NBER WORKING PAPER SERIES

ARE HIGH INCOME INDIVIDUALS BETTER STOCK MARKET INVESTORS?

Martin Feldstein

Shlomo Yitzhaki

Working Paper No. 948

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

July 1982

NBER Working Paper #948
July 1982

Are High Income Individuals Better Stock Market Investors?

ABSTRACT

This paper presents evidence that the corporate stock owned by high income investors appreciates substantially faster than the stock owned by investors with lower incomes. Those with very high incomes enjoy the greatest success on their investments while those with incomes under $20,000 have the least success. The evidence indicates that the differences are large and that they have persisted for a long time.

Martin Feldstein
Shlomo Yitzhaki
National Bureau of
Economic Research
1050 Massachusetts Ave.
Cambridge, MA 02138
(617) 868-3905

Are High Income Individuals Better Stock Market Investors?

Martin Feldstein*

Shlomo Yitzhaki**

This paper presents evidence that the corporate stock owned by high income investors appreciates substantially faster than the stock owned by investors with lower incomes. Those with very high incomes enjoy the greatest success on their investments while those with incomes under $20,000 have the least success. The evidence indicates that the differences are large and that they have persisted for a long time. Since the present paper represents what we believe is the first such evidence of income class differences in the rate of return,[1] our results must be regarded with some caution. But if this apparent fact stands up to further scrutiny, there are important implications for a wide variety of issues including the theory of efficient markets, the rate of saving, and the distribution of wealth.

Our evidence is based on tax return data on individual shareholder transactions. Section 1 describes our data and method of calculating rates of return and presents the basic estimates. The second section then discusses several possible sources of bias. Additional results based on data for a different year are presented in the third section. A final section discusses some of the implications of our finding.

---

*Harvard University and the National Bureau of Economic Research.
**The Hebrew University of Jerusalem and the National Bureau of Economic Research.
We are grateful to the U.S. Treasury for providing the data studied in this paper, to Joel Slemrod for discussion about the statistical analysis, and to him and several other colleagues for discussions about our findings.

[1] In their 1974 study, Blume, Crockett and Friend note that, "There is no evidence that any substantive group of investors, except for exchange specialists and, to some extent, corporate insiders, have out-performed the market consistently over long periods of time." (p. 34).

1. The Data, Method and Results for 1973

The basic data are a sample of 57,676 individual Federal income tax returns that reported corporate stock capital gains in the year 1973.[1] These returns report more than 250,000 corporate stock transactions. For each transaction, we have data on the cost of the asset, the proceeds from the sale, and the dates when the stock was acquired and when it was sold. Associated with each transaction is a variety of information about the taxpayer, including his adjusted gross income, the value of his other capital gains, etc. The sampling weight for each observation makes it possible to estimate population means for any desired subgroup of taxpayers.

Our analysis focuses exclusively on the capital gains realized on corporate stock. Since the rate of gain is likely to be particularly great on stock acquired through employee stock options or by company founders, we have eliminated such transactions from our sample. Our analysis deals only with stock acquired by open market purchase. This also eliminates shares acquired by gift or bequest. The restriction to shares acquired in ordinary market transactions eliminates approximately 87,000 transactions.

We have also eliminated from the sample any transaction for which the holding period is not reported (approximately 44,000 transactions) or for which the total capital gain reported on the tax return is not equal to the sum of the individual transaction gains (approximately 35,000 transactions). These exclusions reduce the final sample to 90,022 transactions.[2]

---

[1] These data were compiled by the Internal Revenue Service for the 1973 Capital Gains Study. The tax returns represent a weighted probability sample of all individual tax returns for that year.

[2] We have repeated some parts of our analysis without imposing the last of these restrictions and found very similar results for the larger sample.

Consider now the way in which we use our data on individual transactions to calculate the average rate of return in each income class. We begin by evaluating for each taxpayer an income defined as adjusted gross income minus the capital gains on corporate stock that are used to calculate the rates of return. Taxpayers are then divided into six broad income groups: less than zero; zero to $20,000; $20,000 to $50,000; $50,000 to $100,000; $100,000 to $200,000; and greater than $200,000. The relatively small number of individuals whose net adjusted gross incomes (i.e., AGI minus the corporate stock capital gains) are negative but who have sold corporate stock are typically investors with large and complex financial transactions whose income reflects large losses in business activities, tax shelters, partnerships, etc. The average value of corporate stock sales among these individuals was more than $11,000 and therefore more than the average sales in any of the other income groups except among those with incomes over $200,000.

Within each income group, transactions are divided according to the time since the stock was purchased. For this purpose, we use 11 time intervals ranging from less than six months to more than 20 years.

For each combination of income class and holding period, we can estimate the value of stock sold per dollar of original cost. More specifically, let S_{jit} be the net sales receipt in transaction j of an individual in income group i of an asset that has been held for time period t. Similarly, let C_{jit} be the original cost of this asset and let w_{jit} be the sampling weight associated with that transaction (i.e., the inverse of the sampling probability). The total value of the stock sold after a holding period of t years by individuals in income class i is thus $S_{it} = \Sigma_j w_{jit} S_{jit}$ and the corresponding initial cost of $C_{it} = \Sigma_j w_{jit} C_{jit}$.

The value of stock sold per dollar of original cost can be used to calculate the rate of return for each combination of income class and holding period.[1] For stock sold after t years by investors in income class i, the rate of return r_{it} is defined by the equation:

$$(1) \qquad S_{it} = C_{it}\, e^{r_{it}t} \quad ,$$

or

$$(2) \qquad r_{it} = \frac{1}{t} \ln \frac{S_{it}}{C_{it}} \quad .$$

The value of t used in this equation is the weighted average of the holding periods for stocks sold within the time interval for the particular holding period.[2]

Table 1 presents the 66 estimated rates of return by holding period and income class as well as for all income classes combined. There is a clear general pattern. The two highest income groups (with incomes over $100,000)

---

[1]Note that the ratio of aggregate sales value divided by initial cost is equivalent to a weighted average of the sales to cost ratios for individual transactions, weighting by initial investment as well as by the sampling weights:

$$\frac{\Sigma_j w_j s_j}{\Sigma_j w_j c_j} = \frac{\Sigma_j w_j c_j \left(\frac{s_j}{c_j}\right)}{\Sigma_j w_j c_j} \quad .$$

[2]Note that our procedure of calculating the rate of return based on aggregate sales and aggregate costs gives an unbiased estimate of the sample rate of return for each income group and time period. The alternative procedure of calculating rates of return for each transaction and then averaging the rates of return gives an underestimate of the rate of return on the total investment made by investors in class i and time period t.

Table 1

Annual Rates of Appreciation of Corporate Stocks
by Income Class and Holding Period

Holding Period (Years)	Income Class					
	All	Zero to $19,999	$20,000 to $99,999	$100,000 to $199,000	$200,000 and over	Less than Zero
Less than 0.5	-25.6	-19.0	-29.8	-31.7	-26.0	21.0
0.5 - 1	-19.1	-6.9	-24.6	-17.0	-26.7	-2.9
1 to 2	-12.0	-3.7	-17.8	-13.4	-12.8	17.1
2 to 3	-0.1	2.8	-2.2	-1.0	0.0	11.9
3 to 4	0.8	2.2	-1.6	-0.3	2.9	18.1
4 to 5	-2.3	1.1	-7.7	-1.0	-3.8	11.9
5 to 7.5	0.9	0.9	-0.3	1.3	1.8	21.0
7.5 to 10	4.0	1.7	4.9	9.1	4.5	8.3
10 to 15	4.1	2.3	5.3	5.4	9.4	13.6
15 to 20	6.7	2.7	9.7	7.9	9.0	13.3
More than 20	5.5	5.0	5.2	6.6	7.4	13.7

All rates of appreciation refer to stock acquired in the open market and sold in 1973. Income is defined as adjusted gross income minus the capital gains on the stocks that were used in calculating the rates of appreciation.

generally have higher rates of return than the two lower income groups. Those with negative net incomes generally have the highest rates of return.

A more direct comparison is facilitated by calculating a simple average rate of return for each income class. With this restriction, equation 1 suggests the regression equation:

$$(3) \qquad \ln \frac{S_{it}}{C_{it}} = a_i + r_i t + e_{it}$$

where e_{it} is a random error. The constant term a_i permits testing whether income group i buys stocks that are "undervalued" or "overvalued" in the sense that their price adjusts up or down independent of the length of time that the asset is held. Note that the 11 values of t are not uniformly spaced but correspond to the weighted average holding periods for the 11 intervals.

Table 2 presents the estimated regression coefficients based on the rates of return for all of the holding periods.[1] The results are striking. In the first four equations, the estimated rate of return rises monotonically from 4.7 percent among investors with adjusted gross incomes between zero and $20,000 to 8.3 percent among investors with incomes over $200,000. The standard errors are quite small and the differences between the low and high values are quite significant.[2]

---

[1]The estimated rate of return coefficients are multiplied by 100 to obtain percentage rates of return.

[2]An explicit F-test rejects the hypothesis that all four groups are equal: the F-statistic is 3.11 and the corresponding critical value at the 0.05 level is 2.85.

Table 2

Estimates of Average Rate of Appreciation by Income Class

Equation	Income Groups and Weighting	Constant Term	Regression Coefficient for Income Class ($ 000)					$\bar{R}^2$	N
			0-20	20-100	100-200	200+	<0		
2.1	Income 0 - $19,999 Unweighted	-15.4 (6.0)	4.7 (0.6)					0.89	11
2.2	Income $20,000-$99,999 Unweighted	-28.9 (12.5)		7.4 (1.2)				0.82	11
2.3	Income $100,000-$200,000 Unweighted	-23.2 (5.5)			7.5 (0.5)			0.96	11
2.4	Income $200,000+ Unweighted	-25.8 (6.0)				8.3 (0.5)		0.96	11
2.5	Income less than zero Unweighted	-1.1 9.8					13.6 (1.0)	0.96	11
2.6	All Income Groups Unweighted	-19.2 (3.9)	4.9 (0.6)	6.7 (0.6)	7.2 (0.6)	7.9 (0.6)	14.9 (0.7)	0.93	55
2.7	All Positve Income Groups Unweighted	-23.3 (3.9)	5.2 (0.6)	7.0 (0.6)	7.5 (0.6)	8.1 (0.6)		0.91	44
2.8	All Positive Income Groups Unweighted	-23.7 (4.6)			7.0 (0.4)			0.87	44
2.9	All Income Groups Unweighted		3.7 (0.7)	5.5 (0.7)	6.0 (0.7)	6.7 (0.7)	13.5 (0.7)	.92	55
2.10	All Income Groups Weighted by Number of Transactions	-19.0 (2.3)	4.8 (2.0)	6.0 (0.8)	6.7 (0.7)	7.5 (0.7)	15.8 (1.8)	0.86	55
2.11	All Income Groups Weighted by Number of Transactions and Holding Period	-26.8 (3.9)	5.5 (1.0)	6.9 (0.4)	7.5 (0.4)	8.3 (0.4)	15.4 (1.0)	0.94	55

Regression coefficients are based on annual rate of appreciation presented in Table 1. All rates of appreciation thus refer to stock acquired in the open market and sold in 1973. Income is defined as adjusted gross income minus the capital gains on the stocks that were used in calculating the rates of appreciation. Standard errors are shown in parentheses.

The anomolous group with negative net income (equation 2.5) has a particularly high rate of return (13.6 percent). As we explain below, this is consistent with the combination of high economic income that appears to characterize many members of this group and their nontaxable status in 1973.

The constant terms of these individual equations are not exactly equal and indicate that the superior performance of the higher income groups is more true for longer holding periods than for shorter ones. However, equation 2.6 confirms the monotonic relation between income and the rate of return when the constant term is constrained to be the same for all groups. Equation 2.7 repeats this with the negative net income group omitted. Moreover, an explicit test of the equality of the constant terms does not reject the hypothesis that all four constant terms (for the groups with positive income) or even all five constant terms are equal.

Constraining the constant terms to be equal has relatively little effect on the estimated rates of return. To show formally that the differences among the rate of return coefficients for the four groups with positive income are statistically significant, we reestimate equation 2.7 with the rates of return constrained to be equal and compare the sums of squared residuals. The constrained estimate appears as equation 2.8. The F-statistic for testing the constraint is 5.92 while the one percent critical value with 3 and 39 degrees of freedom is 4.33. The evidence is thus very strong in indicating that there are significant differences among the rates of return.

The most direct evidence on differences in the rate of return is presented in equation 2.9 where the constant term is constrained to be zero.[1] In

---

[1]The coefficients thus correspond to separate regressions for each income class, each without a constant term.

effect, any differences among the income classes in the timing of purchases and sales is now included in the average rate of return. Now the rate of return increases monotonically from 3.7 for investors with adjusted incomes below $20,000 to 6.7 for investors with adjusted gross incomes over $200,000.

The rates of return presented in the first nine equations of Table 2 are based on regressions that give equal weight to each calculated ratio of sales value to initial cost (S_{it}/C_{it}). In principle, it would be desirable to correct for heteroskedasticity by using generalized least squares. The individual observations represent different numbers of underlying transactions; observations corresponding to a large number of transactions should lie closer to the true regression line and therefore should be weighted more heavily. Similarly, rates of return for longer holding periods provide more opportunity for annual fluctuations to be averaged out and therefore should lie closer to the true regression line. We have therefore reestimated many of the equations with a weighted generalized least squares procedure to correct for heteroskedasticity. The basic estimates are essentially unchanged, regardless of whether we weight observations according to the number of underlying transactions, the holding period, or both.

To save space, we present only two of these weighted estimates. Equation 2.10 repeats the specification of equation 2.6 with each observation weighted by the number of individual transactions. The results are very similar to the unweighted estimates and confirm the monotonic relation between the rate of return and the level of income among individuals with positive incomes. The same confirming estimates are found in equation 2.11 where each observation is weighted by the total holding period represented by the underlying transactions, i.e., the weight assigned to each transaction is the holding period for that transaction and the observations are weighted by the sum of the transaction

weights corresponding to that observation.

The higher rates of return earned by higher income groups is not an artifact of a particular holding period but can be confirmed for a wide range of holding periods. Table 3 presents some of these results. Equation 3.1 repeats the basic results previously shown as equation 2.6. Excluding transactions of less than six months simply adds 0.2 percent to the calculated rate of return in each income class and does not modify the strong pattern of monotonically increasing returns. For investments of more than five years, the relation between income and the rate of return is slightly stronger. Only when investments that have been held for 20 years or more are excluded is the strict monotonicity lost but in this case (equation 3.5) the gap between the rate of return of the lowest income group (4.1 percent for those with net incomes between zero and $20,000) and the rates of return of the higher income groups (an average of 7.9 percent) is even greater than when all holding periods are included.

Eliminating the intercept term in the regression provides estimates of constant average rates of return for each income group. These estimates are presented in equations 3.6 through 3.10, beginning with a repetition of the equation for the entire period. Eliminating short holding periods of six months or a year has no effect at all on the estimated coefficients. For investments of five years or longer (equation 3.9) and for investments of less than 20 years (equation 3.10), the difference in the rates of return between the group with incomes below $20,000 and those with higher incomes is even greater.

The only other evidence that we know on income class differences on rates of return is the very careful study by Blume et al. (1974). They also use tax return data but have additional information on the amount of dividends by company name that each individual reports. With this information, Friend et. al. calculate the rise during a twelve month period in the prices of the shares

Table 3

Estimates of Average Rates of Appreciation for Selected Holding Periods

Equation	Holding Period	Constant Term	Regression Coefficient for Income Class					$\bar{R}^2$	N
			0-20	20-100	100-200	200+	<0		
3.1	All periods	-19.2	4.9	6.7	7.2	7.9	14.9	0.93	55
		(3.9)	(.6)	(.6)	(.6)	(.6)	(.7)		
3.2	More then 6 months	-22.1	5.1	6.9	7.4	8.1	15.1	0.94	50
		(4.3)	(.6)	(.6)	(.6)	(.6)	(.7)		
3.3	More than 1 year	-23.2	5.2	7.2	7.5	8.1	15.1	0.93	45
		(5.0)	(.7)	(.7)	(.7)	(.7)	(.7)		
3.4	More than 5 years	-24.1	5.1	7.2	7.6	8.3	15.0	0.92	25
		(11.4)	(1.0)	(1.0)	(1.0)	(.9)	(1.0)		
3.5	Less than 20 years	-20.4	4.1	8.0	7.6	8.1	15.3	0.88	50
		(4.3)	(.9)	(.9)	(.9)	(.9)	(.9)		
3.6	All periods		3.7	5.5	6.0	6.7	13.5	0.92	55
			(.7)	(.7)	(.7)	(.7)	(.7)		
3.7	More than 6 months		3.7	5.5	6.0	6.0	13.5	0.93	50
			(.7)	(.7)	(.7)	(.7)	(.8)		
3.8	More than 1 year		3.7	5.5	6.0	6.7	13.5	0.93	45
			(.7)	(.8)	(.7)	(.7)	(.8)		
3.9	More than 5 years		3.8	5.9	6.3	7.0	13.5	0.90	25
			(.8)	(.8)	(.8)	(.8)	(.9)		
3.10	Less than 20 years		2.1	6.0	5.6	6.1	13.3	0.86	50
			(1.0)	(1.0)	(1.0)	(1.0)	(1.0)		

Regression coefficients are based on annual rate of appreciation presented in Table 1. Estimates are based on data for all income class and are unweighted. All rates of appreciation refer to stock acquired in the open market and sold in 1973. Income is defined as adjusted gross income minus the capital gains on the stocks that were used in calculating the rates of appreciation. Standard errors are shown in parentheses.

corresponding to these dividends. They report that in 1970 households with AGI less than $25,000 had somewhat greater returns than those with higher AGI while in 1971-72 "individuals with higher AGI averaged marginally higher returns than those with lower AGI" (p. 34).

2. Biases in the Estimated Rates of Return

The rates of return presented in section 1 are not unbiased estimates of the rates of return on investors' entire portfolio of corporate stock. This section discusses several possible sources of bias. We shall explain why we believe that the likely effect of such biases is to cause an underestimate of the observed differences in rates of return among income classes. If this is so, correcting the bias would only strengthen our findings.

The most important source of bias is our use of tax return data. This implies that we look only at the rates of return on the stocks that individuals choose to sell. To understand the likely "realization bias" in the estimated rates of return, it is important to distinguish two reasons why the return on stock that is sold may differ from the return on entire portfolios. First, in the absence of taxes, investors might tend to favor either "winners" or "losers" in their selling decisions. Optimal portfolio behavior requires, ceteris paribus, selling winners to reduce their weight in portfolios. This tendency to make realized gains exceed the overall portfolio return is reinforced by such faulty conventional wisdom as "hold a loser until you get your original investment back" and "you never lose money by taking a profit." Conversely, conventional rules like "cutting your losses quickly" and "backing a winner" would imply that realized gains will be less than gains on stocks not sold. However, none of these selling rules will bias the comparison of returns among different income groups unless the selling rules differ among these income groups.

It is particularly important therefore that our system of taxing capital gains does provide a substantial reason for shareholders to favor selling losers rather than winners and for this bias against selling stocks with large profits to be strongest in high income groups. Since the capital gains tax is due only when a gain is realized by selling the asset, the shareholders can reduce the effective tax rate by postponing the sale of stock on which there is an accrued gain. The advantage of postponement is strengthened by the fact that the unrealized gains that exist when a shareholder dies are subsequently ignored for tax purposes.[1] These tax rules provide an incentive to postpone the sale of stock with gains and, when possible, to sell stocks with losses at the same time that stocks with gains are sold. These incentives are greater the higher the individual's marginal tax rate.[2]

There is therefore a strong reason to believe that realized capital gains represent a greater understatement of the overall capital gains for high income individuals than for lower income individuals.[3] Correcting for this realization bias would therefore raise the rate of return for high income groups by more than it raised the return for lower income groups. This would only strengthen the finding reported in Tables 2 and 3.

The tax rules do not discourage the realization of capital gains by individuals with no other taxable income. Indeed, their incentive to realize

---

[1] The individuals to whom the shares are bequeathed take their market value at the time of death as their "cost" for subsequent capital gains taxation.

[2] For evidence that high tax rates do deter the realization of capital gains, see Feldstein and Yitzhaki (1978), Feldstein, Slemrod and Yitzhaki (1980), Slemrod and Feldstein (1978), and Yitzhaki (1979).

[3] High income individuals also have an incentive to avoid the capital gains tax by giving appreciated stocks to charitable organizations. Such gifts not only avoid the capital gains tax but also earn the donor a deduction for the market value of his gift. In 1977, gifts of appreciated property accounted for approximately 70 percent of total charitable deductions among taxpayers with incomes over $100,000 but less than 5 percent of gifts among taxpayers with incomes under $15,000.

capital gains is especially strong if their nontaxable status is temporary. This strong incentive to realize gains probably accounts for the very high reported rate of return in Tables 2 and 3 for individuals with negative net incomes.

The existence of a group with large capital gains but negative net income points to another source of bias in the estimated relation between income and rates of return. An individual who realizes a large capital gain may engage in tax shelter activities that reduce his income exclusive of those gains or cause that net income to be negative. This implies that some individuals who actually have high incomes will be classified in lower income groups and that this downward misclassification is most likely among individuals with large capital gains. If large capital gains are associated with higher than average rates of return, this tax shelter misclassification tends to weaken the measured association between income and the rate of return.

In addition to the realization bias and the income classification bias, there may also be a bias in the reporting of capital gains. Taxpayers may not report all of their gains honestly. If the extent of under-reporting is the same at all income levels, there will be no bias in the estimated relation between income and rates of return. It is inherently impossible to obtain information on the extent or pattern of such under-reporting. It is worth noting, however, that the probability of an Internal Revenue Service audit rises with the individual's income and that this increasing probability may encourage greater honesty. However, since higher income individuals also have more to gain by under-reporting, it is not possible to make any inference about whether there is a bias in the estimated relation.

A final problem of interpretation is that our evidence relates only to capital gains while the total return to common stock investment includes dividends as well. Because of differences in the taxation of dividends and of capital gains, lower income investors might choose stocks that offer higher dividends. Such behavior would imply that the observed differences in capital gains overstate the differences in the total return on equity. Although the extent of such offsetting differences in dividend rates cannot be determined precisely,[1] we have examined three types of evidence and found that both suggest that the bias from this source is small or non-existent. The first piece of evidence on this question is based on a special survey of income and wealth conducted by the Federal Reserve Board in 1962 and 1963.[2] The survey collected information on the amount of income received from each source and on the value of each type of wealth. The sample was heavily weighted with high income individuals. We have used these data to estimate the average dividend rate on the shares held by investors in four income classes that correspond approximately to the same real incomes as the four income classes that we have used in Tables 1 through 3.[2]

When individuals are classified by our estimate of adjusted gross income,[3] there is virtually no difference in the ratio of dividend income to the market value of stock among different groups of investors with 1962-63 incomes

---

[1]The tax return data on which our study is based gives the amount of each taxpayer's dividends but not the value of the stock on which they are earned.

[2]See Projector and Weiss (1966) for a description of this data.

[3]The Federal Reserve Board survey collected components of economic income from which we estimated adjusted gross income.

over $10,000.[1] More specifically, the dividend return was 3.6 percent for individuals with incomes between $10,000 and $50,000 and 3.8 percent for individuals in the $50,000 to $100,000 income class and in the open-ended class above $100,000. Individuals with positive adjusted gross incomes below $10,000 reported a dividend return of 2.1 percent while those with negative AGI's reported a dividend return of 0.5 percent. There is clearly no suggestion that higher income individuals accepted a lower dividend return in order to obtain greater capital gains.

When individuals are classified by an estimate of adjusted gross income exclusive of capital gains on corporate stock (i.e., by an estimate of the income measure that we used in the rate of return regression), the estimated dividend yields show a clear positive relation to income: a 2.0 percent return for those with positive net AGI below $10,000, 3.5 percent for net AGI between $10,000 and $50,000, 3.6 percent for net AGI between $50,000 and $100,000, and 4.1 percent for net AGI over $100,000. These data suggest that, if anything, the pattern of dividend yields reinforces the disparities in capital gain yields among income classes.

The second source of data on dividend yields comes from a study by Friend and deCani (1966) that used a special sample of individual tax returns

---

[1] The average dividend rate in the income class is defined as $\Sigma_j w_j d_j / \Sigma_j w_j v_j$ where w_j is the sampling weight for household j, d_j is the dividend income of household j and v_j is the value of the household's common stock at the end of the previous period. The summation is over households within a single income class. Note that this measure of the average dividend rate is equivalent to the weighted average of the individual household dividend rates, weighted by the initial value of common stock as well as by the sampling weight.

for 1960. By using the names of the companies for which dividends were reported, Friend and deCani were able to calculate the dividend return on stocks that paid at least some dividend. They reported a small negative relation between the dividend return and income, from 3.5 percent for investors with incomes under $10,000 and between $10,000 and $50,000, to 3.3 percent for investors with incomes between $50,000 and $100,000, 3.1 percent for investors with incomes between $100,000 and $200,000 and 2.1 percent for investors with 1960 incomes over $200,000. These figures suggest that part of the positive association between income and capital gains may be an offset to lower dividend yields. However, the magnitude of the relation between dividends return and income is far too small to account for all of our estimated differences in capital gains. In particular, table 2 shows a 2.8 percentage point increase in the capital gains return between the group with incomes of $100,000 to $200,000. These correspond approximately to the Friend and deCani groups with dividend returns of 3.5 percent and 3.3 percent, a yield difference of only 0.2 percent.[1]

The third set of data are the calculations of Blume et al. (1974) of dividend-value ratios using essentially the same method as Friend and de Cani (1966). They present evidence on the dividend-value ratios at different points in the income distribution for six years between 1958 and 1971. For all years, the ratio for individuals in the top one percent of income is lower than for individuals in the lowest half of the income distribution. The gap is largest in 1971 when the dividend-value ratios in the top group averaged 92 percent of the mean dividend-value ratio while the dividend-value ratio in the bottom group

---

[1]Friend and deCani also provide estimates of the ratio of realized capital gains to the total value of holding and show that this ratio increases with income. This "capital gain yield" is of course not comparable to our estimate of the annual rate of increase in the value of stock between their purchase date and their sale in 1973.

averaged 119 percent of the mean. Applying these percentages to the mean 1971 dividend-value ratio of 3.01 implies that the dividend yield was 2.76 percent in the top group and 3.68 percent in the bottom half of the income distribution. The gap of less than one percentage point is not sufficient to offset the much larger differences in rates of stock appreciation. This conclusion is also supported by the specific dividend-value ratios by income class for 1971 presented by Blume et. al., (1974): the value remained essentially constant at 0.021 or 0.022 in the four income classes beginning with $50,000, $100,000, $200,000, and $500,000. In the three income groups from $10,000 to $50,000 the ratio varied between 0.024 and 0.027. The maximum difference of 0.6 percent per year is again very small relative to the rates of price appreciation.[1]

In short, it seems safe to conclude that the biases caused by differences in realization, in income classification and in dividend yields cannot explain the estimated relation between capital gains and income. The only source of bias that cannot be eliminated is the possibility that honesty and completeness in the reporting of capital gains rises monotonically with income.

3. Additional Evidence for 1962

Although the individual records for 1973 on which our analysis is based are a unique set of publicly available microeconomic data, a similar study of capital gains was conducted by the Treasury some years ago on the basis of tax returns for 1962. The individual records used in that study are no longer available but the published report[2] can be used to analyze the relation between

---

[1]There are two problems that may bias the estimated dividend-value of these ratios: first, individuals do not report all of their dividends; second, high income individuals may own more stock that pays no dividends at all and therefore does not alter the Blume et al. calculation.

[2]Internal Revenue Service (1966). Our analysis uses Table 12, pp. 112-3.

income and the rate of return. Indeed, our interest in the subject of this paper was originally stimulated by studying the published information before the data for 1973 became available.

Our analysis of the published information for 1962 produced results very similar to those reported in section 1. When we initially analyzed the 1962 data, we regarded the results as so striking and so contrary to commonly accepted beliefs that we were reluctant to publish them until we could check them with a second set of data. Moreover, there were some problems with the 1962 data that we could avoid by using the individual records for 1973 and that we feared might have introduced biases that resulted in a spurious relation between income and the rate of return. With the 1973 data we were not only able to avoid these problems but also to assess their importance in the 1962 data by reestimating the results of section 1 with data constructed in the way we found it for 1962; this analysis showed that what we had feared were sources of potential bias actually had little effect on the estimated rates of return.

The published information for 1962 provides the value of stock sold and the original cost (or basis) of those assets classified by income class and by holding period. The income measure used to define income classes was total adjusted gross income and not the net adjusted income that we have used for 1973; AGI has the disadvantage of including the taxable portion of realized capital gains and therefore introducing a spurious positive correlation between measured income and the value of the gains. The capital gains refer to all corporate stock sold, not just to stock originally purchased on the open market; we again feared that this would cause an upward bias in the estimated rates of return for high income investors. Finally, there is no information on the distribution of holding periods within the time intervals; we therefore have

taken the midpoint of each closed interval and assumed 25 years for stock held more than 20 years. As we noted in the previous paragraph, we have used the individual records for 1973 to evaluate the effect of these problems on the results for 1973 and found that their effect was quite small.

Table 4 presents the estimated rates of return for each income class and holding period.[1] The estimated regression coefficients of Table 4 are strikingly similar to the results presented in Table 2. The rate of return rises monotonically from 3.2 percent for investors with incomes below $10,000 to 9.7 percent for investors with incomes over $100,000. The standard errors of each of the estimates is small and the differences among them are clearly significant both economically and statistically. The constant terms are all negative (presumably because 1962 share prices were unusually depressed) but again rise monotonically from -11.4 for the lowest income group to -3.2 for the highest group. This implies that the higher income groups also did better in the timing of their purchases and sales; the estimated regression coefficients therefore understate the net differences among the income classes. This is demonstrated by equation 5 in which a single constant term is estimated for all the classes; the rates of return now rise from 3.0 percent in the lowest income group to 10.0 percent in the highest group.[2]

---

[1] Note that the four income classes for 1962 are those that appear in the published report. The four income classes that we selected for 1973 were chosen to correspond roughly to these 1962 figures with an allowance for the difference between adjusted gross income and our net adjusted income. We have omitted the nontaxable returns for 1962 because it includes both the negative adjusted income individuals and individuals with low but positive income. There is no way to identify from the published data a 1962 group that corresponds to our 1973 negative adjusted income individual.

[2] A standard F test does not reject this constraint; the calculated F value is .94 while the 5 percent critical value for 3 and 72 degrees of freedom is 2.75.

Table 4

Estimates of Average Rates of Appreciation, 1962

Equation	Income Groups and Holding Period	Constant Term	Regression Coefficient for Income Class ($000) 0-10	10-50	50-100	100+	$\bar{R}^2$	N
4.1	0 < AGI < $10,000 All holding periods	-11.4 (2.8)	3.2 (0.4)				0.81	20
4.2	$10,000 < AGI < $50,000 All holding periods	-9.6 (2.8)		4.9 (.36)			0.91	20
4.3	$50,000 < AGI < $100,000 All holding periods	-7.3 (2.9)			6.3 (0.4)		0.94	20
4.4	AGI > $100,000 All holding periods	-3.2 (5.0)				9.7 (.6)	0.93	20
4.5	All income groups All holding periods	-7.9 (1.7)	3.0 (0.4)	4.9 (0.4)	6.3 (0.4)	10.0 (0.4)	0.92	80
4.6	All income groups All holding periods		2.4 (0.4)	4.4 (0.4)	6.0 (0.4)	9.5 (0.5)	0.90	80
4.7	All income groups Stock held more than 1 year		2.5 (0.5)	4.4 (0.5)	5.8 (0.5)	9.6 (0.5)	0.91	32
4.8	All income groups Stock held less than 5 years		-0.7 (1.8)	1.5 (1.8)	4.7 (1.8)	11.0 (1.8)	0.41	64
4.9	All income groups Stock held more than 1 year and less than 20 years		3.2 (0.5)	5.3 (0.5)	6.8 (0.5)	11.0 (0.5)	0.95	28
4.10	All income groups Stock held more than 6 months and less than 20 years		3.1 (0.6)	5.2 (0.6)	6.4 (0.6)	11.3 (0.6)	0.92	48

Regression coefficients for 1962 are based on data published by the Internal Revenue Service on costs and sales revenue for corporate stock sold in 1962 classified by income group and holding period. Income refers to AGI and stocks acquired in all ways are included.

Equation 6 provides the most direct way of comparing average rates of return since the constant term is eliminated. The relative effect of the constant term was of course largest for the lowest income group. In the current equation we see that the net yield really averaged only 2.4 percent for this group and then rose monotonically to 9.5 percent in the highest income group.

The remaining equations of Table 4 show that the same pattern of rates of return are observed for various holding periods. For example, eliminating stocks held less than one year (equation 7) raised the average return for all income groups but did not change the difference between the highest and lowest groups. Eliminating the open ended interval of more than 20 years (equation 8) raised the rates of return further for all groups but significantly increased the yield difference between the lowest and the highest income classes.

4. Explanations and Implications

The evidence that we have developed and presented in this paper indicates that the realized rates of return on corporate stock sold in 1962 and 1973 varied systematically among income groups. In both years, higher income groups experienced substantially greater rates of increase of share values than lower income groups. This surprising finding cannot be explained by the use of realized capital gains or by the other statistical biases in the calculated rates of return.

One possible explanation is that higher income groups enjoy greater share price increases because they take systematically greater risks.[1] Modern

---

[1] Blume et al. (1974) report that tax return filers with larger AGI tended to hold stock with greater nondiversifiable risk but do not provide specific evidence on the extent of differences among income classes. They did, however, present evidence that indicates very little portfolio diversification, suggesting that the beta coefficient might not be the relevant measure of risk bearing.

capital asset pricing theory implies that portfolios with greater nondiversifiable risk (as measured by the beta coefficient) do have higher expected rates of return in an efficient capital market. Because our data contains information only on the realized capital gains, we cannot measure the riskiness of the investors' portfolios and therefore cannot evaluate this explanation directly. It would, however, require very substantial differences in beta coefficients to explain the very large differences in observed rates of return.

Although some theories imply that the beta coefficient is the only factor that influences the yield on a portfolio, a broader view suggests other measures of risk that influence expected asset yields. As Malkiel and Cragg (1980) have recently shown, expected stock returns are also correlated with the sensitivity of the firm's earnings to the financial performance of the aggregate economy and with the degree of uncertainty or unpredictability of the firm's earning. High income investors may be willing to accept such uncertainty in order to obtain higher rates of return.

Another possible explanation is that high income individuals are more likely to have access to inside information that enables them to buy stocks before the market recognizes their value. However, such "early" buying would be reflected in a high constant term in the regressions (indicating an unusual gain soon after purchase) rather than in the high regression coefficient that indicates stocks that cumulatively grow faster year after year.

Several people have suggested that the higher average returns may reflect the fact that higher income individuals generally have larger portfolios and can therefore devote more time or resources to managing their investments. An obvious counterargument to this is that lower income individuals could in

principle achieve the same investment management resources by investing in mutual funds. The view that large individual portfolios have an advantage over both small individual portfolios and mutual funds must therefore rest on the combined effect of being big enough to justify significant portfolio management resources while, unlike mutual funds, being small enough to be able to buy and sell amounts of individual securities that are large relative to the portfolio without altering the market price of those securities.

Finally, it remains possible that our result is a statistical artifact caused by relatively greater under-reporting of gains by lower income individuals or by some source of statistical bias that we have not considered. Since there is no independent basis for either conjecture, we believe that the implications of our finding of significant differences in rates of capital gain accumulation deserve to be considered.

If the observed differences in rates of return are more than a statistical artifact but represent a permanent and stable feature of investments in corporate stock, there are likely to be important implications for the rate of saving, the distribution of wealth and the debt and equity preferences of individual investors.

The differences in rates of capital gain imply that the after-tax rate of return to the highest income group is as high or higher than the after-tax rate of return of lower income (and therefore lower tax bracket) individuals. For example, if individuals in the highest tax bracket have a 7 percent pretax rate of return (subject to a 20 percent capital gains tax) while individuals with incomes under $20,000 have a 4 percent pretax rate of return (subject to a 10 percent capital gains tax), the after tax rates of return are 5.6

percent and 3.6 percent.[1] If individuals in different income classes expect such differences in rates of return, their savings rates might well differ in the same direction. Even if savings rates are the same, the differences in rates of return would imply very substantial differences in accumulated wealth; in 20 years, $100 grows to $203 at 3.6 percent but nearly 50 percent larger ($297) at 5.6 percent. Since the differences in stock returns cannot apply to the returns earned in bank accounts or money market mutual funds, these differences imply that higher income individuals will invest proportionately more in stock.

The surprising nature of our findings and the importance of these results (if they are sustained by future research) for the process of capital accumulation, imply that more analysis of the nature and consequence of interclass differences in rates of return deserves to be done.

---

[1]The fact that capital gains are taxed only when they are realized implies that the effective tax rates are lower and the after-tax yield differences are greater.

Bibliography

Blume, M.E., J. Crockett and I. Friend (1974) "Stock Ownership in the U.S.: Characteristics and Trends" Survey of Current Business, 16-40.

Feldstein, Martin and Shlomo Yitzhaki, (1978) "The Effect of the Capital Gains Tax on the Selling and Switching of Common Stock" Journal of Public Economics, 17- 36.

Feldstein, Martin, Joel Slemrod and Shlomo Yitzhaki, (1980) "The Effects of Taxation on the Selling of Corporate Stock and the Realization of Capital Gains" The Quarterly Journal of Economics, 777-91.

Friend, I. and deCani, J. (1966) "Stock Market Experience of Different Investor Groups" Proceedings of the Business and Economic Statistics of the American Statistical Association.

Internal Revenue Service, (1966) Statistics of Income, Supplemental Report: Sales of Capital Assets Reported on Individual Tax Returns, Publication No. 458, 10-66.

Malkiel, B.J. and J. Cragg (1980) "Expectations and the Valuation of Shares" NBER Working Paper No. 471.

Projector, D. and G.S. Weiss (1966) Survey of Financial Characteristics of Consumers, Federal Reserve Technical Papers, Washington, D.C.

Slemrod, Joel and Martin Feldstein (1978) "The Lock-in Effect of the Capital Gains Tax: Some Time-Series Evidence" Tax Notes VII, 6, 134-5.

Yitzhaki, Shlomo (1979) "An Empirical Test of the Lock-in Effect of the Capital Gains Tax" The Review of Economics and Statistics, 626-29.