32,000 for a joint filer.

[2] 73% of retirees under the age of 72 in the 1982 New Beneficiary Survey reported being aware of the earnings test.

[3] Recent examples include Krueger (1990) on worker's compensation and Meyer (1995) on unemployment insurance.

the analysis of the earnings test. It describes how the earnings test functions. It then provides a brief summary of evidence about elderly labor supply more broadly, follows with a discussion of the earlier papers on the earnings test and considers how many beneficiaries appear to be influenced by the earnings test. Section II begins the inquiry by illustrating how the changes in the earnings test rules shifted the budget set faced by beneficiaries and what the shifts imply for labor supply.

The way in which individuals actually respond to the budget set generated by the earnings test is shown in Section III. I look at patterns of earnings and calculate differences in differences that quantify the comparisons across age groups before and after the changes in the earnings test. This straightforward approach reveals that the elderly bunch in substantial numbers at and just below the earnings exempt amount. The bunching calls into question earlier findings that the earnings test has little impact on behavior and also represents evidence against the view that individuals fail to bunch as predicted at convex kinks in the budget constraint.[4]

In Section IV I proceed to incorporate the piecewise linear budget constraint defined by the earnings test into a theoretical model of labor supply. Estimating the model draws on the reactions to the earnings test along the entire budget set and decomposes those reactions into income and substitution elasticities, permitting predictions about other proposed changes in the earnings test. The elasticities are of more general interest as well in describing the labor supply behavior of the elderly who continue to work, in contrast to the general focus in the literature on retirement. Policies that attempt to make older people more self-sufficient by extending their working lives need to take into account hours elasticities conditional on working, along with participation elasticities.

The typical criticisms of structural estimation are that the estimates arise out of the assumptions imposed on the data and overlook the issue of identification. The changes in the earnings test that are the focus here lead to substantial variation in the budget set both cross-sectionally and over time which reduces the reliance on the assumptions which must be made. So, the modelling contributes an econometric framework for estimating the impact of

[4] Heckman (1983), MaCurdy (1992).

the earnings test, and the earnings test modifications contribute exogenous variation to help identify the labor supply model.

Section V describes the results of estimating the piecewise linear budget set model on the most prominent change in the earnings test using maximum likelihood methods. The estimation yields relatively large income and substitution elasticities that imply substantial deadweight loss from the earnings test, in line with the observed response to the earnings test in Section III. Several validation checks confirm the importance of the exogenous variation in the earnings test for arriving at the structural estimates. The estimates then provide a basis for simulating other policies. In contrast to the intuition that raising the exempt amount to $30,000 would amount to eliminating the earnings test for the most part, simulations show that the overall effect on hours would be only a slight increase because the constraint would simply be shifted onto a new higher earnings group of beneficiaries. On the other hand, eliminating the earnings test would raise hours worked of those currently affected by the earnings test by 13.8%.

I. BACKGROUND

This section begins with a description of how the earnings test works at present and how it has been modified in the last twenty-five years. It follows with a short discussion of post-war trends in retirement and how they might influence the process of estimating the impact of the earnings test. It concludes by summarizing earlier papers on the earnings test, including evidence on the number of beneficiaries affected by the earnings test.

A. The Earnings Test

When a Social Security beneficiary earns more than a threshold amount, his (or her) benefits are reduced at a rate proportional to his additional earnings. The takeaway rate is equivalent to a tax rate applied to wages until benefits are gone. In 1995 a 62-64 year old beneficiary could earn up to $8,160 -- the earnings *exempt amount* -- with no reduction in benefits. When he worked more, he lost $1 in benefits for every $2 he earned -- a 50% *tax rate*. For 65-69 year olds, the test is less restrictive, with a higher exempt amount and a

33% tax rate. The earnings test generates a piecewise linear budget set with one convex kink -- corresponding to the exempt amount -- and one nonconvex kink -- corresponding to the point where all benefits disappear. The exempt amount is raised every year according to the economy-wide increase in the average wage. Individuals who are at least 70 years old are exempt from the earnings test. Nonlabor income is not subject to the earnings test.

The Earnings Test Rules in 1995

	62-64 year olds	65-69 year olds
Exempt amount	$8,160	$11,280
Tax Rate	50%	33%

Changes in the Earnings Test Rules

Year	What changed	Who was	Who was not affecte
1990	Lowered tax rate to 33%	65-69 year olds	62-64 year olds
1983	Eliminated the earnings test	70-71 year olds	62-69 year olds
1978	Raised the exempt amount 25%	65-71 year olds	62-64 year olds
1973	Removed 100% tax "bracket"	62-71 year olds	no one

The earnings test has been altered several times in recent years, as shown above and detailed in Table 1. Most recently, in 1990 the tax rate was reduced for 65-69 year olds from 50% to 33%. Earlier, beginning in 1983, 70-71 year olds were exempted from the earnings test. In 1978 the exempt amount, which used to be the same for everyone, was raised by about 25% for 65-71 year olds. Each of these changes involves a different combination of shifting the location of the convex kink and the slope of the segment above it. A complete picture of how the earnings test affects labor supply emerges from observing the reaction to the set of changes. Furthermore, since each change applied to one age group and not to another, the earnings test response is further isolated by comparing the change in behavior of the affected group relative to the unaffected group as a way to control for unrelated changes in labor supply behavior over time.

An earlier change occurred in 1973, when a second tax rate of 100% which used to apply after a higher earnings threshold was eliminated. That change offers less insight than the later three, though. It applied to all ages, so there is no comparison group available to control for time series changes -- which is particularly critical because retirement rates were rising sharply just then. Section III will demonstrate the response to the 1973 change along with the others, but the analysis will show that individuals apparently did not react to the 100% tax kink before 1973 or to its elimination. The 1973 change is the only one captured by earlier papers on the earnings test, but it is the least useful for identifying its effect.

Another rule potentially complicates the analysis of the earnings test. Beneficiaries are entitled to an increase in future benefits for each month of current benefits foregone, either because of the earnings test or because they initially postpone filing for benefits. For beneficiaries under the age of 65, this feature is called the actuarial adjustment of benefits and amounts to about a 7% increase in future benefits for each year of benefits foregone before 65, which is approximately actuarially fair. For beneficiaries aged 65-69, it is called the Delayed Retirement Credit, which was introduced in 1973 as a 1% annual adjustment in future benefits, raised in 1983 to 3%, and beginning in 1990 has been raised by one-half percent every other year until it reaches 8%.[5] An actuarially fair adjustment of future earnings should make an individual with average life expectancy and no liquidity constraints indifferent to the earnings test. However, there is strong evidence that the earnings test matters and scant evidence that the Delayed Retirement Credit does, which I will discuss in more detail in the earlier literature on the earnings test. Also, in the structural estimation I estimate a model that allows the increase in future benefits to mitigate the impact of the earnings test, but no such effect emerges.

B. What We Know about the Labor Supply of the Elderly

The labor force participation of men aged 60 and older fell from almost 65% in 1930 to about 30% in 1980, with the sharpest declines occurring in the 1950s and 1970s.[6] Figure

---

[5] In 1995 and 1996, the Delayed Retirement Credit is 5½%.

[6] The statistics in this paragraph are reported in Lumsdaine and Wise (1990) and are based on Ransom and Sutch (1988) and Tuma and Sandefur (1988).

1 shows the percentage of men of different ages who reported not working in the previous year in the Current Population Survey. Women of the same ages worked more until around 1970, but since then have worked less as well. Meanwhile, the percentage of persons 65 and older receiving Social Security benefits rose from about 20% in 1940 to 85% in 1960 and about 95% currently, and the size of benefits rose from about 14% of median male income in 1950 to 38% in 1980.

The retirement literature lacks a satisfactory description of what generated the strong trend to earlier retirement. The evidence is mixed on the degree to which Social Security was a factor and suggests a limited role in recent years.[7] In this vein, Krueger and Pischke (1992) found no reversal of the retirement trend during an episode in which real benefits fell for certain individuals instead of rising. They concluded that potentially spurious correlation between retirement rates and real benefits levels make it difficult to decipher the relationship.

Figure 1 notes the strong upward trend in retirement rates from the late 1960s to mid 1970s, the time period covered by most previous papers on the earnings test. As a result, the approach those papers generally take of jointly modelling retirement timing and post-retirement (or post-recipiency) labor supply may be problematic for conclusions about the earnings test. The task of explaining the strong trend in retirement might dominate the model's capacity to explain post-retirement behavior, confounding any identification which is available from the minor time series changes in the earnings test rules before the late 1970s. With this in mind, I will focus most of my efforts on the group of individuals who continue to work at various ages.

C. The Previous Literature on the Earnings Test

Papers by Burtless and Moffitt (1985) and Gustman and Steinmeier (1986) incorporated the earnings test into structural retirement models, while a simpler quantitative approach was taken by several papers in the *Social Security Bulletin*. A final group of

---

[7] In reviewing the literature, Boskin (1986) concluded that the evidence supported a strong link between rising real benefits and falling participation rates in the late 1960s and early 1970s. In their review, Lumsdaine and Wise (1990) concluded that most of the studies in the 1980s "attribute only a modest portion of the early retirement trend to the effect of Social Security provisions."

papers considered how the effect of the age structure of benefits may interact with the earnings test.

Burtless and Moffitt and Gustman and Steinmeier took similar approaches in estimating structural models of the joint decision of retirement age and post-retirement labor supply. The post-retirement hours choice was modelled using the piecewise linear budget set methodology. The authors did not model the Social Security recipiency decision separately, instead assuming it occurs simultaneously with or defines retirement. Each paper used a shortened sample from the Retirement History Survey, a longitudinal data set that lasted from 1969 to 1979.

The previous two subsections point out some potential difficulties with their approach. I first noted that the papers cover only the least interesting of the last four changes in the earnings test. The 1973 elimination of the higher 100% tax kink caused exogenous time series variation in the budget set, but no cross-section variation as the later three changes did. Moreover, the empirical analysis later in this paper finds no apparent response to the 1973 change. What remains to identify the estimates of post-retirement hours elasticities is variation across individuals. However, time series variation in post-retirement labor supply behavior may be confounded by the trend in retirement discussed in the last subsection; while cross-sectional variation in the budget set due to wages and non-labor income is likely to be correlated with omitted personal characteristics, biasing the estimates.

Both papers did estimate relatively elastic labor supply but introduced other considerations that reduced the importance of the earnings test, which would otherwise follow from their estimates. Burtless and Moffitt concluded that the overall impact of eliminating the earnings test would be quite small because only about 12% of retirees worked and earned enough to be affected by the earnings test. Gustman and Steinmeier incorporated the Delayed Retirement Credit into the model, which reduces the impact of the earnings test under certain conditions.

Burtless and Moffitt's argument that few beneficiaries are affected by the earnings test is similar to that made in a series of papers appearing in the *Social Security Bulletin*.[8]

---

[8] Leonesio (1993), Bondar (1990), Packard (1990), Lingg (1986).

Those papers counted up how many individuals are located on different portions of the budget constraint and inferred that the earnings test has a small effect because relatively few beneficiaries are in the regions affected by the earnings test. Nevertheless, even if the percentage affected is small, millions of individuals in the relevant age groups make the number affected large. Table 2 reports data from two articles in the *Social Security Bulletin*. Bondar (1990) calculated that 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, along with 315,000 spouses and dependents, who have never been considered in analysis of the earnings test. Leonesio (1993) also noted the impact of the earnings test on an estimated 174,000 beneficiaries aged 65-69 in the immediate neighborhood of the kink and 582,000 eligibles with positive earnings who did not claim benefits.[9] Over a million individuals appear to be influenced by the earnings test -- a large number in absolute terms, and arguably in relative terms.

Gustman and Steinmeier also estimated relatively elastic labor supply. While they did not mention the earnings test's effect explicitly, they reported results in other papers of simulating a reduction in the earnings test tax rate or of concurrently eliminating the earnings test and making the Delayed Retirement Credit actuarially fair (1985, 1991). However, the potential effect of the Delayed Retirement Credit to wash out the earnings test makes the simulated effect of relaxing the earnings test understandably sensitive to assumptions regarding the impact of the Delayed Retirement Credit.

There is no agreement over the degree to which individuals respond to the Delayed Retirement Credit. Blinder, Gordon and Wise (1980) first used simulations to conclude that the actuarially fair adjustment of benefits for 62-64 year olds can erase the labor supply effect of the earnings test for an individual with average life expectancy, no liquidity constraints, and a full understanding of how benefits are computed. The Delayed Retirement Credit is only now becoming actuarially fair, though. Moreover, Burkhauser and Turner

[9] Individuals could be choosing not to file for benefits because they anticipate earnings high enough that all benefits would be withheld. The decision not to file may be reinforced by the Delayed Retirement Credit.

(1981) disputed the premise of their simulations and also argued that, for many, if not most, individuals, the adjustment of benefits is not actuarially fair even for 62-64 year olds. There is little empirical evidence about the impact of or awareness of the Delayed Retirement Credit. An exception is Reimers and Honig (1993), who found the Delayed Retirement Credit had no effect in a hazard model of labor force reentry among retirees who experienced its introduction. Evidence that the earnings test matters also arises out of the clustering at the convex kink which I will show in Section III. On the strength of such arguments, Burtless and Moffitt omitted the Delayed Retirement Credit from their model. I will also focus on modelling labor supply without considering the Delayed Retirement Credit, but I will estimate a version that can capture whether the Delayed Retirement Credit mitigates the impact of the earnings test.

In sum, there are several reasons to revisit the evidence on the earnings test. Most earlier papers used relatively old data which did not cover any of the substantive changes in the earnings test. The papers with structural models focused primarily on the retirement decision, and their modelling strategy may have hampered their ability to detect the effect of the earnings test. Finally, the arguments made in some papers based on either the small percentage of affected individuals or on the potential impact of the Delated Retirement Credit may be somewhat misleading.

II. MODELLING THE IMPACT OF THE EARNINGS TEST

The theoretical and empirical analysis of utility maximization under a piecewise linear budget constraint provides a useful framework for studying the labor supply of Social Security beneficiaries. This section diagrams the theoretical budget constraints and the implied labor force choices of Social Security beneficiaries under the various changes in the budget set. Afterwards, some possible extensions to the basic framework are discussed. The discussion in this section will suggest a range of possible empirical approaches. One approach I will follow in Section III will be to eliminate all possible confounding influences by focusing on the most unambiguous implications of the theory and by ignoring other available information which may require stronger assumptions to interpret. The other approach undertaken in Sections IV and V will be to incorporate the structure necessary to estimate the theoretical model, which will yield forecasts of the response to a variety of potential policies.

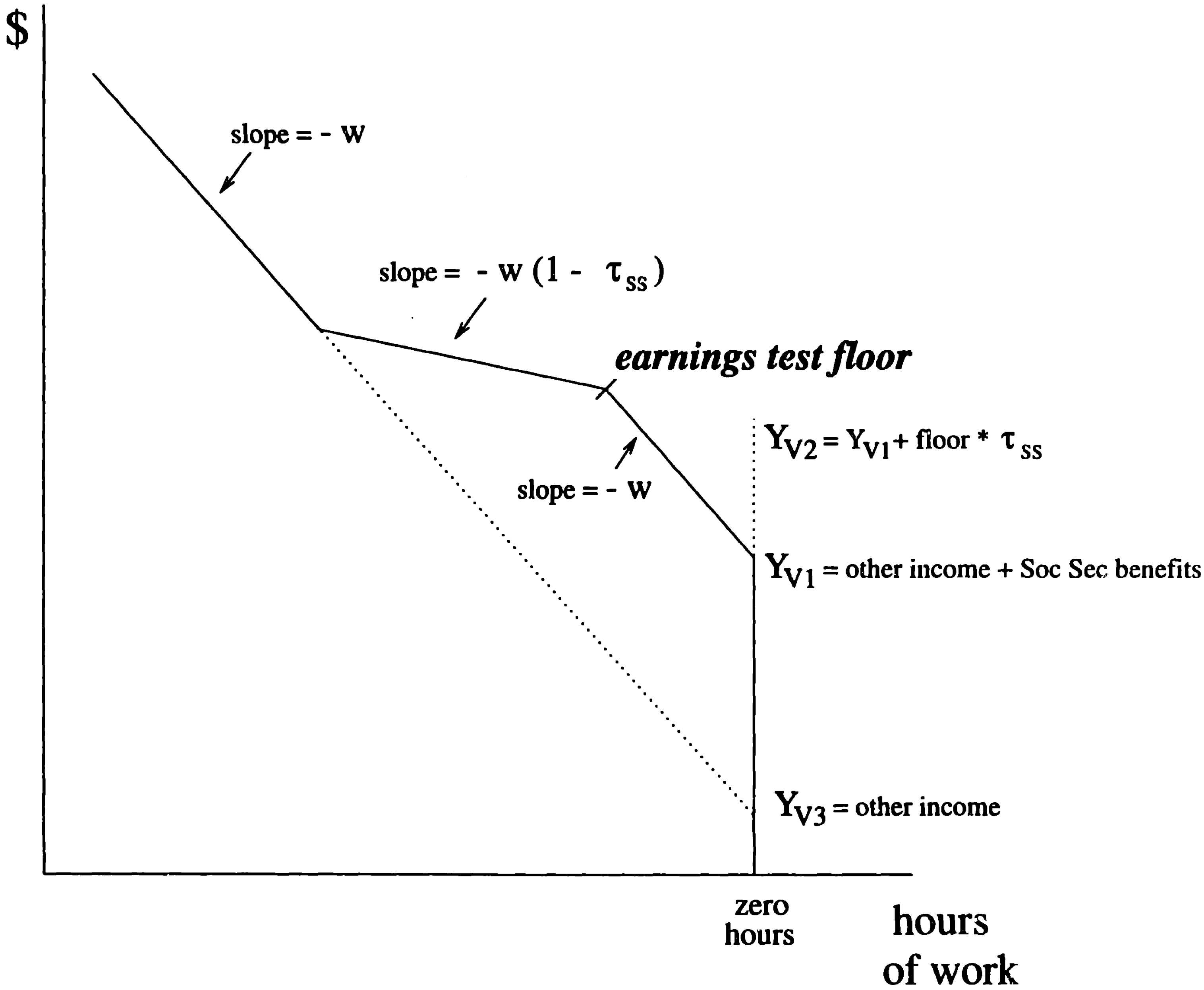
A. The Budget Set Created by the Earnings Test

Figure 2 shows a series of budget sets corresponding to different versions of the earnings test. Figure 2-A displays the budget set created by the earnings test as it functions now. At zero hours of work, the recipient has income which is composed of Social Security benefits and other nonlabor income. He chooses zero hours of work if his indifference curve is tangent to the budget set there or at points to the right, which are infeasible. He joins the labor force when his indifference curve becomes tangent to the lower segment, where he earns gross wage w. As his relative preferences for work increase, he reaches a convex kink at the earnings exempt amount -- $8,160 for a 62-64 year old in 1995. Locating at the kink is compatible with a range of indifference curves, so there should be many individuals bunched there. Then, as the individual prefers more work, he locates on the middle segment, where his benefits are being taxed away as a function of the difference between his earnings and the exempt amount. This makes his net marginal wage equal to $w^*(1-\tau_{SS})$, where τ_{SS} is the earnings test tax rate. The length of the middle segment varies with the size of the individual's initial Social Security benefits and with τ_{SS}. At the point where the individual's benefits disappear, a nonconvex kink appears. No one will locate there; instead, at some point there will be a dual tangency where the individual is indifferent between locating on the middle and upper segments. Finally, if the individual's relative preference for work is strong enough, he chooses to locate on the upper segment where his marginal wage is once again the gross wage and his benefits are gone.

The strongest implication of the theory outlined above is that we should observe individuals massed at the kink. While the theory predicts a cluster exactly at the kink, measurement error or the presence of small rigidities in labor supply choice would plausibly

FIGURE 2 - A

The Earnings Test



spread out the cluster in the neighborhood of or just below the kink.[10] With no other explanation for individuals to mass at a particular point, a pattern of clustering near the kink will be convincing evidence that individuals react to the earnings test.

B. The Comparative Statics of the Four Changes in the Earnings Test

Figure 2-B illustrates how a change in the earnings test alters the budget constraint. It uses as an example the 1990 decrease in the tax rate from 50% to 33% for 65-69 year olds. The lower tax rate made the slope of the middle segment steeper and also lengthened it. How does this affect how much an individual would choose to work? Intuitively, one expects labor supply to rise when the penalty for working is lowered. In fact, some individuals have an incentive to increase their labor supply, but some have an incentive to decrease their labor supply and others face conflicting incentives. There are several distinct situations, corresponding to a person's initial location:

• *People at the kink.* Some people previously at the kink will increase their hours and move onto the middle segment, since the disincentive to working more is now smaller. The greater is the elasticity of substitution between leisure and consumption -- the flatter the indifference curve -- the more people will move off the kink.

• *People on the middle segment.* People who are already losing benefits to the earnings test become wealthier and enjoy a higher marginal wage. If leisure is a normal good, then the income and substitution effects conflict and their relative size determines whether people decrease or increase their hours. The impact of the income effect grows along the middle segment.

• *People on the upper segment below the new nonconvex kink.* They will move onto the middle segment mechanically because they will now be eligible to keep some benefits while working the same hours. They may further respond behaviorially by decreasing their

[10] The notion of optimization error captures small rigidities that prevent an individual from working exactly his desired hours. Optimization error or observationally-equivalent measurement error is modelled in the piecewise linear budget set literature as symmetrically distributed around desired hours. However, in a dynamic setting with frequent small increases in the floor, small rigidities may lead to asymmetric bunching concentrated just below the kink.

FIGURE 2 - B
1990: Lowered Tax Rate for 65-69 Year Olds

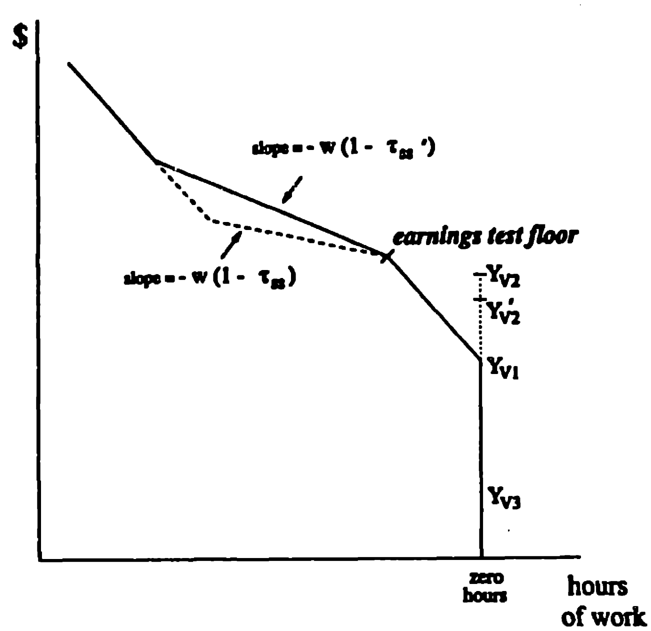


FIGURE 2 - C
1983: Abolished Earnings Test for 70-71 Year Olds

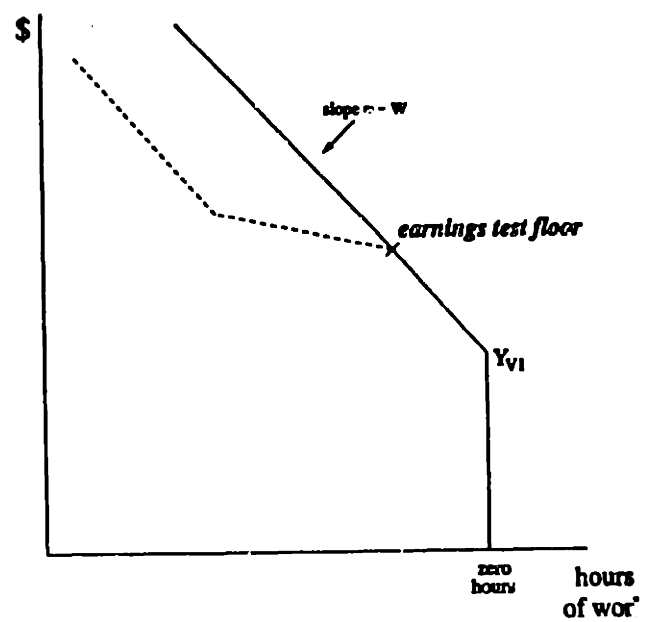


FIGURE 2 - D
1978: Raised Floor for 65-71 Year Olds

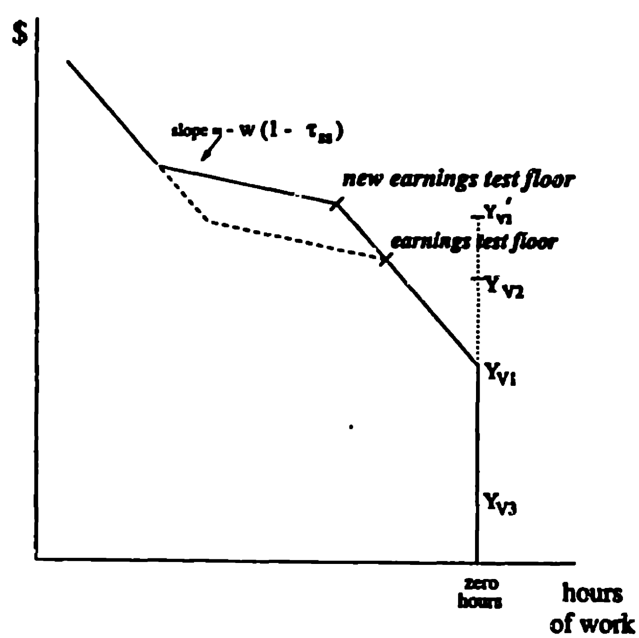
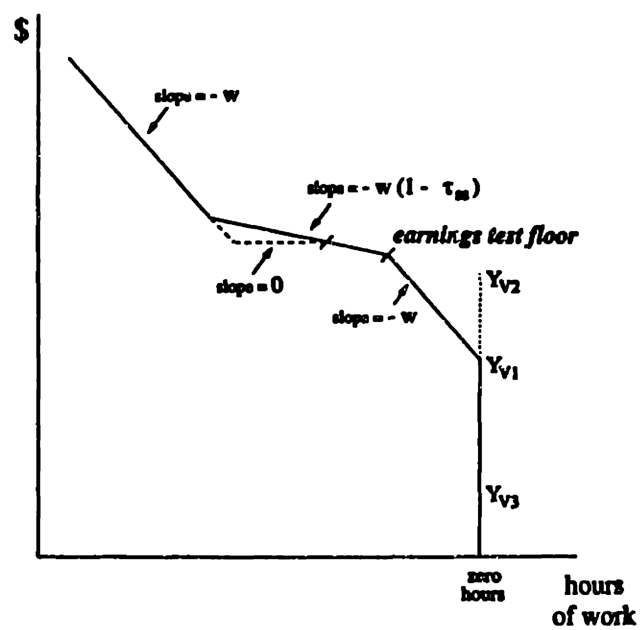


FIGURE 2 - E
1973: Removed a 100% "Tax Bracket" for Everyone



hours because the income and substitution effects work in the same direction. The distinction over whether the movement onto the middle segment involves a behavioral response is important for policy analysis.

- *People on the upper segment a little above the new nonconvex kink.* Some individuals who were on the upper segment are now close enough to the new nonconvex kink that they find themselves on a dominated part of the budget set. They will move to a higher indifference curve with a dual tangency on the middle and upper segments or to an indifference curve on the middle segment, causing a small change in hours.

- *People on the upper segment far above the new nonconvex kink.* Above some point on the upper segment, individuals will be unconcerned about the shift in the budget set.

- *People on the lower segment.* They also will prefer to remain where they were -- barring the existence of some type of rigidity in labor supply choice. If such distortions are present, then a big enough change in the earnings test will induce some individuals who located below the kink initially to increase their hours and move above the kink. It will be more likely the closer the individual was to the kink initially, the bigger the decline in the tax rate, and the smaller the rigidity.

The other changes in the earnings test are also illustrated in Figure 2. Figure 2-C demonstrates the elimination of the earnings test in 1983 for 70-71 year olds. Before 1983 70-71 year olds were subject to the same nonlinear budget constraint as 65-69 year olds; afterwards their budget constraint consisted of a single segment rising from the zero hours of work point with a slope of $-w$. Those 70-71 year olds initially at the kink will raise their labor supply following the elimination of the earnings test. Those located on the middle segment will face conflicting income and substitution effects, with the income effect growing stronger along the segment. Those located on the upper segment will unambiguously work less because there is no conflicting substitution effect. The 1983 experiment is even more useful than the 1990 one for empirical purposes. It causes a more substantial change in the opportunity set, combining movements of both the exempt amount and the tax rate. It also creates two natural comparison groups: a younger group that faces the same budget set as the 70-71 year olds before 1983, and an older group that faces the same budget set after 1983.

Figure 2-D demonstrates the increase in the earnings exempt amount for 65-71 year olds in 1978, while 62-64 year olds experienced no change. The new kink shifts upwards, generating a positive income effect for everyone initially located at or above the kink and also a positive substitution effect for those initially located at or just above the kink.

Finally, before 1973, the earnings test had a second tax rate of 100% triggered above the 50% tax rate, as shown in Figure 2-E. There will be less bunching at the second kink than at the first because of its proximity to the nonconvex kink. Intuitively, a recipient has less benefits to lose and will be more disposed to prefer the upper segment with no benefits instead, relative to being at the first kink when he still has all his benefits. This factor makes the elimination of the 100% tax "bracket" in 1973 more similar to the reduction of the initial tax rate in 1990 than to the elimination of the earnings test for 70-71 year olds in 1983. It also lacks the treatment-control approach afforded by the other experiments.

C. Extensions

Integrating federal income taxes, Social Security payroll taxes, and benefit taxation into the budget set approximates more closely the real world. The federal tax system generates an initial convex kink of 11-15% and lesser convex kinks above it, which were numerous and small before the late 1980s.[11] For a married person with no dependent children and no non-labor income, the initial tax kink -- occurring at the sum of the standard deduction plus two personal exemptions -- was located above the earnings test kink until the early 1980s, but never by more than $2000. Individuals with a little taxable family income faced the tax kink first, while individuals with more than a few thousand dollars in taxable family income did not face the initial income tax kink. In the middle 1980s, the initial tax kink was below the earnings test kink, while it reversed again in the late 1980s. Social Security taxes apply to all earnings along the budget set up to a maximum, where it

---

[11] During the 1970s, the tax schedule had twenty-five brackets, with a one to four percentage point difference in tax rates from one bracket to the next. The number of brackets was reduced to fifteen in 1979, twelve in 1982, and three in 1987.

introduces a nonconvex kink in the budget set.[12] Benefit taxation, which was instituted in 1984, creates an additional convex kink at a relatively high level of earnings.[13]

The earnings test kink is bigger than any involved in the income tax system, so it stands out no matter how the tax system is treated. Also, none of the details about taxes interfere with the comparisons across age groups that is part of the earnings test analysis because everyone faces the same tax system. For individuals with certain small levels of nonlabor income, though, the location of the initial tax kink could overlap with the earnings test tax kink in some years and not in others, potentially interfering with the comparisons over time. In the empirical analysis I will account for the location of the initial tax kink to avoid any confounding influence with the earnings test kink.

Incorporating fixed costs of work, which cause a discontinuity in the budget set at zero hours, would also make the model more realistic. The earnings test can affect the participation decision in that case, and liberalizing the earnings test will encourage labor force participation if fixed costs are not too big or too small.[14] An individual who is induced to enter the labor force after the earnings test is relaxed would select a point on the budget set at or above the initial earnings test kink.

Trying to detect a participation effect is a more ambiguous exercise than looking to see if individuals respond to the earnings test kink. Numerous factors influence the decision not to work, and some are likely to be age-specific, confounding the use of the earnings test rules changes which distinguish by age. Besides, there is little observable variation in fixed

---

[12] The maximum earnings level for Social Security taxes was raised from $7,800 in 1971 to $25,900 in 1980 and over $50,000 by the middle 1990s. The tax rate was 5.85% for both employee and employer for much of the 1970s. After 1977, it was raised gradually to 7.0% in 1984 and 7.65% in 1990.

[13] When modified Adjusted Gross Income (MAGI), measured as Adjusted Gross Income plus benefits plus tax-free interest income, exceeds $25,000 for a single filer or $32,000 for a joint filer, the smaller of either half of MAGI minus the threshold or half of benefits is counted as taxable income. Since 1993, if MAGI also exceeds $34,000 for a single filer or $44,000 for a joint filer, 85% of benefits are taxable.

[14] The other way in which the earnings test can affect the participation decision is by influencing the timing of retirement because it affects how much wealth is accumulated each year to finance post-retirement consumption. Thus, liberalizing the earnings test might cause a wealth effect for a working beneficiary that allows him to retire earlier than planned. This impact would take more than one year to be felt fully.

costs among the elderly that would help to get a handle on the issue.[15] In the empirical analysis, I will present slight evidence of a participation effect. However, lacking a convincing identification strategy, I will omit it in the structural modelling.

D. Measuring the Response to a Change in the Earnings Test

How can one determine empirically the influence of the budget set on labor supply? This section described how a recipient's reaction to external incentives depends on the location and shape of his indifference curves. We have information about his wage and income, which determine his budget set, and about some individual characteristics, which govern his choices. However, the individual preferences for work which generate the indifference curves may have had a role in determining his wage, Social Security benefits and other income. Therefore, measuring the effect of the budget set on labor supply choice may capture the effect of intrinsic characteristics that shape the indifference curves as well. The earnings test is an exogenous feature which has a major role in determining the budget set for recipients. Even better, the changes in the earnings test caused substantial external variation in the budget set over time and across age groups who should otherwise be similar. That is the identifying assumption for the empirical approaches used in this paper.

A variety of possible empirical strategies for quantifying the response to the earnings test arise from reviewing the budget set. The methods vary in the degree to which they formalize that response: the more structured the approach, the more closely the answers resemble theoretical concepts of interest; but also the more the accuracy of the answers depends on the formalization itself being a good description of reality.

In the rest of this paper, I will rely on two distinct approaches. I begin in Section III with a visual presentation of earnings distributions relative to the earnings test kink, thus focusing on the strongest implication of the theory. Earnings patterns show that individuals react to the earnings test in the neighborhood of the kink. Comparing earnings around the kink before and after the different earnings test changes decomposes how individuals respond

---

[15] As an example of a successful identification strategy, Hausman (1980) used the number of children to help explain the participation decision of prime-age women through its effect on child care costs.

to the exempt amount and the tax rate, while comparing the earnings changes to earnings for other individuals unaffected by the modifications controls for aggregate trends which also move the earnings distributions over time.

Quantifying the income and substitution elasticities that underlie the observed reactions, and also capturing the responses of individuals along other parts of the budget set, requires substantially more structure. Section IV incorporates the piecewise linear budget constraint described in this section into a theoretical model of labor supply. With the formalization of the utility maximization approach, the problem of distinguishing the effect of the budget set on labor supply choice from the effect of preferences comes to the fore. Inconsistent estimates result when the wage and income variables capture omitted observable or unobservable influences on an individual's choice of hours. Here, the substantial exogenous variation caused by the earnings test in non-labor income and in the net wage both across groups and over time will be used to identify the estimated elasticities.

III. EARNINGS DISTRIBUTIONS AND THE EARNINGS TEST

The general conclusion in the economics literature is that the earnings test has a small impact on behavior. On the other hand, recipients are informed from many popular sources -- such as *Money* and *Modern Maturity* -- about how and when the earnings test applies. A popular book called *Paying People Not to Work: The Economic Costs of the Social Security Earnings Limit* (Robbins and Robbins 1989) claiming that 700,000 people would return to work following the elimination of the earnings test was widely cited in the popular press, although their methods are reported to be in error.[16] As we shall see, neither the academic nor the popular view fully captures how the earnings of recipients actually respond to the earnings test. There is an observable reaction to the earnings test: individuals pile up below the kink and follow the kink as it moves, while the earnings distribution of similar individuals who are not subject to the earnings test declines smoothly over the same range.

[16] Leonesio (1993) cited a Social Security Administration document which determined that the Robbins and Robbins book contained "serious theoretical and methodological shortcomings."

Yet, the clustering at the kink does not appear to diminish when the tax rate declines, which places a bound on the responsiveness of beneficiaries.

In most of this section, I examine graphs of earnings distributions before and after each of the four changes in the earnings test. Each graph shows how many men had earnings in each $1000 interval above and below the exempt amount for a few years before and after the changes in the earnings test. Afterwards, I will analyze participation as well. The data are taken from the March Current Population Surveys (CPS). Earnings are measured as earnings from wages and salaries plus earnings from self-employment, which is what the earnings test covers. I chose to focus on men because the rules applied to wives' benefits are more complicated and because women's labor force participation rate, and hence their sensitivity to the earnings test, is lower than men's.[17] I restrict the age groups slightly because the CPS lacks data on birthdays. For instance, I only look at 67-69 year olds because I do not know when someone who reports being 65 in March had his birthday during the previous year, when the earnings occurred.

A. Earnings Before and After the Earnings Test Changes

Figure 3 shows earnings distributions *relative to the earnings exempt amount* before and after the 1978 increase in the exempt amount for 65-71 year olds. The graph compares the earnings of 67-69 year olds to the earnings of 63-64 year olds, who did not experience the increase in the exempt amount. Thus, Figure 3-A shows, before the treatment, how many of the comparison group and the treated group had earnings in each $1000 interval above and below the level of the earnings exempt amount divided by the total number of individuals in that age group.

Individuals in both age groups bunch just below the earnings test kink in Figure 3-A. Roughly the same number of individuals 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

[17] Depending on which alternative yields bigger benefits, wives may receive benefits as dependents or as retirees. Dependent benefits are subject to an earnings test applied to both the retiree's and the dependent's earnings, while retiree benefits depend on the retiree's earnings alone.

FIGURE 3-A

1975-77

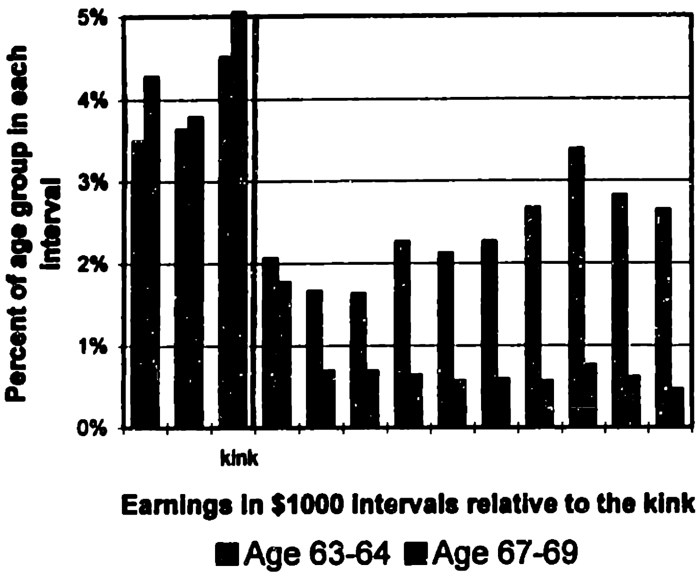


FIGURE 3-B

1979-81

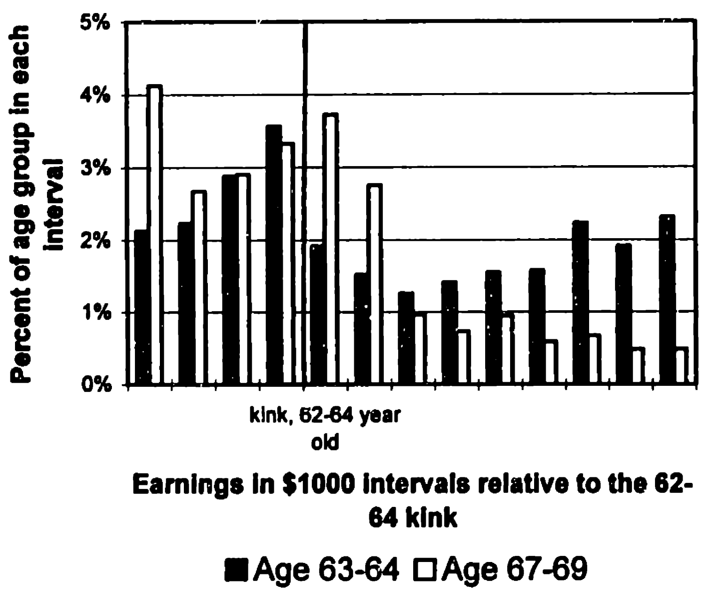
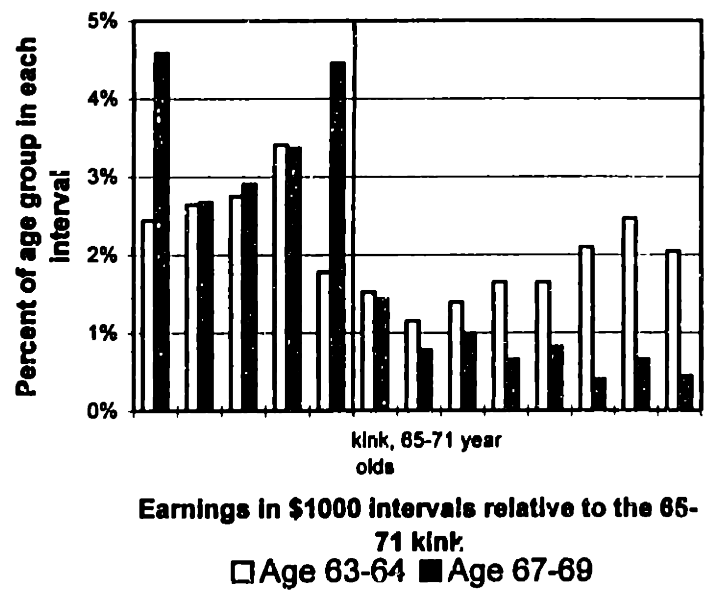


FIGURE 3-C

1979-81



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 the previous wisdom that the earnings test has little impact, and also to the assertion that people do not react to convex kinks in the budget set as theory predicts.

After 1978, the clustering of 67-69 year olds moves up to their new kink, as shown in Figures 3-B and 3-C. Figure 3-B shows earnings of both 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 67-69 year olds no longer react at that point in the distribution. Instead, in Figure 3-C, they cluster at their new higher kink, which, similarly, no longer governs the behavior of the 63-64 year olds.

Those conclusions are quantified in the first part of Table 2, which reports the same data as Figures 3-A, 3-B and 3-C. The first column of each panel reports the proportion of individuals of each age group in a few intervals close to the kink before 1978 for each age group. The last row computes the difference between the percentage of people just below and above the kink and indicates that the bunching is both significant and significantly different from the behavior across any of the other intervals. The other two columns in each panel compare the pattern of earnings at the new separate kinks and confirm that each group now responds strongly to its own kink. Together, Figure 3 and the first part of Table 2 provide strong evidence that individuals react to the earnings test at the kink by holding down their labor supply.

Figures 4 and 5 and the rest of Table 2 make the same types of comparisons of earnings around the kink for the 1983 and 1990 changes in the earnings test. Figures 4-A and 4-B compare the earnings patterns of individuals before the 1983 elimination of the earnings test for 70-71 year olds. Figure 4-A compares the earnings of the treated group to that of the younger control group, who face the earnings test both before and after 1983.[18] The clustering of individuals at the kink is, again, substantial. Figure 4-B approximates the counterfactual of no earnings test for those individuals by comparing the treated group to the

[18] Figure 4 actually uses 71-72 year olds as the treatment group, since they were 70 and 71 years old when the earnings were earned. Also, Figure 4-A omits 1982 in the pre-treatment period because the earnings test was scheduled to be eliminated in 1982 until it was postponed in 1981 for one year.

FIGURE 4-A
1980-81

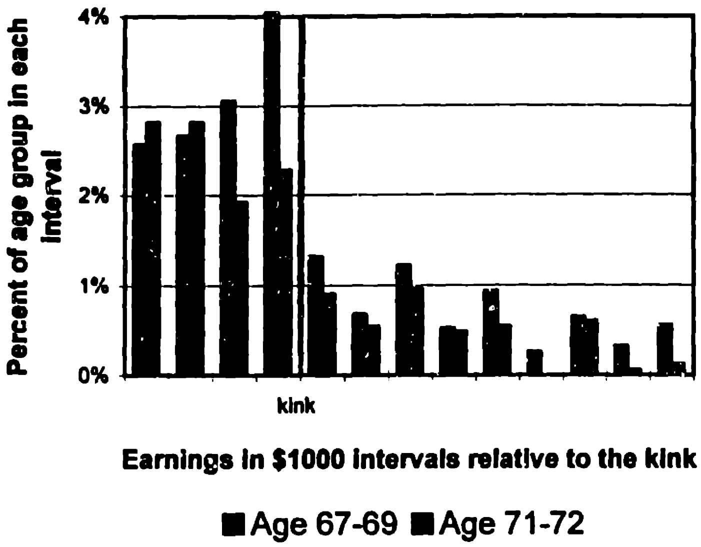


FIGURE 4-B
1980-81

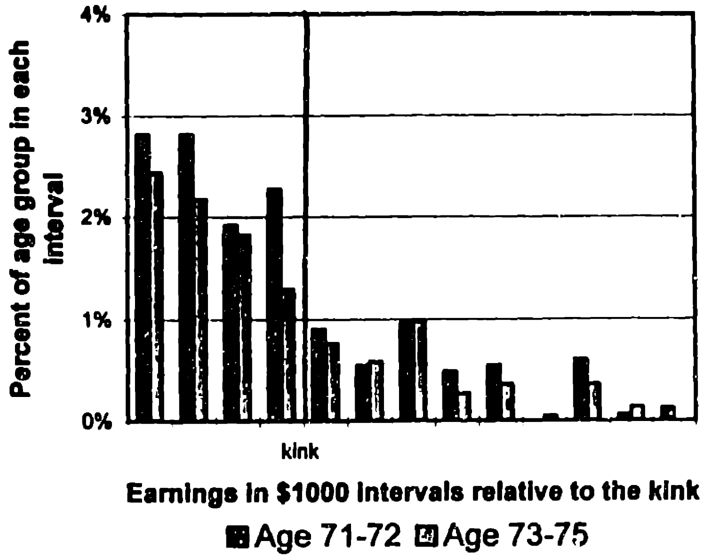


FIGURE 4-C
1984-86

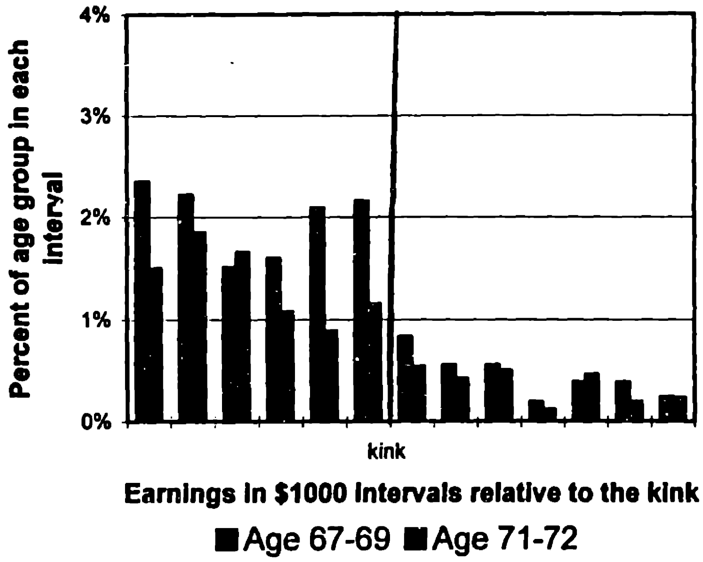
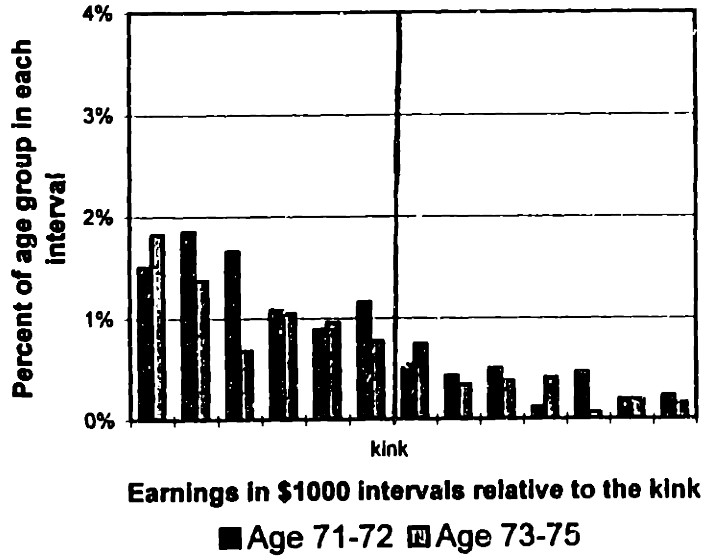


FIGURE 4-D
1984-86



older control group who do not face the earnings test and whose earnings decline smoothly over the same range.

Figures 4-C and 4-D make the same comparisons after 1983. Now, the earnings of the treated group decline much more smoothly over the range of the earnings test kink. The pattern of their earnings closely resembles that of the older group, while the younger group displays continued, though reduced, bunching at the kink. The reaction to the 1983 change shows as dramatically as the 1978 change how the kink governs people's behavior.

The same pattern is quantified in the second section of Table 2. The table compares the treatment group (in the middle panel) with the younger and older groups. In the first three columns, representing those who face the earnings test, the percentage located just before the kink is large: 4.1% of the younger group before 1983, 2.2% of the younger group after 1983, and 2.3% of the treatment group before 1983, versus 0.8-1.3% for the groups not subject to the earnings test. The treatment group after 1983 closely resembles the older control group, who display a smooth decline in earnings over the entire range of the earnings test kink both before and after 1983. In this case, the importance of using both comparison groups becomes clear. Comparing the treatment group to the 67-69 year olds could yield an ambiguous result because the younger group reduced their bunching as well. On the other hand, after 1983 the 71-72 year olds much more closely resemble the older group than the younger group as before.

One complication is that the location of the tax kink has the potential to interfere with the earnings test kink over this period. Before 1983, the initial tax kink for a married person with zero nonlabor income was located at $5400, while the location of the earnings test kink was $5000 in 1980 and $5500 in 1981. Therefore, some of the bunching for all age groups before 1983 might be bunching at the income tax kink by individuals with very low nonlabor income.[19] By 1984 the earnings test kink had risen to $6960, more than $1,000 above the initial tax kink, so the problem no longer arises.

There are two ways to handle this issue. One way is to rely on the differences in

[19] In 1980, with the initial income tax kink at $5400 and the earnings test kink at $5,000, the range of overlap occurs when nonlabor income is between $400 and $999. In 1981, with the earnings test kink at $5500, the range occurs when nonlabor income is between $0 and $899.

differences analysis described above. Since the income tax treatment is the same across age groups, the comparisons will difference out the effect of the tax kink. Another way is to be more precise about accounting for the influence of the tax kink, although the degree of precision is limited by the greater potential for measurement error in imputing the location of the tax kink.[20] Nevertheless, several results taken together make a convincing case that taxes do not confound the analysis above. First, there does not appear to be significant bunching at the initial tax kink -- the number of individuals within $1000 below the initial tax kink is never more than one-third of the number within $1000 below the earnings test kink and causes no visible spike in the distribution. Second, 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 the kinks might overlap. Third, only a small portion of the sample have small enough nonlabor income that the initial tax kink remained within $1000 below the earnings test kink -- 2% of those who worked in 1980 and 9% in 1981, with workers making up about 20% of this sample. Therefore, the reaction to the tax kink does not appear to interfere with the earnings test analysis in any way.

The analysis of earnings patterns continues with Figure 5, focusing on the 1990 decline in the tax rate for 65-69 year olds. The earnings of 63-64 year olds and 67-69 year olds in 1988-89 appear in Figure 5-A, while their earnings in 1991-92 are in Figure 5-B. In both, we see the familiar piling up at the kink. Yet, comparing the two graphs, it is difficult to detect a change in the degree of bunching by the treated group relative to the younger control group, who face unchanged conditions. The decline in the tax rate to 33% represents a smaller change in incentives for those at the kink then the 1983 and 1978 changes, when the tax rate at that point went to zero. Still, the apparent lack of reaction places a bound on the underlying elasticity of labor supply. The third section of Table 2 tabulates this data, again displaying the lack of a significant reaction on the part of the 67-69 year olds to the change in the tax rate.

Figure 6 examines the 1973 experiment when the 100% "tax bracket" of the earnings

---

[20] The location of the income tax kinks for an individual depends on other family income, deductions, and filing status. I assume an individual takes the standard deduction and that he files jointly if he is married.

FIGURE 5-A
1988-89

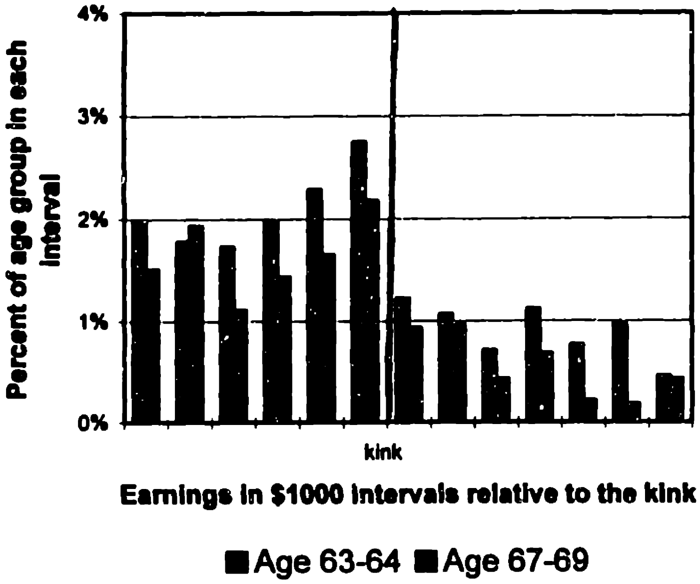
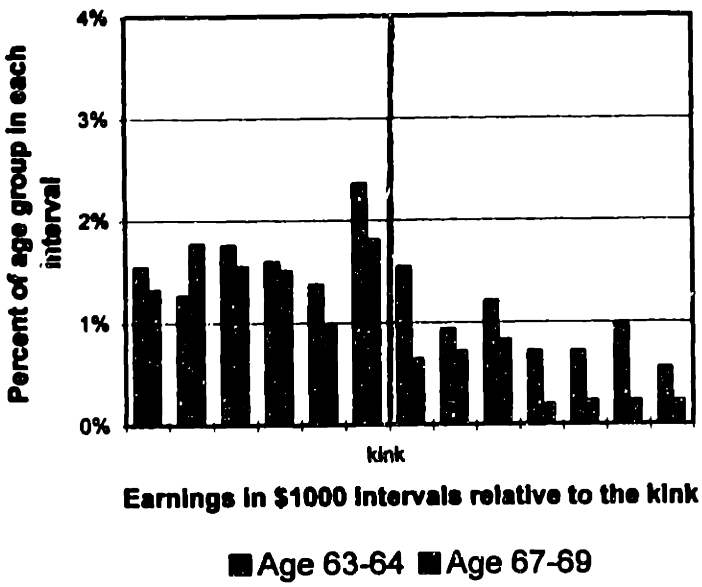


FIGURE 5-B
1991-92



test was abolished. It compares the earnings of everyone aged 63-71 before 1973 in Figures 6-A and 6-B and after 1973 in Figure 6-C. The two pre-1973 figures illustrate the response to the two different kinks in the budget set. Individuals bunched at the first kink of 50%, but apparently not at the higher kink of 100%, which is consistent with the theoretical prediction of less bunching at the second kink if preferences are the same. Figure 6-C shows that individuals continued to react to the 50% kink after the 100% kink was eliminated. It is apparent that the bunching is smaller after 1973, which might suggest that liberalizing the earnings test made it less binding.[21] However, retirement rates increased sharply for all ages over this period, as was shown in Figure 1. Normalizing the distributions by the number who remain in the labor force instead of by the entire sample actually shows that the magnitude of clustering was unchanged among those who work.

The absence of a reaction to the 100% kink or to its elimination in 1973 shows that this episode does not help to identify the estimates of the earnings test's impact in earlier papers that used the Retirement History Survey. Also, the sensitivity of the 1973 figures to the underlying trend in retirement points up how previous estimates of the earnings test's impact may have been clouded by other factors related to retirement -- thus the usefulness of examining the other earnings test experiments, which in turn necessitates using other data besides the Retirement History Survey.

B. The Participation Decision

The bulk of the analysis in this section focused on the earnings distribution of working individuals, for whom the predictions arising out of the earnings test are sharpest. In this section, I examine whether the earnings test has an effect on participation. If there are fixed costs of work causing a discontinuity in the budget set at zero hours, then liberalizing the earnings test would encourage labor force participation.[22] Furthermore, for

---

[21] Removing the 100% tax kink will cause less bunching at the 50% tax kink only in the presence of fixed costs or some other type of rigidity in labor supply choices.

[22] Reimers and Honig (1993) used the Retirement History Survey to study reentry among retirees and found some indirect evidence in favor of fixed costs.

FIGURE 6-A
1971-73

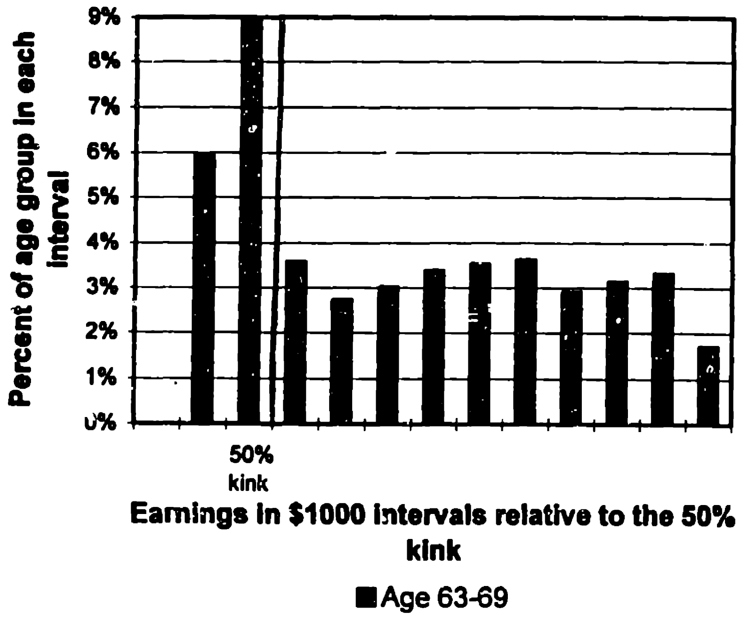


FIGURE 6-B
1971-73

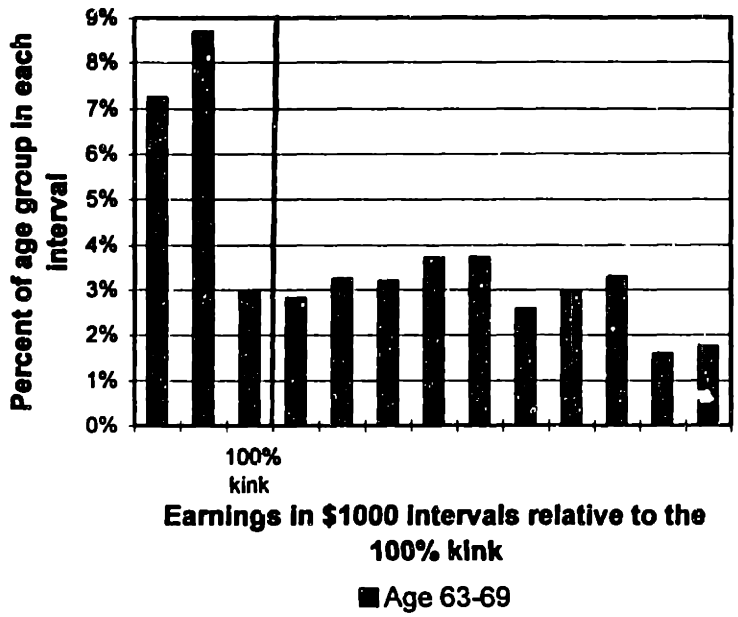
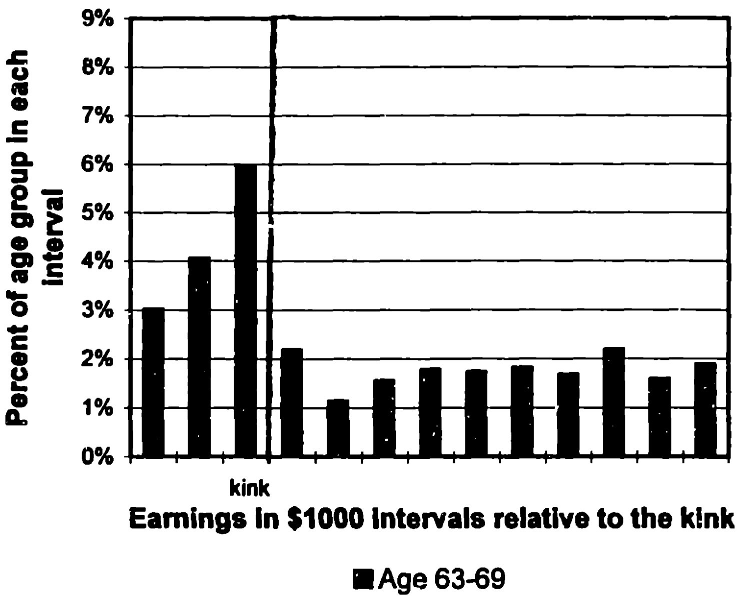


FIGURE 6-C
1974-76



a given size of fixed costs, the earnings test became less binding over time as the exempt amount was raised, so later liberalizations might have a smaller effect on participation.[23]

I compare the percentage participating before and after the different changes. The prediction is that those induced to participate will move to a point above the previous kink. This exercise is not guaranteed to isolate the effect of the earnings test because it does not indicate where people moved to on the budget set.[24] Nevertheless, if there is a pattern of increased participation following a liberalization of the earnings test only for those who experience it, then it may arise form the interaction of the earnings test with fixed costs.

Table 4 presents the percentage with no earnings before and after the different changes in the earnings test. The hypothesis that the group which experiences a liberalization of the earnings test will participate more is observed only for the 1978 change. After 1978 participation rose for both age groups, but it rose more for 67-69 year olds (5.6 percentage points) than for 63-64 year olds (3.0 percentage points), although the difference is not very significant statistically. After 1983 participation fell for all age groups. It fell slightly less for the affected 71-72 year olds (3.5 percentage points) than for the unaffected 67-69 year olds (3.7 percentage points), as it should, but more than for the unaffected 73-75 year olds (2.4 percentage points), the opposite of what it should. All those differences are far from statistically significant. In 1990, participation fell by 2.0 percentage points for 67-69 year olds, with virtually no change for the unaffected 63-64 year olds. The evidence is consistent with the existence of fixed costs of a particular size: big enough that the earnings exempt amount of around $5000 (in 1986 dollars) in the late 1970s affected the participation decision, but small enough that the higher exempt amount of around $7000 in the 1980s did not. Alternatively, labor market conditions might be changing in a way that encourages part-time work and reduces the fixed costs of working; this is an area that bears further analysis.

---

[23] It is difficult to know if fixed costs changed over time.

[24] We cannot infer whether individuals above the kink after a change arrived there from zero or from positive hours.

IV. PIECEWISE LINEAR BUDGET SET MODELLING

The previous section demonstrated substantial bunching at the convex kink generated by the earnings test. This section describes the theoretical model of labor supply choice that arises out of utility maximization subject to a piecewise linear budget constraint, which will capture the bunching at the kink as well as reactions along other parts of the budget set which are more difficult to detect in the raw data. This will permit broader conclusions about the earnings test, while providing labor supply estimates for the elderly that are of more general interest.

The section begins with the background of the piecewise linear budget set approach. It then develops the underlying model of labor supply choice as subject to the same piecewise linear budget set that was illustrated in Section II and discusses the implementation of the maximum likelihood estimation of the model. The final subsection briefly describes the data used in the estimation.

A. Background

Most government tax and transfer schemes, such as progressive tax schedules and means-tested income programs, create nonlinearities in the budget set and discontinuities in the labor supply schedule. Ignoring them in econometric models causes misspecification, but simply using the net wage instead of the gross wage as the regressor makes the net wage endogenous because it is correlated with hours through the nonlinear tax and transfer schedule. Incorporating tax and transfer programs into the estimation is not only necessary but also useful because policy changes provide generally exogenous variation that identifies the estimates. Correctly accounting for the nonlinearities demands maximum likelihood estimation in general.[25]

The consistent treatment of the nonlinear budget set in labor supply models was

[25] Instrumental variables methods are an alternative, but no satisfactory way has been developed to treat kink observations, who face two different marginal tax rates and virtual incomes.

introduced by Burtless and Hausman (1978).[26] The methodology involves selecting a model of labor supply, incorporating stochastic variation into the model, and forming the appropriate likelihood function for each observation. The likelihood function takes into account the choice of hours over the entire exogenous tax schedule -- in this way removing the endogeneity embodied in location on a particular segment which involves a joint choice of hours and tax rate.

Among numerous other uses, the model naturally applies to the labor supply of Social Security beneficiaries. Burtless and Moffitt (1985) extended the framework by jointly modelling the decision of retirement age and post-retirement labor supply in order to estimate the impact various provisions of Social Security, including the earnings test. Gustman and Steinmeier (1986) took a similar approach, but added an option of partial retirement entailing a lower wage and incorporated a more complex utility function with more intricate sources of heterogeneity.

Even though the piecewise linear budget set methodology is a natural extension of the neoclassical theory of labor supply, its econometric application has been questioned in recent years following the work of MaCurdy, Green and Paarsch (1990). They showed that proper specification of the log likelihood over the region of a convex interior kink is only possible if the estimated coefficients obey the condition that the compensated substitution effect is positive. They found that the data tended to have problems meeting that condition, in which case the income coefficient must be constrained to be negative.[27] This, they argued, is a sign of misspecification that they attribute to oversimplifying the labor supply choice. On the other hand, if the condition is legitimately met by the data, then the estimation procedure imposes nothing on the data, a point made by Blomquist (1995). MaCurdy, Green and Paarsch did not examine whether the data always fail to meet the necessary conditions, and if so what elements of the theoretical setup are in disagreement with the data; or whether the problems generally lie in the implementation of the estimation, and if so what are reasonable

---

[26] Moffitt (1986) reviews the technical details of the maximum likelihood procedure.

[27] If the compensated substitution effect is not positive for one or more observations, then they will have a negative probability of locating at the kink.

remedies.[28]

There are reasons to suspect that some applications are more problematic than others, and that the earnings test may be less problematic than most. The earnings test, in contrast to the federal tax system, creates a simple budget set with a clearly-defined kink that is apparently well-known. As evidence, we observe much more bunching in reaction to the earnings test than in most other instances when the budget set is piecewise linear. Also, the location of the kink is easily determined, while the location of the kinks induced by the federal tax schedule must be imputed as a function of other family income, marital status, and potential deductions. Finally, there is considerably more variation in the labor supply of the elderly than in the labor supply of prime-age individuals, which aids in identification.

As it turns out, estimating the model for the earnings test produces no such problems as MaCurdy, Green and Paarsch described. The compensated substitution elasticity is positive for all observations without having to impose any restrictions on the estimated coefficients. This suggests that their findings do not condemn all applications of the piecewise linear budget set method.

B. The Labor Supply Model

A beneficiary's choice of hours H is assumed to obey the standard linear model:

$$
\begin{aligned}
(1)\quad H(w,Y,X) &= \kappa + X\beta + \gamma w(1-\tau) + \delta Y_V + \alpha \\
&= Z\theta + \alpha
\end{aligned}
$$

where w is the gross wage, $w(1-\tau)$ is the net wage, Y_V is virtual income, X are demographic characteristics that influence labor supply; κ, β, γ, and δ are parameters; and α is a random

[28] One alternative employed by Blundell, Duncan and Meghir (1995) avoids requiring the condition of a positive compensated substitution elasticity by excluding observations near the convex kink and adding a selection term in a linear instrumental variables regression to account for the exclusion. Their solution is unappealing in this case because the observations at the kink provide a lot of information about the earnings test. Another alternative being explored by Blomquist and Newey (1996) retains the implications of utility maximization theory but dispenses with functional form and distributional assumptions by using nonparametric estimation methods.

variable which represents unobserved heterogeneity in preferences and is assumed to be uncorrelated with the other right-hand side variables.

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

$$\text{(2)}\quad \begin{array}{ll} \tau = \tau_1 \equiv 0 & \text{if earnings} < \text{E;} \\ Y_V = Y_{V1} \equiv \text{non-labor income} & \\ & \\ \tau = \tau_2 \equiv \tau_{SS} & \text{if E} < \text{earnings} < \text{E} + \text{SS}/\tau\text{;} \\ Y_V = Y_{V2} \equiv Y_{V1} + E^{*}\tau_{SS} & \\ & \\ \tau = \tau_3 \equiv 0 & \text{if E} + \text{SS}/\tau < \text{earnings.} \\ Y_V = Y_{V3} \equiv \text{non-labor, non-SS income} & \end{array}$$

The structure in (2) makes apparent the correlation of the net wage and virtual income terms with hours as hours rise along the budget constraint.

Hausman and others formulated a maximum likelihood procedure for estimating (1) and (2) by adding distributional assumptions. If α in equation (1) is assumed to be fixed for an individual but distributed over the population as $N(0,\sigma_\alpha^2)$, then the model can be estimated using a tobit procedure. The log likelihood function expresses the probability of α occurring such that the individual wishes to locate on the segment or kink where he is observed, as follows:

$$\text{(3)}\quad \log l(H_i) = S1_i * \log\left[\frac{1}{\sigma_\alpha} \phi\left(\frac{\alpha_i = H_i - Z_{1i}\theta = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_1) - \delta Y_{V1\,i}}{\sigma_\alpha}\right) \right] \qquad \textit{(lower segment)}$$

$$+ K_i * \log\left[\int_{Hi-Z1i\theta}^{Hi-Z2i\theta} \frac{1}{\sigma_\alpha} \phi\left(\frac{\alpha_i}{\sigma_\alpha}\right) d\alpha_i \right] \qquad \textit{(kink)}$$

$$+ S2_i * \log[\frac{1}{\sigma_\alpha} \phi(\alpha_i = \frac{H_i - Z_{2i}\theta}{\sigma_\alpha})] \qquad \textit{(middle segment)}$$

$$+ S3_i * \log[\frac{1}{\sigma_\alpha} \phi(\alpha_i = \frac{H_i - Z_{3i}\theta}{\sigma_\alpha})] \qquad \textit{(upper segment)}$$

$$+ K0_i * \log[\int_{-\infty}^{H_i - Z_{1i}\theta} \frac{1}{\sigma_\alpha} \phi(\frac{\alpha_i}{\sigma_\alpha}) d\alpha_i] . \qquad \textit{(zero hours)}$$

$\phi(.)$ is the standard normal probability density function. Each element of (3) is multiplied by an indicator function identifying the part of the budget set where an individual is located and expresses the likelihood of the stochastic term α conditional on the individual's choice of hours and characteristics Z and on the particular budget parameters denoted by the subscript on Z. Desired hours are observed when they occur along one of the budget segments, but not at the convex kink or at zero hours. The term for an individual with zero hours describes the likelihood of desired hours occurring anywhere below zero. The term for an individual at the kink expresses the condition that desired hours are below the kink given the budget parameters of the middle segment and are above the kink given the budget parameters of the lower segment, neither of which is feasible.

A few more details will make (3) operational. To begin, the major features of the income tax system can be incorporated into (2) and (3) in a straightforward fashion. The initial income tax kink is the only one of a comparable magnitude to the earnings test during the period when the model is estimated, while adding all the additional small tax kinks would increase the demands on the maximization procedure and subject it to considerably more measurement error without adding a great deal of meaningful information.[29] Since bunching at the initial tax kink is minor at best, and bunching at any other kink is not discernible, only the initial tax kink is modelled. FICA taxes are added to the formulation as well and assumed to apply to all earnings, a minor simplification in this sample that includes few high earners. The details of implementing the estimation with taxes are discussed in

[29] Earnings and other income and demographic variables which determine the tax bracket each person is in become part of the log likelihood, making the model sensitive to measurement error, as noted by Heckman (1982).

Appendices 1 and 2.

Another issue to resolve is whether to estimate (3) on all individuals or whether to truncate nonworkers. Truncating is desirable in this situation because it focuses the estimation on the exogenous modifications in the budget set, all of which occur away from the zero hours point, and it reduces the burden of explaining the decision to be at zero hours. Truncating is also practical because observed instead of imputed wages may be used. According to Heckman and MaCurdy (1981), the usual sample selection procedure to impute the wage yields inconsistent estimates in the piecewise linear budget setting.[30] The modification to the log likelihood to account for truncation is also presented in Appendix 2.

A final issue involves determining who is at the kink in order to properly specify their log likelihood. Individuals are massed around the earnings test kink, but only some are located exactly on the kink, with many more in a small range below. The practice in such a situation is to assign observations from a small band on both sides of the kink to the kink itself as a way of accounting for the slight degree of optimization or measurement error.[31] The earnings distributions in Section III, however, showed no sign of bunching just above the convex kink, so guided by that analysis, I assign observations occurring slightly below the kink to the kink itself using the $1000 interval suggested by the earlier analysis.

Equation (3), with the enhancements just mentioned, will be estimated on a sample of working married men aged 66-75 in the years before and after the 1983 elimination of the earnings test for 70-71 year olds. The sample includes the treated group along with two comparison groups -- younger men who face the same budget set as 70-71 year olds before 1983, and older men who face the same budget set after 1983. The availability of two

---

[30] The correct way to handle missing wages in the nonlinear budget set model, according to Heckman and MaCurdy, is to integrate over the stochastic variation in the imperfectly observed wage as well as over the other sources of randomness. However, this method is never carried out in practice.

[31] Burtless and Moffitt assigned individuals within one hundred annual hours of the earnings test kink to the kink. Other papers without obvious bunching, such as those focusing on income taxes, instead formalize measurement/optimization error by assuming that (1) expresses desired hours, with observed hours equal to desired hours plus another normally distributed stochastic term. Estimation of the two-error version was attempted here but did not succeed because there is not enough behavior that resembles a distinction between desired and observed hours to identify its variance.

comparison groups serves to identify the two elasticities of interest.

C. The Data

The data are extracted from March Current Population Surveys for the three years before and after the 1983 change. Only men are included, since their labor force participation rate is higher than that for women. The sample excludes individuals who report self-employment income and individuals who report negative non-labor non-Social Security income because their labor supply might be quite dissimilar. It also excludes men who are not working but report receiving no benefits, since they do not face the earnings test and must be postponing recipiency for a reason unrelated to the earnings test.

An issue in specifying the correct log likelihood is that the data is annual, but individuals are likely to retire during the year. A mid-year retiree locates on different parts of the budget set before and after retirement, but his annual earnings and benefit levels will determine where on the budget set I assign him for the entire year. Most analyses on the earnings test and on retirement have been carried out on an annual basis. Few data sets follow the monthly sequence of work and Social Security receipt, making it difficult even to document what is typical behavior.[32] Until more evidence can be gathered, the possibility of misassigning some individuals must be accepted. The misassignment will only lead someone who is not at the kink to appear there with a very small probability. It will not be systematically related to the earnings test rules changes and should not alter the conclusions about the earnings test, which are shown to rely heavily on the response to the rules changes.

The measurement of the wage and other income terms used in the estimation are described in Appendix 1, and sample statistics for the variables of interest are reported in Table 5. 80% of this sample does not work. Within the subsample that works, 57.4% located below the earnings test kink, 8.5% on the kink, 16.7% on the middle segment above the kink, and 17.4% on the upper segment. The annuals hours they report is the variable to be explained by the maximum likelihood model as a function of their reported net wage,

[32] Some evidence might be obtained by matching individuals across March Current Population Surveys. Another possibility is to use the Survey of Income and Program Participation.

virtual income, and other personal characteristics.

V. RESULTS OF THE STRUCTURAL ESTIMATION

A. Estimates of the Labor Supply Model

The estimates for the model are reported in the first column of Table 6. The coefficients in Table 6 have the predicted signs, and they are estimated quite precisely, except for the constant. The estimated wage coefficient indicates that a $1 increase in the real wage would lead to an increase of 52 hours worked per year, which implies an uncompensated wage elasticity of 0.357 at the sample mean real net hourly wage of $7.94 and mean annual hours of 1156 for married men. The estimated income coefficient implies that a $1000 increase in non-labor income would lower annual hours by 37, yielding an income elasticity of -0.758 at the sample mean real virtual income of $23,543. The elasticity estimates are fairly large compared to many in the labor supply literature. This is to be expected, since the elderly display more variation in hours than prime-age individuals, and since the earnings distributions in Section III demonstrate that beneficiaries react to the earnings test.

The large elasticities indicate significant deadweight loss from the distortions imposed by the earnings test. Exact measures of deadweight loss are available by using the unique indirect utility function corresponding to the linear labor supply model to calculate the consumer surplus and compensated labor supply, as in Hausman (1981). The estimates imply a deadweight loss for the average individual on the kink of $3087.[33] The deadweight loss for the average individual on the middle segment, who is facing a 50% tax

[33] To calculate deadweight loss, first compute each individual's utility without the earnings test, given the parameter estimates; then compute each individual's utility with the earnings test and compute how much income each would have to receive to get the same utility with the earnings test as without, which is the consumer surplus; and consumer surplus minus the compensated tax revenue (benefits lost because of the earnings test according to the compensated substitution elasticity) for each individual is the deadweight loss.

rate, is $8502, compared to revenue collected from him (i.e., reduced benefits) of $4636.[34] Thus, the earnings test imposes substantial efficiency costs relative to the government's savings in the form of reduced benefits.

The estimates in Table 6 control for age and education, finding the expected negative effect of the former and the usual positive effect of the latter. The estimates also condition on marital status by restricting the sample to married men. Estimating the model on the entire sample and including a dummy variable for marital status led to substantial correlation among the covariates that made it difficult to converge to estimates for their coefficients.[35] On the other hand, the estimates of the wage and income terms are insensitive to whether the model is estimated on the married or on the whole sample.

The basis for identifying the structural estimates was the exogenous variation in the earnings test over time and across age groups. I performed a number of exercises to verify the importance of the earnings test variation in determining the estimates. Estimating labor supply with the gross wage and other income -- instead of the net wage and virtual income -- using OLS yields insignificant estimates with the wrong signs for both the wage and income terms, so accounting for the shape of the budget set is crucial for arriving at the estimates. A similar point is demonstrated by the estimates in the second column of Table 6. The estimation uses other income instead of virtual income, which, in effect, turns off the income variation that the kinked budget set generates. In this case, the coefficient on the other income term is close to zero, while the influence of the earnings test variation gets channelled into the wage term. As a final check, I estimated the piecewise linear budget set model on the sample separated before and after 1983 to gauge the importance of the earnings test variation over time. The estimates for the pre-1983 sample failed to converge, indicating insufficient variation among the explanatory variables, even though the variation of the budget set across age groups remains. The estimates for the post-1983 sample did converge,

---

[34] Individuals on the upper segment of the budget set experience no deadweight loss because they face only an income effect from losing all their benefits, with no substitution effect.

[35] Highly correlated explanatory variables cause big changes in the affected coefficient estimates accompanied by very small improvements in the log likelihood as the algorithm attempts to find a maximum on a very flat region of the parameter space.

but the wage term was substantially smaller while the income term was similar but less precisely estimated.[36] Those exercises reveal the importance of the earnings test variation in determining the estimates.

An additional specification check was performed by allowing the Delayed Retirement Credit to mitigate the effect of the earnings test. This was undertaken by separating out the earnings test tax rate from the net wage and virtual income terms to allow a distinct coefficient that would be less than one if the Delayed Retirement Credit matters. The estimated coefficient did not drop below one, so the Delayed Retirement Credit does not appear to play a role in the response of individuals to the earnings test.

To get an idea of the model's fit, Table 7 compares actual location and hours for 66-70 year olds in the first column to predicted location and hours in the second column.[37,38] I found that the model correctly predicted segment location for 66% of the sample, which indicates a good fit overall. It predicted location below the earnings test kink and onto the upper segment well.[39] However, it captured poorly the decision to locate on the middle segment instead of the kink. Why? The model is identified primarily from the hours choices of those who faced the elimination of the earnings test and clearly responded to its elimination, so this finding indicates that individuals who remain subject to the earnings test are evidently less responsive. Nevertheless, the estimates predict well the response to changes in the budget set, as shown by the comparison of actual and predicted hours in the last rows. Hours are somewhat underpredicted for 66-70 year olds located above the kink

---

[36] The wage coefficient was 39(5) versus 52(3) in Table 6, while the income coefficient was -32(8) versus -37(3).

[37] To make predictions, α must be computed for each individual not at the kink and drawn from a truncated normal distribution for the individuals at the kink. Then, the estimated coefficients and α are used to compute an individual's utility and desired hours on each of the segments and kinks of the budget set. Predicted hours are the desired hours that offers the highest utility in the set of feasible hours choices.

[38] Table 7 focuses on the 66-70 year olds in the sample, who were 65-69 in the year when the earnings occurred, because the results will be compared to simulations of policy changes that would apply to 65-69 year olds.

[39] Among those located below the kink, the model correctly predicted for 94% of the sample the observed choice relative to the initial tax kink.

who do not face a change in the earnings test, with 1746 predicted versus 1812 in reality. In contrast, hours are predicted quite closely for 71-72 year olds, with 1027 hours predicted before the earnings test was eliminated versus 1063 actually and 956 hours predicted after the earnings test was eliminated versus 960 actually. This suggests that the estimates will be able to predict quite reliably the response to substantive changes in the budget set.

B. Simulations

An important goal of this exercise was to be able to predict the impact of other changes in the earnings test. The labor supply estimates in Table 6 allow various policies to be simulated. The simulations also give a more complete sense of what the coefficient estimates imply, since elasticity calculations pertain only to marginal changes away from the kinks. I will use the estimates to simulate the effect of removing the earnings test for 65-69 year olds and the effect of raising the exempt amount to $30,000, which has been proposed recently in Congress. The first simulation simply repeats the 1983 change for another age group, while the second simulation amounts to doing so for many of them, so the estimates reported for the 1983 sample should be quite applicable. The results of the simulations are shown beginning in the right panel of Table 7.

Since the earnings test has been altered since 1983, the simulations are best compared to those identified as the benchmark predictions, where the earnings test parameters are updated to resemble the rules in 1995.[40] The benchmark predicts mean hours of 1197 for all working individuals and 1706 for individuals located at or above the kink.[41]

Simulation A predicts the impact of eliminating the earnings test. When the earnings test is removed, individuals at the kink will increase their hours because of the positive substitution effect; the impact of individuals on the middle segment may shift from positive to negative as the income effect increases along the segment; and individuals on the upper

[40] In 1995, 65-69 year olds faced a tax rate of 33% instead of 50%, while the floor was about 20% higher in real terms.

[41] The separate effects of lowering the tax rate and raising the floor are predicted to have opposite effects on hours, with the former predicted to raise hours 1.4% and the latter predicted to lower hours 1.9%.

segment will lower their hours because of the positive income effect. In the aggregate, removing the earnings test would raise mean hours for individuals at or above the kink initially from 1706 to 1941, a 13.8% increase. Individuals at the kink and on the middle segment are increasing their hours by more than that, since the individuals on the upper segment are decreasing their hours. On net, the positive substitution effect from removing the earnings test dominates the positive income effect. Also, the positive effect on hours grows for younger individuals: the model predicts a bigger positive effect for 66-70 year olds relative to 71-72 year olds, and therefore 62-64 year olds might be expected to display an even larger reaction to eliminating the earnings test. Moreover, 62-64 year olds still face a more restrictive earnings test, so the overall effect for them could be considerable.

The fiscal cost of eliminating the earnings test, while substantial in the first year, will be mostly made up in the medium-run.[42] As Honig and Reimers (1989) pointed out, the offset arises because the actuarial adjustment for 62-64 year olds and the Delayed Retirement Credit for 65-69 year olds would not be granted to beneficiaries who lose benefits to the earnings test, so their future benefits would be smaller than they would be with the earnings test. 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 fiscal cost of eliminating the earnings test today will be cancelled out within a few years by the absence of an upward adjustment in future benefits.[43] In conclusion, eliminating the deadweight loss experienced by working beneficiaries because of the earnings test will soon be possible at almost no cost to the Social Security Administration.

---

[42] Leonesio (1993) reported Social Security Administration forecasts that eliminating the earnings test for 65-69 year olds would raise benefit payouts to current beneficiaries by $4.3 billion in the first year, while income, payroll and benefits taxes due to higher earnings would offset 14.8% of the cost. The labor supply elasticities they used were taken from Hanoch and Honig (1983) and are 0.17 for the uncompensated wage elasticity and virtually zero for the income elasticity. The elasticities estimated here would lead to a substantially larger offset through taxes paid because of a bigger boost to labor supply.

[43] The only thing that will prevent a complete offset once the Delayed Retirement Credit is actuarially fair on average is adverse selection: people with a shorter than average life expectancy will be less constrained to postpone filing for benefits than they were with the earnings test in place.

Lastly, Table 7 reports the results of simulating an increase in the exempt amount to \$30,000. Interestingly, aggregate hours for those at or above the kink are predicted to rise by only 2.3%, from 1706 to 1746. Contrary to the intuition that raising the exempt amount so much would eliminate the earnings test in practice for most individuals, it would simply make it binding for many with higher earnings. For those at the kink, hours would rise exactly as they would if the earnings test is eliminated, but the situation changes for those initially on the upper segment. While eliminating the earnings test produces a negative effect on their hours through the income effect, raising the exempt amount -- which would induce many of them to locate on the middle segment -- would make them suffer a negative substitution effect as well, so they would lower their hours by more, and many would end up bunched at the kink. This result may be preferable from a distributional perspective but would not reduce aggregate deadweight loss by much.

VI. CONCLUSIONS

The earnings test has been the subject of substantial popular attention, but less academic interest. 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 over the last twenty years altered the budget set for beneficiaries of certain ages and not for beneficiaries of 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 forces behind the bunching. This strategy is implemented in a number of ways.

The first approach is to focus on the strongest implication of the theory -- that we should observe bunching at the convex kink induced by the earnings test -- in order to gauge the earnings test's effect. The data showed that individuals 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 individuals to hold down their labor supply.

As strong as the behavior around the exempt amount is, it is only one way that the earnings test may affect labor supply. The second approach of this paper is to develop and estimate a structural model of labor supply that characterizes behavior along the entire budget set. The estimates yield income and substitution elasticities that are informative about the labor supply of the working elderly more generally, which is a distinct focus compared to most of the previous research on the elderly emphasizing retirement behavior.

A typical criticism of structural estimation, however, is that it pays too little attention to how the parameter estimates are identified. This is where the variation in the earnings test rules, causing substantial changes in the budget faced by working beneficiaries, is most important. These policy shifts amplify the exogenous variation in the right-hand side variables. The substitution and income elasticities that result are relatively large, suggesting substantial deadweight loss from taxation of older workers. The estimates also imply that policies which affect hours choice conditional on working will influence the success of other policies that aim to increase the self-sufficiency of the elderly by encouraging them to stay in the labor force longer.

Another benefit of estimating a formal model paper is to use the estimates to predict the effect of various policies which might be enacted to liberalize the earnings test. Thus, it was interesting to find that the effect on hours of almost tripling the level of the exempt amount, as Congress is considering, would be quite small in the aggregate because the positive effect on hours for low earners would be mostly offset by a severe negative effect for high earners. On the other hand, the model predicts that eliminating the earnings test for 65-69 year olds would raise aggregate hours substantially -- by 13.8% -- with an even greater increase for individuals at and near the kink. Furthermore, the indications are that younger beneficiaries would respond even more favorably. The fiscal cost of such a measure over the medium run would be relatively low because the Delayed Retirement Credit would not be activated, and the cost will approach zero as the Delayed Retirement Credit becomes actuarially fair on average. It should also lead to an increases in the lifetime resources of those who continue to work into their relative old age because of a strong substitution effect in favor of work.

The results show the importance of the earnings test in determining the labor supply behavior of older workers and have further implications for the design of means-tested benefits, which raise the marginal income tax rate at certain income levels. The results also show the usefulness of the substantial changes in the earnings test rules in exposing the underlying behavior. This source of variation may illuminate some other questions as well. The earnings test may interact with fixed costs and the availability of part-time work to influence the decision to participate in the labor force, so the variation in the earnings test can provide some insight into the size of fixed costs and into the evolution of more flexible work arrangements in recent years. Another area to explore is the interaction between the earnings test and the optimal timing of the decision to file for benefits, which has received much less attention than the decision to retire but is the direct link between the retirement decision and the fiscal burden of Social Security. In sum, the earnings test provides a useful tool when considering the decisions of the elderly who continue to work.

TABLE 1: Parameters of the Earnings Test						
	Floor, in 1986 $			Tax Rate		
Age:	62-64	65-69	70-71	62-64	65-69	70-71
1972	4408 7551	4408 7551	4408 7551	50% 100%	50% 100%	50% 100%
1973	5184	5184	5184	50%	50%	50%
1974	5335	5335	5335	50%	50%	50%
1975	5134	5134	5134	50%	50%	50%
1976	5316	5316	5316	50%	50%	50%
1977	5426	5426	5426	50%	50%	50%
1978	5446	6724	6724	50%	50%	50%
1979	5254	6793	6793	50%	50%	50%
1980	4948	6650	6650	50%	50%	50%
1981	4919	6631	6631	50%	50%	50%
1982	5043	6815	6815	50%	50%	50%
1983	5414	7263	-	50%	50%	-
1984	5443	7342	-	50%	50%	-
1985	5500	7456	-	50%	50%	-
1986	5760	7800	-	50%	50%	-
1987	5789	7873	-	50%	50%	-
1988	5670	7782	-	50%	50%	-
1989	5727	7849	-	50%	50%	-
1990	5736	7849	-	50%	33%	-
1991	5697	7822	-	50%	33%	-
1992	5812	7968	-	50%	33%	-
1993	5825	8010	-	50%	33%	-
1994	5938	8242	-	50%	33%	-
1995	5868	8112	-	50%	33%	-

* The dashed lines indicate that the earnings test rules were changed. The 1983 change was originally slated for 1982 but was delayed in 1982 for one year.
Source: Annual Statistical Supplement to the Social Security Bulletin.

TABLE 2: Number of Individuals Affected by the Earnings Test, 1989	
Number of retired worker beneficiaries aged 62-69	9,778,596
- who had benefits withheld	926,342
aged 62-64	168,782
aged 65-69	757,560
- who had earnings within 90-110% of earnings test floor	
aged 65-69	173,700
Number who have not claimed benefits and still work	
aged 65-69	582,000
Number of dependents and survivors who had benefits withheld	314,938

Source: Leonesio, *Social Security Bulletin*, 1990. Bondar, *Social Security Bulletin*, 1993.

TABLE 3: Percentage of Sample in Earnings Intervals

1978 Change	63-64 year olds			67-69 year olds		
		1978-81			1979-81	
Earnings intervals	1975-77	62-64 Kink	65-71 Kink	1975-77	62-64 Kink	65-71 Kink
-$3000 to -$2000	.035	.022	.027	.043	.027	.029
-$2000 to -$1000	.036	.029	.034	.038	.029	.034
-$1000 to kink	.045	.036	.018	.058	.033	.045
kink to +$1000	.021	.019	.015	.018	.037	.014
+$1000 to +$2000	.017	.015	.012	.007	.028	.008
+$2000 to +$3000	.016	.013	.014	.007	.010	.010
Difference at kink	.024 (.006)	.016 (.005)	.003 (.004)	.041 (.006)	-.004 (.005)	.030 (.005)

1983 Change	67-69 year olds		71-72 year olds		73-75 year olds	
Earnings intervals	1980-81	1984-86	1980-81	1984-86	1980-81	1984-86
-$3000 to -$2000	.027	.016	.028	.011	.022	.010
-$2000 to -$1000	.031	.021	.019	.009	.018	.010
-$1000 to kink	.041	.022	.023	.012	.013	.008
kink to +$1000	.013	.008	.009	.005	.008	.007
+$1000 to +$2000	.007	.006	.005	.004	.006	.003
+$2000 to +$3000	.012	.006	.010	.005	.010	.004
Difference at kink	.028 (.006)	.016 (.003)	.014 (.006)	.006 (.004)	.005 (.004)	.001 (.003)

1990 Change	63-64 year olds		67-69 year olds	
Earnings intervals	1988-89	1991-92	1988-89	1991-92
-$3000 to -$2000	.020	.016	.014	.015
-$2000 to -$1000	.023	.014	.016	.010
-$1000 to kink	.027	.024	.022	.018
kink to +$1000	.012	.015	.009	.006
+$1000 to +$2000	.011	.009	.010	.007
+$2000 to +$3000	.007	.012	.004	.008
Difference at kink	.015 (.006)	.008 (.006)	.012 (.005)	.012 (.004)

Each cell shows the percentage of people in that age group who have earnings in that $1000 interval. Standard errors are in parentheses or, where omitted, they are 0.001 to 0.004. The data is from March Current Population Surveys.

TABLE 4: Percentage of Sample Participating in the Labor Force						
1978 change						
	63-64 year olds		73-75 year olds			
	1975-77	1980-81	1975-77	1979-81		
Participation Rate	.525 (.009)	.555 (.005)	.257 (.007)	.313 (.007)		
Change		.030 (.012)		.056 (.010)		
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
Did not participate	.311 (.008)	.275 (.007)	.216 (.010)	.181 (.008)	.180 (.008)	.156 (.006)
Change		-.037 (.010)		-.035 (.012)		-.024 (.010)
1990 change						
	63-64 year olds		67-69 year olds			
	1988-89	1991-92	1988-89	1991-92		
Did not participate	.541 (.011)	.545 (.012)	.304 (.009)	.284 (.009)		
Change		.003 (.016)		-.020 (.012)		

Participation is defined as having positive earnings last year. Standard erros are in parentheses. The data is computed from March Current Population Surveys.

TABLE 5: Summary Statistics

	whole sample	hours>0
Number of observations	23,889	4,876
Number aged 66-70*	13,453	3,435
Number aged 71-72*	4,544	676
Number aged 73-75*	5,892	765
Annual hours	237 (604)	1162 (844)
Gross hourly wage	-	10.68 (10.34)
Net hourly wage	-	7.69 (8.04)
Non-labor income	21,052 (17,773)	20,802 (19,131)
Social Security income	6,080 (2,692)	5,201 (3,339)
Location on the budget set: zero hours	0.796	-
below initial tax kink	0.017	0.081
at initial tax kink	0.005	0.024
lower segment above initial tax kink	0.096	0.469
earnings test kink	0.017	0.085
middle segment	0.035	0.167
upper segment	0.035	0.174
Education, less than high school	0.529	0.421
completed high school	0.262	0.390
college or more	0.210	0.289
Married	0.811	0.842

Men, March Current Population Surveys, 1981-83 and 1984-87. The earnings data is retrospective for the previous year. The sample excludes those with self-employment income or negative non-labor income non-Social Security income, those with earnings less than the earnings test floor but with no Social Security income, and those with real wage less than 1 or more than 100. Variable construction is described in Appendix 1. Wage and income data are reported in 1987 dollars.

* The age pertains to March of the year following the earnings. Therefore, the treatment group of 70-71 year olds are approximately the 71-72 year olds in the data.

TABLE 6: Maximum Likelihood Estimates		
1983 change		
Dependent Variable: Annual Hours		
Dummy for completed high school	889 (136)	517 (105)
Age-65	-173 (26)	-161 (20)
Net wage	52 (3)	71 (3)
Virtual income, $1000	-37 (3)	
Other income, $1000		-4 (2)
Constant	-144 (175)	-476 (145)
Heterogeneity standard deviation, σ_α	1820 (63)	1650 (50)
Uncompensated wage elasticity*	0.357 (0.022)	0.488 (0.021)
Income elasticity*	-0.758 (0.057)	
Log likelihood	-13844	-13917
Observations	4112	4112

The estimates are derived from maximum likelihood estimation of a tobit of hours on the variables listed in this table. The log likelihood is reported in equation (3) but augmented, as described in Appendix 2. Asymptotic standard errors are in parentheses.

* The elasticities are calculated at the sample means of 1156 hours, 7.94 net wage, and 23.543 virtual income.

The sample: Married men aged 66-75 from the March CPSs covering 1980-82 and 1984-86. The sample excluded those with self-employment income, negative non-labor non-Social Security income, and those with earnings below the earnings test floor but with no Social Security income. Sample statistics are reported in Table 5. Wage and income variables are expressed in 1987 dollars.

TABLE 7: Simulation Results

2907 66-70 year olds.*	Actual	Predicted	Benchmark Floor raised 20%, tax rate 33%	Simulation A Earnings test eliminated	Simulation B Floor raised to $30,000
Located on: budget set below earnings test kink	1261	1550	1671	2907	2383
kink	322	618	233	-	360
middle segment	620	58	965	-	161
upper segment	704	681	38	-	3
Below kink, predicted to be on: kink	-	0	0	-	0
middle segment	-	0	0	-	0
upper segment	-	0	0	-	0
On kink, predicted to be on:					
middle segment	-	0	265	-	0
upper segment	-	2	1	-	0
On middle segment, predicted to be on:					
kink	-	389	0	-	47
upper segment	-	150	0	-	0
On upper segment, predicted to be on:					
kink	-	154	2	-	290
middle segment	-	0	642	-	156
Average hours (Standard error)	1226 (16)	1216 (16)	1197 (14)	1307 (16)	1216 (15)
Average hours at or above kink** (Standard error)	1812 (20)	1746 (21)	1706 (16)	1941 (16)	1746 (16)
Average hours, 71-72,* before 1983 (Standard error)	1063 (47)	1027 (45)	-	-	-
Average hours, 71-72, after 1983 (Standard error)	960 (50)	966 (51)	-	-	-

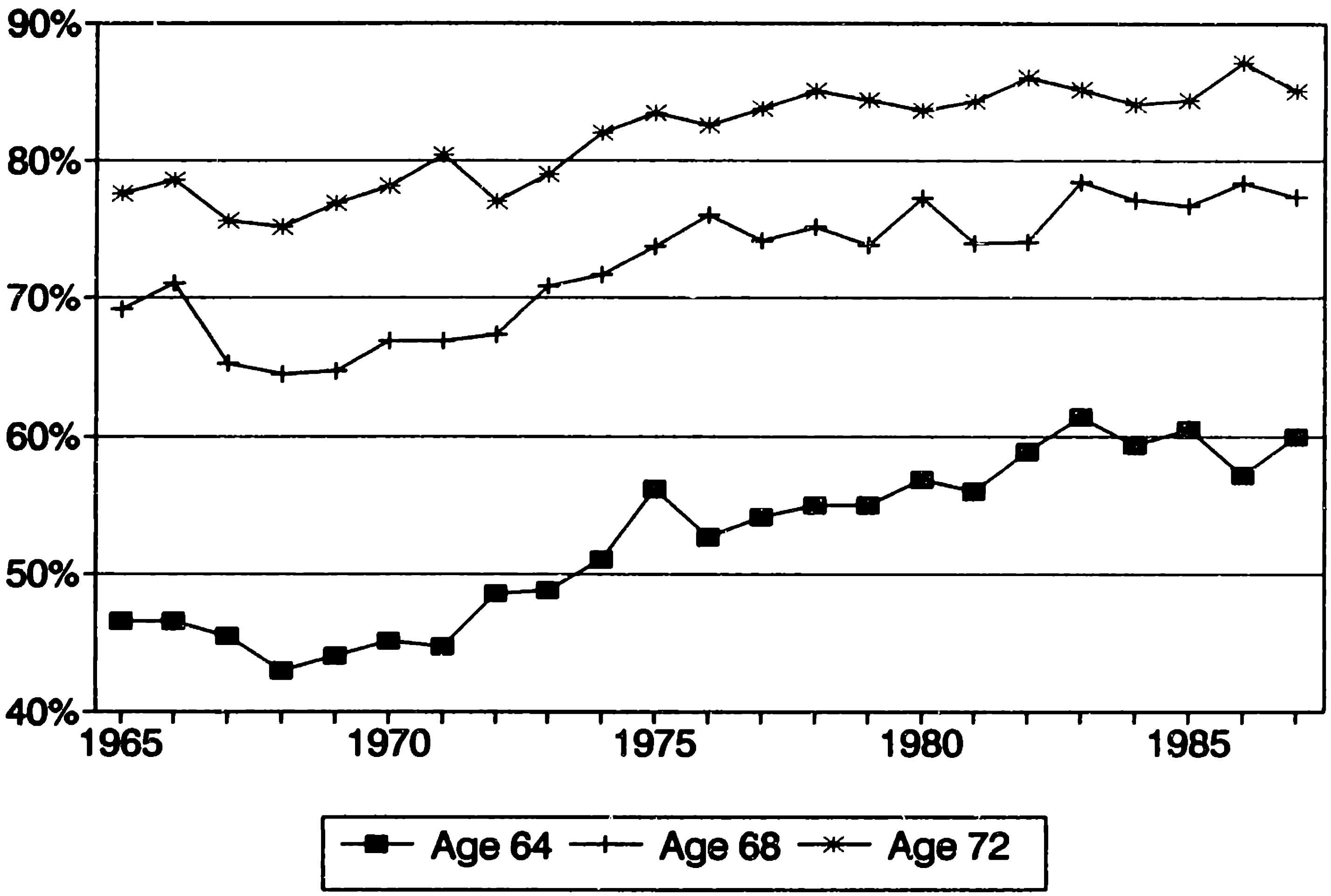
The predictions and simulations are based on the estimates in Table 6, first column. For everyone except those at the kink, the disturbance α was calculated from the estimates. For those at the kink, the limits of a truncated normal distribution were calculated and a random draw was made from that distribution. Then the coefficients and α were used to compute each individual's desired hours, earnings and utility on each segment as the parameters of the earnings test were changed. The hours choice of each individual was the feasible hours that gave the highest utility.

* Recall that those who are 66-70 (71-72) in the sample were 65-69 (70-71) when their earnings occurred. There are 573 71-72 year olds in the sample.

** For the simulations, average hours are reported for the 1200 individuals located at or above the kink, since they are the ones predicted theoretically to be affected by the simulated changes in the earnings test.

FIGURE 1

Percentage Who Report Not Working, CPS



APPENDIX 1: Data and Budget Set Parameters

The data used in this paper comes from March Current Population Surveys. The variables are defined as follows:

Variable	*Detail*	*Definition*
earnings	Dollars earned last year	Earnings from wages and salaries plus earnings from self-employment
hours	Hours worked last year	Usual hours worked last year times weeks worked last year
education, completed high school	Dummy variable	Has completed 12 years of education or more
married	Dummy variable	Married or separated
w	Real gross hourly wage	Annual earnings divided by annual hours h and deflated by the CPI to 1987 dollars
other income	Real non-labor non-Social Security income	Family income minus earnings and Social Security benefits of the individual, deflated by the CPI to 1987 dollars
y_v	Real virtual income	Defined as a function of other income, Social Security benefits, and the earnings test as detailed in equation (2) in the text

The earnings test parameters are listed in Table 1.

The tax system parameters are the following:

Location of the initial tax kink, for a married individual. The maximum of zero or the standard deduction plus two personal exemptions minus other income.

Initial tax rate. The marginal tax rate applying in the first bracket.

FICA tax rate. The employee's share of the Social Security payroll tax.

APPENDIX 2: Modifications to the Log Likelihood Function

This appendix describes some modifications to the log likelihood function defined by equation (3) in the text.

Taxes

The first modification is to add a simplified version of the income tax system. First, FICA taxes are assumed to apply to all earnings, so the FICA tax rate times the gross wage is subtracted from the wage term at each location. The initial income tax rate is modelled as well, so the income tax rate times the gross wage is subtracted from the wage term where income taxes apply. To determine that, the location of the initial tax kink is calculated for each individual, as described in Appendix 1. For individuals who face income taxes for the first hour of work, the log likelihood requires no additional alterations. For individuals who face the initial tax kink after some hours of work, an additional segment where the individual pays only FICA taxes and an additional kink where the individual chooses whether to enter the income tax system are added at the beginning of (4). In a few cases -- for those with a certain small amount of nonlabor income in certain years -- the location of the initial tax kink and the earnings test kink coincide, so for them the kinks are modelled as occurring at the same point.

Truncation

The next modification involves accounting for truncation. If the log likelihood is only estimated for working individuals, then their α's are drawn from a truncated normal distribution. The log of the probability of choosing positive hours is subtracted from each individual's log likelihood to describe the truncation.

Accounting for those who do not face the earnings test

The older individuals in the data set who are exempted from the earnings test face a different budget set. Their log likelihood omits the terms that are specific to the earnings test. They face either a single budget segment and pay FICA and income taxes, or they also face a pre-income tax segment and an income tax kink.

Estimating a two-error model

Besides estimating the model just described, a version that formalized optimization error was estimated as well. In this formulation, equation (1) in the text represents desired hours, and observed hours equal desired hours plus a normally distributed error term ϵ, which can be allowed to be correlated with α with correlation coefficient ρ, although the log likelihood below abstracts from this possibility. Under this scheme, individuals cannot be assigned to a particular segment or kink of the budget set, because their observed hours may occur on a different part of the budget set than their desired hours. The log likelihood for each observation, therefore, accounts for choices onto all parts of the budget set. The elements describe all possible outcomes of the joint probability of α and of $\alpha + \epsilon$, i.e., of desired hours and of observed hours. Thus, there is an element that describes the likelihood of desired hours occurring on each particular segment and convex kink of the budget set jointly with the likelihood of the individual's observed hours. The likelihood for zero hours is a special case. It is assumed that, if desired hours are zero, then observed hours are zero as well. Additionally, there is a probability that observed hours are zero when desired hours exceed zero because of optimization error. In this way, an individual with positive α may have a big negative draw of ϵ, and he works zero hours.[1] This formulation involves additional terms, again one for each segment and convex kink. Thus, the log likelihood of observing hours H_i, abstracting from income taxes, is:

$$
\begin{aligned}
\log l(H_i) = \eta_i * \log\Big[& \int_{v1i}^{u1i} f(\epsilon_i + \alpha_i = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_1) - \delta Y_{V1\,i},\ \alpha_i)\, d\alpha_i && \textit{(segment 1)} \\
& + \int_{u1i}^{u2i} \frac{1}{\sigma_\epsilon} \phi\left(\frac{\epsilon_i = H_i - H'}{\sigma_\epsilon}\right) \phi\left(\frac{\alpha_i}{\sigma_\alpha}\right) d\alpha_i && \textit{(kink)} \\
& + \int_{u2i}^{\alpha'} f(\epsilon_i + \alpha_i = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_2) - \delta Y_{V2\,i},\ \alpha_i)\, d\alpha_i && \textit{(segment 2)} \\
& + \int_{\alpha'}^{\infty} f(\epsilon_i + \alpha_i = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_3) - \delta Y_{V3\,i},\ \alpha_i)\, d\alpha_i \Big] && \textit{(segment 3)} \\
+ (1 - \eta_i) * \log\Big[& \int_{-\infty}^{v1i} \frac{1}{\sigma_\alpha} \phi\left(\frac{\alpha_i = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_1) - \delta Y_{V1\,i}}{\sigma_\alpha}\right) d\alpha_i && \textit{(}H^* = 0\textit{)} \\
& + \int_{v1i}^{u1i} \int_{-\infty}^{v1i} f(\epsilon_i + \alpha_i = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_1) - \delta Y_{V1\,i},\ \alpha_i)\, d\epsilon_i\, d\alpha_i && \textit{(}H=0,\ H^* \textit{ on segment 1)} \\
& + \int_{u1i}^{u2i} \int_{-\infty}^{-H'} \frac{1}{\sigma_\epsilon} \phi\left(\frac{\epsilon_i = H_i - H'}{\sigma_\epsilon}\right) \phi\left(\frac{\alpha_i}{\sigma_\alpha}\right) d\epsilon_i\, d\alpha_i && \textit{(}H=0,\ H^* \textit{ on kink)} \\
& + \int_{u2i}^{\alpha'} \int_{-\infty}^{v2i} f(\epsilon_i + \alpha_i = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_2) - \delta Y_{V2\,i},\ \alpha_i)\, d\epsilon_i\, d\alpha_i && \textit{(}H=0,\ H^* \textit{ on segment 2)} \\
& + \int_{\alpha'}^{\infty} \int_{-\infty}^{v3i} f(\epsilon_i + \alpha_i = H_i - \kappa - X_i\beta - \gamma w_i(1-\tau_3) - \delta Y_{V3\,i},\ \alpha_i)\, d\epsilon_i\, d\alpha_i \Big] && \textit{(}H=0,\ H^* \textit{ on segment 3)}
\end{aligned}
$$

where

$$
\eta_i = \begin{cases} 1 & \text{if } H_i > 0 \\ 0 & H_i = 0\,; \end{cases}
$$

$$u_{ji} = H' - \kappa - X_i\beta - \gamma w_i(1-\tau_j) - \delta Y_{Vj\,i} \qquad j = 1,2\,;$$

$$v_{ji} = -\kappa - X_i\beta - \gamma w_i(1-\tau_j) - \delta Y_{Vj\,i} \qquad j = 1,2,3\,;$$

H' is hours at the convex kink;
α' is that α which causes the individual's indifference curve to be tangent to the two segments connected by the nonconvex kink;
$f(.)$ is the joint density function of $\omega = \epsilon + \alpha$ and α, which are distributed joint normal with zero means, variances $\sigma_\omega^2 = \sigma_\alpha^2 + \sigma_\epsilon^2$ and σ_α^2, and correlation $\rho = \sigma_\omega/\sigma_\alpha$; and
$\phi(.)$ is the standard normal density function.

The first group of terms, multiplied by η, describe the probability of having positive observed hours. So, the

[1] Optimization error is meant to capture the notion that someone wants to work particular hours H^*, but can only find work that offers somewhat similar hours H. Thus, an individual might want to work a little, but he is unlucky and cannot find a job near enough to his desired hours of work, so he does not work at all. The alternative explanation of ϵ as measurement error is more consistent with the idea that H is positive if H^* is positive. That version of the model was attempted as well.

first term in the group expresses the probability of individual i's observed hours, if i's desired hours are on the first segment; the next term expresses the probability of i's observed hours, if i's desired hours are on the convex kink; etc. The second group of terms, multiplied by $1 - \eta$, describe the probability of i's having observed hours equal to zero. It begins with a term that is the probability of i's having desired hours equal to zero. The remaining terms describe the probability of i's having desired hours on a given segment or kink and observed hours equal to zero.

To make this operational, the log likelihood function must be expressed only in terms of observables and parameters to be estimated. To start, each term, which describes the joint probability of $\alpha + \epsilon$ and of α, is transformed into the marginal probability of $\alpha + \epsilon$ times the conditional probability of α given $\alpha + \epsilon$. Then, in the first group, the conditional probability of α is evaluated at the given limits of integration, which allows each term to be simplified to a normal density function times the difference of two normal distribution functions. In the second group, both the marginal probability and the conditional probability are evaluated at the appropriate limits, yielding expressions that consist of a normal distribution function times the difference between two normal distribution functions. The final step is to replace α' with an expression that involves only observables and parameters. This is done by solving for α' as that value of α which equates the indirect utility functions on the adjoining segments. Hausman (1980) derives the (unique) indirect utility function which yields the linear labor supply model.

REFERENCES

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

Blomquist, Soren. 1995. "Restrictions in Labor Supply Estimation: Is the MaCurdy Critique Correct?" *Economics Letters* 47: 229-35.

Blomquist, Soren, and Whitney Newey. 1996. "Nonparametric Estimation of Labor Supply Functions Generated by Piecewise Linear Budget Constraints." Massachusetts Institute of Technology. Mimeo.

Blundell, Richard, Alan Duncan and Costas Meghir. 1995. "Estimating Labour Supply Responses Using Tax Reforms." Institute for Fiscal Studies. Mimeo.

Bondar, Joseph. 1993. "Beneficiaries Affected by the Annual Earnings Test, 1989." *Social Security Bulletin* 56: 20-8.

Boskin, Michael J. 1986. *Too Many Promises: The Uncertain Future of Social Security*. Homewood, IL: Dow Jones-Irwin.

Burkhauser, Richard V., and John Turner. "Can Twenty-Five Million Americans Be Wrong: A Response to Blinder, Gordon and Wise." *National Tax Journal* 34: 467-72.

Burtless, Gary, and Jerry Hausman. 1978. "The Effect of Taxes on Labor Supply." *Journal of Political Economy* 86: 1103-30.

Burtless, Gary, and Robert A. Moffitt. 1985. "The Joint Choice of Retirement Age and Post-Retirement Hours of Work." *Journal of Labor Economics* 3: 209-36.

Gustman, Alan L., and Thomas L. Steinmeier. 1985. "The 1983 Social Security Reforms and Labor Supply Adjustments of Older Individuals in the Long Run." *Journal of Labor Economics* 3: 237-53.

---------- and ----------. 1986. "A Structural Retirement Model." *Econometrica* 54: 555-84.

---------- and ----------. 1991. "Changing the Social Security Rules for Work After 65." *Industrial and Labor Relations Review* 44: 733-45.

Hanoch, Giora, and Marjorie Honig. 1983. "Retirement, Wages, and Labor Supply of the Elderly." *Journal of Labor Economics* 1: 131-51.

Hausman, Jerry. 1980. "The Effect of Wages, Taxes, and Fixed Costs on Women's labor Force Participation." *Journal of Public Economics* 14: 161-94.

----------. 1981. "Exact Consumer's Surplus and Deadweight Loss." *American Economic Review* 71: 662-76.

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

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

Honig, Marjorie, and Cordelia Reimers. 1989. "Is It Worth Eliminating the Earnings Test?" *American Economic Association Papers and Proceedings* 79: 103-7.

Krueger, Alan B. 1990. *Worker's Compensation Insurance and the Duration of Workplace Injuries*. NBER Working Paper No. 3253.

Krueger, Alan B., and Jorn-Steffen Pischke. 1992. "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation." *Journal of Labor Economics* 10: 412-37.

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

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

Lumsdaine, Robin L., and David A. Wise. 1990. *Aging and Labor Force Participation: A Review of Trends and Explanations*. Cambridge, Mass.: NBER Working Papers No. 3420.

MaCurdy, Thomas. 1992. "Work Disincentive Effects of Taxes: A Reexamination of Some Evidence." *American Economic Association Papers and Proceedings* 82: 243-9.

MaCurdy, Thomas, David Green and Harry Paarsch. 1990. "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." *Journal of Human Resources* 25: 415-90.

Meyer, Bruce. 1995. "Lessons from the U.S. Unemployment Insurance Experiments." *Journal of Economic Literature* 33: 91-131.

Moffitt, Robert. 1986. "The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method." *Journal of Business and Economic Statistics* 4: 317-28.

National Center for Policy Analysis. 1995. "Tax Fairness for the Elderly: Eliminating the Social Security Earnings Penalty." Policy Backgrounder No. 137. Dallas, TX, September 14.

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

Ransom, Roger L., and Richard Sutch. 1988. "The Decline of Retirement and the Rise of Efficiency Wages: U.S. retiRement Patterns, 1870-1940," in *Issues in Contemporary Retirement*, R. Ricardo-Campbell and E. Lazear, eds. Stanford, CA: Hoover Institution.

Reimers, Cordelia, and Marjorie Honig. 1993. "The Perceived Budget Constraint under Social Security: Evidence from Reentry Behavior." *Journal of Labor Economics* 11(1): 184-204.

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

Triest, Robert K. "The Effect of Income Taxation on Labor Supply in the United States." *Journal of Human Resources* 25: 491-516.

Tuma, Nancy Brandon, and Gary O. Sandefur. 1988. "Trends in the Labor Force Activity of the Aged in the United States, 1940-1980," in *Issues in Contemporary Retirement*, R. Ricardo-Campbell and E. Lazear, eds. Stanford, CA: Hoover Institution.

CHAPTER 2

The Effect of Old Age Assistance on Retirement

The labor force participation of older men has been falling for several decades. While there is some controversy over whether the decline began as early as the late 1800s or as late as the late 1930s, the decline since the 1930s has been dramatic. Ransom and Sutch (1986) calculated that the labor force participation rate of men over 60 fell from 64.5% in 1930 to about 32.2% in 1980.

Economists have devoted considerable attention to trying to explain increased retirement rates over recent decades -- since the 1970s or perhaps the 1960s. Much of the focus has been on Social Security, but many papers have found it has a small or negligible role.[1] Other researchers have argued that private pensions create strong incentives for older workers to withdraw from the labor force.[2] However, while Social Security benefits and private pensions are the main source of income for the elderly today, another public transfer program dwarfed Social Security before 1950, as Parsons (1991) pointed out.[3] Old Age

---

[1] In reviewing this literature, Boskin (1986) concluded that the evidence supported a strong link between rising real benefits and falling participation rates in the late 1960s and early 1970s. In their review, Lumsdaine and Wise (1990) concluded that most of the studies in the 1980s "attribute only a modest portion of the early retirement trend to the effect of Social Security provision."

[2] Notable papers in this literature include Kotlikoff and Wise (1985, 1987) and Stock and Wise (1990).

[3] Parsons also argued that Social Security and pensions were "quite modest" (p. 658) between 1930 and 1950. However, Lumsdaine and Wise (1990) reported that Social Security recipiency expanded significantly during the 1940s from 20% of individuals 65 and over to 60% in 1950, while pension coverage climbed from 24% in 1950 to about 40% in 1960. They do not report pension coverage before 1950. Meanwhile, the Social Security replacement rate fell during the 1940s to a little less then 20% in

Assistance (OAA) was a means-tested program for those 65 and older established by the same legislation that created Social Security in 1935. By 1940, 22% of the aged population was receiving OAA.[4] The size of annual Social Security transfers only began to exceed that of OAA during the 1950s. Thus, the heyday of OAA coincides with the early stages of the modern decline in the elderly's labor force participation and may have played a significant role.

Another distinctive feature of OAA was that benefit levels were determined by the states. This is an advantage for studying OAA compared to Social Security because differences in benefit generosity -- either across states at a point in time or, even better, within a state over time -- can be associated with the corresponding differences in labor supply. In constrast, Social Security is a national program, with all individuals facing the same rules at a point in time.[5] Variation across individuals in Social Security benefits arises because of past earnings histories, which may well be correlated with present labor supply and thus bias estimates of the influence of Social Security.[6] Social Security rules changed over time, but so did many other factors influencing labor supply, some of which might be difficult to control for. Thus, studying OAA may be one of the best opportunities for identifying how income flows for the elderly that are tied to low earnings or to retirement affect labor supply, besides yielding insight into early retirement trends.

Parsons focused on the generosity of OAA and its possible relationship to early changes in labor supply. He employed state-level aggregate data from the Censuses of 1930, 1940, and 1950 and estimated that 50% of the decline in those twenty years could be attributed to OAA. He was compelled to use aggregate data because he chose to analyze

---

1950 and then rose to around 30% in 1960.

[4] OAA recipiency gradually dropped off in the 1950s and 1960s and was replaced in 1974 by Supplemental Security Income. 11.1% of the elderly population received SSI at its inception, declining to 6.5% by 1992.

[5] Supplemental Security Income benefits vary across states, but, as already noted, the program is a fraction of OAA's size.

[6] This is the case for pensions as well.

pre-OAA labor force behavior in relation to post-OAA behavior, and individual records from the Census of 1930 no longer exist. The same conclusions should result from forgoing the comparison to 1930 and using microdata from the 1940 and 1950 Censuses, provided enough variation in OAA benefit levels exists between 1940 and 1950. Using microdata avoids certain other difficulties. Foremost, it is not clear how much Parsons' results depend on the much sharper decline of labor force participation in the 1930s than in the 1940s. As Parsons and others have acknowledged, the decline between 1930 and 1940 may be an artifact of the data to an unknown extent because the questions about labor force activity were modified in the interim. Also, it may be that using aggregate data obscures other features of the relationship between OAA and labor force participation. For instance, the influence of OAA should vary with some observable demographic characteristics like household structure and education. Those interactions are much easier to pinpoint with individual data. In the same vein, it also becomes easier to distinguish between the effects of individual and state characteristics on participation. Lastly, aggregation to the state level is easily interpretable as a linear probability model at the micro level, but not as a probit or logit, which can be readily estimated with individual level data.

In this paper, I will estimate the relationship between state benefit levels and the labor force participation of the elderly using individual data from the Censuses of 1940 and 1950. Accounting for the role of individual level covariates that affect the participation decision along with state level covariates that may influence earning opportunities and benefit determination, I will estimate a quite strong role for Old Age Assistance. In the next section I will describe the nature of the OAA program and the variation across states and over time in its generosity. In Section II I will discuss trends in labor force participation over a similar period and the problems in measuring those trends. I will discuss the estimation method in Section III and the data in Section IV and present and analyze the estimation results in Section V.

I. OLD AGE ASSISTANCE

By 1934, 27 states had established some type of assistance for the elderly, offering average annual benefits of $174 to 235,000 recipients.[7] Still, the Social Security Act of 1935 represented the foundation of the modern welfare state.[8] Besides establishing Old Age and Survivor's Insurance, as Social Security was known, and disability and unemployment insurance, it mandated that each state initiate a program to distribute aid to the elderly, to dependent children, and to the blind. The states had to meet certain standards related, for example, to residency, age, and citizenship requirements -- in return for which the government matched state contributions for assistance. In 1940, the federal matching rate for OAA was 50% up to a maximum of $40 per month in 1940.

OAA grew quickly, as Table 1 demonstrates. By 1940 fully 22.0% of the population 65 and over, 2.0 million individuals, were recipients of OAA. The size of the transfer was substantial -- the average nationwide benefit in June 1940 was $241, while per capita personal income in 1940 is estimated at $593.[9] Table 1 also shows that variation across states was significant. Average benefits ranged from $91 in Arkansas to $455 in California, with a standard deviation of $94, and recipiency rates ranged from 8.1% in Washington, D.C., and 11.2% in New Jersey to 50.2% in Oklahoma, with a standard deviation of 7.7%.

In 1950, the national average benefit had risen to $515, or $289 in 1940 dollars, a 20% real increase. Most states boosted their benefit levels. The correlation between the change in the benefit and the 1940 benefit level was -0.50, so the states with the biggest increases tended to be the states with relatively low 1940 benefits, such as Texas and Florida. Some of the increase was no doubt spurred by more generous federal matching:

---

[7] *Monthly Labor Review*, 1936.

[8] While the Great Depression was the catalyst, Baack and Ray (1988) noted that other contributing factors to the establishment of national social insurance were the decades-old lobbying of progressive social movements and the experience with the national veterans' pension program, along with state relief programs established in the 1920s.

[9] Parsons, p.660.

the federal matching rate had risen to 75% for the first $20 per month and 50% for the next $30 per month. The recipiency rate remained quite high at 22.7%. Lastly, the variation across states in benefit levels remained sizeable, with a standard deviation of benefits of $103, as did the variability in the change in benefits between 1940 and 1950, with a standard deviation of 12.9% -- which is key for the econometric analysis here. The changes in benefit levels within a state between 1940 and 1950 (in the fourth column of Table 1) as it is correlated with the change in labor force participation at the same time, will be used in this paper to determine the effects of OAA on labor supply.

Recipiency should rise with benefits if OAA is to affect labor force participation. Table 2 demonstrates this relationship. Regression 2-A indicates no relationship in the pooled data. Regression 2-B adds state and year effects, which uses the change in benefits within states between 1940 and 1950 to identify the impact on recipiency. Benefit changes within states had a strong positive effect: a 10% rise in real benefits led to a 2.9 percentage point increase in recipiency.[10] There was also a 5.8 percentage point overall drop in recipiency between 1940 and 1950 after controlling for benefit levels -- indicated by the coefficient on the year effect -- perhaps arising because of improved economic opportunities in 1950 relative to 1940.

In this study, it would be valuable to have information on the rules states used to determine benefits as a function of household characteristics. Unfortunately, I have been unable to unearth very definitive rules. In fact, it appears that the process of benefit determination was somewhat arbitrary and also that an individual's benefit level could fluctuate from month to month for exogenous reasons. In principle, benefits were determined by subtracting an applicant's income from what was called a monthly budget, or standardized living requirement as a function of family size.[11] This implies a 100% tax rate on benefits as earnings rose. Other rules are not specified, however. The *Social Security*

[10] The estimates in regression 2-B are significant, but their standard errors are fairly large nonetheless.

[11] This is described in various issues of the Social Security Bulletin as being the generally applicable procedure. It is not made clear whether the method was mandated federally or simply the common practice among most states.

Bulletin mentions that the determination of expected support from other kin as a component of income was somewhat arbitrary, and that states readily changed the size of the household budget.[12] Furthermore, the actual amount paid out depended on the state's monthly fiscal situation. It appears that it was not uncommon for transfers to vary from month to month, generally taking the form of a reduction in the monthly budget. On other occasions, states gave a higher benefit to make up for previous months' cutbacks. Therefore, the average monthly benefit will be the best available way to measure expected program generosity.[13]

II. THE LABOR FORCE PARTICIPATION OF THE AGED

Measuring labor force participation before 1940 is a controversial exercise. The questions asked in the Census about labor force activity before 1940 related to whether the respondent was "gainfully employed". In 1940, the questions were modified to yield information corresponding to the modern economic concept of participation.

For some time, the accepted estimates of labor force participation showed a decline at least from the late 1800s through to 1930. Durand (1948) computed a series which is representative of such estimates. Durand developed a set of "crude adjustment factors" for sex and age groups which he applied to the figures on the percentage gainfully employed from the Censuses from 1890 to 1930.[14] The resulting series of labor force participation for the elderly declines over the entire era, from 68.2% for men 65 and over in 1890 to 54.0% in 1930 and 42.2% in 1940.

Durand's method was disputed by Ransom and Sutch (1986). They went back to the raw data and took into account the instructions to enumerators in the different Census years.

---

[12] At one point in the late 1940s, New York determined that the prices it had been using for the standard set of monthly needs was too high. It lowered them, causing benefit obligations to fall substantially.

[13] Parsons also used the average benefit level as the independent variable in his regressions.

[14] The quotes indicate Parsons's description of Durand's approach. Parsons used Durand's data in order to demonstrate the decline in participation between 1930 and 1940.

Their resulting series for the participation rate for older men is much flatter between 1870 and 1930. When they adjust this series for the secular shift out of agricultural employment, which exhibits substantially higher participation rates, participation rose until 1930. Their unadjusted participation rate for men 60 and over was in the neighborhood of 64-66% between 1870 and 1930. It was 64.5% in 1930 and 61.5% in 1937, and for men 65 and over it was 53.9% in 1930 and 49.1% in 1937. However, their methods have since come under criticism as well.[15]

Still, both series -- that of Durand and the newer one of Ransom and Sutch -- exhibit a break between 1930 and 1940, when the questions about labor force activity were altered. The debate was not able to clarify how much of the measured decline in labor force participation in the 1930s was a result of the modified Census questions. That issue raises doubts about using 1930 data together with 1940 and later data to draw conclusions about changes in labor force participation since 1930.

There is no doubt, meanwhile, that labor force participation declined only slightly between 1940 and 1950. In between, older individuals were temporarily drawn into or kept in the labor force because of World War II.[16] Summary reports of the Censuses indicate that the participation rate was 54.7% for men 60 and over in 1940 and 54.5% in 1950.[17] I have calculated participation rates for elderly men in the Public Use Microdata Samples from the Censuses of 1940 and 1950. The Censuses ask respondents what their labor force status was last week, and the possible answers are at work, with a job but not at work, in the armed forces, unemployed, and out of the labor force. Counting all but the last as participating in the labor force, the participation rates for the subsample of men who were 65 and older and were not institutionalized or living in group quarters are shown by state in Table 3. Their participation rate was 42.0% in 1940 and 40.2% in 1950. Nevertheless, the small change in the national average masks considerable variation across states in each year

---

[15] Moen (1987).

[16] There is no information on how much of that was reversed during demobilization in the late 1940s.

[17] Reported in Ransom and Sutch, p.14.

and within states between 1940 and 1950. The standard deviation of the change in participation across states is 15.0%.

It is not clear in Parsons' work how much of the estimated effect of OAA on labor force participation is driven by the big decline in the 1930s.[18] He recognized the possibility of mismeasurement, but pointed out that including year effects in the regressions will absorb any level differences in the participation rate, leaving a consistent estimate of the effect of OAA. However, if the nature of the distortion is not constant across states, then the estimates will not be consistent. More broadly, the regressions may not have much power to distinguish whether the relationship between OAA and labor force participation is causal because of the strong time series trend.[19] If Parsons' approach is legitimate, on the other hand, then similar answers will emerge when the analysis uses state-level variation in benefits between 1940 and 1950. Thus, this exercise will afford a check on whether the earlier results depend on the changing definition of labor force participation.

III. EMPIRICAL APPROACH

The empirical model which I consider will be of this form:

$$(1)\quad LFP_{ist} = \alpha + X_{it}*\beta + Y_{st}*\gamma + \delta*\log(OAA_{st}) + state_s + year_t + \epsilon_{ist}\ .$$

The index *i* refers to the individual, *s* to the state he lives in, and *t* to the year in which he is observed. LFP is a dummy that is one if the individual participates in the labor force and zero if he does not. Equation (1) is expressed as a linear probability model, but I will estimate a probit in Section V. Labor force participation may be influenced by a vector of

---

[18] Parsons used labor force participation data derived by Lee *et.al.* (1957) from the Censuses because Durand's adjustment factors are not available by state.

[19] A similar issue arises in the Social Security literature. Generally, real Social Security benefits have risen and labor force participation has fallen over time, but Krueger and Pischke (1992) argued that the correlation between them is spurious after examining an episode when real benefits fell, but retirement continued to rise.

individual characteristics X which may or may not vary over time; a vector of time-varying state characteristics Y which may consist of state-level data which is aggregated over individuals or of time-varying characteristics of the state itself; state-varying time-constant characteristics, which must be subsumed in a state-specific intercept; time-varying national characteristics, which must be subsumed in a year-specific intercept; and the generosity of OAA within each state in each year.

Parsons approach is equivalent to adding up equation (1) over all individuals within the state in each year and dividing by state population.[20,21] In this way, the left-hand side variable becomes the labor force participation rate for the state in that year. On the right-hand side, the vector of X variables becomes isomorphic to the Y variables. The interpretation of the OAA variable and the time effect remain the same. In the vector of state-level covariates, Parsons included the share of full-time workers below the poverty level, the percentage of individuals 65 and over who are 75 and over, the growth rate of the male 25-35 year old population, the percentage of workers employed in agriculture, and the average value of farms.

Equation (1) provides a context for comparing estimates obtained using state-level data to estimates from individual data. In either case, a consistent estimate of δ can be obtained only if all omitted variables are uncorrelated with the level of OAA benefits. Omitted variables include any observable or unobservable personal or state characteristics, either time-varying or constant, which influence labor force participation. Three advantages of using individual data emerge: first, it is easier to measure elements of X at the individual level than at the state level; second, the effect of OAA can be allowed to vary with the observable personal characteristics X_{it} in a much more straightforward fashion; and third, using individual level data yields gains in efficiency when including state effects to control for constant state characteristics that may be correlated with OAA benefits.

[20] Equation (1) does not aggregate neatly if it takes the form of a logit or probit.

[21] This interpretation of estimating (1) on aggregate state data implies that the error term becomes heteroscedastic. Thus, the equation should be estimated with weighted least squares where the weights are the state's annual population, which Parsons does not do.

The first advantage of estimating (1) with microdata means that I can use explicit controls for age and for living on a farm instead of the aggregate data on the age structure and agricultural employment that Parsons must use. Other variables included in X are dummies for education, living in a Standard Metropolitan Area (SMA), being the head of the household, being foreign born, being non-white, and being married. Meanwhile, Y will include the state per capita income and unemployment rate. Second, I can also insert interaction terms of the form $X_{it}*\log(OAA_{st})*\phi$, allowing OAA to have a different effect across groups in the population, captured by the vector ϕ.

The third advantage of using microdata is embodied in the state-specific intercepts. The state intercepts provide a very flexible functional form that make it unnecessary to specify precisely what are the time-constant state characteristics that may be correlated with OAA benefits and with labor force participation, which is more important when using state-level data. Unmodelled state characteristics could include relatively concrete concepts associated with the state's industrial structure or the role of the state government in economic life, or very abstract concepts such as social attitudes and community and family structure. Including state intercepts in the regression of individual data means that the effect of OAA is identified from the variation in OAA benefits within states across years -- the hypothesis being that states with a bigger increase in OAA benefits should have a bigger drop in labor force participation. The drawback is that the variation in OAA benefits within states might be too small, in which case the state intercepts would absorb much of the effect of OAA.

IV. DATA

To estimate equation (1), I use individuals from the Public Use Microdata Samples (PUMS) of the Censuses of 1940 and 1950.[22] Each PUMS contains a 1% sample of all individuals enumerated in the Censuses and reports information such as marital status, race, age, sex, nativity, household size and composition, farm residence, employment status, industry, occupation, and hours worked last week. The 1940 Census also contains

[22] I would like to thank Dora Costa for making this data available to me.

information for all individuals on education, individual earnings last year, and migration since five years earlier; the 1950 Census has that information for the subset of individuals who answered the extended survey and are referred to as sample line individuals. The 1940 Census further contains information on family earnings and house tenure which is not available in the 1950 Census. The 1950 Census contains more detailed information on non-labor income which is not available in the 1940 Census. However, because the earnings and income data are not reported very broadly, I will not use that information.

From each PUMS, I selected men between the ages of 66 and 85 who are not in group or institutional quarters. Sample statistics for these individuals are reported in Table 4. The sample consists of 35,799 observations for 1940 and 40,074 observations for 1950. One-third of the 1950 group consists of sample line individuals, who were asked many more questions than non-sample line individuals. Compared to today, the elderly men in 1940 and 1950 are somewhat less racially mixed, more foreign born, less educated and live more rurally. The sample characteristics evolved between 1940 and 1950, as well. The 1950 sample has more very old men, more racial diversity and lives less on farms and more in SMAs. The 1950 sample line individuals are more educated than the 1940 individuals, which may be due to differences between sample line and non-sample line individuals in 1950 and to differences between elderly men in 1940 and 1950. The sample line individuals in 1950 are less racially mixed, more likely to be head of household, and less likely to live on a farm. The regressions will control for differences between the 1940 and 1950 individuals and differences between the sample line and non-sample line individuals in 1950 by allowing both separate intercepts for the different groups and separate slopes for their demographic characteristics. Any remaining differences between the subsamples will not bias estimates of the OAA effect as long as there remain no systematic differences across states that are correlated with OAA generosity.

The dependent variable will be whether an individual was in the labor force. An individual was in the labor force if he reported being at work last week, with a job but not at work, in the armed forces, or unemployed. The unemployment rate was higher in 1940 than in 1950 for this sample, and indeed it was so for all men in the PUMS. The percentage of men who report being out of the labor force for another reason besides being unable to work increased substantially between 1940 and 1950, which is an indication that the modern concept of retirement -- retirement for leisure, or retirement as a planned stage of life -- was proliferating. Most of the respondents who worked last week worked at least 40 hours, so it is reasonable to ignore part-time work as a step towards retirement in these years.

Table 5 reports the statistics on labor force participation for different groups of the population by year. Participation declines with age and increases with education, but less sharply than today in both cases. Participation is higher for those who are married, household heads, nonwhites, natives and farm and non-SMA dwellers. The differences in participation across these groups generally narrow in 1950.

The data on OAA benefits was collected from various issues of the Social Security Bulletin. They calculate the average benefit by dividing total dollar obligations, which excludes administrative and overhead costs, by the total number of recipients. OAA recipiency rates are calculated by dividing the number of recipients by the total population aged 65 and over in each state. The CPI is used to deflate the 1950 benefit levels. Finally, per capita income by state was taken from the *Statistical Abstract*.

V. EMPIRICAL RESULTS

This section will present estimates of the effect of OAA on labor force participation using several different specifications. I begin with a basic version that includes OAA benefits along with demographic characteristics, and I then allow the effect of OAA to vary across demographic groups. After, I add state effects and some state-level variables to control for additional sources of cross-state heterogeneity.

A. Estimates with Demographic Characteristics

Different versions of equation (1) estimated as probits are reported in Tables 6 and 8. Table 6 shows how the results change when person level covariates -- X_{it} in (1) -- are added. Table 8 will include state level covariates as well -- $state_s$ and Y_{st} in (1). Instead of reporting the probit coefficients, the tables show the marginal effects calculated from the probit coefficients. The marginal effects measure the impact on participation of small changes in the continuous variables, or changes from zero to one in the discrete variables, at the sample means of the right-hand side variables.

In Table 6, regression 6-A reports the univariate relationship between labor force participation and the log OAA benefit, demonstrating the strong negative correlation of OAA and labor force participation in the raw data. Regression 6-B adds controls for education, marital status, nativity, race, household headship, living on a farm, and living in a Standard Metropolitan Area. The demographic characteristics are allowed to have a different effect for individuals in 1940, for individuals in 1950, and for sample line individuals in 1950.

The demographic variables explain part of the raw correlation between OAA and labor force participation, but the semielasticity of labor force participation with respect to log benefits of -0.026 (0.006) at the sample means remains quite negative and signficant. It suggests that a 1% increase in benefits would lead to a 0.026 percentage point decline in the probability of labor force participation. However, the aggregate effect of the observed, nonmarginal 20.5% increase in real benefits between 1940 and 1950 is not predicted accurately by the marginal effect. Instead, I computed the change in the probit index and resulting predicted probability of participation for each individual in 1940 if real benefits were at their 1950 level.[23] Actual participation in 1940 was 42.04%, while the mean of the predicted probability of participation for the sample is 42.07% (0.12). Raising benefits to their 1950 level results in mean predicted participation of 41.59%, a 0.47 percentage point decline. The actual decline in participation was 1.82 percentage points to 40.22%, so OAA appears to explain about a quarter of the decline. In comparison, Parsons' estimates attributed one-half of the decline between 1930 and 1950 to OAA.

The coefficients on the demographic characteristics are virtually all significantly different from zero and often are significantly different for different subsamples.[24] The

[23] I also performed the calculation using the individuals in 1950 and benefit levels in 1940, which yields almost identical results.

[24] Jointly restricting the coefficients on the demographic characteristics to be the same across subsamples is rejected at better than a 99.9% confidence level. Taken individually, restricting the coefficients for each characteristic to be the same is not rejected for living in a SMA and is rejected for all others at a 97% confidence level or better.

marginal effects conform to the differences in participation demonstrated in Table 5. More education is associated with greater labor force participation, but less so in 1950 than in 1940.[25] Being a household head and being married are associated with greater labor force participation, particularly for sample line individuals in 1950. Being foreign born is associated with lower labor force participation except for sample line individuals in 1950, while being nonwhite is associated with higher labor force participation except for non-sample line individuals in 1950. Finally, living on a farm raises labor force participation a great deal and living in a SMA raises it a little. The coefficients on the age dummies are not shown but are highly significant. Age has a negative and declining effect on participation, as it does today.

It would be interesting to understand the sensitivity of different types of individuals to OAA. In particular, theory and policy design suggest certain relationships between potential earnings, other family income, household structure and the importance of OAA. OAA should have a bigger impact on individuals with weaker earning opportunities. Also, because it is means-tested and not just earnings-tested like Social Security, those with greater non-labor income would be less likely to receive OAA even if they did retire. However, the PUMS data is less well-suited than modern datasets to understand these interactions. Non-labor family income is only available for a small part of the sample.[26] Earnings information is available for those who work in both years, but fewer covariates and probably more measurement error make it more troublesome than usual to impute earnings for those who do not work. A feasible alternative is to allow OAA to vary with the demographic characteristics that are available in the data and that we expect to be correlated with earnings and with other family income. Table 7 offers evidence in support of this approach. It shows a significant relationship between individual demographic characteristics and earnings for those who work in regression 7-A and between individual characteristics and the existence of another earner in the family -- which is the most widely available measure of other family

[25] Recall that in 1950 education is reported only for sample line individuals.

[26] Non-labor income is reported only for sample-line individuals in 1950 and for no one in 1940.

income -- in regression 7-B.

Regression 6-C relaxes the restriction that OAA have the same effect for all individuals by allowing its impact to depend on demographic characteristics. Now, the coefficient on log benefits with no interactions describes the effect of OAA on an individual with most of the average sample characteristics: less than nine years of schooling, a household head, not married, native born, white, and not living on a farm or in a SMA. Such a person who experiences a 1% increase in real benefits will have a 0.045 percentage point decline in his chance of participating in the labor force -- almost twice as big as the constrained effect estimated in regression 6-B.

The OAA interaction terms explain the effect of OAA on individuals with other demographic characteristics. The coefficients on the interactions are not all individually significant but are jointly highly signficant at better than a 99.9% confidence level. The education terms are of particular interest, because we believe that education has a strong economic relationship to earnings, and consequently to non-labor income too. The expected relationship is borne out for someone with twelve years of school, who is much less sensitive to OAA than someone with less schooling and, we expect, poorer earnings opportunities. Those with twelve years of school are in fact virtually unaffected by OAA, since the positive marginal effect from the OAA interaction term is almost the same as the negative marginal effect from the main effect for OAA. However, the marginal effect of 9-11 year of school is negative, showing that they are more sensitive to OAA than those with 0-8 years of school -- which we do not expect to be the case -- but the coefficient is imprecisely estimated.[27] Meanwhile, the marginal effects from the terms for having more than 12 years of school become smaller again, where we expect them to be larger than the others, but they are very imprecisely estimated.[28]

---

[27] Such an effect is consistent, to an extent, with the insignificant coefficient on 9-11 years of school in regression 7-B, showing that those with 9-11 years of school exhibit no higher probability of having another family member with positive earnings than those with 0-8 years of school.

[28] Those individuals are a quite small proportion of the sample -- 7-8% of those for whom we know schooling, so attending college or even more around the turn of the century was an unusual event and may signal substantial idiosyncrasy in their labor supply behavior.

The marginal effects on the terms that interact OAA with the other demographic characteristics are reported as well. It is interesting that being a household head leads to a larger (more negative) impact of OAA. Someone who is not a household head might live with his children or other relatives, and the available support from other kin was taken into account in determining the eligibility for and size of benefits. Someone who lives on a farm, on the other hand, is less sensitive to OAA, which may point to higher assets or other household income that preclude eligibility. The other interaction terms are not significant.

The predicted decline in participation induced by the 20.5% increase in real benefits between 1940 and 1950 is 0.75 percentage points for the part of the sample represented by the OAA main effect. The effect of OAA is significantly less negative for individuals with 12 years of schooling, non-household heads, and farm dwellers. Using all of the interaction terms to compute the effect of the change in OAA results in a 0.58 percentage point aggregate decline, closer to the 0.47 percentage point decline predicted by the constrained estimation in regression 4-B. Thus, freeing the OAA variable to have different effects on different population groups does not change the aggregate prediction by much, but it reveals a great deal about how different types of individuals reacted.

The effect of the different characteristics on participation now depends on benefit levels as well, with the coefficients on the demographic characteristics reporting the effect on participation when benefits are zero. Now, the impact of education on participation is not very monotonic until the interaction with benefits is considered. This suggests that education patterns vary across states in a way that is correlated with benefit levels. Similarly, part of the positive relationship between farm dwelling and participation is correlated with variation across states in benefit levels. A stronger positive relationship between SMA dwelling and participation emerges, with SMA dwellers appearing to be more sensitive to OAA than others. Also, the negative effect of being foreign born on participation in regression 4-B appears here to be entirely explained by living in states that have high levels of OAA.

The estimates in Table 6 report the relationship among individual characteristics, benefit levels and participation. Still, state-level variables may be involved in the interaction between OAA and participation -- either directly, through economic or political conditions, or indirectly, through other unmeasured individual characteristics which vary across states.

B. Estimates Adding State Characteristics

The regressions in Table 8 report the estimates when state effects and state per capita income are added to the specification in 6-C. State effects are a flexible way to control for factors which vary across states and which influence labor force participation and may be correlated with benefits. Instead of having to specify all such potential factors, the state effect captures them jointly by allowing a different intercept for each state. However, there may still remain some omitted characteristics which vary within a state over time. For this reason, I also add state per capita income in 1940 and 1950 to the regressions. Per capita income may capture a number of different factors. Since the PUMS lack information on family resources, per capita income can control in a rough way for differences in both family wealth and earning opportunities, which certainly vary across states and affect the retirement decision. Additionally, a state's prosperity might affect how it sets benefits, either by loosening fiscal constraints or by reducing the perceived necessity for income support. In that case. leaving out a measure of state prosperity would bias the estimate of the OAA effect if state prosperity also captures factors which influence elderly labor force participation.

Regression 8-A adds state fixed effects to the specification last seen in regression 6-C. The fixed effects lower the log likelihood significantly and change the coefficient on the OAA main effect substantially, although not significantly, from -0.045 (0.015) previously to -0.068 (0.025), a 50% increase. Thus, cross-state heterogeneity in participation behavior had been obscuring part of OAA's effect on participation -- states with bigger increases in benefits had a bigger drop in participation, but other factors about those states smoothed the participation decline. The coefficients on the OAA interaction terms are mostly unchanged, so the OAA effect is strengthened for almost everyone. The only exceptions are that the negative coefficients on being foreign born and on living in a SMA are quite a bit smaller in absolute value than in regression 6-C, though not significantly so. This probably signals a relative lack of variation in those characteristics across states, so that their influence on participation cannot be identified separately from the state effects.

Regression 8-B adds state per capita income, which makes the OAA effect yet more negative.[29] The marginal effects of log benefits and log per capita income are relatively precisely estimated in spite of substantial positive correlation.[30] Log state per capita income has a strong positive influence on labor force participation -- with a marginal effect of 0.054 (0.037) -- so income appears to be picking up improved labor market conditions that encourage individuals to postpone retirement, and omitting the income term had caused the OAA variable to absorb part of that positive effect. According to the estimates, the 35.8% increase in per capita income would have raised participation by 1.66 percentage points if nothing else had changed. However, that interpretation is problematic because rising income may well have had a role in increasing OAA generosity as well, which lowered participation. Log benefits regressed on log per capita income and state effects yields a coefficient of 0.576 (0.162). Interpreted causally, the mean increase in per capita income accounts for a 20.6% increase in mean benefits and a net increase in participation of 0.1 percentage points. More light might be shed on the role of state income by exploring its relationship to both benefit determination and labor force participation.

Meanwhile, the marginal effect of OAA has now been strengthened to -0.085 (0.028). Again, the coefficients on the OAA interaction terms are virtually unchanged. The estimates suggest that the 20.5% increase in benefits between 1940 and 1950 would have lowered participation by 1.13 percentage points in the aggregate if nothing else changed. Thus, estimating the full specification reveals an even stronger role for OAA, contributing to over 60% of the decline in participation.

The main effects of the demographic characteristics are mostly unchanged from regression 6-C. The coefficients on being foreign born and on living in a SMA were altered along with the OAA interactions terms when state effects were added. Meanwhile, the same pattern persists for the education interactions, with a bigger impact of OAA for those with 9-11 years of school compared to those with less schooling, and a smaller impact of OAA on

---

[29] The state unemployment rate, calculated for men aged 18-64 in the PUMS, was also tried in the regression. Its coefficient was not significant, nor did it alter the other coefficients noticeably.

[30] The coefficient on per capita income is significant only at the 86% confidence level, while the coefficient on log benefits is significant at better than 99%.

those with 12 years of school. The overall effect of different demographic characteristics on participation depends on the benefit level. Table 9 reports the impact of these demographic characteristics on participation at the mean benefit levels in 1940 and 1950, but the effects are not measured very precisely.

VI. SUMMARY AND CONCLUSION

This paper examines the response of labor force participation to the expansion of the Old Age Assistance program in the 1940s. Using microdata from the 1940 and 1950 Censuses, I estimated a very substantial effect of OAA on participation. Yet, economists have had trouble attributing the much bigger decline in participation that followed to Social Security, even though Social Security came to be received by a much larger proportion of the population. A large part of the problem in quantifying the impact of Social Security and pensions on retirement arises because the income flows they produce do not vary across the population randomly but depend on past labor supply. Old Age Assistance gets around this difficulty because benefit levels varied across states. The impact of OAA is estimated here from the relationship of changes in participation to changes in benefit levels within states between 1940 and 1950, a powerful identification strategy which avoids confounding the effect of benefits with other factors that vary across states and that may be difficult to specify. The conclusion from this approach is that the growth of transfers for the elderly that are tied to labor force withdrawal has played a significant role in the trend towards earlier retirement. It remains to analyze further the implications of the results on Old Age Assistance for the later growth in Social Security.

The estimates suggest that rising real OAA benefits between 1940 and 1950 explained as much as 60% of the decline in labor force participation. Specifications without state-level variables indicated that one-quarter to one-third of the decline was attributable to OAA. However, the specification with state effects and state income appears to explain labor force participation better and suggests further work in understanding the relevance of the state-level variables, which might result in a more precise estimate of the OAA effect. It would be useful to understand why the variable for state per capita income made a big impact in the regressions, in part because it is highly correlated with the OAA variable and therefore raises its standard error. The income variable should be capturing both a wealth effect discouraging work and a substitution effect encouraging work. Further investigation might also be directed into the role of time-varying state level characteristics which affect participation and might still be captured by the OAA variable in the current specification.

The approach used in this chapter could be extended in a number of directions. The availability of large samples and a reasonable amount of information in the Public Use Microdata Samples of the Censuses since 1940 have hardly been exploited. For example, the information on employment and income could be used to analyze the role of increasing postwar economic prosperity on retirement. Further analysis of social programs would also be fruitful, combining the PUMS with the wealth of program data available in the monthly *Social Security Bulletin* beginning in 1938. The political economy of the programs themselves suggests another area for research. Considerable work analyzing the genesis of Social Security has been undertaken, but much less is known about the other programs established under the same legislation, which were the first programs that were federally mandated and state run and represent most of today's social insurance besides Medicare and Medicaid. The conclusions from studying Old Age Assistance suggest many avenues for further research.

TABLE 1: OAA Data by State

	1940		1950		
	Average Annual Benefit	Recipiency Rate	Average Annual Benefit	Real Increase	Recipiency Rate
New England					
Maine	258.83	17.3%	496.08	7.4%	16.1%
New Hampshire	257.22	11.9%	510.46	11.2%	12.9%
Vermont	191.84	15.8%	407.71	19.1%	17.3%
Massachusetts	344.42	23.1%	743.62	21.0%	21.5%
Rhode Island	233.90	12.5%	538.48	29.0%	14.7%
Connecticut	322.60	13.3%	663.05	15.2%	11.1%
Middle Atlantic					
New York	306.69	12.9%	627.99	14.7%	9.6%
New Jersey	248.52	11.2%	564.95	27.4%	6.2%
Pennsylvania	264.64	14.5%	460.42	-2.5%	10.4%
East North Central					
Ohio	279.58	22.8%	554.34	11.1%	17.7%
Indiana	216.07	23.0%	421.80	9.4%	14.5%
Illinois	251.47	24.3%	520.62	16.0%	16.7%
Michigan	202.83	22.2%	542.92	50.0%	21.7%
Wisconsin	268.89	21.3%	494.91	3.1%	17.0%
West North Central					
Minnesota	257.71	29.6%	537.46	16.9%	20.7%
Iowa	251.07	24.2%	578.54	29.1%	18.1%
Missouri	193.11	28.3%	505.51	46.7%	32.3%
North Dakota	201.31	22.4%	563.23	56.8%	18.6%
South Dakota	236.71	33.2%	462.57	9.5%	22.0%
Nebraska	198.24	26.3%	520.52	47.1%	18.3%
Kansas	234.28	17.3%	591.97	41.6%	20.1%
South Atlantic					
Delaware	134.03	13.1%	327.24	36.8%	6.4%
Maryland	211.92	15.0%	424.96	12.4%	7.4%
D.C.	307.78	8.1%	485.41	-11.6%	5.0%
Virginia	117.67	11.3%	242.03	15.3%	9.2%
West Virginia	166.37	17.6%	293.01	-1.3%	19.3%

	1940		1950		
	Average Annual Benefit	Recipiency Rate	Average Annual Benefit	Real Increase	Recipiency Rate
North Carolina	121.70	22.8%	248.02	14.2%	27.3%
South Carolina	98.81	24.3%	262.40	48.8%	36.6%
Georgia	96.02	17.8%	256.65	49.8%	46.1%
Florida	144.11	26.8%	464.15	80.5%	29.2%
East South Central					
Kentucky	104.90	25.7%	231.59	23.7%	28.7%
Tennessee	120.88	23.5%	354.10	64.2%	27.9%
Alabama	112.54	14.6%	240.34	19.7%	40.9%
Mississippi	99.86	18.6%	216.87	21.7%	42.9%
West South Central					
Arkansas	90.51	18.4%	271.89	68.3%	44.8%
Louisiana	142.65	26.8%	565.45	122.1%	68.1%
Oklahoma	212.62	50.2%	599.84	58.1%	52.0%
Texas	123.77	34.1%	399.53	80.9%	43.8%
Mountain					
Montana	216.84	33.6%	609.49	57.5%	23.2%
Idaho	263.29	28.2%	545.39	16.1%	26.3%
Wyoming	285.63	27.1%	652.58	28.0%	23.3%
Colorado	405.03	47.6%	812.06	12.4%	43.7%
New Mexico	173.99	19.3%	415.46	33.8%	30.2%
Arizona	336.28	34.0%	571.84	-4.7%	30.2%
Utah	254.27	45.1%	550.32	21.3%	23.9%
Nevada	316.90	33.4%	619.73	9.6%	24.2%
Pacific					
Washington	265.77	27.2%	767.48	61.8%	34.9%
Oregon	256.61	20.7%	610.04	33.2%	17.9%
California	455.48	25.5%	843.36	3.8%	29.9%
Weighted Nat'l Average	241.36	21.8%	515.42	19.7%	22.7%
Standard Deviation	93.95	7.7%	183.22	12.9%	26.2%

The average benefit is transfers divided by recipients. The recipiency rate is recipients divided by population 65 and over. The national average and standard deviation of the average benefit is weighted by recipients per state; and the recipiency rate is weighted by state population. *Source:* Social Security Bulletin, various issues.

TABLE 2: The Effect of a State's Benefits on Recipiency		
Dependent variable: Recipiency rate	2-A	2-B
Log annual OAA benefit	.004 (.029)	.294 (.077)
Year 1950		-.058 (.024)
Constant	.215 (.159)	-1.045 (.358)
State dummies?	No	Yes
Adjusted R^2	-.0102	.6012

N=98. The recipiency rate and annual OAA benefit are defined in the notes to Table 1. Standard errors in parentheses.

TABLE 3: Labor Force Participation Rates			
	1940	1950	Change
New England			
Maine	47.8%	36.6%	-11.2
New Hampshire	37.6%	40.6%	3.1
Vermont	52.1%	34.8%	-17.4
Massachusetts	40.1%	36.4%	-3.6
Rhode Island	42.1%	40.1%	-2.0
Connecticut	45.1%	42.9%	-2.2
Middle Atlantic			
New York	40.4%	41.4%	1.0
New Jersey	39.0%	38.1%	-0.9
Pennsylvania	41.4%	37.7%	-3.7
East North Central			
Ohio	36.8%	41.2%	4.5
Indiana	42.0%	47.0%	5.0
Illinois	37.5%	40.6%	3.1
Michigan	42.8%	40.1%	-2.8
Wisconsin	39.0%	41.7%	2.7
West North Central			
Minnesota	38.8%	43.6%	4.8
Iowa	39.3%	42.4%	3.1
Missouri	45.1%	42.3%	-2.7
North Dakota	48.5%	50.8%	2.3
South Dakota	47.1%	53.4%	6.4
Nebraska	38.9%	42.2%	3.3
Kansas	42.4%	45.4%	3.0
South Atlantic			
Delaware	47.1%	47.7%	0.6
Maryland	41.2%	42.7%	1.5
D.C.	31.1%	36.6%	5.6
Virginia	53.6%	47.6%	-5.9

	1940	1950	Change
West Virginia	44.2%	37.0%	-7.2
North Carolina	57.6%	45.8%	-11.8
South Carolina	56.1%	46.3%	-9.8
Georgia	51.6%	43.5%	-8.1
Florida	35.0%	28.8%	-6.2
East South Central			
Kentucky	49.3%	42.4%	-6.9
Tennessee	47.7%	39.7%	-8.0
Alabama	51.8%	44.4%	-7.4
Mississippi	58.5%	47.8%	-10.7
West South Central			
Arkansas	53.4%	44.4%	-9.0
Louisiana	44.7%	27.6%	-17.1
Oklahoma	38.9%	33.0%	-5.8
Texas	42.4%	40.8%	-1.6
Mountain			
Montana	45.0%	36.1%	-8.9
Idaho	52.8%	44.1%	-8.7
Wyoming	42.6%	64.9%	22.3
Colorado	34.6%	33.7%	-0.9
New Mexico	42.3%	40.5%	-1.8
Arizona	38.4%	36.2%	-2.2
Utah	42.1%	44.0%	1.9
Nevada	45.2%	48.8%	3.6
Pacific			
Washington	37.3%	36.8%	-0.5
Oregon	42.0%	43.4%	1.5
California	34.2%	31.1%	-3.0
National	42.0%	40.2%	-1.8

Men 65 and older, not institutionalized or in group quarters, from the Public Use Microdata Samples for the 1940 and 1950 Censuses. Labor force participation is calculated as at work, had a job but not at work, in the armed forces, or unemployed last week, divided by the total.

TABLE 4: Sample Means (Standard Deviations)				
	1940	1950		
		All	Sample Line*	Non-Sample Line*
Number of Observations	35,799	40,074	13,347	26,727
Age between:				
66-70	0.463 (0.499)	0.443 (0.497)	0.438 (0.496)	0.445 (0.497)
71-75	0.294 (0.450)	0.294 (0.456)	0.301 (0.459)	0.290 (0.454)
76-80	0.165 (0.371)	0.181 (0.385)	0.181 (0.385)	0.181 (0.385)
81-85	0.077 (0.267)	0.085 (0.279)	0.080 (0.272)	0.083 (0.276)
Non-white	0.072 (0.258)	0.085 (0.279)	0.073 (0.259)	0.092 (0.289)
Married	0.662 (0.473)	0.683 (0.465)	0.686 (0.464)	0.682 (0.466)
Head of household	0.800 (0.400)	0.726 (0.446)	0.812 (0.391)	0.683 (0.465)
Foreign-born	0.243 (0.423)	0.244 (0.429)	0.230 (0.421)	0.250 (0.433)
Yrs of schooling complete	6.79 (3.75)		7.11 (3.84)	
0-8 years	0.818 (0.386)		0.761 (0.427)	0
9-11 years	0.061 (0.238)		0.083 (0.275)	0
12 years	0.057 (0.232)		0.079 (0.270)	0
13-15 years	0.030 (0.171)		0.035 (0.185)	0
16 or more years	0.035 (0.184)		0.042 (0.201)	0
Yrs of schooling	0	0.667 (0.471)	0	1
Lives on a farm	0.271 (0.444)	0.203 (0.402)	0.181 (0.385)	0.214 (0.410)
Lives in an SMA	0.469 (0.499)	0.519 (0.500)	0.506 (0.500)	0.525 (0.499)
Hours worked last week	16.5 (23.8)	15.6 (22.6)	15.6 (22.6)	15.6 (22.7)
Hours worked | hrs>0	46.0 (15.0)	42.5 (15.9)	42.1 (16.1)	42.7 (15.9)
Worked nonzero hours	0.359 (0.480)	0.367 (0.482)	0.370 (0.483)	0.365 (0.481)
Labor force activity:				
In the labor force	0.420 (0.494)	0.402 (0.490)	0.406 (0.491)	0.400 (0.490)
At work	0.389 (0.488)	0.388 (0.487)	0.391 (0.488)	0.387 (0.487)
Unemployed	0.031 (0.173)	0.014 (0.117)	0.015 (0.123)	0.013 (0.113)
Not in the labor force	0.580 (0.494)	0.598 (0.490)	0.594 (0.555)	0.600 (0.529)
Because unable to work	0.424 (0.494)	0.336 (0.472)	0.321 (0.467)	0.344 (0.475)
For another reason	0.156 (0.363)	0.262 (0.439)	0.273 (0.445)	0.256 (0.436)
Consumer price index	100.2	178.8		

* Sample line individuals answered a more detailed questionnaire.
The sample consists of men aged 66-85 who were not in institutional or group quarters, as reported in the PUMS of the Censuses of 1940 and 1950.

TABLE 5: Labor Force Participation (Standard Deviations)				
	1940	1950		
		All	Sample Line*	Non-Sample Line*
Number of Observations	35,799	40,074	13,347	26,727
Labor Force Participation	0.420 (0.494)	0.402 (0.490)	0.406 (0.491)	0.400 (0.490)
Age between:				
66-70	0.569 (0.495)	0.556 (0.497)	0.558 (0.497)	0.554 (0.497)
71-75	0.366 (0.482)	0.359 (0.480)	0.357 (0.479)	0.360 (0.480)
76-80	0.236 (0.424)	0.226 (0.419)	0.243 (0.429)	0.218 (0.413)
81-85	0.133 (0.340)	0.118 (0.323)	0.128 (0.334)	0.114 (0.317)
Non-white	0.483 (0.500)	0.425 (0.494)	0.436 (0.496)	0.421 (0.494)
Married	0.486 (0.500)	0.463 (0.499)	0.461 (0.499)	0.464 (0.499)
Head of household	0.474 (0.499)	0.467 (0.499)	0.445 (0.497)	0.480 (0.500)
Foreign-born	0.363 (0.481)	0.371 (0.483)	0.384 (0.486)	0.365 (0.481)
Years of schooling completed				
0-8 years	0.404 (0.491)		0.389 (0.487)	
9-11 years	0.488 (0.500)		0.452 (0.498)	
12 years	0.514 (0.500)		0.460 (0.499)	
13-15 years	0.504 (0.500)		0.505 (0.501)	
16 or more years	0.591 (0.492)		0.539 (0.499)	
Lives on a farm	0.606 (0.489)	0.570 (0.495)	0.607 (0.489)	0.555 (0.497)
Lives in an SMA	0.386 (0.489)	0.387 (0.487)	0.396 (0.489)	0.382 (0.486)

* Sample line individuals answered a more detailed questionnaire.
The sample consists of men aged 66-85 who were not in institutional or group quarters, as reported in the PUMS of the Censuses of 1940 and 1950.

TABLE 6: Probit Results with Demographic Characteristics

Dep variable: In the labor force	6-A	6-B			6-C			
Log annual OAA benefit	-.091 (.005)	-.026 (.006)			-.045 (.015)			
Interactions:		x1940	x1950	xSL x1950	xOAA	x1940	x1950	xSL x1950
9-11 years of schooling		.089 (.012)	.060 (.017)		-.024 (.024)	.225 (.133)	.204 (.140)	
12 years of schooling		.148 (.012)	.087 (.017)		.041 (.024)	-.075 (.130)	-.135 (.130)	
13-15 years of schooling		.126 (.016)	.133 (.025)		-.003 (.034)	.144 (.189)	.151 (.199)	
16+ years of schooling		.238 (.015)	.168 (.023)		-.028 (.033)	.379 (.177)	.325 (.189)	
Married		.074 (.007)	.041 (.009)	.035 (.014)	-.005 (.013)	.100 (.070)	.068 (.072)	.034 (.014)
Foreign born		-.011 (.007)	-.033 (.008)	.035 (.014)	-.021 (.018)	.108 (.101)	.088 (.104)	.036 (.014)
Nonwhite		.026 (.011)	-.032 (.011)	.049 (.021)	-.000 (.002)	.029 (.018)	-.029 (.019)	.047 (.021)
Not a household head		-.198 (.007)	-.188 (.008)	.064 (.016)	.115 (.015)	-.461 (.043)	-.503 (.049)	.061 (.016)
Lives on a farm		.299 (.007)	.239 (.009)	.064 (.015)	.028 (.013)	.148 (.069)	.082 (.071)	.061 (.015)
Lives in an SMA		.039 (.006)	.046 (.007)	-.002 (.012)	-.013 (.012)	.114 (.068)	.122 (.070)	-.000 (.012)
Fixed effect			.092 (.012)	-.088 (.015)			.096 (.013)	-.087 (.015)
Log likelihood	-51183	-43726			-43675			
Age dummies?	No	Yes			Yes			

This table reports the marginal effects calculated from the estimated probit coefficients at the mean of the right-hand side variables. Standard errors are in parentheses. The dependent variable is whether the individual participates in the labor force. Sample: men aged 66-85 not living in institutional or group quarters from the 1940 and 1950 PUMS. N=75873.

TABLE 7: Regressions of Own and Other Family Earnings

	7-A	7-B
Dependent Variable:	Own earnings last year	Other family members had positive earnings last year
Sample	All in 1940, sample line individuals in 1950; those with nonzero earnings	All in 1940, sample line individuals in 1950
N	12,073	45,035
Independent Variables:		
9-11 years of schooling	431 (39)	.006 (.009)
12 year of schooling	762 (39)	.043 (.009)
13-15 years of schooling	716 (56)	.031 (.012)
16+ years of schooling	1465 (51)	.081 (.011)
Married	174 (27)	-.025 (.005)
Foreign born	-96 (26)	-.070 (.005)
Nonwhite	-552 (40)	-.047 (.008)
Not head of household	-188 (35)	-.350 (.007)
Lives on a farm	-346 (33)	.097 (.005)
Lives in an SMA	526 (23)	-.072 (.005)
Year 1950	814 (23)	.093 (.005)
Constant	582 (31)	.733 (.006)
Adjusted R^2	.2732	.1110

Sample: men aged 66-85 not living in institutional or group quarters from the 1940 and 1950 PUMS. The dependent variable is whether the individual participates in the labor force. Standard errors in parentheses.

TABLE 8: Probit Results Adding State-Level Variables

Dep variable: In the labor force	8-A				8-B			
Log annual OAA benefit	-.068 (.025)				-.085 (.028)			
Log state per capita income					.054 (.037)			
Interactions:	xOAA	x1940	x1950	xSL x1950	xOAA	x1940	x1950	xSL x1950
9-11 years of schooling	-.028 (.024)	.243 (.134)	.223 (.141)		-.030 (.024)	.253 (.134)	.235 (.141)	
12 years of schooling	.043 (.025)	-.079 (.131)	-.143 (.130)		.041 (.025)	-.069 (.131)	-.133 (.132)	
13-15 years of schooling	.006 (.034)	.101 (.189)	.104 (.199)		.003 (.034)	.114 (.190)	.118 (.200)	
16+ years of schooling	-.023 (.033)	.361 (.180)	.304 (.191)		-.025 (.033)	.371 (.179)	.316 (.190)	
Married	-.003 (.013)	.092 (.070)	.063 (.073)	.033 (.014)	-.003 (.013)	.089 (.070)	.059 (.073)	.033 (.014)
Foreign born	-.011 (.019)	.041 (.106)	.015 (.109)	.036 (.014)	-.011 (.019)	.042 (.106)	.018 (.109)	.036 (.014)
Nonwhite	.002 (.002)	.037 (.018)	-.024 (.019)	.049 (.021)	.002 (.002)	.038 (.018)	-.025 (.019)	.049 (.021)
Not a household head	.112 (.015)	-.458 (.044)	-.499 (.050)	.060 (.016)	.113 (.015)	-.459 (.044)	-.500 (.049)	.059 (.016)
Lives on a farm	.033 (.013)	.129 (.069)	.060 (.072)	.058 (.015)	.032 (.013)	.134 (.070)	.064 (.072)	.058 (.015)
Lives in an SMA	-.001 (.013)	.045 (.073)	.049 (.075)	-.001 (.012)	-.001 (.013)	.041 (.073)	.049 (.075)	-.001 (.012)
Fixed effect			.102 (.013)	-.085 (.015)			.085 (.018)	-.085 (.015)
Log likelihood	-43437				-43436			
Age dummies?	Yes				Yes			
State dummies?	Yes				Yes			

This table reports the marginal effects calculated from the estimated probit coefficients at the mean of the right-hand side variables. Standard errors are in parentheses. The dependent variable is whether the individual participates in the labor force. Sample: men aged 66-85 not living in institutional or group quarters from the 1940 and 1950 PUMS. N=75873.

TABLE 9: Effect of Demographic Characteristics on Participation		
	Marginal effect of demographic characteristics at the mean log benefits:	
	1940	1950
Mean real OAA benefits	239.1	287.9
Mean log real OAA benefits	5.403	5.608
9-11 years of schooling	-.146 (.114)	-.151 (.118)
12 years of schooling	.216 (.114)	.224 (.118)
13-15 years of schooling	.016 (.157)	.017 (.163)
16+ years of schooling	-.097 (.157)	-.101 (.163)
Married	-.022 (.059)	-.022 (.062)
Foreign born	-.022 (.086)	-.022 (.090)
Nonwhite	-.005 (.022)	-.006 (.022)
Not a household head	.551 (.065)	.572 (.067)
Lives on a farm	.168 (.059)	.174 (.062)
Lives in an SMA	-.000 (.059)	-.000 (.062)

The coefficients from regression 8-B reported in Table 8 are used to calculate the effect of the characteristics reported above at the mean log OAA benefit. Standard errors calculated according to the delta method are in parentheses.

REFERENCES

Baack, Ben, and Edward John Ray. 1988. "Federal Transfer Payments in America: Veterans' Pensions and the Rise of Social Security." *Economic Inquiry* 26: 687-702.

Durand, John D. 1948. *The Labor Force in the United States, 1890-1960*. New York.

Friedberg, Leora. 1996. "The Labor Supply Effects of the Social Security Earnings Test." Massachusetts Institute of Technology. Mimeo.

Gordon, Roger H., and Alan S. Blinder. 1980. "Market Wages, Reservation Wages, and Retirement Decisions." *Journal of Public Economics* 14: 277-308.

Hurd, Michael, and Michael J. Boskin. 1981. "The Effect of Social Security on Retirement in the Early 1970s." *Quarterly Journal of Economics* 96: 314-8.

Kotlikoff, Laurence B., and David A. Wise. 1985. "Labor Compensation and the Structure of Private Pension Plans: Evidence for Contractual versus Spot Labor Markets," in *Pensions, Labor, and Individual Choice*, D. Wise, ed. Chicago: University of Chicago Press.

Kotlikoff, Laurence B., and David A. Wise. 1987. "The Incentive Effects of Private pension Plans, " in *Issues in Pension Economics*, Z. Bodie, J. Shoven, and D. Wise, eds. Chicago: University of Chicago Press.

Krueger, Alan B., and Jorn-Steffen Pischke. 1992. "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation." *Journal of Labor Economics* 10: 412-37.

Lee, Everett S., Ann Ratner Miller, Carol P. Brainerd, and Richard Easterlin. 1957. *Population Redistribution and Economic Growth: United States, 1870-1950.* Volume 1: *Methodological Considerations and Reference Tables*. Philadelphia.

Lumsdaine, Robin L., and David A. Wise. 1990. *Aging and Labor Force Participation: A Review of Trends and Explanations*. Cambridge, Mass.: NBER Working Papers No. 3420.

Moen, Jon. 1987. "The Labor of Older Men: A Comment." *Journal of Economic History* 47: 761-67.

Monthly Labor Review. 1936. "State Old-Age Pensions in 1935." October.

Parsons, Donald O. 1991. "Male Retirement Behavior in the United States, 1930-1950." *Journal of Economic History* 51: 657-74.

Ransom, Roger L., and Richard Sutch. 1988. "The Labor of Older Americans: Retirement of Men On and Off the Job." *Journal of Economic History* 46: 1-30.

Stock, James, H. and David A. Wise. 1990. "Pensions, the Option Value of Work, and Retirement." *Econometrica* 58: 1151-80.

CHAPTER 3

Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data

Since the beginning of 1996, several states have proposed rolling back unilateral and no-fault divorce, a major reversal of a trend towards relaxed divorce laws which began in 1970.[1] No-fault divorce allows divorce without blame, while unilateral divorce requires only one party's willingness to divorce. The stated aim of repealing such measures is to strengthen families by making divorce more difficult. However, the evidence is decidedly mixed over whether no-fault and unilateral divorce contributed to the sharp increase in divorce rates in the U.S. over the last thirty years.

While the early analysis of this issue was undertaken by legal scholars, the more thorough empirical work has been conducted by economists. The principal paper in this literature, by Peters (1986), began by framing the issue within an economic model of marriage. She showed that, in theory, unilateral divorce will not have an impact on divorce rates if the conditions required for the Coase theorem apply. If so, then the efficient outcome of getting divorced or staying married will occur whatever the divorce law, which will only have an impact on the distribution of gains. Peters tested this hypothesis empirically by estimating the impact of state divorce regimes on divorce probabilities and found that the effect was essentially zero. However, Allen (1992) disputed her results and estimated a positive effect, which Peters (1992) rebutted in turn. Peters and Allen used the same data and estimation method, differing principally on controls for geographical

[1] *The New York Times*, February 12, 1996.

heterogeneity in divorce propensities.[2]

In this paper I use a panel of state-level divorce rates to analyze the impact of changes in state divorce laws. The sample includes virtually every divorce in the U.S. over the entire duration of the law changes and allows me to control thoroughly for heterogeneity across states and over time. The controls turn out to be crucial to the results, as suggested by the disagreement between Peters and Allen. First, in a regression that includes state and year effects to control for differences across states in the propensity to divorce and for differences over time in the national propensity to divorce, the law change appears to have no impact on divorce rates. However, further investigation reveals that state fixed effects alone misspecify the dynamic structure of the unobservables, camouflaging the variation induced by the law changes. Including state-specific trends, which allow the state divorce propensities to change linearly and even quadratically over time, reveals that switching to unilateral divorce law raised a state's divorce rate by about 7%, which comprised 16% of the increase in the U.S. divorce rate between 1970 and 1985.[3]

I. BACKGROUND

The divorce rate in the United States soared in the 1960s and 1970s, as demonstrated in Figure 1 by the upper line.[4,5] During that time, states substantially liberalized and simplified their divorce laws, which can be seen by the population-weighted percentage of states with unilateral divorce, shown in the lower line. The "no-fault revolution" began in 1970 when California replaced all fault grounds for divorce -- such as adultery, cruelty, etc.

---

[2] Peters and Allen also disputed the classification of the legal regime in certain states. The issue turns out not to be important to my results, as discussed later.

[3] These are the dates thought of as the beginning and end of the "no-fault revolution".

[4] The divorce rate is defined as the number of divorces per thousand people.

[5] Cherlin (1992) documented that the divorce rate rose steadily in the late nineteenth century and early twentieth. After a sharp peak during World War II, the divorce rate dropped below trend in the 1950s. The rise in the 1960s and 1970s looks especially stark compared to the low level in the 1950s but also was steeper than the century-long trend.

-- with the sole ground of irretrievable breakdown of the marriage. Over the following decade and a half, every state adopted no-fault grounds for divorce. Most states legalized unilateral divorce as well, under which the no-fault grounds do not require the agreement of both spouses.

Casual observation suggests attributing at least some of the increasing prevalence of divorce to the widespread shift to unilateral divorce, which appeared to make divorce simpler. On the other hand, lawmakers and legal experts at the time argued that the motivation for changing the laws was to bring law books into alignment with actual practice and to eliminate the need to assign blame during a divorce, but not to make divorce easier.[6]

A theoretical model of divorce is laid out in Peters (1986) and Becker (1981). Peters showed that, under perfect information and costless bargaining, a marriage will be dissolved when it is efficient to do so: "when the joint value of the marriage is less than the sum of the values of opportunities that face each spouse at divorce" (p.438). Barring asymmetric information and transactions costs, a change in the law from mutual to unilateral divorce would not make divorce more likely. By the Coase theorem, it would simply redistribute property rights from the spouse who wished to remain in the marriage to the spouse who wished to leave. Thus, under mutual divorce the spouse who wished to leave the marriage had to compensate the spouse who wished to stay in order to obtain the efficient outcome of divorce, while under unilateral divorce the spouse who wishes to stay in the marriage must compensate the one who wishes to leave. The law change would alter the compensation scheme between the spouses without making them more likely to divorce.

[6] This intention is stated in the letter of the law in several states, such as Wisconsin. With respect to actual practice at the time, studies showed that many divorces were in fact by mutual agreement, regardless of legal requirements (Freed and Foster [1979]). As an example, news stories some years ago reported that, while the former wife of Governor Douglas Wilder of Virginia had divorced him on the grounds of cruelty, they had fabricated the grounds in order to obtain a divorce which they both wanted. Such occurrences suggested to the legal and legislative community that blame was not necessary to determine whether a couple should divorce.

II. MEASURING THE EFFECT OF UNILATERAL DIVORCE

A. Approaches in the Previous Literature

Peters and Allen discussed several ways in which asymmetric information or transaction costs could destroy the Coase theorem result and cause a switch to a unilateral divorce law to encourage divorce. In order to determine the relevance of the Coase theorem, Peters endeavored to measure the impact of a state's divorce law on divorce probabilities. In his *Comment*, Allen used an almost identical regression but arrived at a different conclusion. Peters and Allen both used a supplementary survey of individuals' marital history between 1975 and 1978 from the 1979 Current Population Survey. Both estimated a logit on a sample of women who reported being married in 1975, where the left-hand side variable was one if a woman became divorced by 1978 and zero if she stayed married. Controlling for demographic variables believed to affect the likelihood of divorce, they used the variation of divorce regimes across states at that time to identify whether unilateral divorce raised divorce rates.

Peters (1986) also included on the right-hand side a state's 1970 divorce rate as a proxy for the state's unobserved propensity for divorce. She added four region dummies as well, which, she argued in Peters (1992), served the same purpose in controlling for unobserved divorce propensities in addition to controlling for other unobserved regional characteristics correlated with divorce. The coefficient she estimated on the unilateral dummy was virtually zero. Allen argued against using the region dummies and also omitted the 1970 divorce rate. He found a significant positive impact: living in a state with a unilateral divorce law raises the probability of divorce by 1.4%.

B. Estimating the Effect with Panel Data

The debate about unobserved state heterogeneity in divorce propensities can be addressed by using a panel. A panel also reveals, through year effects, how evolving nationwide characteristics, such as attitudes, affect divorce. Finally, a panel is better able to reveal a gradual impact of the law change.[7]

The relationship which Peters estimated can be easily adapted for a panel of state divorce rates. Her regression took the following form:

$$(1) \quad divorce_{ist} = a_0 + a_1 * unilateral_{st} + a_2 * Z_{ist} + e_{ist}\,,$$

which describes a population of women who are married at the beginning of the period of observation.[8] Each observation is subscripted for the individual i, the state s in which she lives and the time period t in which she is observed, although t is degenerate when the data is cross-sectional. The variable *divorce* is one if the woman is divorced in year t and zero if she is not. We are interested in the variable *unilateral*, which is one if the woman's state has a unilateral divorce law in that year and zero if it has a mutual divorce law. Her demographic characteristics, such as education, number of children, and age, are summarized in Z.[9]

To estimate the relationship expressed in (1) with state-level aggregate data, add up equation (1) over observations within a state and divide by the population to get the following:

$$(2) \quad divrate_{st} = b_0 + b_1 * unilateral_{st} + b_{2s} * state_s + b_{3t} * year_t + u_{st}\,.$$

Now, the dependent variable *divrate* is the state level divorce rate, which is the number of divorces that occur within a state each year divided by the state population in thousands. The disturbance u is equal to $(\sum_{i \epsilon s} e_{ist})/pop_{st}$, the sum of all individual disturbances within a state each year divided by the population. This makes u heteroscedastic, so population is

---

[7] Peters' and Allen's cross-sectional data was collected during the middle of the transition to unilateral divorce, which would lead them to underestimate its true effect if it only emerges gradually.

[8] Peters estimated a logit, not a linear probability model as expressed in (1).

[9] Peters used age, education, race, number of kids under 18 and number of kids squared, an SMSA dummy, and a dummy for being in the labor force.

used to perform weighted least squares.[10] Using a panel allows the inclusion of *year*, which are year effects that control for evolving unobserved national attributes that affect the likelihood of divorce. Finally, the demographic characteristics Z from (1) become, for example, the number of urban residents or the number of children in the state.

Equation (2) assumes that such demographic characteristics are unchanging within a state over time. The coefficients cannot be separately identified, so the state characteristics are subsumed in the fixed effect *state*. The use of state effects, which is only possible with the panel, can be advantageous: it is a more flexible way to explain divorce patterns, rather than attempting to include all relevant divorce covariates on the right-hand side. However, the factors which influence divorce may vary within a state over time, which will confound the estimates of the state effects. That will also bias the estimate of the coefficient on *unilateral* if the changing factors are correlated with the law changes across states, and if such factors do not change at a national level uniformly and get picked up by the year effects. The specification in (2) can be relaxed further, then, to capture such changing influences within a state over time, provided there is enough variation in *unilateral*. For instance:

$$(3)\quad divrate_{st} = b_0 + b_1{*}unilateral_{st} + b_{2s}{*}state_s + b_{3t}{*}year_t + b_{4s}{*}state_s{*}time + b_{5s}{*}state_s{*}time^2 + u_{st}\,.$$

The variable *time* is a time trend, so the interaction terms *state*time* and *state*time*^{2} are linear and quadratic trends for each state.[11]

This specification approximates nonlinear trends in state-level characteristics that influence divorce, with the patterns allowed to vary across states. Thus, (3) represents an extremely flexible way to control for heterogeneous divorce behavior. While it remains possible that this specification still does not capture all of the factors explaining *divrate*,

[10] The intuition behind weighted least squares is that a positive effect of *unilateral* in California will carry more weight than, say, a positive effect in New Hampshire.

[11] The use of state-specific linear trends follows Jacobson, LaLonde and Sullivan (1993).

additional evidence against the endogeneity of *unilateral* will be presented after the results of estimating (1), (2) and (3) are discussed.

III. DATA

The panel used in this paper covers all fifty states and the District of Columbia from 1968 to 1987. Most states legalized unilateral divorce during this time period, as reported in Table 1. These data are based on Freed and Foster (1977, 1979, 1981), Freed and Walker (1990), Sepler (1981) and on states' legal codes themselves.[12] Peters and Allen disagreed over how to classify states whose only unilateral ground is living separate and apart -- from six months to five years, depending on the state and the year.[13] I test how the living separate ground affects the results and find that allowing it a separate effect strengthens the measured effect of unilateral divorce, but not significantly. Also, too few states have the living separate ground to determine whether it more closely resembles mutual or unilateral divorce. In the results presented in the rest of this paper, I classify the living separate ground as unilateral divorce.

Divorce rates by state, reported in *Vital Statistics of the United States*, are shown in Table 1. The sample covers virtually every divorce in the U.S. for twenty years. It is not complete because, in certain states in certain years, *Vital Statistics* reports that the data was collected in an irregular or incomplete fashion. Eleven dummies, described in the Appendix, control for those instances. There are a few missing values as well, which are listed in the footnotes to Table 1.

---

[12] The classification of states is not without controversy. Part of the confusion arises over the distinction between no-fault and unilateral divorce, which led me to consult the state laws themselves.

[13] Peters claimed that the living separate and apart ground requires a large enough cost that it would be cheaper for a couple to agree on a compensation scheme than to divorce unilaterally, so she classified those states as having mutual divorce.

IV. RESULTS

The coefficient on *unilateral* turns out to be quite sensitive to the nature of state and time controls used, as Allen demonstrated. I first report the results from the standard panel data approach, where including state and year fixed effects wipes out most of the effect of *unilateral*. After adding state-specific trends as well, the coefficient on *unilateral* becomes significant and large and continues to be so under further modifications. Following that, I consider the issue of endogenous legislation and then review some other features of the results. Finally, I discuss the implications of the estimations and conclude.

A. Results with State and Year Effects Alone

The results of estimating the basic regression and then adding state and year effects, as in equation (2), are shown in Table 2. Regression 2.1 has no state or year effects, and the estimated coefficient on *unilateral* is statistically significant and very large. However, the explanatory power of the regression is low. Adding year effects in regression 2.2 boosts the explanatory power of the regression from 0.188 to 0.863, while adding state effects in regression 2.3 raises it to 0.948. Adding the year and state effects also eliminates most of the effect of *unilateral*. The resulting coefficient of 0.075 (0.048) is small, but significantly different from zero at a 12% confidence level. The state and year fixed effects -- which encompass everything affecting divorce behavior that is, respectively, constant within a state over time and uniform each year across the nation -- explain most of the patterns of divorce in the U.S. The additional effect of adopting a unilateral divorce law, according to regression 2.3, was to raise divorces by less than 0.1 per thousand people, or 1.5% of the average national divorce rate of 5.0.

B. Adding State-Specific Trends

It is clear from the results above that unobserved covariates and unobservable divorce propensities -- which may include, for instance, social attitudes, religious beliefs and family size -- are the main determinants of divorce. However, assuming that they are constant within a state over twenty years or changing over time but uniform across the nation is restrictive. State-year interactions are completely unrestrictive but infeasible. State-specific trends, though, are a feasible alternative that allows the unobserved state factors influencing divorce to have a linear trend, with the trend allowed to vary across states. The estimated coefficient on *unilateral* retains the same interpretation as before: a positive coefficient measures by how much a unilateral divorce law raises a state's divorce rate, which now might also have a positive or negative trend.

Regression 2.4 includes the state-specific trends. The regression has a higher adjusted R^2 than 2.3, and the state trends are jointly significant. More importantly, the coefficient on *unilateral* is larger and has a smaller standard error. The estimate of 0.331 (0.039) means that unilateral divorce raised the national divorce rate by 0.3 divorces per thousand people, which is 6.7% of the sample mean of 5.0.

Why the big change in the estimate of *unilateral*'s coefficient when state trends are added? It is because imposing a flat state divorce propensity misspecifies the underlying divorce behavior and in turn camouflages the variation induced by the law change. Naturally, imposing constant state divorce propensities when they are really trending will bias the estimates of the constants -- the estimated intercepts will reflect the average of the trend instead of the true intercept, which is the state's initial divorce propensity only. If, further, the omitted slopes are correlated with the law changes, the result is omitted variable bias. The problem is most easily understood visually, as in Figure 2-A. The upper line in Figure 2-A is California's actual divorce rate, with the box signalling the year of the law change. The lines near the bottom show what remain of the divorce rate to be explained by the law change in each regression. The lower dashed line removes the constant and the state and year effects from regression 2.3, and the solid line removes the state trend as well from regression 2.4. Both lines jump in the year of the law change, but the dashed line from regression 2.3 ultimately confounds the influence of *unilateral* with the omitted downward trend in California's divorce propensity, biasing downward the estimated effect of *unilateral*. The solid line, on the other hand, demonstrates that, controlling for all other influences, the law change raised the divorce rate permanently. Figure 2-B illustrates the same point for Michigan. While the effect of *unilateral* in Michigan is smaller, the solid line still shows that the law change raised the divorce rate, while the dashed line obscures the effect of *unilateral*.

The possibility remains that unobservables exhibit more complex dynamic behavior than can be captured by the linear trend. Therefore, regression 2.5 includes a quadratic in the time trend for each state as well. Now, the effect of *unilateral* is measured as a break in the pattern of divorce at a certain moment not resembling a smooth quadratic. The state quadratic terms in regression 2.5 are jointly significant, raising the adjusted R^2 once more. Meanwhile, the estimated coefficient on *unilateral* declines slightly and the standard error increases, with an estimate of 0.304 (0.053) compared to 0.331 (0.039) without the quadratic. The point estimate is a little less than 10% smaller but not significantly different, so the impact of *unilateral* remains large even when the underlying divorce behavior is paramet zed very flexibly, which places strong restrictions on the kind of variation in the divorce rate that can be attributed to the law change.

The strongly significant coefficient on *unilateral* when state trends are included indicates that both Peters' and Allen's regressions were misspecified. Unless the individual-level demographic controls he used fully explain the divorce rate, Allen -- by omitting any controls for state divorce propensities -- was naturally more likely to find an impact of the law change, as regression 2.1 demonstrated. Yet, Peters' approach of including the state's 1970 divorce rate was also inadequate. The effect of including the divorce rate from several years earlier is similar to that of including only the state fixed effect as in regression 2.3, obscuring the impact of *unilateral*.

C. Endogenous Legislation

Any analysis of how legal changes affect behavior raises the question of endogeneity. In the case of divorce, it might be that rising divorce rates caused the law to be changed or that increasing permissiveness towards divorce, which in turn raised divorce rates, did -- not that the law changes caused divorce rates to rise. Several authors argue against this.[14] Also, Peters asserted that interest groups of divorcing individuals were not part of the debate

[14] For example, legal scholars Freed and Foster (1979) and Sepler (1981).

over no-fault divorce laws and that proponents of the change were primarily legal scholars. Unfortunately, it is difficult to come by instruments in order to test properly for endogeneity.[15]

As an alternative, I have explored the correlation across states of initial divorce rates and the chronology of the law changes. As a rough indicator, I ranked all the states by their 1968 divorce rates and then by the year in which they switched to unilateral divorce. The correlation coefficient between the two rankings was 0.09, indicating almost no relationship.

As an additional check, I employed a strategy suggested by Gruber and Hanratty (1993). They added to their model a dummy variable of one if the law change they studied was to occur in the following year. Under the null hypothesis in which the law change has a causal effect on the left-hand side variable, the dummy should have a coefficient of zero. Otherwise, if there is reverse causality or some other type of endogeneity, then the coefficient will be different from zero. Regression 4.1 in Table 4 shows that the coefficient on the lead dummy is indistinguishable from zero.

D. Additional Analysis

Because they generate most of the explanatory power in the regressions, the estimated coefficients on the fixed effects and trends are of interest in themselves. The coefficients on the year effects, which control for unobserved factors that affect divorce nationally, are significant and quite big, although adding state trends in regressions 2.4 and 2.5 reduces their magnitude a little and raises their standard errors. The pattern of the year effects mirrors the trend of divorce in Figure 1, with the coefficients growing bigger through the 1970s, peaking around 1980, and then stabilizing. Thus, whatever were the factors common across the nation that affected divorce -- more liberal attitudes towards divorce, for example -- they increasingly raised divorce rates until 1980, after which their influence steadied.

---

[15] An instrument must be correlated with the law changes but not with the divorce rate independently and must also, in this case, exhibit state and time variation. In a time series study, Sepler used the number of articles about no-fault divorce in legal journals as a control; but that does not vary across states. Other potential instruments relate to the composition of the legislature (number of women, number of divorced people, etc.), but those data are costly to obtain.

The coefficients on the state effects and trends from regression 2.4, shown in Table 3, tell us about geographical patterns in divorce propensities. Large negative intercepts are prevalent in the Northeastern and North Central states, indicating that their residents had the least propensity to divorce at the beginning of the sample. But while starting low, the propensity to divorce rose somewhat in parts of the region, as indicated by the positive coefficients on the slopes. Residents of Southern states generally began with a moderate propensity to divorce that rose over the period, as indicated by intercept coefficients close to zero and positive slope coefficients. Meanwhile, residents of Western states started with the greatest propensity to divorce, which moderated in some states.[16]

Given the emphasis above on the dynamics of the unobserved factors affecting divorce, it is instructive to explore whether the estimation properly describes the dynamics of how *unilateral* affected divorce. For instance, it might be that the effect of switching to a unilateral divorce law is temporary, not permanent.[17] To evaluate the duration of *unilateral*'s impact, I added dummies that are one in the years of and immediately following the law change. If the effect of the law change were actually temporary, we would expect a positive coefficient on these lags and an insignificant coefficient on *unilateral*.

The results are shown in regressions 4.2-4.4 of Table 4, of which the most illustrative is the last, with dummy variables for the year of and two years following the law change. The more general dynamic structure actually strengthens the effect of *unilateral*, although the estimate of 0.431 (0.051) for the long-run effect is not statistically different from the estimate

---

[16] The intercept and slope coefficients on Nevada, while noticeably large, are actually not very important. Nevada is known as a haven for both marriage and divorce, the latter because of very short residency requirements. This would bias upward the coefficient on *unilateral* if a couple that would have divorced in Nevada now divorces at home -- a story corroborated by Nevada's large negative slope coefficient. However, the very high divorce rate in Nevada, shown in Table 1, is misleading because Nevada has an extremely small resident population. Divorces in Nevada have ranged between nine and fifteen thousand since the late 1960s, while nationwide divorces hit one million by 1975. The percentage of nationwide divorces obtained in Nevada, while closer to 2% in the 1960s, has been in the neighborhood of 1% since then. Those magnitudes will not have much of an impact in estimating the coefficient on *unilateral* with weighted least squares.

[17] For instance, the switch to unilateral divorce may have led to a "stock adjustment", where a backlog of bad marriages were ended without raising the long-term probability of divorce.

in regression 2.5. The lag dummies have negative coefficients, revealing that the long-run effect was attained gradually. Thus, in the year of the law change, the point estimate of the impact of *unilateral* was to raise the divorce rate by 0.255 [0.431-0.176], and in the following year by 0.297 above the level two years earlier. This result also gives another reason why the estimated magnitude of *unilateral* is much larger than Peters' and Allen's, because their estimates pertain to a single cross-section observed in the middle of the transition period from mutual to unilateral divorce in many states.

Finally, the results are robust to the classification of states whose only unilateral ground is living separate and apart for a specified time. The regressions reported in the text classify such states as having unilateral divorce. Peters argued instead that those states, in effect, have mutual divorce, leading to a reclassification of fifteen states.[18] Using this tighter definition, the estimated coefficient of *unilateral* increases to 0.429 (0.050), and then adding dummy variables for the disputed states according to the required length of separation yields a coefficient of 0.432 (0.050). The coefficient on the dummy for requiring a twelve month separation is 0.516 (0.104), while the coefficients on the other dummies are in the range of 0.05 to 0.20 but with big standard errors. The point estimates suggest that the separation requirement may indeed invoke a different response than other unilateral grounds but still encourages divorce, but the standard errors make it impossible to state anything conclusively,

V. DISCUSSION AND CONCLUSION

Using panel data sheds new light on the impact of the no-fault revolution on divorce rates. Previous studies by economists found conflicting evidence about the impact of unilateral divorce -- attributing to it at most a relatively small increase in the divorce rate. This has been interpreted as affirming the relevance of the Coase theorem in explaining the absence of an effect of the divorce regime on marriage outcomes.

The earlier studies used cross-sectional data, exploiting the cross-state variation in

---

[18] Two of the cases actually involve reclassifications after the period which Peters considers.

divorce laws. However, unobserved attributes of state populations might be correlated with both divorce behavior and the divorce law, leading to a biased estimate of the law's influence. In this paper, I use a panel of state-level divorce data, which allows me to identify the impact of divorce laws off of the legal variation within states over time. This approach reveals that controls for cross-state heterogeneity affect the conclusions about unilateral divorce substantially. However, simply adding state and year effects does not absorb all of those effects. The impact of the law change on divorce becomes evident by freeing up the state controls to capture smooth trends in divorce behavior. Using such flexible controls shows that unilateral divorce raised divorce rates by about 7%, which amounted to 16% of the overall increase between 1970 and 1985. Moreover, this estimate is robust (and even increases) when the impact of the law change is permitted to be either temporary or permanent; when the possibility of endogenous legislation is considered; and when a distinct effect for living separate and apart requirements is allowed.

These results demonstrate that the Coase theorem does not apply: transaction costs in marriage and divorce were altered by the shift in the divorce regime, and/or asymmetric information caused property rights to shift when the law did. The only way to get a clearer picture of how this occurred would be to examine a longitudinal data set of individuals that included information on demographic and economic variables affecting divorce, the cost of divorcing, and the asymmetric information and property rights it involves.

In sum, the current move to tightening divorce requirements should indeed lower the number of divorces -- although without necessarily raising welfare. It is also important to note that the results demonstrate that other factors besides the switch to unilateral divorce had a great deal to do with the increase in divorce in the U.S.

TABLE 1: Divorce Rates and Law Changes, by State						
State	Average Divorce Rate: Number of Divorces per 1000 people 1968-72	1973-77	1978-82	1983-87	Year in which unilateral divorce law was passed[a,b]	Details
Alabama	4.4	6.1	6.7	6.3	1971	
Alaska	5.6	7.6	8.5	7.4	before 1968	
Arizona	7.0	7.4	7.3	6.8	1973	
Arkansas	5.5	6.9	6.9	6.8	never	separation is grounds, but only if mutually agreed on
California	4.9	5.9	5.6	5.2	1970	
Colorado	4.7	6.4	6.3	5.9	1971	
Connecticut	2.1	3.6	4.3	3.7	1973	
Delaware	2.9	4.6	5.0	4.8	never	
District	3.3	4.6	6.2	4.4	1977	one year separation required
Florida	5.8	7.2	7.2	6.7	1971	
Georgia	4.0	5.7	6.2	5.6	1973	
Hawaii	3.6	4.9	4.7	4.4	1973	
Idaho	4.9	6.2	6.8	6.1	1971	
Illinois	3.4	4.3	4.5	4.2	1984	two year separation required
Indiana[1]	4.6	6.1	7.0	6.3	1973	
Iowa	2.6	3.5	3.9	3.6	1970	
Kansas	3.8	5.2	5.5	5.1	1969	
Kentucky	3.3	4.2	4.3	4.9	1972	
Louisiana[2]	2.7	3.2	3.8	3.6	before 1968	one year separation required
Maine	3.7	4.9	5.5	5.0	1973	
Maryland	2.4	3.7	4.0	3.6	before 1968	five year separation required; then two years
Massachusetts	2.1	2.8	3.1	3.2	1975	
Michigan	3.4	4.6	4.7	4.2	1972	
Minnesota	2.2	3.3	3.8	3.5	1974	
Mississippi	3.6	5.0	5.4	4.9	never	
Missouri	3.9	5.0	5.5	5.0	1973	two year separation required
Montana	4.4	5.7	6.2	5.3	1975	
Nebraska	2.5	3.6	4.1	3.9	1972	

State	1968-72	1973-77	1978-82	1983-87	Year	Details
Nevada	20.9	17.2	16.2	14.4	1973	
Nw Hampshire	3.5	5.0	5.5	4.9	1971	
New Jersey	1.7	2.8	3.6	3.7	1971	eighteen month separation required
New Mexico[3]	4.3	7.3	7.8	8.5	1973	
New York	1.5	3.0	3.5	3.7	never	
North Carolina	2.8	4.0	4.8	4.9	before 1968	one year separation required
North Dakota	1.7	2.7	3.3	3.4	1971	
Ohio	3.6	4.9	5.3	4.9	1974	one year separation required
Oklahoma	6.3	7.6	7.8	7.5	1953	
Oregon	4.7	6.3	6.7	5.9	1973	
Pennsylvania	2.0	2.9	3.3	3.4	1980	three year separation required, and court determines marriage is broken
Rhode Island	1.6	3.0	3.7	3.8	1976	
South Carolina	2.3	3.3	4.3	4.1	1969	three year separation required; later one year
South Dakota	2.1	3.2	3.9	3.6	1985	
Tennessee	4.2	6.1	6.5	6.3	never	two year separation if no kids
Texas	4.7	5.9	6.6	6.0	1974	
Utah	3.8	4.9	5.3	5.2	before 1968	three year separation required
Vermont	2.2	3.8	4.6	4.5	before 1968	six month separation required
Virginia	2.6	3.8	4.5	4.4	before 1968	two year separation required
Washington	5.3	6.9	6.8	6.1	1973	
West Virginia	3.0	4.5	5.3	5.0	before 1968	two year separation required; later one year
Wisconsin	2.0	2.9	3.7	3.5	1977	one year voluntary separation required; if not voluntary and one claims irreconcilable differences, then court decides
Wyoming	5.6	7.1	8.0	7.2	1977	
National	3.5	4.8	5.2	4.9	1977	

[a] Unilateral divorce requires the consent of only one party and may be granted on the grounds of irretrievable breakdown, irreconcilable differences, and/or living separate and apart.
[b] If the law change occurred in the second half of the year, it is attributed it in the regressions to the following year.
[1] Data for 1968-69, 1974, 1976-87. [2] Data for 1971-72, 1976, 1978-80, 1982-83. [3] Data for 1970, 1974-80, 1983-85.
Source: Divorce rates from *Vital Statistics of the United States*, various years. Data on law changes from Sepler (1981) and from annotated legal codes, various states.

TABLE 2: Regression Results

	Dependent Variable: Divorce Rate (Divorces per 1,000 people)				
Independent Variables:	2.1	2.2	2.3	2.4*	2.5
Intercept	3.650 (0.081)	2.787 (0.215)	3.192 (0.177)	3.173 (0.229)	3.191 (0.286)
Unilateral	1.377 (0.101)	0.885 (0.114)	0.075 (0.048)	0.331 (0.039)	0.282 (0.043)
Adjusted R^2	0.188	0.241	0.948	0.976	0.981
State Effects**	No	No	Yes, F=262.9	Yes, F=209.1	Yes, F=125.3
Year Effects**	No	Yes, F=4.6	Yes, F=80.7	Yes, F=99.1	Yes, F=8.8
State Trend, Linear**	No	No	No	Yes, F=22.1	Yes, F=8.2
State Trend, Quadratic**	No	No	No	No	Yes, F=5.9

... *continued on next page* ...

* The coefficients on the state effects and trends are shown in Table 3.
** All reported F-statistics have corresponding p-values that are smaller than 0.00005.
Standard errors are in parentheses. N=994. Regressions are weighted by state population. The 1968 year effect is omitted. The estimated coefficients on datadum are not shown (Appendix 1).

TABLE 2: Regression Results, continued

Independent Variables:	Dependent Variable: Divorce Rate (Divorces per 1,000 people) 2.1	2.2	2.3	2.4*	2.5
1969 year effect		0.221 (0.303)	0.234 (0.079)	0.224 (0.058)	0.221 (0.080)
1970		0.402 (0.305)	0.548 (0.080)	0.508 (0.068)	0.507 (0.132)
1971		0.544 (0.303)	0.760 (0.080)	0.704 (0.081)	0.701 (0.181)
1972		0.756 (0.303)	1.066 (0.080)	0.975 (0.097)	0.978 (0.225)
1973		1.013 (0.305)	1.337 (0.081)	1.219 (0.115)	1.220 (0.264)
1974		1.116 (0.305)	1.572 (0.082)	1.421 (0.133)	1.428 (0.296)
1975		1.419 (0.307)	1.798 (0.083)	1.631 (0.152)	1.640 (0.322)
1976		1.523 (0.305)	1.923 (0.083)	1.751 (0.170)	1.762 (0.343)
1977		1.476 (0.305)	1.878 (0.083)	1.697 (0.189)	1.711 (0.357)
1978		1.579 (0.304)	2.000 (0.083)	1.814 (0.208)	1.821 (0.366)
1979		1.743 (0.304)	2.162 (0.083)	1.974 (0.227)	1.984 (0.370)
1980		1.669 (0.305)	2.136 (0.084)	1.933 (0.247)	1.947 (0.370)
1981		1.751 (0.305)	2.166 (0.084)	1.965 (0.266)	1.981 (0.365)
1982		1.463 (0.304)	1.923 (0.083)	1.721 (0.286)	1.740 (0.357)
1983		1.385 (0.303)	1.822 (0.083)	1.616 (0.305)	1.638 (0.348)
1984		1.342 (0.307)	1.853 (0.085)	1.575 (0.324)	1.605 (0.339)
1985		1.303 (0.308)	1.857 (0.085)	1.563 (0.344)	1.608 (0.333)
1986		1.256 (0.306)	1.767 (0.085)	1.477 (0.364)	1.533 (0.332)
1987		1.246 (0.303)	1.633 (0.084)	1.429 (0.383)	1.499 (0.340)

TABLE 3: Coefficients on State Fixed Effects and Trends in Regression 2.4

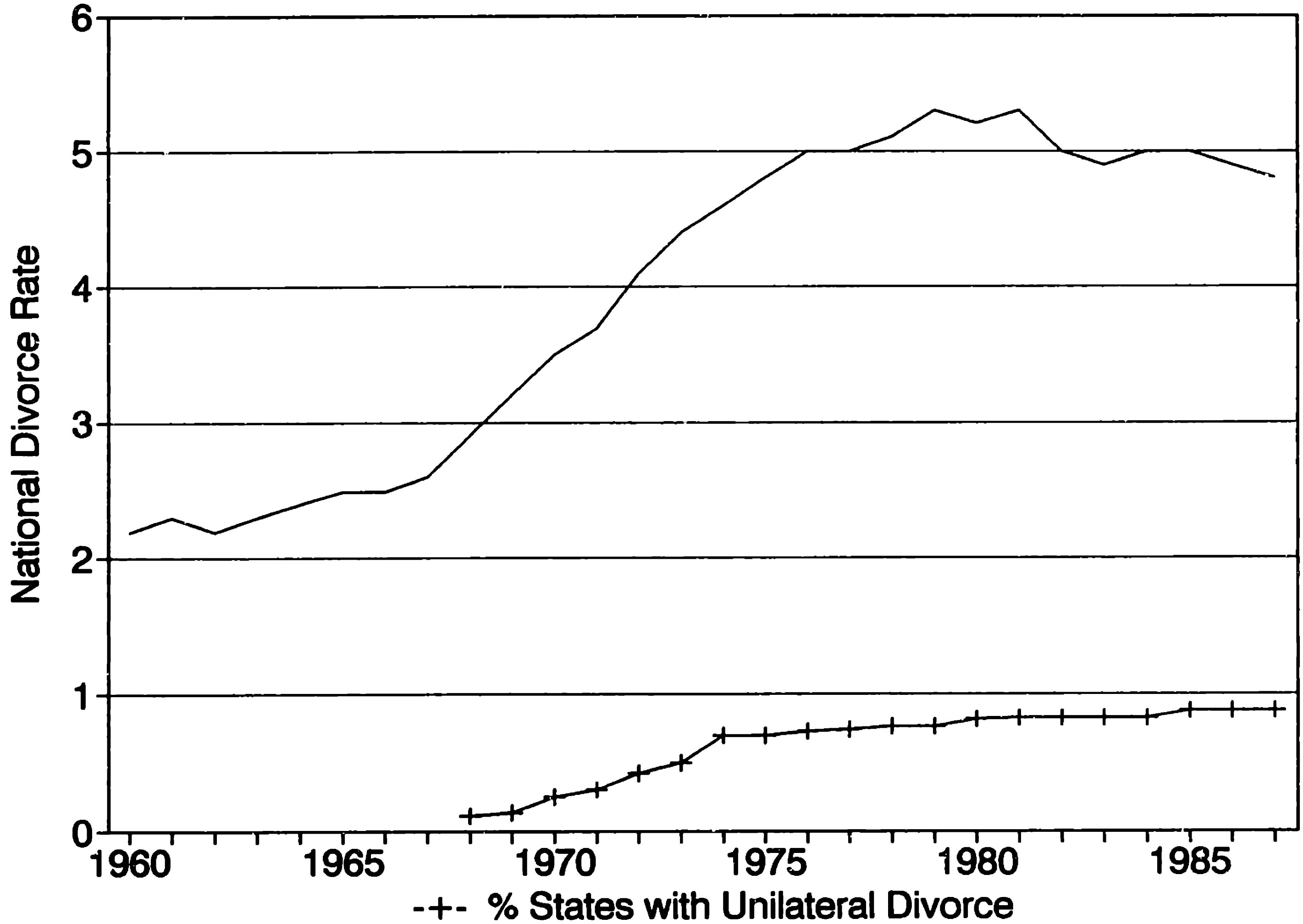
	Coefficient (Std Error) on:			Coefficient (Std Error) on:	
State	Fixed Effect	Trend	State	Fixed Effect	Trend
NORTHEAST			West Virginia	-1.037 (0.286)	0.065 (0.025)
Maine	(omitted)	(omitted)	Nth Carolina	-1.467 (0.250)	0.075 (0.022)
Nw Hampshire	-0.099 (0.345)	0.004 (0.030)	Sth Carolina	-1.863 (0.267)	0.052 (0.023)
Vermont	-1.846 (0.409)	0.082 (0.036)	Georgia	0.446 (0.249)	0.013 (0.022)
Massachusetts	-1.758 (0.246)	-0.016 (0.022)	Florida	2.250 (0.241)	-0.032 (0.021)
Rhode Island	-2.149 (0.327)	0.047 (0.029)	EAST SOUTH CENTRAL		
Connecticut	-1.553 (0.262)	0.020 (0.023)	Kentucky	-0.811 (0.259)	0.043 (0.023)
MIDDLE ATLANTIC			Tennessee	0.599 (0.253)	0.066 (0.022)
New York	-2.316 (0.233)	0.074 (0.021)	Alabama	0.745 (0.257)	0.034 (0.022)
New Jersey	-2.359 (0.242)	0.051 (0.021)	Mississippi	0.070 (0.273)	0.014 (0.025)
Pennsylvania	-1.772 (0.236)	-0.002 (0.021)	WEST SOUTH CENTRAL		
EAST NORTH CENTRAL			Arkansas	1.838 (0.279)	0.003 (0.025)
Ohio	0.041 (0.237)	-0.008 (0.021)	Louisiana	-1.963 (0.337)	0.030 (0.031)
Indiana	1.403 (0.351)	-0.012 (0.023)	Oklahoma	2.207 (0.299)	0.015 (0.025)
Illinois	-0.183 (0.237)	-0.033 (0.021)	Texas	0.998 (0.236)	0.004 (0.021)
Michigan	-0.267 (0.239)	-0.031 (0.021)	MOUNTAINS		
Wisconsin	-1.774 (0.251)	0.010 (0.022)	Montana	0.907 (0.353)	-0.027 (0.031)
WEST NORTH CENTRAL			Idaho	1.262 (0.347)	-0.003 (0.030)
Minnesota	-1.530 (0.255)	-0.004 (0.022)	Wyoming	2.136 (0.446)	0.014 (0.038)
Iowa	-1.346 (0.265)	-0.008 (0.023)	Colorado	1.145 (0.271)	-0.014 (0.024)
Missouri	0.252 (0.250)	-0.015 (0.022)	New Mexico	1.030 (0.669)	0.143 (0.064)
Nth Dakota	-2.335 (0.368)	0.030 (0.032)	Arizona	2.516 (0.344)	-0.045 (0.028)
Sth Dakota	-1.633 (0.360)	0.025 (0.032)	Utah	-0.315 (0.313)	0.027 (0.027)
Nebraska	-1.385 (0.294)	0.011 (0.026)	Nevada	16.800 (0.382)	-0.467 (0.032)
Kansas	-0.030 (0.274)	0.009 (0.024)	PACIFIC		
SOUTH ATLANTIC			Washington	1.888 (0.257)	-0.043 (0.023)
Delaware	-0.711 (0.380)	0.053 (0.034)	Oregon	1.241 (0.274)	-0.013 (0.024)
Maryland	-1.460 (0.256)	0.003 (0.022)	California	1.215 (0.233)	-0.073 (0.021)
Dist. Columbia	-0.148 (0.351)	0.004 (0.032)	Alaska	2.060 (0.459)	0.038 (0.039)
Virginia	-1.528 (0.252)	0.050 (0.022)	Hawaii	0.089 (0.340)	-0.045 (0.030)

TABLE 4: Alternative Specifications of Regression 2.4

Independent Variables	Dependent Variable: Divorce Rate (Divorces per 1,000 people)				
	2.4	4.1	4.2	4.3	4.4
Intercept	3.173 (0.229)	3.173 (0.229)	3.179 (0.228)	3.185 (0.228)	3.188 (0.228)
Unilateral	0.331 (0.039)	0.330 (0.044)	0.368 (0.042)	0.413 (0.046)	0.431 (0.051)
Dummy for Law Change Occurring:					
1 Period Later		-0.001 (0.053)			
In This Period			-0.130 (0.049)	-0.162 (0.051)	-0.176 (0.054)
1 Period Ago				-0.120 (0.051)	-0.134 (0.054)
2 Periods Ago					-0.042 (0.052)
3 Periods Ago					
Adjusted R^2	0.976	0.976	0.976	0.976	0.976

Standard errors are in parentheses. N = 994. Regressions are weighted by state population. The estimated coefficients on the constant, state effects and trends, and year effects are not shown.

Figure 1



National Divorce Rate
6
5
4
3
2
1
0
1960
1965
1970
1975
1980
1985
-+- % States with Unilateral Divorce

Figure 2-A
California

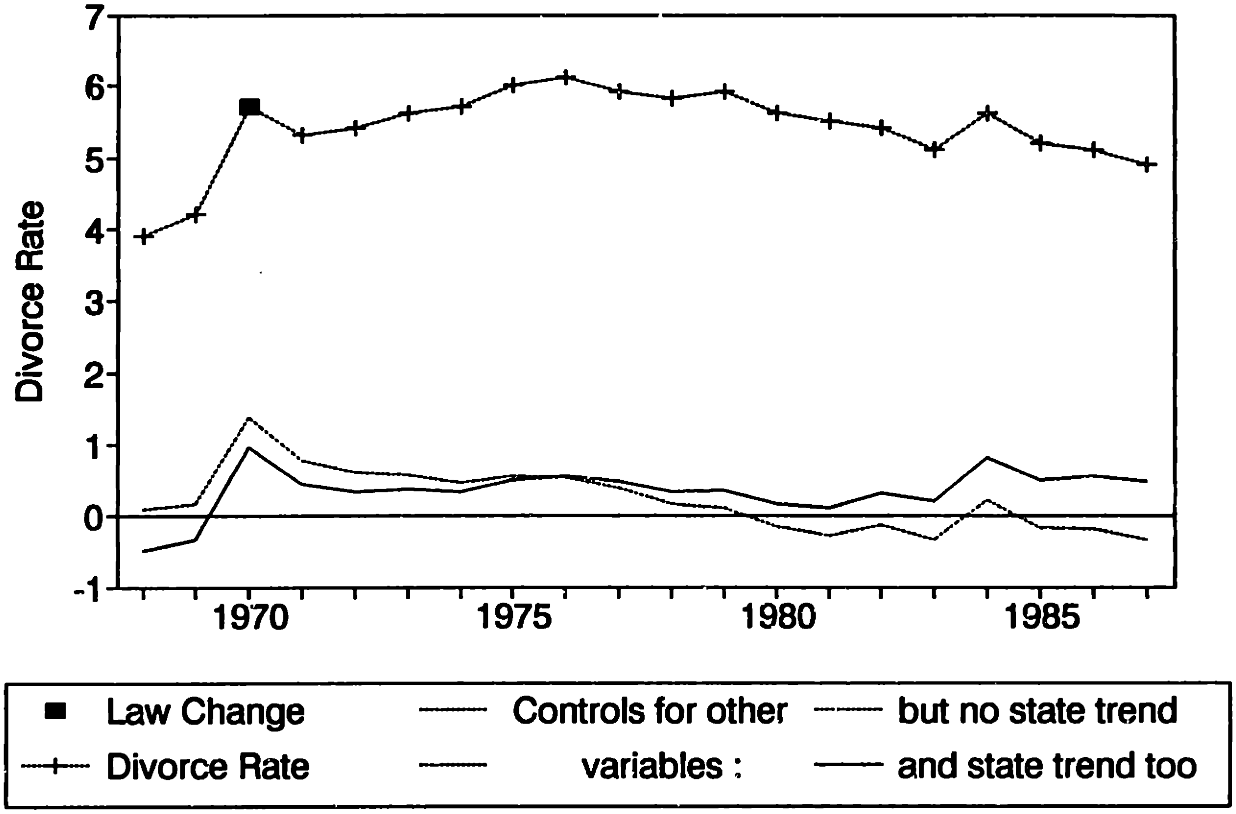
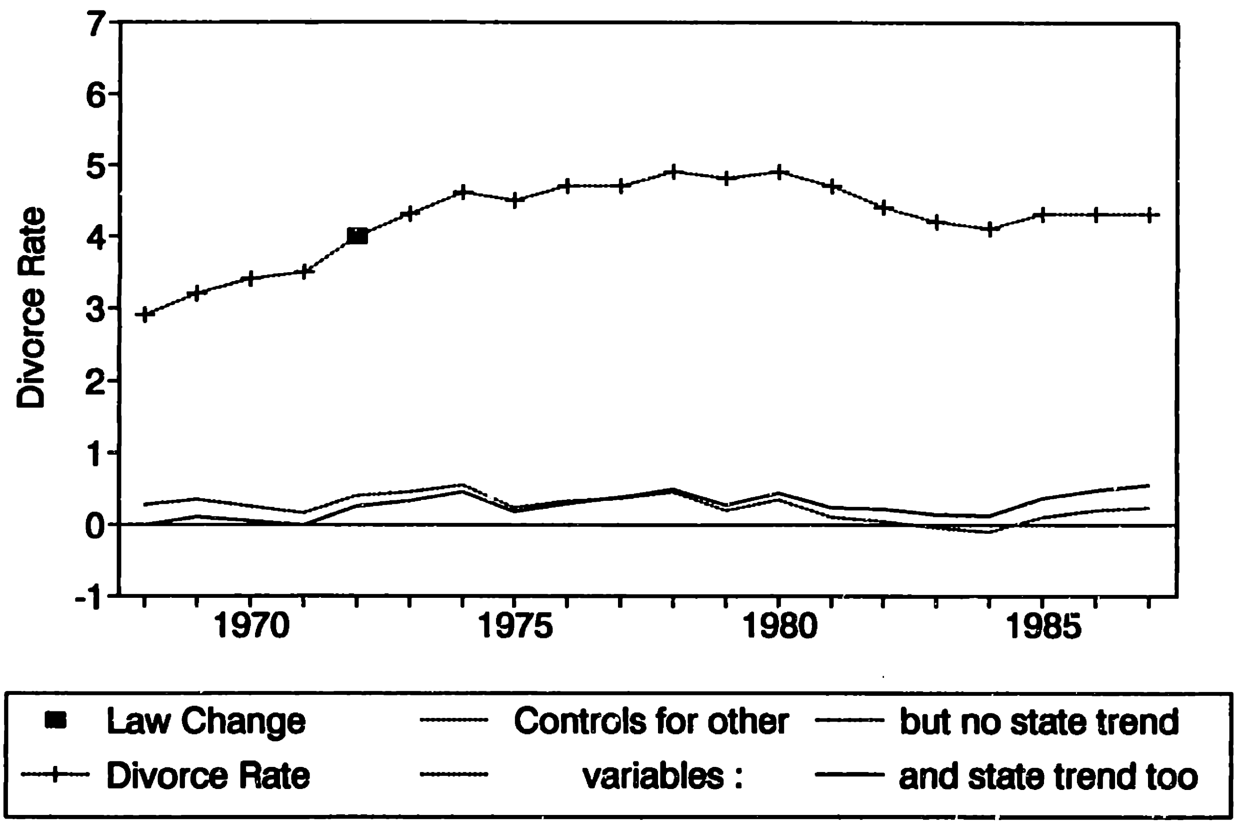


Figure 2-B
Michigan



APPENDIX

Dummy variables were included in the regressions to account for the following irregularities in the data on divorce rates:

State	*Time Period*	*Reason*
Kentucky	1978-1984	incomplete
Mississippi	1984-1985	incomplete
Arkansas	1973-1974, 1980, 1982-1985	incomplete
New Mexico	1968-1973, 1981-1982, 1986-1987	incomplete
New Mexico	1981-1982, 1984-1985	divorce petitions filed
California	1984-1986	include legal separations
Ohio	1981-1985	incomplete
Indiana	1974,1977-1986	include divorce petitions
Indiana	1968-1977	incomplete
Oklahoma	1970-1974	include divorce petitions
Arizona	1968-1972	include divorce petitions

Louisiana had incomplete data for the entire series, which is accounted for in the estimation by the Louisiana state effect. The regression results are quite similar when these observations are omitted.

REFERENCES

Allen, Douglas, W. 1992. "Marriage and Divorce: Comment." *American Economic Review* 82: 679-85.

Becker, Gary S. 1981. *A Treatise on the Family*, Cambridge: Harvard University Press.

Cherlin, Andrew J. 1992. *Marriage, Divorce, Remarriage*, Cambridge: Harvard University Press.

Freed, Doris Jonas, and Henry H. Foster, Jr. 1977. "Divorce in the Fifty States: An Outline." *Family Law Quarterly* 11: 297-313.

____________, and ____________. 1979. "Divorce in the Fifty States: An Overview as of 1978." *Family Law Quarterly* 13: 105-28.

____________, and ____________. 1981. "Divorce in the Fifty States: An Overview." *Family Law Quarterly* 14: 229-47.

____________, and Walker, Timothy R. 1990. "Family Law in the Fifty States: An Overview." *Family Law Quarterly* 23: 495-608.

Gruber, Jonathan, and Maria Hanratty. 1993. "The Labor Market Effects of Introducing National Health Insurance: Evidence from Canada." Mimeo, Massachusetts Institute of Technology.

Jacobson, Louis S., Robert J. LaLonde and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review* 83: 685-709.

Peters, H. Elizabeth. 1986. "Marriage and Divorce: Informational Constraints and Private Contracting." *American Economic Review* 76: 437-54.

Peters, H. Elizabeth. 1992. "Marriage and Divorce: Reply." *American Economic Review* 82: 686-93.

Sepler, Harvey J. 1981. "Measuring the Effects of No-Fault Divorce Laws Across Fifty States: Quantifying a Zeitgeist." *Family Law Quarterly* 15: 65-102.

Statistical Abstract of the United States. 1954-60.

Vital Statistics of the United States. 1960-1987.