

From the
Office of Tax Policy Research

WORKING PAPER SERIES

It's Not About the Money: Why Natural Experiments Don't Work on the Rich

by

Austan Goolsbee

University of Chicago and NBER

The Office of Tax Policy Research, established in 1987, promotes policy-oriented research in taxation, and serves as a liaison between the academic, business, and policy making communities. We are committed to using state-of-the-art methods to analyze tax policy issues, and to disseminate our findings, and those of a broader academic community, to people in the policy making community.

LEADING IN THOUGHT AND ACTION

**IT'S NOT ABOUT THE MONEY:
WHY NATURAL EXPERIMENTS DON'T WORK ON THE RICH**

Austan Goolsbee

University of Chicago, GSB,

American Bar Foundation,

and N.B.E.R.

Abstract

An influential literature on the effects of marginal tax rates on the behavior of the rich has claimed that the elasticity of taxable income with respect to the net of tax share is very high, possibly exceeding one. These high estimated elasticities imply that cutting taxes on the rich does not lose much revenue—possibly increases it—and that progressivity generates a large amount of deadweight loss. To identify this elasticity, these studies have conducted “natural experiments” comparing the rich to other income groups and assuming that they are the same except for changes in their tax rates. This paper tests the natural experiment assumption using alternative data on the compensation of a panel of several thousand corporate executives and finds it clearly to be false. Compared to others, the very rich have incomes which trend upward at a faster rate, are much more sensitive to demand conditions, and are much more likely to be in a form whose timing can be shifted in the short run. If representative, these three facts may reduce previously estimated elasticities by up to 75 percent or more. The paper also suggests ways of improving existing methods.

Section I: Introduction

One of the liveliest areas of debate of the last twenty years in public economics has been the argument over the behavioral effects of marginal tax rate changes. Whereas an early literature based on conventional measures of labor supply found very small effects of tax rates on hours worked, a new literature looking at total taxable income and using a more subtle “natural experiments” approach has usually found large elasticities of taxable income with respect to the net of tax share. Because of such findings, this literature has emphasized the futility of tax progressivity and the benefits of reducing marginal rates on high income people.

The methodological approach in what I term the New Tax Responsiveness (NTR) literature is to control for the unobserved determinants of taxable income by using “natural experiments,” in other words by comparing different income groups and assuming that unobserved variables are identical for individuals in those groups. If true, changes to the progressivity of the rate structure can be used to identify the effect of taxation. The NTR literature simply compares the changes in taxable income of the different groups to the changes in the tax rates of the groups to determine what is commonly referred to as the “difference-in-differences” estimate of the elasticity of taxable income. While many of the papers simply compare the groups without using statistical estimation, the same assumptions are required to identify regression estimates of the elasticity of taxable income.

The influence of the NTR literature such as Lindsey (1987) and Feldstein (1995a) is undeniable and, if correct, has profound implications for tax policy and revenue estimation. The backbone of the NTR approach, however, is this assumption that lower income people are a valid control group for higher income people—that the change in income of the two groups would have been identical if there were no change in taxes. If this assumption is false, existing estimates may have significant biases. Little is known about the natural experiment assumption, however, because the NTR literature has been based on tax return data that contain little information about the individuals and whether they differ in important ways.

This paper examines the natural experiment assumption by looking in other data to see if there are differences between the very rich and the “somewhat” rich and, if so, how much bias such differences may create for the NTR literature. The paper does this by examining data on prominent high-income people, primarily panel data on thousands of high-income executives from the detailed compensation data reported by companies in proxy statements and 10-Ks and compiled for all companies in the S&P 500, S&P Mid-Cap 400 and S&P Small-Cap 600 from 1991-1995 but also general data on prominent high-income people like professional athletes. The results from these data raise serious concerns with natural experiment methods comparing the very rich even to the somewhat rich. These publicly identified rich individuals may not be representative of other high-income people, but if they are, existing elasticities in the NTR literature may be overstated by 75 percent or more.

Three basic differences between the very rich and other groups in these data account for the bias: First, cash compensation of very high income people relative to even moderately high income people has risen substantially, creating a spurious correlation between tax cuts and relative income growth of the wealthy. As an example, subtracting out these income trends could cut the elasticity estimated from TRA86 in half or even 3/4.

Second, the compensation of very high income people is much more responsive to demand conditions than others' and the state of the business cycle during the tax bills of the last two decades may have biased the estimated elasticities upward—20-25 percent for TRA86, for example. For executives, this is true based simply on cash compensation. The increasing importance of stock options and other types of capital income makes this responsiveness even more pronounced.

Third, the share of income in a form whose timing can be shifted in the short run (e.g., bonus, options, long-term incentive payouts) rises dramatically with income in these data. This means that natural experiments, especially those based on the tax increase of 1993, may detect large temporary changes in the timing of transactions but classify them as behavioral responses. The true long-run responses are small. The increasing use of stock options at the high end has

also generated a substantial mean reverting component of income which makes the choice of sample problematic in NTR studies since they use tax return data.

By documenting these three facts and exploring the biases they create, this paper provides quantitative evidence against the natural experiments approach of the NTR when applied to very rich people. The very rich really are different. The results are in the tradition of papers such as Blundell, Duncan, and Meghir (1995) or Heckman (1996) which criticize the difference-in-difference methodology, but the results here are specifically directed toward the top of the income distribution. For each of the objections presented, the paper also presents methods which could be used at least to improve the estimates in existing NTR research, even when tax return data are the only data available.

The current sample of several thousand very high income people is substantially larger and more detailed than almost all samples in the previous literature but may not be representative of the full universe of high-income people, particularly the high salary income.¹ The paper does not examine the interest income or portfolio decisions of the rich, only wages and salaries. The theory implicit in the NTR literature, however, is really about the decisions people make about how to take their salary income and not about when they choose to realize capital gains so, if anything, these executives are of *greater* interest for tax policy research than is the average rich person. For example, while these executives make up a small fraction of the very rich (1% or less), Goolsbee (1997) has shown that they can account for up to 20% of the total wage and salary decline of the top 1,000,000 taxpayers in response to the tax increase of 1993.

This paper will proceed in several sections. Section II will outline the NTR literature and its findings. Section III will present an overview of the new data on the rich. Section IV will present three empirical facts from the data and explain how they may bias conventional NTR estimates and what could be done about them. Section V will conclude.

Section II: The New Tax Responsiveness Literature

The NTR literature arose in response to the literature on taxes and labor supply. In the conventional literature (see Pencavel, 1986, Heckman, 1993, or Moffitt and Wilhelm, 1998), tax responsiveness was measured as tax induced changes to hours worked. Supporting the general notion that labor supply is close to zero, taxes seemed to have little impact, especially for men (see Eissa, 1996).²

The NTR literature looks at changes to total taxable income rather than hours worked. Individuals can control not just their hours of work but also how much of their income they take in nontaxable form or report at all. Tax increases can reduce taxable income dramatically even if the individuals work the same number of hours. Between the two, it is taxable income that revenue authorities should care about when calculating deadweight loss or making revenue forecasts and therefore when calculating the optimal size of the government (see Feldstein, 1995b, or Slemrod and Yitzhaki, 1996).

The origin of the NTR literature is really Lindsey (1987) who showed that the tax cuts of the early 1980s at the high end of the income distribution corresponded to a large increase in the share of total taxable income going to high-income people. In essence he compared high-income people to people in other brackets over the 1982-1984 period, used the tax cut as a natural experiment, and found substantial implied elasticities. This work was based on cross-sectional tax return data as was the later work of Feenberg and Poterba (1993) which looked at the share of total income going to very high income people over time and found an important increase around TRA86. In response to criticism of cross-sectional data and the potential for “rank-reversal” problems (see Slemrod, 1992a), later efforts turned to data on panels of tax returns.

Feldstein (1995a), which explored the tax cuts of TRA86 with panel data, is really the canonical NTR paper. It explicitly described TRA86 as a natural experiment and used a “difference-in-differences” methodology to identify the taxable income elasticity. By assuming that people in the 49-50 percent brackets, 42-45 percent brackets, and 22-38 percent brackets would have had the same income growth but for the tax changes, Feldstein compared the relative income changes of the various tax bracket groups with their relative tax changes and found that

incomes of the very rich rose the most and that they were also the group with the biggest tax cut. The resulting elasticities of taxable income exceeded one.

Because I will revisit these estimates later, I present the exact calculation in the top panel of Table 1. The first three rows in the top panel list the change in the net of tax share and adjusted taxable income plus gross loss for each of the three groups as calculated on a panel of tax returns. The second three rows show the difference-in-differences for various combinations of the three groups. The implied elasticities are then listed in the rightmost column of the second three rows. They are all greater than one, indicating substantial shifts of taxable income in response to rate changes and possibly large enough to trigger what Slemrod (1994a) has called the high-income Laffer curve where cutting taxes on the rich would actually raise revenue.

Feldstein's data were criticized (see Slemrod, 1995) for having only a small number of observations at the high end (57 in the top two brackets) and his results for not using statistical methods which could indicate the precision of the estimates (Gravelle, 1993). Auten and Carroll (1995, 1997), however, using an internal treasury sample of thousands of high-income tax returns and a regression methodology were again able to find significant elasticities, though smaller than those in Feldstein. With this nonpublicly-available data, they also had information on occupation and other nontax factors as reported on the tax returns and they found that the controls for other factors and the weighting of the sample did make some difference to the results. Their best estimate of the elasticity was around 2/3.

One lingering question in the NTR literature has remained the critique of Slemrod (1992b, 1994a, 1995) that the elasticities estimated in these natural experiments are not behavioral elasticities at all, nor even permanent changes to the form of compensation but rather are simply changes to the timing of transactions in the short run. This timing issue returned to the forefront when Clinton raised taxes on the rich in 1993. Feldstein and Feenberg (1996), doing a preliminary analysis of the increase with cross-sectional tax return data, find that incomes of approximately one million richest taxpayers fell significantly from 1992-1993 when

taxes rose while incomes of lower income groups rose. Their calculations indicate a large elasticity. Others have responded that the change was temporary (Parcell 1996, Toder, 1996).

As the discussion below will indicate, the tax increase of 1992-1993 is extremely important for evaluating the NTR literature and the debate over its effect has resisted resolution with tax return data.³ Goolsbee (1997), however, using the detailed compensation data on high-income executives, shows that more than 20 percent of the total wage and salary decline of the top 1,000,000 taxpayers can be accounted for by around 10,000 executives and that this change was entirely temporary, driven by a dramatic rise in the exercise of options in late 1992 in anticipation of higher rates. In these data, the long-run elasticity of taxable salary income is at most .4 and probably close to zero.

In summary, the NTR literature consistently makes the natural experiment assumption, uses tax return data to estimate effects, and usually finds a large elasticity. There is little evidence, however, about whether tax changes really are natural experiments. Do the rich and the moderately rich really differ only because of taxes or is the difference about more than just the money? With tax return data, it is difficult to know. To whatever extent the groups differ systematically in a way that is correlated with tax changes, the existing estimates will be biased. Data with more information on the rich and their money are required to compare these groups.

Section III: Independent Data on High-Income People

For an independent source of data which allows for tests of the natural experiment assumption, this paper primarily uses information on individual corporate executives of 1500 companies from 1991-1995. By law, all publicly traded firms are required to report the total compensation in a variety of forms for their top five highest paid employees in various SEC documents like 10-Ks and proxy statements. Standard and Poor's compiles the information for all companies in the S&P 500, the S&P Mid-Cap 400 and the S&P Small Cap 600 in their EXECUCOMP database. These are the same data used in Goolsbee (1997). The paper will also use data on the incomes of other prominent high income people such as professional athletes.

The data on executive pay are much more detailed in their description of income variables than are the data from a tax return. In addition to total compensation, the executive compensation data follow the same individuals over time and separately report their income from salary, bonus, Long-Term Incentive Plan (LTIP) payout, options exercised, "other" income, and the Black-Scholes value of options granted. The executive compensation data also include information on the financial and accounting performance of the executive's firm as well as other information on the individual. I restrict the sample to people who had data in at least four years and whose firms have fiscal years ending in December so as to coincide with the tax year. Using executives regardless of the number of years they were in the sample made no difference to the results.

A summary of these executives is listed in table 2. Clearly these are very high-income people. The average total income in taxable forms for these executives in 1992 dollars is \$877,000 per year and the median is \$490,000. Salary and bonus make up the predominant share of taxable compensation (about 2/3). Additional forms of income not currently taxable like options granted and "other" compensation are also important. Because the choice of companies does not restrict the sample to large firms and because the data are not restricted to CEOs, about 20 percent of the executives have relatively "low" incomes of less than \$250,000.

One major advantage of using nontax return data on income is that there is no need to worry about making income comparable across years with different tax laws, as is necessary for taxable income. The meaning of each of the categories is unambiguous. One disadvantage of using independent income data, however, is that no information is available about income not coming from the firm, such as interest income. Slemrod (1994b), however, has shown that among high income taxpayers the share of income coming from wages and salaries is highly bimodal with modes at 90-100 percent from wages and at 0-10 percent from wages. Looking at these executives is obviously picking up one side of the distribution, so nonfirm income may be relatively unimportant.

I define an individual's taxable income to be the sum of salary, bonus, options exercised in the year, and LTIP payments. LTIP payouts are predominantly, but not always, cash. Sometimes they are shares and therefore not taxable in the current year which will tend to bias the results toward finding no response of this form of compensation to tax rates. A similar problem exists for bonus income in that some firms report bonuses for the current year but technically pay the bonus in an adjoining calendar year. This should not affect the estimates of long-run elasticities but will tend to bias downward any temporary responses.

I classify exercised options as taxable wage income because this is the tax treatment of the most common type of option, the Non-Qualified option (NQO). At exercise, the difference between the stock price and the option strike price is treated as wage income to the individual and deductible for the firm. The Incentive Stock option (ISO) is not treated as income until the shares are actually sold, at which time the income is treated as capital gains and not deductible for the firm. NQOs are by far the most common, however, and are, generally, the overall tax advantaged form (see Scholes and Wolfson, 1992 or the data in Conference Board, 1992) so this assumption is fairly accurate.

Nontaxed forms of compensation are reported as "other compensation." This is defined to include anything not included in the other categories; while this is not entirely nontaxable income, in practice it includes mainly perquisites, other personal benefits, and premiums paid for split-dollar life insurance policies. By law, firms are required to report perquisites that total more than \$50,000 or 10% of the executive's annual base salary. The compensation data, starting in 1992, also include the Black-Scholes value of options granted in a year which are not reported on a tax return until exercised.

To separate people into income groups, the NTR literature usually takes incomes in the year before the tax change and divided individuals by tax bracket using that income. As the results below indicate, mean reversion has increasingly made this a poor way to divide an income sample. Instead, I use a measure of permanent income to divide the groups. This permanent income is the average sum of salary, bonus, LTIP payouts, the Black-Scholes value of options by

individual, and the value of “other” compensation. This is, in some sense, the average amount of money the individual receives in a given year regardless of whether the income is reported on the tax return. This is obviously unavailable with data from tax returns because individuals do not have to report any nontaxable income they receive such as the value of options granted but not exercised or the value of non-taxed perquisites. I will use different cutoff values in different circumstances but previous work has shown that the choice of income cutoffs does not matter in these data (Goolsbee, 1997).

Section IV: Are the Rich Systematically Different?

The NTR assumption that the taxable income of different groups would grow at the same rate except for the changes in tax rate may be violated for many reasons and most of these reasons can lead to bias in the estimated tax price elasticities. The data on executives show that three differences in particular may be important.

Difference 1: Relative Incomes of the Very Rich Are Increasing Greatly Over Time

Rising wage inequality in the 1980s has been well documented in the labor literature and explained by a variety of nontax factors (see Katz and Murphy, 1992 or Levy and Murnane, 1992). Obviously if taxes are cut for the rich when the incomes at the top are growing faster than those lower down, this will create a bias in the natural experiment. All the relative increase in income of the rich will be incorrectly attributed to the tax change. This is potentially quite important for evaluating the tax cuts of the 1980s, which were largest for the highest group, and this critique of such work has been made from the beginning (see Gravelle, 1993 or Slemrod, 1996). In fact, Slemrod (1996) shows that including a simple measure of inequality removes all evidence that taxes matter for the incomes of the very rich from 1954-1985. He regresses the data of Feenberg and Poterba (1993) listing the share of total income going to the top .25 percent of the distribution on marginal rates but controls for wage inequality trends by including the log

difference of the 90th percentile and 10th percentile of male weekly wages. The results show virtually no effect of tax rates. Taxes matter only in the years around TRA86.

Some of the literature has argued that the inequality trend itself is the result of falling marginal rates (see discussions in Feenberg and Poterba, 1993, Slemrod 1994b, or Auten and Carroll 1997). The argument is summarized by Slemrod using figure 1 from Feenberg and Poterba (1993) which shows the share of wages received by the top .25 percent of the adult population. The trend begins at around the time that high marginal rates started a monotonic fall through the 1970s and 1980s, suggesting a tax based source.

Several observations, however, cast serious doubt that taxes are the source of the rise in inequality (as the footnote below indicates, the doubts remain looking at AGI rather than just wages). First, by far the largest increase in U.S. history of the net of tax share at the high end of the income distribution was the reduction in the top marginal rate from 91 percent to 70 percent following the 1964 Revenue Act. The net of tax share more than tripled--a vastly larger tax cut even relative to other income groups than anything in the 1980s. There is, however, no evidence of an increase in the wage share at the high end.⁴ Second, the income trend at the high end remains almost exactly constant from the mid-1970s to the mid-1980s despite major tax cuts in the early 1980s and no tax cuts in the late 1970s or the mid-1980s up to TRA86.⁵ Third, as the data below will demonstrate, relative incomes of the very rich have continued to rise through the 1990s despite rising marginal rates.

The trend in figure 1 is precisely the type of bias that will lead the NTR literature to overstate the elasticities using tax cuts as a natural experiment. Basically, the NTR literature attributes the entire increase in the share going to the rich in figure 1 to taxes when they should attribute only the amount going to the rich over and above the trend line. The latter is obviously much more modest (this is really just a graphical demonstration of the results in Slemrod (1996) showing that except for TRA86, taxes have little effect).

The most common response of NTR practitioners to this critique has been to tighten the natural experiment and argue that comparing the rich to the moderately rich should eliminate

most of the inequality trend bias because high-income groups are similar. The idea is that someone in the top 1% of the distribution may face a very different environment or income trend than someone at the median but probably the same environment as someone in the top 10%. Thus comparing those two with a natural experiment avoids the problem. The data on corporate executives, however, show that even this natural experiment assumption is not true and that correcting it can reduce the estimated elasticities greatly.

To estimate the true effect of tax rates, any analysis must subtract out the nontax-related trends in income for each group. Table 3 presents the average growth rate of salary and bonus from 1991-1995 for various income classes. The early 1990s is a period with no tax cuts so tax induced growth is not an issue. I take as a group comparable to the “highest” group in Feldstein (1995a), people with permanent income greater than \$177,000 in 1985 dollars and as comparable to the “high” group executives below \$177,000.⁶ There are not enough executives with incomes substantially below \$100,000 so I do not calculate a growth rate for the “medium” group.

It is very clear that the growth rates differ substantially by income class. Incomes of individuals in the highest group grew at 9.1 percent annually and in the high group at 2.4 percent. This may be seem like an extreme case but the trends are similar to those listed in Hall and Liebman (1997) for other high-income people from 1982-1994. They find that average incomes in high-income occupations (CEOs, NBA players, MLB players) rose between 7.7 and 12.5 percent per year while the income cutoff to be in the top ½ percent of AGI (which had a starting income in Feldstein’s high group) rose 3.4 percent per year. They also list the growth rate of professors and of average workers (who are around Feldstein’s medium group) of between 0.5 and 1.3 percent.

Any growth rates of these magnitudes can easily explain the majority of the estimated elasticities of taxable income in natural experiments even if restricted to the relatively rich. To see the potential importance, the bottom panels of table 1 redo the difference-in-differences calculation of Feldstein (1995a) but subtract out, using various income growth rates, the amount that income would have risen *without any tax effect whatsoever*. With growth rates of 8.9

percent, 3.4 percent, and .5 percent for the three groups as in the middle panel, the actual elasticity of taxable income is not in the 1-1.5 range originally estimated but instead falls to slightly over 1/3. Presumably these estimates would be even smaller using the data in Auten and Carroll (1997), whose sample is one year longer and whose estimated elasticities were not as large to begin with.

The relative trends actually observed in the data are even larger than the trends subtracted from the middle panel, as are trends in the income of the other rich people. Even taking the lowest growth rate for the highest group (7.7 percent annually) and the highest rates for the other groups (3.4 percent and 1.3 percent), as in the bottom panel of table 1, the elasticities are still cut by 40 to 50 percent.

These results suggest that inequality trends may account for much of the estimated income response of the wealthy to the tax cuts of the 1980s. Future work may establish that the differences in trend rates are not as large as they are among these individuals but, at the very least, the data here indicate that there is a clear difference between the very rich and other groups in this attribute and this difference will bias the work of the NTR. These same trends should work the opposite way for tax *increases* on the rich which makes OBRA 1993 a key piece of evidence for evaluating the elasticities in the NTR (for evidence on the topic see Goolsbee, 1997).

The results also suggest that NTR work should certainly allow for income trends that vary by income class. This means that they cannot evaluate tax changes with a before/after comparison of two individual years of data. With multiple years of data, regression methods could control for differing trends by income class and still identify the tax effect, even if only tax return data are available.

Difference 2: The Income of the Very Rich Is More Sensitive to Demand Conditions

The NTR literature has used, almost exclusively, tax return data to analyze income elasticities, but it is important to note how little information is provided on a tax return. The

individual's economic environment is almost totally unknown. Auten and Carroll (1997), using their nonpublic tax return samples, have done careful work using occupation, age, region and other controls available on the returns and find that these factors are important for income growth. We know painfully little, however, about how the pay of different groups responds to economic conditions and firm performance.⁷ The natural experiment approach is simply to assume that the responses are the same for the rich as for the somewhat rich and therefore conclude that the unobserved factors don't bias the results. Unfortunately, the data on executives indicate this is false.

Without even looking at stock options, which are taken up in the next section, table 4 presents a regression of the log of cash compensation (salary, bonus, and LTIP) from 1991-1995 with individual fixed effects and separate trends, but allows different coefficients by income class on the log of the market value of the individual's firm, the ratio of earnings to the book value of capital for the firm, and the growth rate of GDP. The coefficients show that the income of the rich group is about twice as responsive to real market value and to national growth and 50 percent more responsive to earnings than is the income of the lower group.

This makes theoretical sense if one believes that the leaders of firms are more responsible for performance. It is possible, of course, that the differences are greater among executives than among other high income people. It is likely, however, that the same effects are at work among other groups such as doctors, lawyers, athletes, and investment bankers--anywhere the highest paid individuals are more responsible for firm performance.

The tax bills of the last two decades have taken place at particular points in the business cycle which make this issue relevant for the NTR literature. The most famous NTR papers, Lindsey (1987) and Feldstein (1995a), look across samples (1982-84 for Lindsey and 1985-1988 for Feldstein) where the stock market increased dramatically, corporate profits rose and GDP growth increased. The regression of table 4 suggests that these conditions will tend to increase the incomes of the rich more than the moderately rich even without tax changes.

To illustrate its importance, I reexamined TRA86 but subtracted out the effect of stock prices, GDP growth, and corporate profits on pay and recalculated the elasticity. To do so, I made the simplifying assumption that all firms had increases in real market value equal to the increase in real market value of the S & P market index and increases in earnings to capital ratios equal to the increase in the profits before taxes to shareholder equity ratio for manufacturing, both of which, as well as the real GDP growth rate, came from the Economic Report of the President. Due to these economic conditions, the highest group's income should have risen by 12.2 percent and the high group by 6.7 percent without any tax changes. Taking these growth rates out and recalculating the tax response reduces the estimated elasticity by about 25 percent. These market effects are also potentially important when analyzing the tax increases of 1993 and 1994 because of the high variability of the stock market and the business cycle from 1991-1995.

Unfortunately for the NTR, truly correcting for this bias is impossible using tax return data because they lack any information about firm performance. But even with tax return data, gains could be made by including more than just two years of data and allowing national economic conditions like unemployment or GDP growth to vary by income class as a control. With the detailed tax return data, industry performance could also be used.

Difference 3: The Share of Time-Shiftable Compensation Rises Dramatically with Income

Slemrod (1992, 1994a, 1994b, 1995) has consistently remarked that we must be careful not to interpret short-run changes to the timing of transactions as long-run elasticities. NTR practitioners such as Feldstein and Feenberg (1996) have argued that natural experiments comparing the very rich to the somewhat rich probably do not suffer from this problem because there must be differences in the timing sensitivity of pay between the rich and very rich for the natural experiment to be biased. Evidence of precisely such a difference is the final important element found in the data on executive compensation.

Certain forms of income are much easier to retime than others (exercising stock options or receiving bonuses in December rather than January, for example). The data suggests that the

share of income in these shiftable forms may be much higher for the very rich. Figure 2 presents a nonparametric kernel regression of the share of total compensation coming as bonus or LTIP payout by income level.⁸ These payments are usually made in lump sums, often at the end of the year or quarter, and so retiming these payouts is a relatively easy way to shift income in the short run from one year to another. The figure clearly shows that the share of income in shiftable form rises dramatically with permanent income, more than doubling when income rises from \$100,000 to \$1,000,000. Figure 3 presents a similar kernel regression but for the share of compensation coming as stock options (valued by the Black-Scholes formula). There is, again, a dramatic rise in the share as income rises. The share almost triples when income rises from \$100,000 to \$1,000,000. Note that this is not the value of options *exercised* (which would usually appear on a tax return as wage income) but the value of options *granted* (which does not appear anywhere on a tax return). As such, these regressions show that the *potential* for short-run shifting is much greater for the very rich than for the moderately rich. Since NTR papers do not separate long and short run effects, the natural experiments will tend to confuse the two and overestimate the long-run elasticities.

The increasing use of options and the large outstanding stock of unexercised options has also created a large mean reverting component to the income of many of the very rich. Any sampling strategy, such as commonly used in the NTR, of dividing up individuals based on their income in a year before a tax change will, at least in the 1990s, tend to disproportionately choose precisely those people who are cashing in their stock options and whose income will fall in the future back to a more “normal” level.⁹ There is positive serial correlation in salary compensation for individuals in this sample but very strong negative serial correlation in exercised option income. Since exercised options are largely just included as part of wages and salaries on a tax return, it means that, again, tax return data are becoming less useful over time. At the very least it is imperative to get some measure of permanent income when breaking up groups for a natural experiment.

To see the importance of temporary shifts versus long-run changes, table 5 conducts a simple difference-in-differences analysis of these executives using short-run data from 1992-1993 and then longer-run data from 1992-1995 to see if there is a difference in responsiveness.¹⁰ The short-run analysis is basically a counterpart of the work of Feldstein and Feenberg (1996) who look at the 1992 to 1993 change and find a large elasticity although the same issues would apply in a regression framework.

In the first panel we see that high-income people had a 12.5 percent reduction in their net of tax share from 1992 to 1993 while the next group had a decrease of 7.2 percent (below \$140,000 per year there was no change but this is a small fraction of the sample). The real incomes for the panel of individuals with data in all of the years from 1991-1995 fall by 11.8 percent for executives in the highest group but rise by 6.9 percent for the lower group. This indicates an elasticity of taxable income of 3.6(!).

The second panel, however, looks at the longer-run change from 1992-1995 for the same people. Now we see that the change in the net of tax shares for the two groups (counting the change to the payroll tax in 1994) is 5.2 percent but their income growth differs by only 1.9 percent. The elasticity is slightly less than .4 and almost ten times smaller than the short-run response. Even the longer-run elasticity of .4 is biased upward by the fact that income in the base year of 1992 is artificially high due to timing shifts. In other words, incomes of the rich and the very rich responded much differently in the short run but very similarly over the longer run--totally consistent with a simple retiming story. Another fact really drives this point home. In 1991, 27 percent of the total taxable compensation to executives came as options exercised, in 1992 this rose to 42 percent and in 1993 it returned to 27 percent.

For the 1993 tax increase, at least, timing issues seem to be quite important and they may also be important for other tax changes. This problem would be easy to address in future NTR work given sufficient data. Regression based methods could just include future tax rates as well as current rates to separate out anticipatory and transitory reactions from permanent responses in the tradition of the work on permanent versus transitory tax effects in other contexts found in

Auten and Clotfelter (1982), Burman and Randolph (1994), or Randolph (1995). By doing so, however, the conclusions of the NTR might be changed substantially. In the more detailed regression based analyses found in Goolsbee (1997), including future tax terms show large short-run responses to taxation but more permanent responses as much as 20 times smaller. The true elasticity of taxable income is between 0 and .4 and means that the deadweight loss from the Clinton tax increase, for example, rather than being almost 200 percent of the revenue collected as claimed in Feldstein and Feenberg (1996), is more like 20-25 percent.

V. Conclusion

After ten years of finding large elasticities of taxable income with respect to the net of tax share by using tax changes as natural experiments, the NTR literature has had a large impact on the conventional wisdom regarding progressivity and the efficiency loss of high marginal tax rates. I argue, however, that the results from these papers are based on the faulty assumption that the very rich differ from other income groups only because they have different tax rates. Independent data on several thousand high-income executives as well as on other prominent rich people show that even the moderately rich are not a valid control group. Three facts lead to problems.

1) Secular trends in income inequality which are not related to taxation dramatically alter estimated elasticities from the 1980s. The executive data, confirming other data in the literature, show that income for individuals in the highest income class grows significantly faster than for those in lower income classes. Correcting for this trend could reduce the elasticities based on TRA86, for example, from between 1 and 1.5 to around 1/2 or 1/3.

2) The pay of the very rich is about twice as responsive to demand conditions and firm performance as is the pay of the moderately rich. There is little information available on a tax return to correct for this problem and the natural experiments in the NTR literature, particularly of the tax cuts of the 1980s, examine years where this is specifically problematic. For TRA86, accounting for these factors may reduce the estimated elasticity by a further 20-25 percent.

3) The very rich have a much greater share of income in forms whose timing can be shifted thus making their relative responsiveness in a natural experiment much higher in the short run than in the long run. Higher income classes have much larger fractions of their incomes coming as stock options, bonuses, or LTIP payouts. In the 1990s, this difference has become even more important. The higher share means that at least part of the relative response to taxation picked up in a natural experiment is likely to be temporary. Looking at the executives following the tax increase of 1993, the short-run elasticity is almost ten times larger than the long-run elasticity.

Taken together, these three facts suggest that the elasticities of the NTR literature may be significantly overstated. Future research may find more areas where the rich differ from other income classes but it is clear from these data that, as they say, the rich really are different and it is not just because they have more money. It is possible, as suggested in this paper, for the NTR to at least partially correct for these problems using regression methods on more years of data and including more controls but the results here suggest that the conclusions from such an exercise might change substantially.

For some of the problems raised, however, tax return data will never be able to provide the information necessary to estimate the tax responsiveness correctly. Perhaps the most accurate thing to conclude is that we need more data sources on the rich which include information allowing us to correct and control for the types of issues raised in this paper, data such as the information on executive compensation. Only these types of data will allow us to estimate the true behavioral response to taxation.

BIBLIOGRAPHY

- Auten, G. and R. Carroll (1995), "Behavior of the Affluent and the 1986 Tax Reform Act," in *Proceedings of the 87th Annual Conference on Taxation of the National Tax Association*, Columbus, Ohio: pp.7-12.
- Auten, Gerry and Robert Carroll (1997), "The Effect of Income Taxes on Household Behavior" Mimeo, Office of Tax Analysis, U.S. Department of the Treasury.
- Auten, G. and C. Clotfelter (1982), A Permanent Versus Transitory Tax Effects and the Realization of Capital Gains @ *Quarterly Journal of Economics*, 97(4) pp. 613-632.
- Blinder, Alan and Howard Esaki (1978), "Macroeconomic Activity and Income Distribution in the Postwar United States," *Review of Economics and Statistics*, 60: 604-609
- Blundell, Richard, Alan Duncan, and Costas Meghir (1995), "Estimating Labour Supply Responses Using Tax Reforms" Institute for Fiscal Studies, Mimeo.
- Burman, L. and W. Randolph (1994), A Measuring Permanent Responses to Capital-Gains Tax Changes in Panel Data, @ *American Economic Review*, 84(4), pp. 794-809
- Carroll, Robert (1997), "Taxes and Household Behavior: New Evidence from the 1993 Tax Act" Mimeo, Office of Tax Analysis, U.S. Department of the Treasury.
- Conference Board, *Top Executive Compensation*, various years. Conference Board, New York.
- Cutler, David and Larry Katz (1991), "Macroeconomic Performance and the Disadvantaged," *Brookings Papers on Economic Activity*, 1991:2, pp. 1-74.
- Eissa, N. (1996), A Tax Reforms and Labor Supply @ in *Tax Policy and the Economy v. 10*, J. Poterba, ed. pp. 119-151. MIT Press, Cambridge.
- Feenberg, D. (1998), "Comment" on J. Alm and S. Wallace, "Are the Rich Different?" This volume.
- Feenberg, D. and J. Poterba (1993), "Income Inequality and the Incomes of Very High Income Taxpayers," in *Tax Policy and the Economy v. 7* J. Poterba ed., MIT Press, Cambridge.

- Feldstein, M. (1995a), "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *Journal of Political Economy*, June 103(3) pp. 551-572.
- Feldstein, M. (1995b), "Tax Avoidance and the Deadweight Loss of the Income Tax," NBER Working Paper #5055.
- Feldstein, M. and D. Feenberg (1996), "The Effect of Increased Tax Rates on Taxable Income and Economic Efficiency: A Preliminary Analysis of the 1993 Tax Rate Increases" in *Tax Policy and the Economy* v. 10 J. Poterba, ed. pp. 89-117. MIT Press, Cambridge.
- Goolsbee, Austan (1997), "What Happens When You Tax the Rich? Evidence From Executive Compensation" Mimeo, University of Chicago.
- Hall, B. and J. Liebman (1997), "Are CEOs Really Paid Like Bureaucrats?" Mimeo Harvard University
- Hardle, W. (1990), *Applied Nonparametric Regression*. Cambridge University Press; Cambridge.
- Heckman, J. (1993), "What Has Been Learned About Labor Supply in the Past Twenty Years?" *American Economic Review* 83(2): 116-121.
- Heckman, J. (1996), AComment@ on Eissa in *Empirical Foundations of Household Taxation*, M. Feldstein and J. Poterba, eds. University of Chicago, Chicago, Illinois.
- Katz, L. and K. Murphy (1992), "Changes in Relative Wages, 1963-1987: Supply and Demand Factors" *Quarterly Journal of Economics* 107(1): 35-78.
- Levy, Frank and Richard Murnane (1992), "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations" *Journal of Economic Literature*, 30: 1333-1381.
- Lindsey, L. (1987), "Individual Taxpayer Response to Tax Cuts: 1982-1984, with Implications for the Revenue Maximizing Tax Rate," *Journal of Public Economics* 33: 173-206.
- Lindsey, Lawrence (1990), *The Growth Experiment*. Basic Books; New York.
- Moffitt, R. and M. Wilhelm (1998), "Labor Supply Decisions of the Affluent," this volume.

- Parcell, A. (1996), "Income Shifting in Response to Higher Tax Rates: The Effects of OBRA 93" Mimeo OTA, presented at ASSA meetings, San Francisco, 1996.
- Pencavel, John (1986), "Labor Supply of Men: A Survey" in *Handbook of Labor Economics*, Vol. 1, Orley Ashenfelter and Richard Layard, eds. Amsterdam: Elsevier.
- Randolph, W. (1995), "Dynamic Income, Progressive Taxes, and the Timing of Charitable Contributions," *Journal of Political Economy*, 103(4), pp. 709-38.
- Sammartino, Frank and David Weiner (1997), "Recent Evidence on Taxpayers' Response to the Rate Increases of the 1990's" *National Tax Journal*, 50(3).
- Scholes, M. and M. Wolfson (1992), *Taxes and Business Strategy*, Prentice Hall, New Jersey.
- Silverman, B. (1986), *Density Estimation for Statistics and Data Analysis*. London; Chapman and Hall.
- Slemrod, J. (1992), "Do Taxes Matter? Lessons From the 1980's," *American Economic Review*, 82(2), pp. 250-256
- Slemrod, J. (1994a), "A General Model of the Behavioral Response to Taxation": Mimeo, University of Michigan
- Slemrod, J. (1994b), "On the High-Income Laffer Curve" in *Tax Progressivity and Income Inequality* J. Slemrod, ed. Cambridge: Cambridge University Press.
- Slemrod, J. (1995), "Income Creation of Income Shifting? Behavioral Responses to the Tax Reform Act of 1986" *American Economic Review*, 85(2), pp. 175-180
- Slemrod, J. (1996), "High Income Families and the Tax Changes of the 1980s" in *Empirical Foundations of Household Taxation* M. Feldstein and J. Poterba, eds. University of Chicago, Chicago.
- Slemrod, J. and S. Yitzhaki (1996), "The Cost of Taxation and the Marginal Efficiency Cost of Funds" *IMF Staff Papers* 43(1).
- Toder, E. (1995), "Comments" Mimeo, NBER Conference November, 7 1995.

TABLE 1: DIFFERENCE IN DIFFERENCE ESTIMATES BASED ON TRA86

| <u>Taxpayer Group</u> (by 1985 Rate) | <u>Percent Change</u> 1985-1988 | | |
|---|------------------------------------|--|------------|
| | Net of Tax Rate | Adj. Taxable Income Minus Gross Loss | Elasticity |
| Medium (22-38 percent) | 12.2 | 6.4 | |
| High (42-45 percent) | 25.6 | 20.3 | |
| Highest (49-50 percent) | 42.2 | 44.8 | |
| | <u>Difference-in-Differences</u> | | |
| High minus Medium | 13.4 | 13.9 | 1.04 |
| Highest minus High | 16.6 | 24.5 | 1.48 |
| Highest minus Medium | 30.0 | 38.4 | 1.25 |
| <u>With Income Trends</u> | Net of Tax Rate | ATI Minus Gross Loss Minus Time Trend | Elasticity |
| Medium (0.5 percent annual) | 12.2 | 4.9 | |
| High (3.4 percent annual) | 25.6 | 9.7 | |
| Highest (8.9 percent annual) | 42.2 | 15.7 | |
| | <u>Difference-in-Differences</u> | | |
| High minus Medium | 13.4 | 4.8 | .36 |
| Highest minus High | 16.6 | 6.0 | .36 |
| Highest minus Medium | 30.0 | 10.8 | .36 |
| <u>With Income Trends</u> | Net of Tax Rate | ATI Minus Gross Loss Minus Time Trend | Elasticity |
| Medium (1.3 percent annual) | 12.2 | 2.4 | |
| High (3.4 percent annual) | 25.6 | 9.7 | |
| Highest (7.7 percent annual) | 42.2 | 24.9 | |
| | <u>Difference-in-Differences</u> | | |
| High minus Medium | 13.4 | 7.3 | .54 |
| Highest minus High | 16.6 | 15.2 | .92 |
| Highest minus Medium | 30.0 | 22.5 | .75 |

Notes: The top panel of the table comes from the calculations in Feldstein (1995a). The first three rows list the percent change in net of tax share and of adjusted taxable income minus gross loss (as calculated by Feldstein). The next three rows show a differences-in-differences calculation comparing the three groups. The next two panels repeat the same procedure but subtract out from the income growth the three year time trends for each group whose annual rate is listed in parentheses in the first column. This yields the pure effect of the tax.

TABLE 2: DESCRIPTIVE STATISTICS FOR EXECUTIVE INCOME, 1991-1995

| Type of Income | Average Value (in 1992 dollars) |
|------------------------------------|---------------------------------|
| Salary | \$320,000 |
| Bonus | \$219,000 |
| LTIP Payout | \$59,000 |
| Options Exercised | \$279,000 |
| Total Taxable Compensation: | \$877,000 |
| Median Total Taxable Compensation: | \$490,000 |
| Options Granted | \$364,000 |
| Total "Other" Compensation | \$41,000 |
| Executive Years of Data | 17,240 |
| Number of Executives | 4,231 |

Notes: Authors calculations based on data described in the text. The total number of non-missing observations differs by type of income. The number listed is based on total taxable income.

TABLE 3: GROWTH RATES OF INCOME FOR VARIOUS GROUPS

| Group/Occupation | Time Period | Average Income (Start of period) | Average Annual Growth |
|--|-------------|-------------------------------------|--------------------------|
| Cash Income of Executives with Permanent Income | | | |
| Greater Than \$177k | 1991-95 | \$514,000 | 9.1 |
| Less Than \$177k | 1991-95 | \$154,000 | 2.4 |
| CEOs | 1982-94 | \$945,110 | 7.7 |
| MLB Players | 1982-94 | \$376,300 | 9.0 |
| NBA Players | 1982-94 | \$325,600 | 12.8 |
| Top .5 percent AGI | 1982-94 | \$180,900 | 3.4 |
| Professors | 1982-94 | \$40,700 | 1.3 |
| Average Workers | 1982-94 | \$30,400 | 0.5 |

Notes: The first two rows of data are based on the author's calculations using the data described in the text. The other data comes from Hall and Liebman (1997).

TABLE 4: SENSITIVITY OF CASH INCOME TO ECONOMIC CONDITIONS

| Vars. | (1) |
|---|---|
| Ln (Market Value)*[High Income] | .275 (.008) |
| Ln (Market Value)*[Moderate Income] | .146 (.017) |
| Earnings to Capital Ratio [High Income] | .334 (.033) |
| Earnings to Capital Ratio [Moderate Income] | .217 (.090) |
| GDP Growth Rate [High Income] | 1.625 (.212) |
| GDP Growth Rate [Moderate Income] | .836 (.561) |
| Controls | Individual Fixed Effects Time Trends By Income Class |
| n | 18618 |
| R ² | .90 |

Notes: The dependent variable is the log of real cash compensation as defined in the text. Standard errors are in parentheses. The sample covers executives with four or more years of data and December fiscal years from 1991-1995.

**TABLE 5:
SHORT- VERSUS LONG -UN ELASTICITY OF TAXABLE INCOME IN THE 1990s**

| <u>Short Run</u> | | <u>Percent Change 1992-93</u> | |
|----------------------|-----------------|----------------------------------|-------------------|
| | Net of Tax Rate | Taxable Compensation | Elasticity |
| High | 7.2 | 6.9 | |
| Very High | 12.5 | -11.8 | |
| | | <u>Difference-in-Differences</u> | |
| Very High minus High | 5.2 | 18.7 | 3.6 |
| <u>Long Run</u> | | <u>Percent Change 1992-95</u> | |
| | Net of Tax Rate | Taxable Compensation | Elasticity |
| High | 11.5 | 16.5 | |
| Very High | 16.7 | 14.6 | |
| | | <u>Difference-in-Differences</u> | |
| Very High minus High | 5.2 | 1.9 | 0.4 |

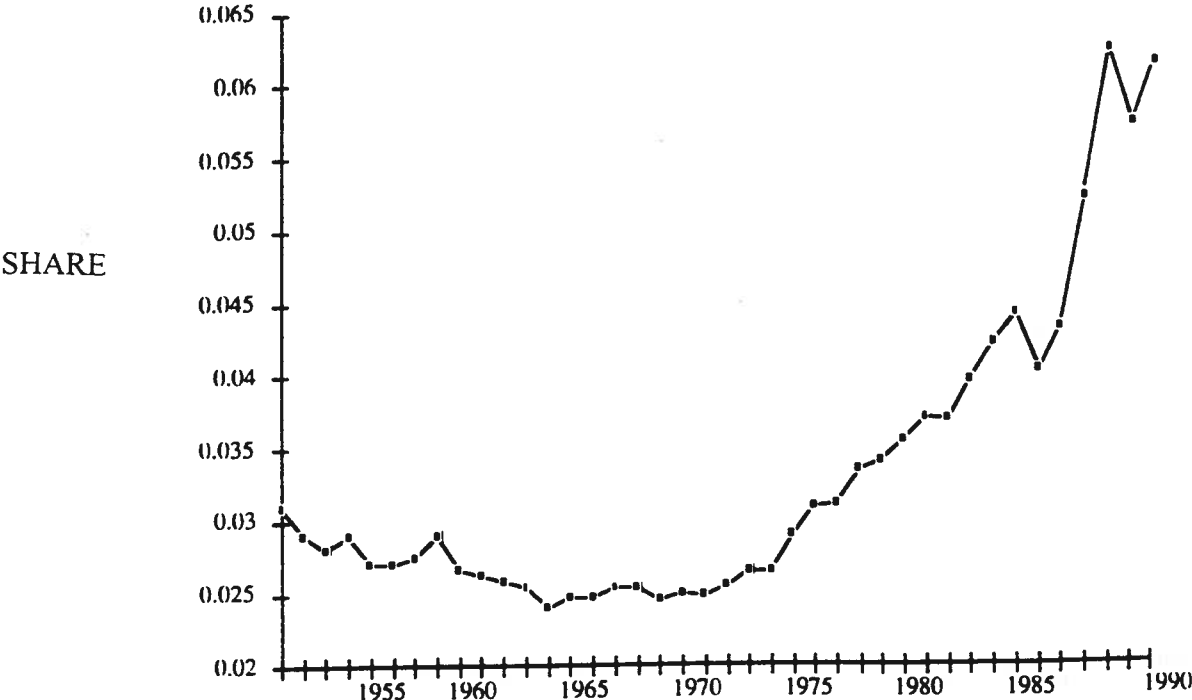
Notes: Author's calculations based on executive compensation data described in the text. The "high" income groups is defined as executives with permanent income less than \$275,000 and "very high" as executives with more than \$275,000. The sample period in the first panel is 1992-93 and in the second panel is 1992-95.

I wish to thank Steve Levitt, Maggie Newman, James Poterba, Terry Shevlin, and Joel Slemrod for helpful comments and the University of Chicago, GSB for financial support.

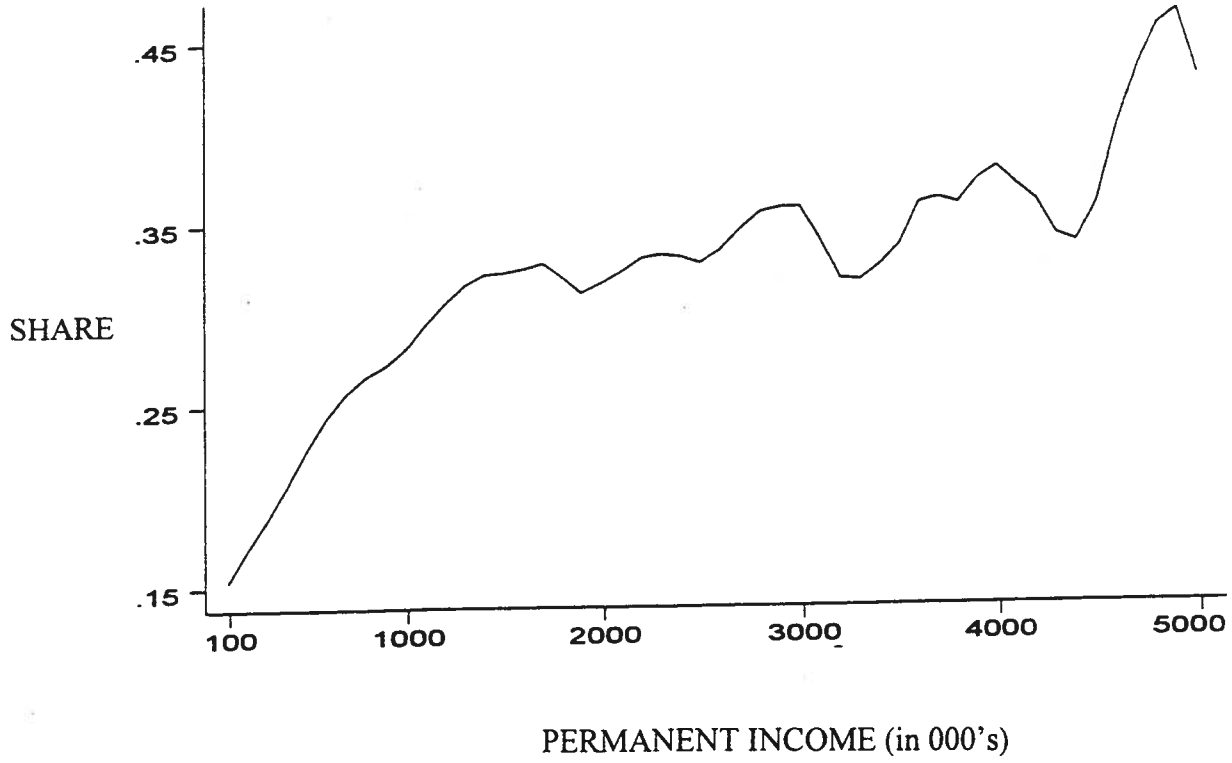
1. The only sample of comparable size is the internal (not publicly available) U.S. Treasury sample in Auten and Carroll (1995, 1997) which has about the same number of high-bracket individuals as the current sample but much less information about forms of compensation or the individual's economic environment.
2. This has always been less true for women, see Eissa (1996).
3. The debate continues with contemporaneous work by Sammartino and Weiner (1997) arguing that the long-run elasticity is small and Carroll (1997) arguing it is large. The results depend critically on how the sample of high-income people is chosen and whether the "control" groups are actually comparable to the "treatment" groups for high-income people.
4. There is evidence that incomes rose for very high income people in Lindsey (1990) and repeated in Feenberg (1998) but a simple "natural experiment" on these data comparing the relative income changes of the rich to their relative tax rates changes implies an elasticity of taxable income of only about 0.2 or less.
5. Feenberg and Poterba (1993) also look at the share of total AGI going to high income and find that the rise starts around 1981. There is, again, however, no rise in income around the massive tax cut of 1964. In addition, once they take out capital gains income, there is very little evidence of any increasing share in the 1980s until TRA86. Slemrod (1996) found that the trend removed all influence of tax rates from 1954-85 for both wages and for AGI.

6. This is the average salary in the lowest bracket of Feldstein's "highest" group but the exact cutoff was unimportant
7. There is a related macroeconomic literature on the effect of business cycle fluctuations on income distribution (see Blinder and Esaki, 1978 or Cutler and Katz, 1992 for a discussion).
8. A kernel regression is just a type of local averaging. The most important choices to be made in such methods are the weighting scheme for the averaging (the choice of kernel) and the size of the local window over which to average (bandwidth selection). The regressions here use an Epanichnekov kernel with a bandwidth chosen according to the plug-in method of Silverman (1986). Details on nonparametric regression can be found in Hardle (1990).
9. This is probably less of a problem for the work on the 1980s when stock options were less common then.
10. To be conservative, the cutoffs between groups is set at permanent income of \$275,000 per year even though the cutoff for the tax increase was \$250,000. Here, too, the cutoff point does not change the conclusions at all.

FIGURE 1: SHARE OF WAGES RECEIVED BY TOP .25% OF ADULTS
(Source: Feenberg and Poterba, 1993)



**FIGURE 2: BONUS AS A SHARE OF TOTAL COMPENSATION
BY INCOME LEVEL**



**FIGURE 3: OPTIONS GRANTED AS A SHARE OF TOTAL COMPENSATION
BY INCOME LEVEL**

