PET
7/27/06
9:00 a.m.

Investment Taxes and Equity Returns

Clemens Sialm*
University of Michigan and NBER

July 9, 2006

Abstract

This paper investigates whether investors are compensated for the tax burden of equity securities. Effective tax rates on equity securities vary over time due to frequent tax reforms and cross-sectionally due to persistent differences in propensities to pay dividends. The paper finds an economically and statistically significant relationship between risk-adjusted stock returns and effective personal tax rates using a new data set covering tax burdens on a cross-section of equity securities between 1927 and 2004. Consistent with tax capitalization, stocks facing higher effective tax rates tend to compensate taxable investors by generating higher before-tax returns.

JEL Classification: G12, H20, E44

Keywords: Tax Capitalization; Dividend Payments; Asset Pricing.

*I thank Gene Amromin, Alan Auerbach, Long Chen, Dhammika Dharmapala, Bob Dittmar, Ken French, Michelle Hanlon, Jim Hines, Li Jin, Marcin Kacperczyk, Gautam Kaul, Han Kim, Nellie Liang, Roni Michaely, Lubos Pastor, Monika Piazzesi, Jeff Pontiff, Jim Poterba, Kristian Rydqvist, Steve Sharpe, John Shoven, Joel Slemrod, Ivo Welch, Moto Yogo, Lu Zheng, and seminar participants at the Federal Reserve Bank of New York, the National Bureau of Economic Research, Stanford University, the Texas Finance Festival, the Western Finance Association, the University of California at Berkeley, the University of Michigan, and the Wharton Conference on Household Portfolio Choice for helpful comments and suggestions. I am grateful to Zhuo Wang for outstanding research assistance and to the Mitsui Life Center at the University of Michigan for financial support. The paper was previously circulated under the title "Tax Changes and Asset Pricing: Cross-Sectional Evidence." Address: Stephen M. Ross School of Business; University of Michigan; 701 Tappan Street; Ann Arbor, MI 48109-1234; Phone: (734) 764-3196; Fax: (734) 936-0274; E-mail: sialm@umich.edu.

1 Introduction

Asset returns should reflect the tax burden if the marginal investor is subject to taxes. In particular, assets facing higher tax burdens should offer higher risk-adjusted returns than less highly taxed assets. On the other hand, taxes should not be related to before-tax returns if the marginal investor is effectively tax-exempt, as discussed by Miller and Scholes (1978). This study tests empirically whether before-tax equity returns are related to their effective tax rates using a new data set covering personal tax burdens on a cross-section of equity securities between 1927 and 2004.

My paper demonstrates that effective tax burdens vary substantially over time and cross-sectionally. Based on the average marginal tax rates derived by Poterba (1987b), I compute effective tax rates on equity securities. The aggregate tax burden on equity securities has declined over the last couple of decades as tax reforms reduced the statutory tax rates on dividends and capital gains and as corporations replaced a significant fraction of relatively highly taxed dividends with share repurchases. In addition to the time-series variation in tax burdens, there is also a significant cross-sectional variation in tax burdens. Due to differences in the taxation of dividends and capital gains, stocks that distribute a larger fraction of their total returns as dividends tend to be taxed more heavily than stocks that distribute a smaller fraction of dividends. Dividend paying stocks faced, on average, an effective tax rate that is more than three times higher than the effective tax rate of non-dividend paying stocks.

The empirical test presented in this paper uses the time-series and the cross-sectional variation in tax burdens to determine whether the average returns of different equity securities depend on their tax burden. To obtain the effective tax burdens, I sort all publicly traded common stocks in the U.S. into portfolios according to the lagged annual dividend yield and compute effective tax rates on these portfolios. The results indicate that the average returns of these stock portfolios are positively related to their effective tax rates even after controlling for size, book-to-market, and momentum effects.

The impact of taxes on asset returns is economically and statistically significant and remains robust over several sub-periods, using various measures of the effective tax rate, and using different econometric specifications. Although the dividend yield is highly correlated with the effective tax rate, I demonstrate that the effective tax burden and not the dividend yield is the driving force behind the higher expected returns. In particular, regressing the abnormal returns on the dividend yield and on the interaction term between the dividend yield and the dividend tax rate results in a significantly positive coefficient on the interaction term and in an insignificantly negative coefficient on the dividend yield. Thus, stocks paying high dividend yields tend to have relatively high risk-adjusted returns particularly in periods when taxes are high. This result indicates that the reported effects are likely due to taxes and not due to the fact that dividend yields might proxy for additional risk or style effects not captured by the common factors of Carhart (1997).

Over the last several decades, the empirical effect of personal dividend and capital gains taxes on stock prices and stock returns has received a lot of attention in the finance, economics, and accounting literatures.[1] Research papers have investigated whether dividend yields are related to stock returns[2], whether the ex-dividend day price behavior is caused by tax effects[3], and whether individual tax reforms have an

[1]See Auerbach (2002), Poterba (2002), Allen and Michaely (2003), and Graham (2003) for literature reviews.

[2]The papers in this literature include, for example, Brennan (1970), Black and Scholes (1974), Litzenberger and Ramaswamy (1979, 1980, 1982), Blume (1980), Gordon and Bradford (1980), Miller and Scholes (1982), Auerbach and King (1982, 1983), Poterba and Summers (1984), Keim (1985), Chen, Grundy, and Stambaugh (1990), Naranjo, Nimalendran, and Ryngaert (1998), Kalay and Michaely (2000), Dhaliwal, Li, and Trezevant (2003), and Dhaliwal, Krull, Li, and Moser (2005).

[3]The papers in this literature include, for example, Elton and Gruber (1970), Kalay (1982), Eades, Hess, and Kim (1984, 1994), Barclay (1987), Michaely (1991) Bali and Hite (1998), Frank and Jagannathan (1998), Green and Rydqvist (1999), Graham, Michaely, and Roberts (2003), Elton, Gruber, and Blake (2005), and Chetty, Rosenberg, and Saez (2005).

impact on asset valuations[4].

My paper belongs to the first group of papers that analyze whether asset returns depend on dividend yields, based on the after-tax version of the CAPM by Brennan (1970). Black and Scholes (1974) test dividend tax capitalization by adding the dividend yield during the prior 12 months as an independent variable to the market model. They do not find a significant relationship between asset returns and dividend yields over the period between 1936 and 1966. On the other hand, Litzenberger and Ramaswamy (1979) find a positive and statistically significant dividend yield coefficient using a different test design focusing on months in which companies pay dividends. Kalay and Michaely (2000) point out that the effect identified by Litzenberger and Ramaswamy occurs only for short-run returns around ex-dividend dates and not for long-run returns. Naranjo, Nimalendran, and Ryngaert (1998) use an alternative measure of the dividend yield that is based on declared dividend distributions and correct for the three Fama and French (1993) factors to demonstrate convincingly that stock returns are positively related to the dividend yield during the period from 1963 to 1994. However, they do not find a statistically significant relationship between equity returns and the tax burden on dividends. Several additional studies have investigated this issue in more detail. However, the results are sensitive to how dividends are measured and whether some omitted risk factors that are correlated with dividend yields do explain the results.

Fama and French (1998) study whether the market value relative to the book value of a firm is related to dividends. They find a positive relationship between dividends and relative asset valuations and argue that the information about the profitability obscures any tax effects of financing decisions. Despite the numerous papers in this area, Graham (2003) argues that "the profession has made only modest progress

[4]The papers in this literature include, for example, Guenther (1994), Lang and Shackelford (2000), Ayers, Cloyd, and Robinson (2002), Sinai and Gyourko (2004), Amromin, Harrison, and Sharpe (2005), Auerbach and Hassett (2005), and Dai, Maydew, Shackelford, and Zhang (2006).

documenting whether investor taxes affect asset prices" (p. 1120).

In my paper, I relate the average returns on equity portfolios directly to their tax burdens and not just to their dividend yields. The tax burdens on different stock portfolios are computed using actual dividend and capital gains realizations from the U.S. Internal Revenue Service over the period between 1927-2004. By computing effective tax rates of equity securities, I take advantage of the substantial time-series variation in tax burdens of equity securities. I also use a long-term definition of the dividend yield that is based on the lagged dividend distributions and that is unaffected by potential informational biases. Furthermore, since dividend yields might proxy for risk and style factors that are not related to tax effects, I present a test that distinguishes between dividend yield and tax effects.

The cross-sectional results are consistent with the time-series results of McGrattan and Prescott (2005) and Sialm (2005). McGrattan and Prescott (2005) show that tax and regulatory changes can explain the large secular movements in corporate equity values relative to GDP between 1960 and 2001. They base their inferences on a carefully calibrated growth model. However, they do not perform an econometric analysis of the relationship between tax rates and asset valuations. Sialm (2005) finds a statistically significant negative relationship between effective tax rates and the aggregate valuation level on equity securities after controlling for several macro-economic variables. However, these two papers do not take into account cross-sectional variations in tax burdens. Adding a cross-sectional dimension to the data increases the power of the econometric tests significantly.

The paper is structured as follows: Section 2 derives the effective tax rates on various portfolios of equity securities. Section 3 reports the results of the empirical test investigating whether there is a relationship between average asset returns and effective tax rates. Section 4 shows that the tax capitalization results are robust to numerous alternative specifications. Section 5 concludes.

2 Derivation of Effective Tax Rates

This section describes the derivation of the effective tax rate of equity securities between 1927 and 2004. I compute the average tax rate faced by domestic taxable investors and investigate whether this measure of the aggregate tax burden is related to the abnormal returns of equity securities.

2.1 Definition of the Effective Tax Rate

The effective tax rate of an equity portfolio depends not only on the statutory tax rate but also on the management style of the stock portfolio. The tax burden on a stock portfolio can be reduced by holding assets with low dividend yields, by deferring the realization of capital gains, and by accelerating the realization of capital losses. The expected effective tax yield of portfolio k at time t is given by:

$$\hat{\kappa}_{k,t} = \hat{y}_{k,t}^{div}\tau_t^{div} + \hat{y}_{k,t}^{scg}\tau_t^{scg} + \hat{y}_{k,t}^{lcg}\tau_t^{lcg}. \tag{1}$$

The expected effective tax yield depends first on the marginal tax rates on dividends τ^{div} and short- and long-term capital gains τ^{scg} and τ^{lcg}, which are simply the statutory marginal tax rates for specific income brackets.[5]

Furthermore, the composition of the sources of income from equity investments has an important impact on the tax burden of a portfolio. The expected dividend yield $\hat{y}^{div}$ is defined as the sum of the expected taxable dividends divided by the initial value of the portfolio.[6] Similarly, the expected short- and long-term capital gains yields $\hat{y}^{scg}$ and $\hat{y}^{lcg}$ are defined as the anticipated proportions of the portfolio

[5]This section assumes that the statutory tax rates are known to investors at the beginning of each year. However, Section 4 shows that the empirical results are not affected qualitatively if the current tax rate is replaced by the lagged tax rate.

[6]Expected variables include a "hat" to distinguish between expected and actual variables. For example, $\hat{y}_{k,t}^{div}$ denotes the expected or anticipated dividend yield of portfolio k during year t, whereas $y_{k,t}^{div}$ denotes the actual realized dividend yield of portfolio k during year t.

values that will be realized either as short- or long-term capital gains.[7] The tax yield $\hat{\kappa}_{k,t}$ denotes the ratio between the total anticipated taxes and the current portfolio value. The effective tax rate $\hat{\tau}_{k,t}$ is the ratio between the tax yield $\hat{\kappa}_{k,t}$ and the expected return of the portfolio.

To illustrate the various definitions, I report a simple example. Suppose that the marginal tax rates on dividends and long-term capital gains are $\tau^{div} = 0.4$ and $\tau^{lcg} = 0.2$ and that the expected dividend and capital gains yields are $\hat{y}^{div} = 0.04$ and $\hat{y}^{lcg} = 0.02$. In this case, the investor expects to pay taxes on dividends equal to 1.6 percent of the initial portfolio value (0.04×0.4) and taxes on long-term capital gains equal to 0.4 percent of the initial portfolio value (0.02×0.2). Thus, the expected tax yield $\hat{\kappa}$ is 2 percent of the initial portfolio value and the expected effective tax rate $\hat{\tau}$ is 20 percent using a 10 percent expected return.

2.2 Dividend and Capital Gains Tax Rates

Marginal statutory tax rates on dividend and long-term capital gains income have fluctuated considerably, as shown in Figure 1 and Table 1. The figure shows the statutory federal marginal dividend and long-term capital gains tax rates for households in three different income brackets. Two income brackets correspond to real income levels of \$100,000 and \$250,000 expressed in 2004 consumer prices. The third bracket corresponds to the marginal tax rate for the top income bracket. Generally, dividend taxes are considerably higher and more volatile than long-term capital gains tax rates. The marginal short-term capital gains tax rates are not depicted separately

[7] For several reasons, capital gains are not taxed completely. First, the realization of capital gains can be deferred indefinitely. The deferral of the realization of capital gains is beneficial because the present value of the tax liabilities decreases if the tax payments are postponed. Second, the taxation of capital gains can be avoided completely due to the "step-up of the cost basis" at the time of death, which eliminates the taxation of all unrealized capital gains. Third, tax evasion is more prevalent for capital gains realizations than for dividends. Poterba (1987a), Auerbach, Burman, and Siegel (2000), Ivkovich, Poterba, and Weisbenner (2005), and Jin (2006) analyze the capital gains realization behavior of individual and institutional investors.

since they are very similar to the marginal dividend tax rates. The construction of the time series is explained in more detail in Appendix A.1.

To compute the average tax rates on dividends and capital gains, I follow Poterba (1987b) and construct dollar-weighted average tax rates for dividends τ^{div} and short- and long-term capital gains τ^{scg} and τ^{lcg}. The Internal Revenue Service publishes annually since 1917 the distribution of income sources of taxpayers in different income brackets. The marginal tax rate can be determined for each of these income brackets. The value-weighted mean of the marginal tax rates of investors in the different income brackets is called the "average marginal tax rate." This tax rate will represent the average tax burden of investors in various tax brackets. Prior to 1965, I hand-collected tax distribution data from different issues of the Statistics of Income of the IRS. Since 1965, the NBER publishes the average marginal tax rates on an annual basis. Additional details on the construction of the dividend and capital gains tax rates are summarized in Appendix A.2.

Figure 2 depicts the average marginal tax rates of dividend income and long-term capital gains between 1927 and 2004 based on IRS data for taxable investors. The average marginal tax rate on realized long-term capital gains is generally less than the average marginal dividend tax.[8]

2.3 Dividends and Capital Gains Realizations

The sources of investment income for equity securities varied considerably between 1927 and 2004. Dividend income was the dominant source of income for stock holders during most of the period. In the 1980s and 1990s, dividend yields decreased substantially as companies retained a larger proportion of their earnings and as they increased share repurchases.

[8] See Gordon and Bradford (1980), Auerbach and King (1982), and Fama and French (2005) for a discussion of how heterogeneous taxes are aggregated in equilibrium.

The actual dividend yield of a stock portfolio $y^{div}_{k,t}$ is defined as the ratio between the actual dividends paid during a 12-month period divided by the initial price. Dividends are defined here as taxable dividends according to the CRSP distribution codes. The detailed codes are listed in Appendix A.5:

$$y^{div}_{k,t} = \frac{\sum_{i=1}^{12} d_{k,t-i}}{p_{k,t-13}}. \tag{2}$$

The annual Statistics of Income of the Internal Revenue Service report the aggregate short- and long-term capital gains and the dividends declared by individuals. The average propensities to realize short- and long-term capital gains are assumed to equal the average propensities obtained from the IRS between 1927 and 2004.[9] Appendix A.3 explains the construction of the expected capital gains yields in more detail.

I assume that investors anticipate to realize a fixed proportion of capital gains out of the total expected returns net of expected dividend payments.[10] Thus, investors expect to realize larger capital gains for stock portfolios that are anticipated to pay smaller dividend yields. This specification results in average expected capital gains realizations that fit the average realization behavior over the whole sample period according to the IRS.

[9]Between 1927 and 2004, the average aggregate short- and long-term capital gains yield based on IRS data are -0.11 percent and 2.05 percent, respectively. Although the yearly propensities to realize capital gains fluctuate because of changes in stock market performance and because of changes in tax rates, the propensity to realize capital gains out of total capital appreciations remains remarkably stable over longer time periods.

[10]The assumption that only a fraction of the total capital gains are realized results in a lower effective tax rate on capital gains, as discussed by Bailey (1969), Feldstein, Slemrod, and Yitzhaki (1980), and Poterba (1987a). This assumption implicitly takes into account the present value of future tax liabilities. My estimation method results in a ratio between the effective accrual rate of capital gains and the statutory rate of 26.7 percent, which is very close to the 25 percent used by Poterba (1987b). However, this capital gains realization behavior is more tax-efficient than the implied valuations of capital gains taxes from Green and Hollifield (2003) and Chay, Choi, and Pontiff (2006). The robustness tests in Table 7 demonstrate that the results are not affected if I use alternative annual capital gains realizations ranging between zero and 100 percent.

2.4 Dividend Portfolios

To obtain some cross-sectional variation in the effective tax burdens, the common domestic stocks in the CRSP database are divided into portfolios according to the lagged dividend yield and the lagged market capitalization of the companies. The portfolios are formed monthly for three sorting criteria. The first criterion forms 30 portfolios according to the dividend yield and the size of the underlying stocks. All the common stocks in the CRSP database are first sorted monthly into six groups based on their lagged one-year dividend yields (one group corresponds to non-dividend paying stocks and the other five groups correspond to quintile dividend yield portfolios). Subsequently, each of the six dividend yield groups is further divided into five quintile portfolios according to the lagged market capitalization. As commonly done in the asset pricing literature, the cutoff levels for the market capitalizations of the five quintile portfolios are based only on the distribution of the market capitalization on the NYSE. The second sorting criterion forms 11 portfolios based on the lagged one-year dividend yield. One of the 11 portfolios includes non-dividend paying stocks, and the other 10 portfolios are dividend yield decile portfolios. The third criterion forms two portfolios based on whether companies paid taxable dividends in the prior year. The portfolio returns are computed using value weights within each portfolio. However, the results do not differ much if I use equal weights instead, as explained in more detail in the empirical section.

Table 2 shows the current and the future value-weighted dividend yields of the 11 dividend yield portfolios. Although dividend yields revert towards the mean, dividend payments are relatively persistent. Thus, the tax properties of the 11 dividend yield portfolios do not tend to change dramatically over short time periods.

The computation of the effective tax rate as described in equation (1) requires the anticipated dividend yield $\hat{y}^{div}$, which is not observable. The base case assumes that the dividend yield of each portfolio anticipated during the next month equals the dividend yield during the prior 12 months. This definition ensures that the dividend yield is based strictly on past data and is not affected by informational biases as described by Miller and Scholes (1982). Robustness tests summarized in Section 4 analyze the impact of using alternative anticipated dividend yields. One alternative measure defines the anticipated dividend yield as the fitted value of a partial adjustment model. Another anticipated dividend yield proposed by Naranjo, Nimalendran, and Ryngaert (1998) is based on dividend announcements. Using these alternative measures does not affect the results qualitatively.

2.5 Cross-Sectional Distribution of Effective Tax Rates

Based on these assumptions, it is possible to derive effective tax yields for different portfolios according to equation (1). Table 3 summarizes the moments of three different measures of the tax burden on equity portfolios using the three different portfolio formation criteria between 1927 and 2004. The first row reports the summary statistics of the dividend yield over the previous 12 months. The last two rows summarize the moments of the expected tax yield $\hat{\kappa}$ and the expected effective tax rate $\hat{\tau}$. The effective tax rate $\hat{\tau}$ is simply defined as the ratio between the expected tax yield $\hat{\kappa}$ and the average return on all portfolios over the whole period. The moments of the different portfolio criteria differ slightly since they give different weights to different groups of stocks. The tax yield is highly correlated with the dividend yield, indicating that most of the variation in the tax yield derives from the dividend yield.

Figure 3 summarizes the variation of expected tax rates for two value-weighted portfolios. The first portfolio includes all stocks that do not pay any taxable dividends and the second portfolio includes all other stocks. The effective tax rate of dividend paying stocks is substantially larger and more volatile than the effective tax rate of non-dividend paying stocks. The difference in tax burdens is particularly pronounced in the 1940s and 1950s and in the late 1970s. On average, dividend paying stocks face taxes that are more than three times higher than non-dividend paying stocks.

Due to data limitations it is necessary to make several simplifying assumptions about the equity holdings of investors. In particular, it is not possible to observe the identify of the investors. This could be problematic since clientele effects might induce highly taxed investors to avoid high dividend yield stocks.[11] Furthermore, corporations might also adjust their payout policies depending on their clienteles.[12] Thus, the derived effective tax rate might be a noisy measure for the tax rate of the marginal investor. However, measurement error in the effective tax rate should bias the results against finding an impact of taxes on asset returns. I show in Section 4 that the tax capitalization results are robust using numerous alternative assumptions for computing effective tax rates.

3 Taxes and Asset Returns

This section presents the main test of the capitalization of personal taxes taking advantage of both the time-series and the cross-sectional variation in tax rates.

3.1 Empirical Specification

The empirical estimation of the tax effects on equity returns is done in the base case in two stages. In a first stage, abnormal asset returns are computed based on conventional factor pricing models, such as the one-factor CAPM, the three-factor Fama

[11] For example, Allen, Bernardo, and Welch (2000) and Baker and Wurgler (2004) discuss why firms might pay dividends. Grinstein and Michaely (2005), Graham and Kumar (2006), and Dahlquist, Robertsson, and Rydqvist (2006) are recent empirical studies of dividend clientele effects for institutional and individual investors.

[12] Brav, Graham, Harvey, and Michaely (2005) document that tax considerations are secondary for financial executives when deciding on payout policies.

and French (1993) model, and the four-factor Carhart (1997) model. The empirical specification of the Carhart model is as follows:

$$
\begin{aligned}
r_{k,t} - r_{F,t} &= \alpha + \beta^{M}_{k,t}(r_{M,t} - r_{F,t}) + \beta^{SMB}_{k,t}(r_{S,t} - r_{B,t}) \\
&\quad + \beta^{HML}_{k,t}(r_{H,t} - r_{L,t}) + \beta^{UMD}_{k,t}(r_{U,t} - r_{D,t}) + \epsilon_{k,t}. \qquad (3)
\end{aligned}
$$

The return of portfolio k during time period t is denoted by $r_{k,t}$. The subscript M corresponds to the market portfolio and the subscript F to the risk-free rate. Portfolios of small and large stocks are denoted by S and B; portfolios of stocks with high and low ratios between their book values and their market values are denoted by H and L; and portfolios of stocks with relatively large and small returns during the previous year are denoted by U and D. The Carhart model nests the CAPM model (which includes only the market factor) and the Fama-French model (which includes the size and the book-to-market factors in addition to the market factor).[13]

The factor loadings β denote the sensitivities of the returns of a portfolio to the various factors. The factor loadings are estimated during a rolling window using data over the previous 60 months. This rolling factor regression reduces the length of the sample of abnormal returns by five years.

To determine the abnormal return $\tilde{\alpha}_{k,t}$ at time t of portfolio k, I subtract the expected portfolio return based on the previously estimated factor loadings from the expected portfolio return in excess of the risk-free rate:

$$
\begin{aligned}
\tilde{\alpha}_{k,t} &= r_{k,t} - r_{F,t} - \beta^{M}_{k,t-1}(r_{M,t} - r_{F,t}) - \beta^{SMB}_{k,t-1}(r_{S,t} - r_{B,t}) \\
&\quad - \beta^{HML}_{k,t-1}(r_{H,t} - r_{L,t}) - \beta^{UMD}_{k,t-1}(r_{U,t} - r_{D,t}). \qquad (4)
\end{aligned}
$$

[13] The market, size, book-to-market, momentum factors and the risk-free rate are obtained from Ken French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html).

In a second stage, the abnormal return is regressed on the tax yield $\hat{\kappa}$:

$$\tilde{\alpha}_{k,t} = \gamma + \delta \hat{\kappa}_{k,t} + \epsilon_{k,t}. \tag{5}$$

The coefficient δ should be positive if investors are compensated for the personal taxes by obtaining higher before-tax returns for assets facing higher tax burdens, particularly in periods where taxes are relatively high. A coefficient of one implies that the abnormal return increases exactly by the amount of the tax. However, the coefficient can differ from one because the marginal investor might differ from the average investor used to compute the effective tax yield and because of general equilibrium effects.

The two-stage estimation method ensures that the dividend yield coefficient does not capture risk effects that are included in the Carhart factors. However, the two-stage methodology might bias against finding an impact of taxes on asset returns, because some factors in the first-stage regression might proxy for the tax burden of the different portfolios. For example, the market or the size factors might be related to the dividend yield since small and high-beta stocks are less likely to pay dividends. As a robustness test, I also report in Table 8 a one-stage regression that estimates the tax yield coefficient δ simultaneously with the factor loadings.

3.2 Dividend Portfolios

Figure 4 depicts the rolling coefficient estimates of the factor loadings based on the Carhart model for the returns of value-weighted portfolios of dividend paying and non-dividend paying stocks between 1932 and 2004. Since the dividend portfolio accounts for a large fraction of the total market capitalization during most of the sample period, it is not surprising that the market beta is very close to one and that the other factors do not differ much from zero. However, there is an interesting variation in the estimated factor loadings of the non-dividend portfolio. The results indicate that non-dividend paying stocks tend to have a higher exposure to the aggregate market and they tend to be smaller stocks. Furthermore, non-dividend paying stocks tend to be value stocks before 1960 and they tend to be growth stocks after 1980. Non-dividend paying stocks in the early part often were distressed companies with relatively low market values, while non-dividend paying stocks in the latter part were often young companies with favorable growth prospects.[14] Due to this significant variation in the factor loadings, it is crucial to estimate time-varying factor loadings.

Table 4 summarizes the raw and the abnormal returns of the portfolios formed according to the lagged dividend yield. The table lists the averages of the time-series of these monthly excess and abnormal returns. The first column reports the raw returns between 1927-2004, and the other columns report the abnormal returns between 1932-2004. Table 4 demonstrates that stocks paying high dividend yields tend to have significantly higher average abnormal returns than stocks paying no or low dividend yields. For example, stocks in the highest dividend decile outperform non-dividend paying stocks by 38 basis points per month after adjusting for the four-factor Carhart model.[15]

The return differences are slightly less pronounced if the portfolios are equally weighted instead. The Carhart abnormal return difference between the top dividend decile and non-dividend paying stocks amounts to 28 basis points per month, but remains statistically significant at a one percent confidence level. These results are

[14]Fink, Fink, Grullon, and Weston (2005) show that firms started to issue public equity earlier in their life cycle since the early 1960s. This might explain the change in the propensity to pay dividends of newly listed firms.

[15]The fact that asset returns of high-dividend stocks tend to be relatively high seems at first glance to contradict the catering theory of dividends of Baker and Wurgler (2004), which argues that managers pay dividends when investors put a stock price premium on payers. However, the evidence of Baker and Wurgler (2004) is primarily based on the aggregate time-series variation of corporate payout decisions. Furthermore, they focus on dividend initiations and omissions instead of the total dividend payments. Thus, the results in this paper are not directly comparable with their paper.

consistent with the hypothesis that investors require higher expected returns for high-dividend stocks because of their higher tax burden.

3.3 Average Abnormal Returns and Tax Yields

Figure 5 depicts the relationship between average annualized abnormal returns and the average annualized tax yield for the 30 dividend/size portfolios over the whole sample period. For each of the 30 portfolios, I compute the average excess return over the market. In addition, I compute the abnormal returns for the one-, three-, and four-factor models using rolling regressions as summarized in equation (4). The figures show a positive relationship between average tax yields and average equity returns regardless of the risk-adjustment method. This result shows that there is a robust relationship between tax yields and risk-adjusted asset returns even after aggregating all observations over time and ignoring the time-series variation in tax burdens. The relationship is weaker using excess market returns. This occurs primarily because of the very high excess return for the portfolio that includes non-dividend paying stocks that are in the smallest market capitalization quintile. These stocks have the lowest effective tax rates and the highest average excess returns. The high abnormal performance of these stocks is reduced after adjusting the returns for the Fama-French factors.

3.4 Base Case Tax Capitalization Regression

The following results take full advantage of the time-series variation in effective tax rates and regress the excess and abnormal monthly returns of each portfolio on the corresponding tax yields. Table 5 summarizes the regression estimates for equation (5) for the three different portfolio formation criteria. Each column reports the regression coefficients using different dependent variables: The first column reports the results using the return of the portfolio in excess of the market return. The last three columns report the results using the abnormal returns from the CAPM, the Fama-French, and the Carhart factor models based on equation (4). The panel data set exhibits significant cross-sectional correlation. As suggested by Petersen (2005), I use clustered standard errors to adjust for the cross-sectional correlation.

Panel A of Table 5 reports the estimation results based on the 30 dividend/size portfolios. The estimates in Panel A with the excess (abnormal) returns are based on 28,080 (26,280) monthly portfolio observations between 1927-2004 (1932-2004). The tax yield coefficient δ in Panel A is significantly different from zero regardless of the factor model used to adjust the returns. The coefficient estimates become more statistically significant after adjusting for the Fama-French and the Carhart common factors. It is not surprising that the R-squares of the regressions are relatively small, since taxes are not the major determinant of asset returns at a monthly frequency.

Panels B and C summarize the tax yield coefficients based on 11 or two portfolios formed according to the dividend yield. The coefficients on the tax yield variables δ are all significantly positive, except for the two portfolio classification using returns that are not adjusted for risk. By using only two portfolios, the cross-sectional distribution in tax burdens is reduced dramatically and non-dividend paying stocks (which account for less than 10 percent of the market capitalization over the whole sample period) are given substantial weight. All tax yield coefficients using abnormal returns are significantly different from zero at a five percent confidence level.[16]

All the reported results are based on value-weighted portfolios. In unreported results, I show that the results are qualitatively similar if portfolios are formed using equal-weighted portfolios. For example, the coefficient on the effective equally-

[16] Adjusting the returns by introducing the liquidity factor of Pastor and Stambaugh (2003) in addition to the four Carhart factors does not affect the qualitative results of the paper for the period between 1966 and 2004, when the liquidity factor is available. The liquidity factor is obtained from WRDS (http://wrds.wharton.upenn.edu/).

weighted tax yield δ equals 1.22 with a standard error of 0.24 using the four-factor alpha and the 30 dividend/size portfolios. The results are also robust to excluding the smallest companies: The tax yield coefficient is positive and statistically significant using the Carhart abnormal returns for five sub-samples sorted according to size quintiles. Thus, the results remain robust even if the smallest 80 percent of stocks are excluded. Furthermore, the results are also unaffected if the companies that did not pay dividends in the prior year are excluded.

4 Robustness Tests

This section investigates the robustness of the results using alternative tax measures or alternative estimation methodologies.

4.1 Subperiod Evidence

Table 6 reports the tax capitalization coefficients δ from equation (5) for six different subperiods using the 30 dividend/size portfolios. The majority of the coefficient estimates are significantly positive. Using Carhart-adjusted returns, all coefficient estimates are significantly positive, except the coefficient for the period between 1990 and 2004. The tax yield coefficient measures the impact of a fixed change in the tax yield. Whereas the tax yield coefficient is relatively stable over the whole sample period, the standard deviation of the tax yield has decreased gradually over time. For example, the cross-sectional standard deviation in the monthly tax yield ranges between 0.18 percent in 1943 and 0.02 percent in 2004. Thus, the total impact of taxes on asset returns has decreased dramatically over time.

The coefficients tend to be relatively large in the 1960s and the 1980s, time periods where the aggregate tax burden on equity securities decreased substantially. An unexpected reduction in the dividend tax rate results in larger returns of stocks paying high dividends compared to stocks paying low dividends generating relatively high tax yield coefficients.

4.2 Different Tax Measures

To construct the effective tax rate, it is necessary to make some simplifying assumptions. This section shows that the results are robust to alternative definitions of the effective tax rate. Table 7 lists the tax capitalization coefficient δ for alternative measures of the tax burden on equity securities. The first row (Base Case) simply repeats the results from Table 5 for comparison.

4.2.1 Anticipated Dividend Yield

In the base case, the anticipated dividend yield is set equal to the lagged actual dividend yield. This assumption might bias the results due to the mean reversion of the dividend yields as shown in Table 2. To avoid any biases in the anticipated dividend yields, I estimate a partial adjustment model, where the actual dividend yield at time t of the stocks which are included in portfolio k at time $t-1$ is regressed on the lagged dividend yield of the same portfolio and on the lagged dividend yield of the market portfolio.

$$y^{div}_{k(t-1),t} = \alpha_k + \beta_k y^{div}_{k(t-1),t-1} + \gamma_k y^{div}_{M,t-1} + \epsilon_{k,t}. \tag{6}$$

This partial adjustment model allows for persistence in the dividend yield and for a reversion of the dividend yield toward the aggregate market yield. Since the composition of each portfolio changes slightly every year because of changes in lagged dividend yields and market capitalizations, it is important to follow the same set of underlying stocks over time. Thus, equation (6) relates the dividend yield of stocks included in portfolio k at time $t-1$ with the dividend yield of the same portfolio of stocks $k(t-1)$ at time t. Because of auto-correlated error terms and because of a lagged dependent variable, the coefficients of this linear model and the first-order auto-correlation are estimated using maximum likelihood. This partial adjustment model is estimated for each portfolio separately to allow the adjustment coefficients to differ depending on the dividend yield and the size of the stocks included in the portfolio. Furthermore, the estimation uses data at an annual frequency to avoid overlapping observations. I use the fitted values from this partial adjustment model to obtain an estimate of the anticipated dividend yield of portfolio k during the next year $\hat{y}^{div}_{k,t}$. Row (2) of Table 7 shows that the results are not substantially different using the fitted dividend yield based on the partial adjustment model.

Naranjo, Nimalendran, and Ryngaert (1998) propose an alternative method to compute the dividend yield, where the dividend yield is computed as four times the last declared quarterly dividend before the end of month $t-1$ divided by the price at the end of month $t-1$. Furthermore, they exclude stocks with special dividends or other distributions in the prior 12 months and they also exclude stocks with nontaxable quarterly dividends in the prior 24 to 13 months. Row (3) shows that a positive and statistically significant relationship results also using their measure of the dividend yield.[17]

4.2.2 Different Tax Brackets

The base case assumes that the marginal investor faces a tax rate on dividends and capital gains equal to the tax rate of the average investor. Rows (4) to (6) use instead three different federal statutory tax brackets on dividend income and short- and long-term capital gains to compute the tax yield. The tax yield coefficients under these three alternative tax yields are all significantly positive. Whereas the

[17]The relatively high tax yield coefficients in this specification can be explained by the impact of short term reversal since the this month's return is regressed on the expected dividend yield (which equals the expected dividend payment divided by the price at the end of the previous month).

tax capitalization coefficients are larger than one for the $100,000 tax bracket, they are significantly smaller than one for the top tax bracket. This result is consistent with the marginal investor having an intermediate tax bracket. The fact that the tax capitalization coefficient in the base case tends to be larger than one suggests that the marginal investor is in a higher tax bracket than the average investor.

4.2.3 Different Capital Gains Realizations

The base case assumes that investors anticipate to realize a fixed proportion of their capital gains every year and to defer the remaining gains. Sophisticated investors might avoid a significant portion of capital gains taxes by deferring the realization of capital gains and by accelerating the realization of capital losses. On the other hand, Green and Hollifield (2003) and Chay, Choi, and Pontiff (2006) derive less efficient capital gains realization behavior than the behavior implied by the historical IRS realizations employed in the base-case specification. The seventh and eighth rows investigate whether different assumptions on the capital gains realization behavior affect the results. The seventh row assumes that investors completely avoid realizing any capital gains and the eighth row assumes that investors expect to realize all capital gains annually at the long-term capital gains tax rate.[18] The coefficient estimates are only marginally different from the base case, indicating that the results are driven primarily by dividend taxes and not by capital gains taxes.

4.2.4 Tax-Sheltered Environments

During the last several decades there has been a significant increase in equities held in tax-sheltered accounts. Following Chaplinsky and Seyhun (1990), I compute average marginal tax rates on dividends and capital gains taking into account equity securities

[18] Thus, the tax yield in row (7) is $\hat{\kappa}_{k,t} = \hat{y}^{div}_{k,t}\tau^{div}_t$ and the tax yield in row (8) is $\hat{\kappa}_{k,t} = \hat{y}^{div}_{k,t}\tau^{div}_t + (\hat{r}_{k,t} - \hat{y}^{div}_{k,t})\tau^{lcg}_{k,t} = \hat{r}_{k,t}\tau^{lcg}_{k,t} + \hat{y}^{div}_{k,t}(\tau^{div}_t - \tau^{lcg}_{k,t})$.

held in tax-sheltered environments. According to the Flow of Funds of the Federal Reserve, the proportion of corporate equity held by tax-exempt investors, such as non-profit organizations, pension funds, and tax-qualified retirement accounts, increased from less than 10 percent in the 1950s to 45 percent in 2004. The vast majority of other equities are held by taxable investors either directly or indirectly through mutual funds and other intermediaries. Income on stocks held in tax-sheltered accounts generally faces zero dividend and capital gains taxes. The substantial increase in tax-exempt environments results in a significant decrease in the aggregate tax rate on equity securities.[19] Row (9) includes assets held in tax-sheltered accounts and computes the tax yield coefficient for all investors, where stocks in retirement accounts are assumed to face zero taxes. This change in the tax yield has only a minor impact on the results.

4.2.5 Additional Measures of Tax Burdens

The base case assumes that the average returns of the different portfolios are identical, as explained in Appendix A.3. However, stocks with higher mean returns tend to have higher capital gains realizations and higher effective tax yields. In row (10), I use the actual average return for each portfolio to compute the capital gains yield according to equation (1). As expected, the coefficients increase slightly and become more statistically significant.

Investors might not have access to all the available information on current tax rates and income distributions at the beginning of the year. Furthermore, tax rates are endogenous and might depend on the stock market performance. Row (11) uses the lagged tax yield during the previous 12 months as the explanatory variable. The positive relationship between tax yields and risk-adjusted returns remains intact.

[19]The proportion of equity held in taxable accounts is estimated using the Flow of Funds published by the Board of Governors of the Federal Reserve Bank, as explained in more detail in Appendix A.4.

The last row regresses the excess and abnormal returns on the anticipated dividend yield and shows a positive relationship between dividend yields and abnormal returns. Companies paying high dividend yields tend to have higher returns. The results indicate that the tax effect remains robust even if the time-series variation in tax rates is ignored.

4.3 Different Empirical Specifications

Table 8 reports the tax yield coefficients using alternative test methodologies. The results in the first row repeat the coefficient estimates δ in the base case using clustered standard errors.[20]

The regression results in the second row include an indicator variable for each of the 936 months and cluster by the dividend/size portfolios to obtain the standard errors. This time-fixed effect controls for macroeconomic variables that affect all asset portfolios symmetrically and vary over time. The tax yield coefficient increases in three specifications and decreases in one specification relative to the base case.

The Prais and Winsten (1954) regressions summarized in the third and fourth rows estimate a linear regression that is corrected for first-order serially correlated residuals. The estimated auto-correlation is relatively small and none of the coefficient estimates are affected significantly by adjusting for auto-correlation. The third row adjusts for cross-sectional correlation using clustered standard errors by time, whereas the fourth row uses panel corrected standard errors following Beck and Katz (1995). The different methods to adjust for cross-sectional correlation have almost identical standard errors.

[20] Clustering for cross-sectional correlation is important. For example, the standard error of the tax yield coefficient for excess returns in the base case is 0.51 in the specification with clustered standard errors by time. The corresponding standard error would be just 0.24 using regular standard errors without clustering. On the other hand, clustering by portfolio would result in a slightly lower standard error of 0.47. See Petersen (2005) for a comparison of different methods to compute standard errors in panel data.

Similar results occur if I estimate equation (5) using the Fama and MacBeth (1973) approach. The standard errors for the Fama-MacBeth regressions follow the Newey and West (1987) adjustment using a lag length of 60 months. The lag length corresponds to the estimation window for the factor loadings.

The sixth row investigates whether the results are subject to seasonality, which might occur because the factor portfolios are only adjusted once annually or because investors update their portfolios only infrequently, as discussed by Jagannathan and Wang (2005). To address this issue, I run panel regressions separately for each of the 12 months. The table reports the means and the standard errors of the coefficient estimates for these 12 regressions. The mean coefficient estimates and the standard errors are almost identical to the base case. Furthermore, the vast majority of individual coefficients on the monthly data are positive and no seasonal patterns are discernible. For example, all monthly coefficients using both the Fama-French and the Carhart abnormal returns are positive.

The last three specifications in Table 8 report the coefficient estimates using a one-stage method, where the factor loadings are estimated simultaneously with the tax-yield coefficient δ:

$$\begin{aligned} r_{k,t} - r_{F,t} &= \alpha + \beta_{k,t}^{M}(r_{M,t} - r_{F,t}) + \beta_{k,t}^{SMB}(r_{S,t} - r_{B,t}) \\ &\quad + \beta_{k,t}^{HML}(r_{H,t} - r_{L,t}) + \beta_{k,t}^{UMD}(r_{U,t} - r_{D,t}) + \delta \kappa_{k,t} + \epsilon_{k,t}. \end{aligned} \tag{7}$$

This estimation method does not require prior returns to compute factor loadings since they are estimated simultaneously with the tax effect. The seventh and eighth rows use a pooled panel regression with clustered standard errors and the ninth row uses a Seemingly Unrelated Regression (SUR). The SUR methodology estimates the factor loading coefficients for each of the 30 portfolios separately. To allow for time-varying risks, I estimate separate factor loadings for each of the 30 portfolios during each 60-month period in specifications (7) and (9). On the other hand, specification (8) assumes that the factor loadings stay constant over the whole sample period. The estimates for the abnormal returns using this one-stage method do not differ significantly from the base case specification.

4.4 Dividend Yields Versus Effective Taxes

The variable that measures the expected tax yield depends primarily on the interaction effect between the dividend tax rate and the dividend yield, as shown in equation (1). The following specification tests whether this interaction effect remains important after introducing the impact of the two components separately. In particular, I estimate the following regression:

$$\tilde{\alpha}_{k,t} = \beta_0 + \beta_1 \tau_t^{div} \times y_{k,t}^{div} + \beta_2 y_{k,t}^{div} + \beta_3 \tau_t^{div} + \epsilon_{k,t}. \tag{8}$$

A tax capitalization effect should generate a positive coefficient on the interaction term between the dividend tax rate and the dividend yield. On the other hand, the coefficient on the lagged dividend yield should be significant if the dividend yield proxies for additional risk factors or for behavioral biases. This specification also allows the performance of a "horse race" between the dividend yield and the interaction effect. If the results in the previous section are driven by the dividend yield and not by tax effects, then the coefficient on the dividend yield should be significant and the coefficient on the interaction effect should be insignificant.

Table 9 summarizes the results of this specification. The coefficients on the interaction term are always positive, indicating that the dividend yield effect is particularly pronounced in periods where taxes are relatively high. The coefficient on the interaction term is statistically significant using three- or four-factor adjusted returns. The insignificant results for the excess and the CAPM-adjusted returns might be due to multicollinearity, since the correlation between the dividend yield and the interaction term between the dividend yield and the dividend tax rate is 0.89. On the other hand, the coefficient on the dividend yield is negative and not statistically significant under any of the four specifications, indicating that the effect described in the previous section is likely a tax effect and not just a dividend yield effect.[21]

Naranjo, Nimalendran, and Ryngaert (1998) also find a positive relationship between the abnormal returns and the dividend yield. However, they do not find a significant tax effect if they include both the dividend yield and the interaction effect between the dividend yield and a proxy of the tax rate. Their results differ for two main reasons: First, they use only data between July 1963 and December 1994 to estimate the tax effects. As Naranjo, Nimalendran, and Ryngaert (1998) discuss, it is difficult to obtain significant results due to the high correlation between the dividend yield and the interaction effect between the tax rate and the dividend yield. I also cannot find a statistically significant coefficient on the interaction term if I restrict the sample to the period between July 1963 and December 1994 (the coefficient β_1 decreases from 2.67** (1.21) to 1.18 (1.95) using Carhart returns). The longer time period has significantly higher variation in tax rates, which improves the power of the tests.

Second, Naranjo, Nimalendran, and Ryngaert (1998) use the yield difference between Treasury and municipal securities (i.e., the implied tax rate on municipal bonds) as a proxy of the effective tax rate of dividends. The implied tax rate on municipal bonds is a noisy measure of the tax rate on equity securities because investor clienteles differ between municipal bonds and common stocks and because the yield difference also captures default and liquidity differences, which are very significant determinants of municipal bond yields according to Wang, Wu, and Zhang (2005). In unreported

[21] The results are not affected if the interaction effect is replaced by the effective tax yield κ (which is highly correlated with the interaction effect) or after including time-fixed effects.

results, I find that using the implied tax rate on municipal bonds as the proxy for dividend taxes between 1963 and 1994 results in an insignificantly negative coefficient on the interaction term and an insignificantly positive coefficient on the dividend yield.[22]

The coefficient on the dividend tax rate β_3 is significantly negative using the abnormal returns. Since the dependent variable $\tilde{\alpha}_{k,t}$ is an abnormal return, the coefficient β_3 captures the tax effect for a portfolio with a dividend yield of zero *relative* to a benchmark portfolio. A negative coefficient β_3 occurs whenever the dividend yield of the benchmark portfolio is higher than the dividend yield of the zero dividend yield portfolio, because the zero yield portfolio requires a smaller compensation for investment taxes than a benchmark portfolio that has a positive dividend yield.[23]

4.5 Dividends Versus Share Repurchases

Companies can distribute cash to shareholders by either paying dividends or by repurchasing stocks. Boudoukh, Michaely, Richardson, and Roberts (2006) and Lei (2005) show that adding share repurchases to dividends increases the power of predictive regressions. One major difference between the two ways to distribute cash is that dividends tend to be taxed more heavily than the resulting capital gains due to share repurchases. Thus, it should be expected that the tax effect of share repurchases is less pronounced than the tax effect of dividend payments. In unreported results, I form 30 portfolios based on the repurchase yield and the market capitalization of the underlying stocks. Consistent with the tax capitalization hypothesis, the tax yield effect is less pronounced for share repurchases than for taxable dividend payments.

[22]The implied tax rate on municipal bonds is computed as $\tau_t^{muni} = 1 - r_t^{muni}/r_t^{treas}$, where r^{muni} and r^{treas} are the one-year prime municipal bond yield and the one-year Treasury bill yield from Salomon Brother's *Analytical Record of Yields and Yield Spreads*.

[23]The panel regressions summarized above do not allow to determine whether raw returns are positively related with tax rates, since the dependent variable is the difference in the returns of two portfolios. Consistent with tax capitalization, I obtain in unreported results a significantly positive relationship between raw returns and the tax yield.

4.6 Tax Interaction Effects

To investigate whether the results are robust to size, book-to-market, and momentum effects, I include additional characteristics of the portfolios in the second stage of my main specification. I compute in each month for each of the 30 portfolios the value-weighted market capitalization, the value-weighted book-to-market ratio, and the value-weighted prior-year return.[24] I take the logarithm of the average market capitalization because this variable is highly skewed to the right. To facilitate the interpretation of the coefficient estimates and to eliminate time effects, I standardize the variables for each time period by dividing the demeaned portfolio characteristics by the standard deviations across the 30 portfolios.

Table 10 summarizes the results of the second stage regression using the abnormal Carhart return as the dependent variable. The first column corresponds to the fourth column in Panel A of Table 5 in the paper. The second column includes the standardized measures of the size, the book-to-market, and the momentum characteristics of the portfolio holdings. Whereas the size and the momentum coefficients are insignificant, the coefficient on the book-to-market characteristic is positive and significant. However, the coefficient on the tax yield variable only increases marginally from 1.37 to 1.40 and remains highly statistically significant. Thus, the results stand up to including size and book-to-market effects in the second stage.

To investigate whether the tax effects are affected by the characteristics of the stocks, I also include interaction effects between the tax yield and the standardized stock characteristics. As previously shown by Naranjo, Nimalendran, and Ryngaert (1998), the tax yield effect decreases with the size of the companies. On the other hand, the tax effect appears to be stronger for growth and momentum stocks.

[24] The market capitalizations and the one-year returns are obtained from CRSP. The book-to-market values are taken from COMPUSTAT and from Ken French's website based on Davis, Fama, and French (2000) (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html).

4.7 Cross-Sectional and Time-Series Variation of Tax Premia

To investigate whether a trading strategy based on tax burdens is profitable, I plot in Figure 6 the yearly abnormal return spread against the tax yield differential between the highest dividend-yield decile portfolio and the no-dividend yield portfolio. The abnormal returns are computed using the four-factor model of Carhart.

The tax capitalization hypothesis makes three predictions about the relationship between return spreads and tax differentials: First, highly taxed stocks should have a higher average return than less highly taxed stocks. Second, the return spread should be higher in periods where the tax differential is larger. And third, the return spread should be zero if there is no tax differential between the different portfolios.

Figure 6 confirms the three predictions of tax capitalization. First, highly taxed securities tend to pay significantly higher returns than less highly taxed securities. The mean return spread between high-dividend and no dividend stocks equals 4.55 percent (indicated by the dashed horizontal line) with a standard error of 1.35 percent. Thus, dividend paying stocks tend to compensate taxable investors by paying higher abnormal returns. This result is driven by the average cross-sectional difference in tax burdens. Second, the slope of the solid regression line is significantly positive and equals 1.82 with a standard error of 0.90. The slope of the regression line is based on the time-series variation in tax differentials and ignores the level effect due to the average cross-sectional variation in tax differentials. Third, the intercept is not significantly different from zero, indicating that the abnormal return spread would be zero if all equity securities were taxed symmetrically.

Although there is a positive relationship between tax yield differentials and abnormal return spreads, there remains a significant amount of variation in the return differentials. Thus, taxes can only explain a small fraction of the time-series variation in return spreads using annual data. This result is comforting, otherwise it would be puzzling why tax-neutral arbitrageurs would not immediately take advantage of the return differential by going long highly taxed stocks and going short less highly taxed stocks. The existence of such traders would likely eliminate the return differential between the two groups of securities.

The fact that taxes have an impact on asset returns implies that there are important limits to arbitrage as discussed by Pontiff (1996), Shleifer and Vishny (1997) and Fama and French (2005). There are several frictions that prevent investors to take advantage of the before-tax return differentials. First, tax arbitrage can cause trading costs that reduce the return of any strategy that generates high turnover. Second, significant risk remains in a long-term trading strategy that buys high-dividend stocks and shorts non-dividend paying stocks. The significant residual risk decreases the incentives for tax-neutral arbitrageurs to eliminate these price discrepancies.

5 Conclusions

The paper sheds new light on the controversy of whether taxes are capitalized into asset prices taking advantage of both the cross-sectional and the time-series variation in tax burdens. The effective personal taxation of equity securities fluctuated considerably between 1927-2004. Stocks paying a large proportion of their total returns as dividends face significantly higher tax burdens than stocks paying no dividends. The results indicate that there is an economically and statistically significant relationship between before-tax asset returns and effective tax rates. Stocks that tend to have higher tax burdens tend to compensate taxable investors by offering higher before-tax returns.

References

Allen, F., A. E. Bernardo, and I. Welch (2000). A theory of dividends based on tax clienteles. *Journal of Finance 55*(6), 2498–2536.

Allen, F. and R. Michaely (2003). Payout policy. In G. M. Constantinides, M. Harris, and R. M. Stulz (Eds.), *Handbook of the Economics of Finance Volume 1A Corporate Finance*, pp. 337–429. Amsterdam: Elsevier North-Holland.

Amromin, G., P. Harrison, and S. Sharpe (2005). How did the 2003 dividend tax cut affect stock prices? Federal Reserve Discussion Paper 2005-61.

Auerbach, A. J. (2002). Taxation and corporate financial policy. In A. J. Auerbach and M. S. Feldstein (Eds.), *Handbook of Public Economics Vol. 3*, pp. 1251–1292. Amsterdam: North Holland.

Auerbach, A. J., L. E. Burman, and J. M. Siegel (2000). Capital gains taxation and tax avoidance: New evidence from panel data. In J. Slemrod (Ed.), *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*. Cambridge, MA: Harvard University Press.

Auerbach, A. J. and K. A. Hassett (2005). The 2003 dividend tax cuts and the value of the firm: An event study. NBER Working Paper 11449.

Auerbach, A. J. and M. A. King (1982). Corporate financial policy with personal and institutional investors. *Journal of Public Economics 17*, 259–285.

Auerbach, A. J. and M. A. King (1983). Taxation, portfolio choice, and debt-equity ratios: A general equilibrium model. *Quarterly Journal of Economics 98*(4), 587–610.

Ayers, B. C., C. B. Cloyd, and J. R. Robinson (2002). The effect of shareholder-level dividend taxes on stock prices: Evidence from the Revenue Reconciliation Act of 1993. *Accounting Review 77*(4), 933–947.

Bailey, M. J. (1969). Capital gains and income taxation. In A. C. Harberger and M. J. Bailey (Eds.), *Taxation of Income from Capital*, pp. 11–49. Washington, DC: Brookings.

Baker, M. and J. Wurgler (2004). A catering theory of dividends. *Journal of Finance 59*(3), 1125–1165.

Bali, R. and G. L. Hite (1998). Ex dividend day stock price behavior: Discreteness or tax-induced clienteles? *Journal of Financial Economics 47*(2), 127–159.

Barclay, M. J. (1987). Dividends, taxes, and common stock prices: The ex-dividend day behavior of common stock prices before the income tax. *Journal of Financial Economics 19*(1), 31–44.

Beck, N. and J. N. Katz (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review 89*(3), 634–647.

Black, F. and M. Scholes (1974). The effects of dividend yield and dividend policy on common stock prices and returns. *Journal of Financial Economics 1*(1), 1–22.

Blume, M. E. (1980). Stock return and dividend yield: Some more evidence. *Review of Economics and Statistics 62*(1), 1–22.

Boudoukh, J., R. Michaely, M. Richardson, and M. R. Roberts (2006). On the importance of measuring payout yield: Implications for empirical asset pricing. *Forthcoming: Journal of Finance*.

Brav, A., J. R. Graham, C. R. Harvey, and R. Michaely (2005). Payout policy in the 21st century. *Journal of Financial Economics 77*, 483–527.

Brennan, M. J. (1970). Taxes, market valuation, and financial policy. *National Tax Journal 23*, 417–429.

Burman, L. E. (1999). *The Labyrinth of Capital Gains Tax Policy. A Guide for the Perplexed*. Washington: Brookings.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance 52*(1), 57–82.

Chaplinsky, S. and H. N. Seyhun (1990). Dividends and taxes: Evidence from tax-reduction strategies. *Journal of Business 63*(2), 239–260.

Chay, J., D. Choi, and J. Pontiff (2006). Market valuation of tax-timing options: Evidence from capital gains distributions. *Journal of Finance 61*(2), 837–865.

Chen, N.-F., B. Grundy, and R. F. Stambaugh (1990). Changing risk, changing risk premiums, and the dividend yield effects. *Journal of Business 63*(1/2), S51–S70.

Chetty, R., J. Rosenberg, and E. Saez (2005). The effects of taxes on market responses to dividend announcements and payments: What can we learn from the 2003 tax cut. NBER Working Paper 11452.

Dahlquist, M., G. Robertsson, and K. Rydqvist (2006). Direct evidence of dividend tax clienteles. Swedish Institute for Financial Research.

Dai, Z., E. Maydew, D. A. Shackelford, and H. H. Zhang (2006). Capital gains taxes and asset prices: Capitalization or lock-in? NBER Working Paper No. 12342.

Davis, J. L., E. F. Fama, and K. R. French (2000). Characteristics, covariances and average returns 1929-1997. *Journal of Finance 55*(1), 389–406.

Dhaliwal, D., L. Krull, O. Z. Li, and W. Moser (2005). Dividend taxes and implied cost of equity capital. *Journal of Accounting Research 43*(5), 675–708.

Dhaliwal, D., O. Z. Li, and R. Trezevant (2003). Is a dividend tax penalty incorporated into the return on a firm's common stock? *Journal of Accounting and Economics 35*, 155–178.

Eades, K. M., P. J. Hess, and E. H. Kim (1984). On interpreting security returns during the ex-dividend period. *Journal of Financial Economics 13*, 3–34.

Eades, K. M., P. J. Hess, and E. H. Kim (1994). Time-series variation in dividend pricing. *Journal of Finance 49*, 1617–1638.

Elton, E. J. and M. J. Gruber (1970). Marginal stockholder tax rates and the clientele effect. *Review of Economics and Statistics 52*(1), 68–74.

Elton, E. J., M. J. Gruber, and C. R. Blake (2005). Marginal stockholder tax effects and ex-dividend day behavior: Evidence from taxable versus non-taxable closed-end funds. *Forthcoming: Review of Economics and Statistics*.

Fama, E. F. and K. R. French (1993). Common risk factors in the return on bonds and stocks. *Journal of Financial Economics 33*(1), 3–53.

Fama, E. F. and K. R. French (1998). Taxes, financing decisions, and firm value. *Journal of Finance 53*(3), 819–843.

Fama, E. F. and K. R. French (2005). Disagreement, tastes, and asset prices. *Forthcoming: Journal of Financial Economics*.

Fama, E. F. and J. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy 81*, 607–636.

Feenberg, D. and E. Coutts (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management 12*(1), 189–194.

Feldstein, M., J. Slemrod, and S. Yitzhaki (1980). The effects of taxation on the selling of corporate stock and the realization of capital gains. *Quarterly Journal of Economics 94*(4), 777–791.

Fink, J., K. E. Fink, G. Grullon, and J. P. Weston (2005). IPO vintage and the rise of idiosyncratic risk. James Madison University.

Frank, M. and R. Jagannathan (1998). Why do stock prices drop by less than the value of the dividend? Evidence from a country without taxes. *Journal of Financial Economics 47*(2), 161–188.

Gordon, R. H. and D. F. Bradford (1980). Taxation and the stock market value of capital gains and dividends: Theory and empirical results. *Journal of Public Economics 14*(2), 109–136.

Graham, J. R. (2003). Taxes and corporate finance: A review. *Review of Financial Studies 16*(4), 1075–1129.

Graham, J. R. and A. Kumar (2006). Do dividend clienteles exist? Evidence on dividend preferences of retail investors. *Journal of Finance 61*, 1305–1336.

Graham, J. R., R. Michaely, and M. R. Roberts (2003). Do price discreteness and transactions costs affect stock returns? Comparing ex-dividend pricing before and after decimalization. *Journal of Finance 58*(6), 2611–2636.

Green, R. C. and B. Hollifield (2003). The personal-tax advantage of equity. *Journal of Financial Economics 67*, 175–216.

Green, R. C. and K. Rydqvist (1999). Ex-day behavior with dividend preference and limitations to short-term arbitrage: The case of Swedish lottery bonds. *Journal of Financial Economics 53*(2), 145–187.

Grinstein, Y. and R. Michaely (2005). Institutional holdings and payout policy. *Journal of Finance 60*, 1389–1426.

Guenther, D. A. (1994). The relation between tax rates and pre-tax returns: Direct evidence from the 1981 and 1986 tax rate reductions. *Journal of Accounting and Finance 18*, 379–393.

Internal Revenue Service (Ed.) (1954). *Statistics of Income*. Washington D.C.: U.S. Treasury Department.

Ivkovich, Z., J. Poterba, and S. Weisbenner (2005). Tax-motivated trading by individual investors. *American Economic Review 95*(5), 1605–1630.

Jagannathan, R. and Y. Wang (2005). Lazy investors, discretionary consumption, and the cross section of stock returns. *Forthcoming: Journal of Finance*.

Jin, L. (2006). Capital gain tax overhang and price pressure. *Forthcoming: Journal of Finance*.

Joint Committee on Taxation (1988-1998). *General Explanation of Tax Legislation*. Washington D.C.: U.S. Government Printing Office.

Kalay, A. (1982). The ex-dividend day behavior of stock prices: A re-examination of the clientele effect. *Journal of Finance 37*(4), 1059–1070.

Kalay, A. and R. Michaely (2000). Dividends and taxes: A re-examination. *Financial Management*, 55–75.

Keim, D. B. (1985). Dividend yields and stock returns: Implications of abnomral January returns. *Journal of Financial Economics 14*, 473–489.

Lang, M. H. and D. A. Shackelford (2000). Capitalization of capital gains taxes: Evidence from stock price reactions to the 1997 rate reduction. *Journal of Public Economics 76*(1), 69–85.

Lei, Q. (2005). Cash distributions and returns. University of Michigan.

Litzenberger, R. H. and K. Ramaswamy (1979). The effects of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics 7*, 163–195.

Litzenberger, R. H. and K. Ramaswamy (1980). Dividends, short selling restrictions, tax-induced investor clienteles and market equilibrium. *Journal of Finance 35*, 469–482.

Litzenberger, R. H. and K. Ramaswamy (1982). The effects of dividends on common stock prices: Tax effects or information effects? *Journal of Finance 37*, 429–443.

McGrattan, E. R. and E. C. Prescott (2005). Taxes, regulations, and the value of U.S. and U.K. corporations. *Review of Economic Studies 72*(3), 767–796.

Michaely, R. (1991). Ex-dividend day stock price behavior: The case of the 1986 tax reform act. *Journal of Finance 46*(3), 845–859.

Miller, M. H. and M. S. Scholes (1978). Dividends and taxes. *Journal of Financial Economics 6*(4), 333–364.

Miller, M. H. and M. S. Scholes (1982). Dividends and taxes: Some empirical evidence. *Journal of Political Economy 90*(6), 1118–1141.

Naranjo, A., M. Nimalendran, and M. Ryngaert (1998). Stock returns, dividend yields, and taxes. *Journal of Finance 53*(6), 2029–2057.

Newey, W. and K. West (1987). A simple, positive, semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 1987*(55), 703–708.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Pechman, J. A. (1987). *Federal Tax Policy* (5th ed.). Washington D.C.: Brookings.

Petersen, M. A. (2005). Estimating standard errors in finance panel data sets: Comparing approaches. Kellogg School of Management, Northwestern University.

Pontiff, J. (1996). Costly arbitrage: Evidence from closed-end funds. *Quarterly Journal of Economics 111*(4), 1135–1151.

Poterba, J. M. (1987a). How burdensome are capital gains taxes? Evidence from the United States. *Journal of Public Economics 33*(2), 157–172.

Poterba, J. M. (1987b). Tax policy and corporate saving. *Brookings Papers on Economic Activity 1987*(2), 455–503.

Poterba, J. M. (2002). Taxation, risk-taking, and household portfolio behavior. In A. J. Auerbach and M. S. Feldstein (Eds.), *Handbook of Public Economics Vol. 3*, pp. 1109–1171. Amsterdam: North Holland.

Poterba, J. M. and L. H. Summers (1984). New evidence that taxes affect the valuation of dividends. *Journal of Finance 39*(5), 1397–1415.

Prais, S. J. and C. B. Winsten (1954). Trend esimators and serial correlation. Cowles Commission Discussion Paper No. 383.

Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. *Journal of Finance 52*(1), 35–55.

Sialm, C. (2005). Tax changes and asset pricing: Time-series evidence. NBER Working Paper 11756.

Sinai, T. and J. Gyourko (2004). The asset price incidence of capital gains taxes: Evidence from the taxpayer relief act of 1997 and publicly-traded real estate firms. *Journal of Public Economics 88*, 1543–1565.

Wang, J., C. Wu, and F. Zhang (2005). Liquidity, default, taxes and yields on municipal bonds. Federal Reserve Discussion Paper 2005-35.

A Data Appendix

A.1 Statutory Tax Rates

Taxable income is derived for three real income levels after deducting exemptions for a married couple filing jointly with two dependent children from the fixed income levels. The proportion of total deductions relative to the adjusted gross income is assumed to equal the proportion of total deductions in the whole population for each year as reported by the Internal Revenue Service. The marginal income tax brackets and exemptions are determined using the Statistics of Income of the Internal Revenue Service (1954) for the years 1913-1943, Pechman (1987) for the years 1944-1987, and different issues of the Instructions to Form 1040 from the IRS for the remaining years between 1988-2004. The values of the Consumer Price Index from 1913-1957 are taken from the Bureau of Labor Statistics.[25] Total deductions as a proportion of adjusted gross income (AGI) are derived from different issues of the Statistics of Income of the IRS. Marginal income tax rates for individuals in two different tax brackets corresponding to Adjusted Gross Income levels of 100 and 250 thousand U.S. dollars (with 2004 consumer prices), as well as the highest marginal income tax rate are derived.

The long-term capital gains tax rate applies to realized gains with a holding period of more than five years. The data source for the capital gains tax rates for 1927-1950 is the Synopsis of Federal Tax Laws from the Statistics of Income for 1950. The remaining tax rates are taken from different issues of the General Explanations of Tax Legislation by the Joint Committee on Taxation (1998) and Table 2-4 from Burman (1999).

A.2 Average Marginal Tax Rates

The time series for the average marginal tax rates of dividends and short- and long-term capital gains are computed using different annual issues of the Statistics of Income between 1917 and 1964 and the average marginal tax rates from the National Bureau of Economic Research between 1965-2004. Post-2001 data is derived from the 2000 Tax Model, since these years are not yet available from the Statistics of Income of the IRS. However, the data includes the impact of EGGTRA and JGTRRA and reflects therefore the tax changes through January 2005. The NBER publishes average marginal tax rates for selected income sources since 1960 using their Taxsim software.[26]

The NBER publishes average marginal tax rates that include state and local taxes. For the early data, I use the National Income and Product Accounts published by the Bureau of Economic Analysis to determine the state and local tax rates. The BEA summarizes the current personal income tax receipts of state and local governments

[25] Data can be found at http://www.bls.gov/cpi/home.htm.

[26] Additional information on this microsimulation model can be found in Feenberg and Coutts (1993). The time series can be downloaded from http://www.nber.org/∼taxsim.

(Table 3.3) and the federal government (Table 3.2).[27] I assume that the state and local government tax rate is a fixed proportion of the federal tax rate according to the annual revenues.

A.3 Aggregate Capital Gains Yields

The annual Statistics of Income of the Internal Revenue Service report for most years between 1917 and 2004 the total short- and long-term capital gains and the dividends declared by individuals. The capital gains given by the Statistics of Income include capital gains from many sources and not just from stock transactions. The IRS does unfortunately not report every year the proportion of capital gains that result from transactions of corporate equities. However, for eight years between 1959 and 2004, the IRS reports the sources of capital gains in more detail. On average, about 35 percent of the capital gains result from transactions of corporate equity. I interpolate the fraction of stock capital gains using these eight years.

The IRS reports the dollar amount of dividends D_t and short- and long-term capital gains SCG_t and LCG_t. However, the IRS does not report the value of the total taxable assets. I compute the aggregate short- and long-term capital gains yields y_t^{scg} and y_t^{lcg}, by multiplying the aggregate value-weighted dividend yield y_t^{div} with the ratio between the short- and long-term realized capital gains SCG_t and LCG_t divided by the total dividend payments D_t:

$$y_t^{scg} = y_t^{div} \frac{SCG_t}{D_t}, \tag{9}$$

$$y_t^{lcg} = y_t^{div} \frac{LCG_t}{D_t}. \tag{10}$$

I assume that investors anticipate to realize a fixed proportion of capital gains out of the total expected returns net of expected dividend payments. Thus, investors expect to realize larger capital gains for stock portfolios that are anticipated to pay smaller dividend yields. The expected return on portfolio k is given by $\hat{r}_{k,t}$ and the expected market return is given by $\hat{r}_{M,t}$. The time-series of capital gains yields for equity portfolio k are assumed to be as follows:

$$\hat{y}_{k,t}^{lcg} = \hat{y}^{lcg} \frac{\hat{r}_{k,t} - \hat{y}_{k,t}^{div}}{\hat{r_{M,t}} - \hat{y}^{div}} = \eta_0^{lcg} - \eta_1^{lcg} \hat{y}_{k,t}^{div}, \tag{11}$$

$$\hat{y}_{k,t}^{scg} = \hat{y}^{scg} \frac{\hat{r}_{k,t} - \hat{y}_{k,t}^{div}}{\hat{r_{M,t}} - \hat{y}^{div}} = \eta_0^{scg} - \eta_1^{scg} \hat{y}_{k,t}^{div}. \tag{12}$$

This specification results in average expected capital gains realizations that fit the average realization behavior over the whole sample period according to the IRS.

The expected tax yield $\hat{\kappa}$ depends positively on the expected return of the portfolio k, because portfolios with higher expected returns are assumed to generate higher

[27] The data can be downloaded from http://www.bea.gov.

capital gains according to equations (11) and (12). In the empirical section, I relate abnormal portfolio returns to the tax yield coefficient $\hat{\kappa}$. To avoid any spurious correlation between the tax yield and the portfolio return, I adjust the tax yield coefficient by using the same expected return for all portfolios over the whole sample period. As demonstrated in a robustness test in Section 4, this assumption results in more conservative estimates of the tax capitalization coefficient compared to the case where the expected returns of each portfolio are set equal to the sample averages of the portfolio returns. Based on these assumptions, it is possible to derive effective tax yields for different portfolios according to equation (1):

$$\begin{aligned}\hat{\kappa}_{k,t} &= \hat{y}_{k,t}^{div}\tau_t^{div} + \hat{y}_{k,t}^{scg}\tau_t^{scg} + \hat{y}_{k,t}^{lcg}\tau_t^{lcg} \\ &= \hat{y}_{k,t}^{div}\tau_t^{div} + \left(\eta_0^{scg} - \eta_1^{scg}\hat{y}_{k,t}^{div}\right)\tau_t^{scg} + \left(\eta_0^{lcg} - \eta_1^{lcg}\hat{y}_{k,t}^{div}\right)\tau_t^{lcg} \\ &= \hat{y}_{k,t}^{div}\left(\tau_t^{div} - \eta_1^{scg}\tau_t^{scg} - \eta_1^{lcg}\tau_t^{lcg}\right) + \eta_0^{scg}\tau_t^{scg} + \eta_0^{lcg}\tau_t^{lcg}. \end{aligned} \tag{13}$$

A.4 Tax-Exempt Assets

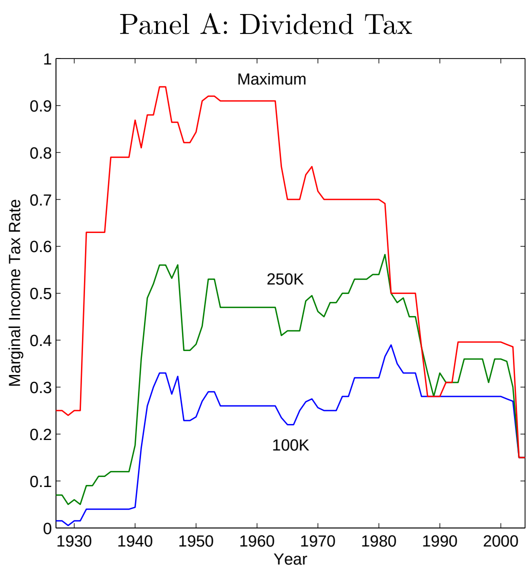
The proportion of equity held in taxable accounts is estimated using the Flow of Funds published by the Board of Governors of the Federal Reserve Bank.[28] The proportion is only computed for equities held by domestic investors, since it would be impossible to determine the marginal tax rates faced by international stock investors. The detailed derivation of the time series is available upon request. The flow of funds publishes this distribution of equity holdings only between 1945 and 2001. The values prior to 1945 and after 2001 are taken from the most recent available year.

A.5 Taxable Dividends

Dividends are defined here as taxable dividends according to the CRSP distribution codes. The following distribution codes correspond to taxable dividends: 1200, 1202, 1212, 1218, 1222, 1228, 1232, 1231, 1238, 1239, 1242, 1248, 1252, 1258, 1262, 1268, 1272, 1278, 1279, 1282, 1292, 1312, 1318, 1332, 1338, 1342, 1348, 1352, 1362, 1368, 1372, 1378, 1412, 1418, 1438, 1712, 1718, 1772, 1812, 1818, 1872, and 1999.

[28] The data can be downloaded from http://www.federalreserve.gov/releases/Z1/.

Figure 1: Statutory Federal Marginal Dividend and Capital Gains Tax Rates
The marginal dividend and long-term capital gains tax rates are depicted over the period from 1927 to 2004 for three different real income levels. Two curves correspond to the marginal income tax rates for households with real income levels of 100 and 250 thousand U.S. dollars expressed in 2004 consumer prices. The third curve corresponds to the marginal income tax rate for the top income tax bracket. The data are based on data from the IRS.



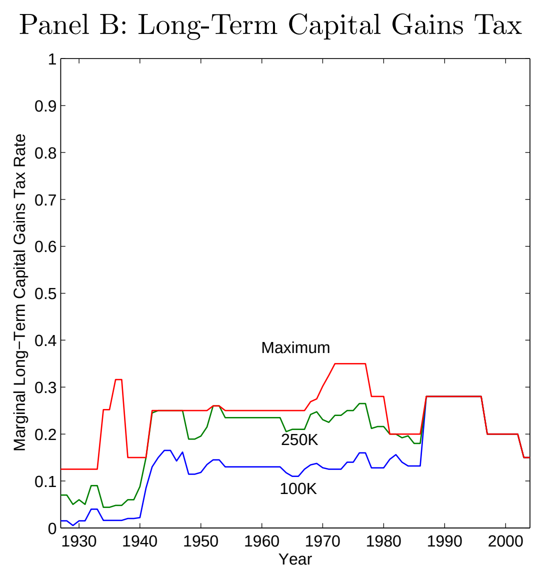


Figure 2: Average Marginal Investment Income Tax Rates
The dollar-weighted average marginal tax rates on dividend income and long-term capital gains are depicted between 1927 and 2004. The tax rates include taxes imposed by state and local governments. The data are obtained from the Statistics of Income of the IRS and from the NBER.

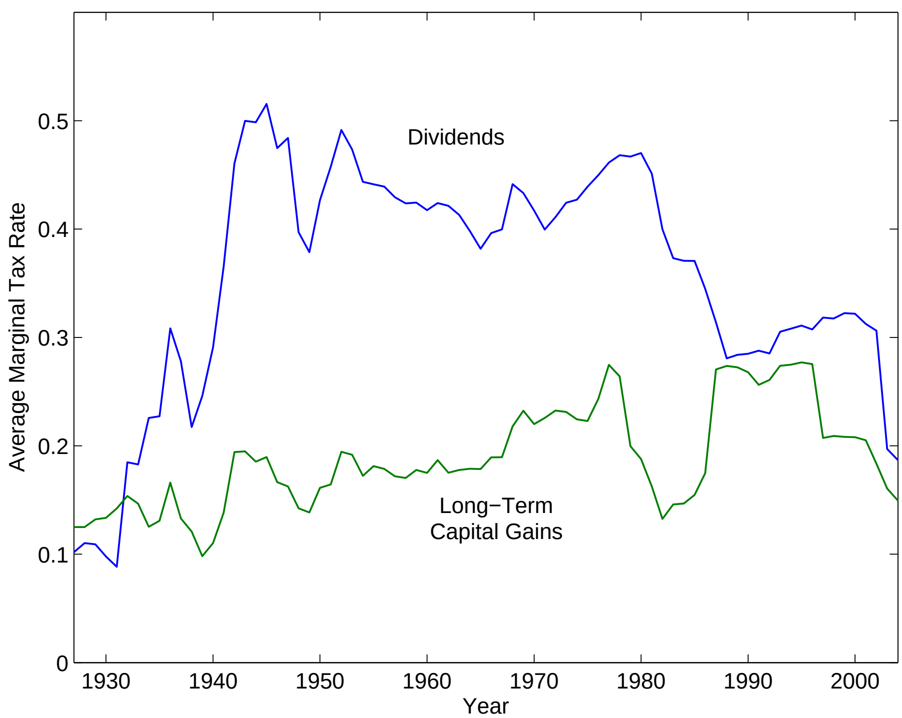


Figure 3: Distribution of Effective Tax Rates
The effective tax rates are depicted for two stock portfolios sorted according to the prior-year dividend yield over the period from 1927 to 2004. The effective tax rate $\hat{\tau}$ is defined as the ratio between the tax yield $\hat{\kappa}$ and the average return over the whole sample period. The tax yield is defined as $\hat{\kappa}_{k,t} = \hat{y}^{div}_{k,t}\tau^{div}_t + \hat{y}^{scg}_{k,t}\tau^{scg}_t + \hat{y}^{lcg}_{k,t}\tau^{lcg}_t$, where τ^{div}_t, τ^{scg}_t, and τ^{lcg}_t are the average marginal tax rates on dividends and short- and long-term capital gains, and $\hat{y}^{div}_{k,t}$, $\hat{y}^{scg}_{k,t}$, and $\hat{y}^{lcg}_{k,t}$ are the expected dividend yields, and the expected long- and short-term capital gains yields. The lower curve corresponds to the portfolio that includes all the stocks that did not pay any dividends in the previous year, and the upper curve corresponds to the portfolio that includes all stocks that did pay dividends in the previous period.

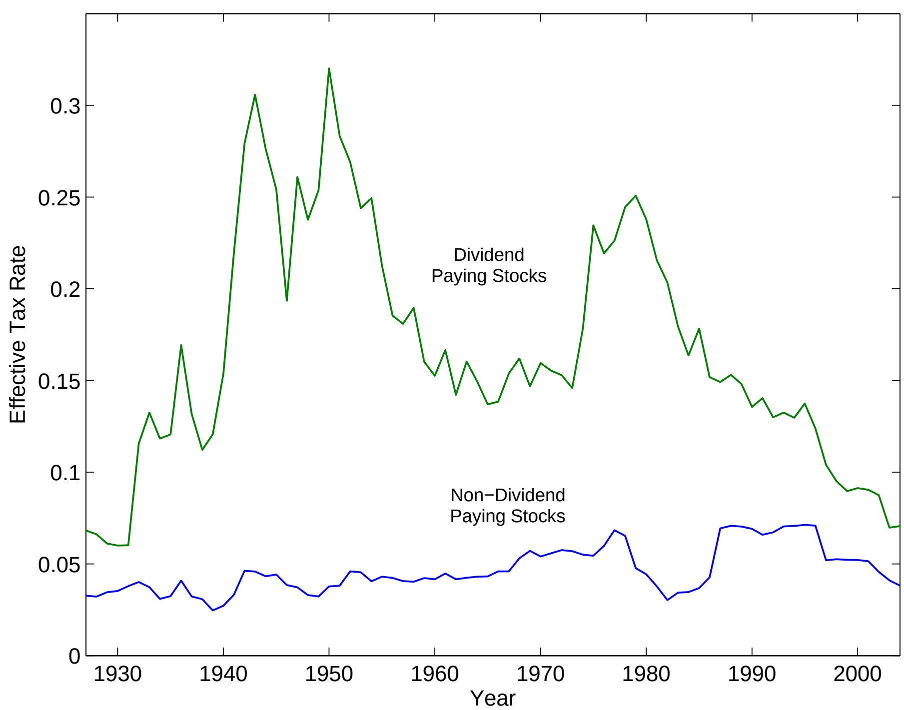


Figure 4: Factor Loadings for Dividend and Non-Dividend Portfolios
The factor loadings for the dividend and the non-dividend portfolios are summarized over the period from 1927 to 2004. The factor loadings of the Carhart (1997) model are computed on a rolling basis using 60 months of prior return data. The portfolios are value-weighted.

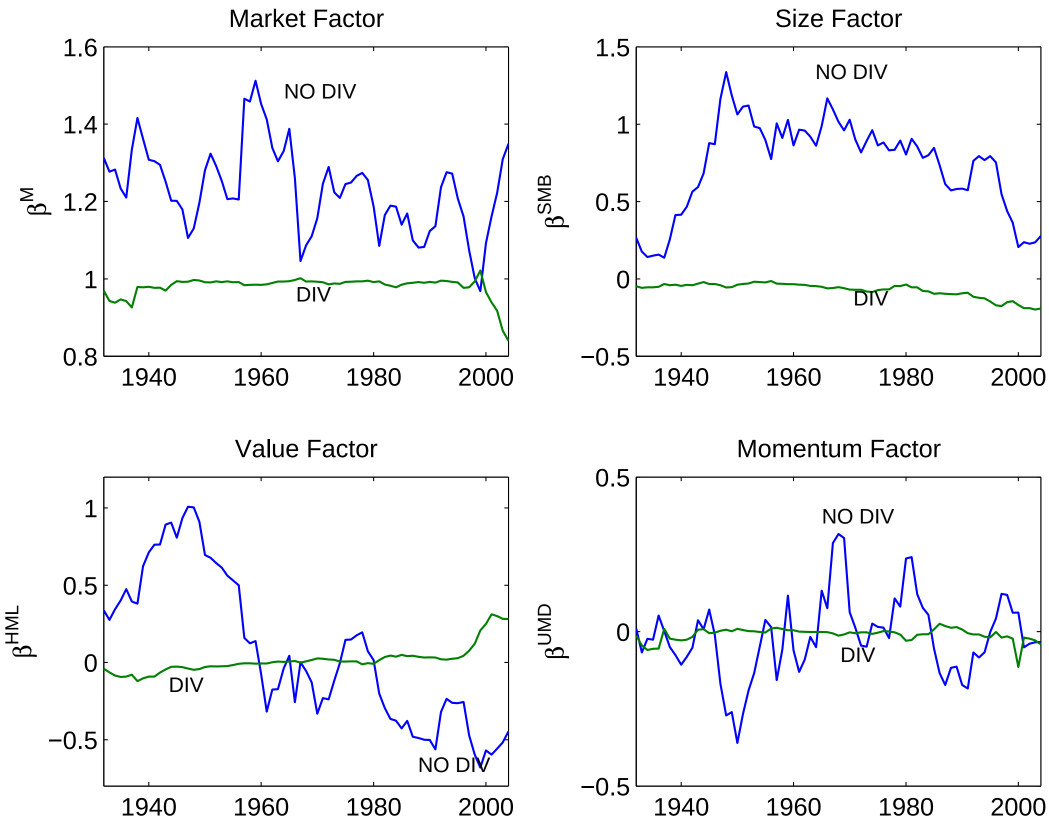


Figure 5: Cross-Sectional Relationship Between Abnormal Returns and Effective Tax Rates

The figure relates average tax yields to the performance of 30 value-weighted portfolios formed according to six dividend yield groups and five market capitalization groups between 1927-2004. The excess return is defined as the difference between the portfolio return and the value-weighted market return. The abnormal returns are computed based on the CAPM, the Fama-French, or the Carhart models.

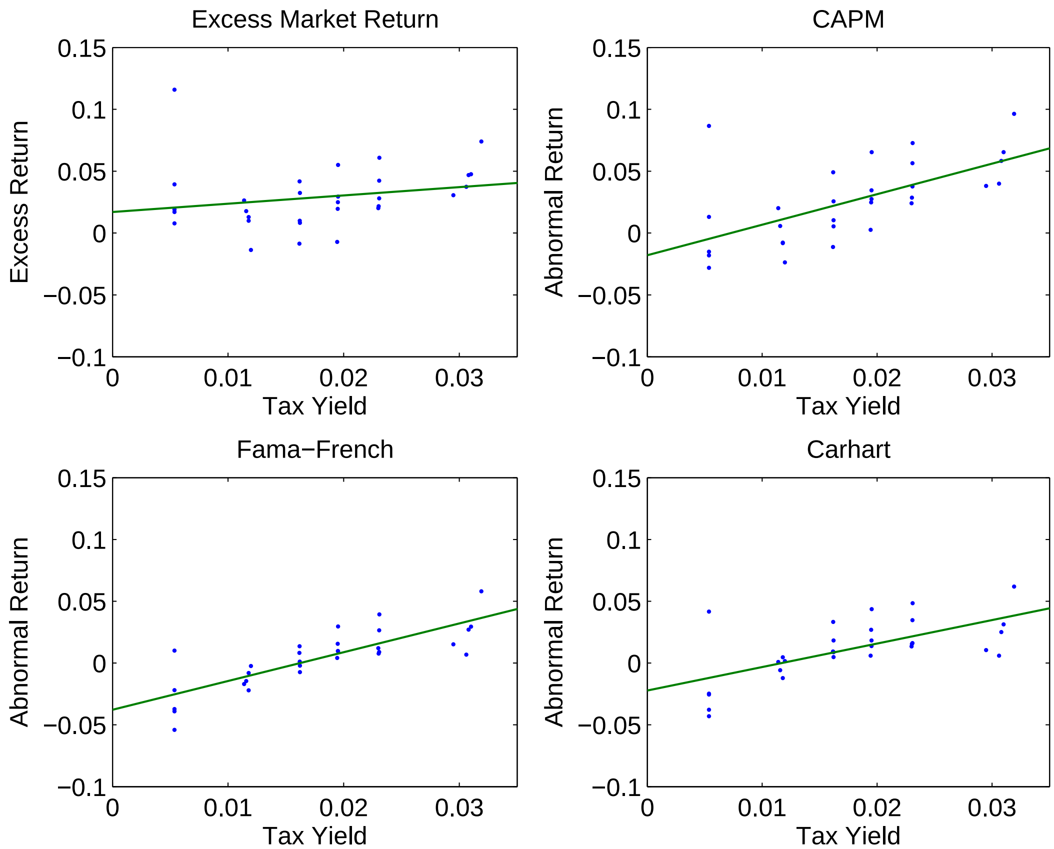


Figure 6: Time-Series Relationship Between Return Differentials and Tax Differentials
The figure relates the annual tax yield differentials to the abnormal performance differentials between stocks in the top dividend-yield decile and non-dividend paying stocks using the Carhart model between 1932-2004. The horizontal axis depicts the difference between the tax yields and the vertical axis depicts the difference between the abnormal Carhart returns of the two portfolios.

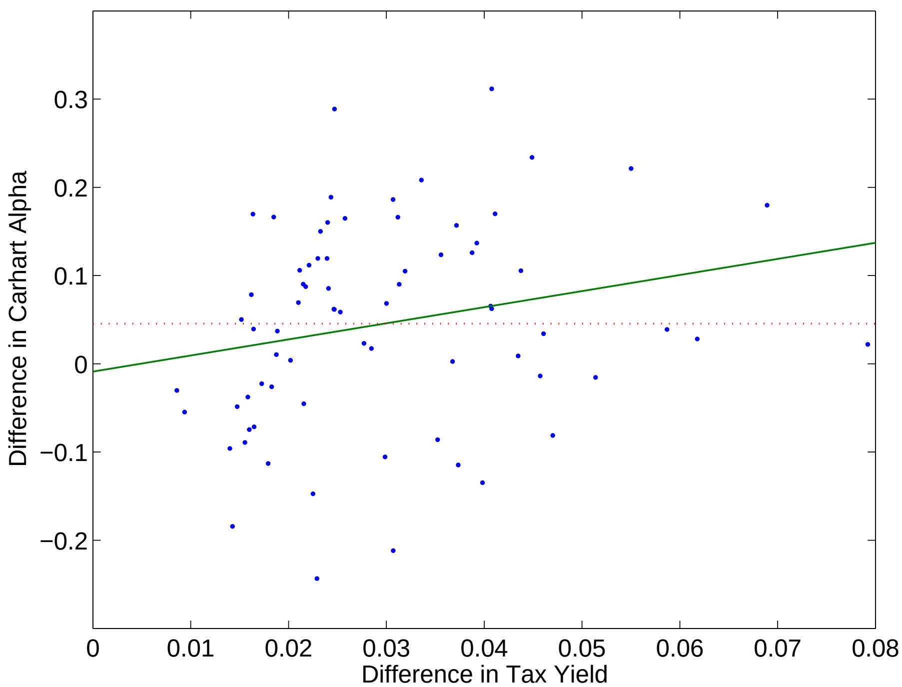


Table 1: Dividend and Capital Gains Tax Rates

This table summarizes the moments of the main variables used in this study over the period from 1927 to 2004. The variables are expressed in percent.

Panel A: Moments

		Mean	Std. Dev.	Min	Max
(1)	Statutory Dividend Tax (100K)	23.16	10.18	0.50	39.00
(2)	Statutory Dividend Tax (250K)	37.31	15.45	5.00	58.26
(3)	Statutory Dividend Tax (Max)	63.71	23.63	15.00	94.00
(4)	Statutory LT Capital Gains Tax (100K)	13.72	7.47	0.50	28.00
(5)	Statutory LT Capital Gains Tax (250K)	19.95	7.15	4.40	28.00
(6)	Statutory LT Capital Gains Tax (Max)	24.03	6.01	12.50	35.00
(7)	Average Dividend Tax	29.02	11.10	8.37	48.82
(8)	Average ST Capital Gains Tax	28.62	11.37	6.77	48.82
(9)	Average LT Capital Gains Tax	14.90	3.05	8.30	21.17

Panel B: Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	1								
(2)	0.90	1							
(3)	0.18	0.53	1						
(4)	0.73	0.44	−0.29	1					
(5)	0.88	0.80	0.17	0.84	1				
(6)	0.52	0.58	0.34	0.42	0.64	1			
(7)	0.47	0.77	0.91	−0.04	0.45	0.49	1		
(8)	0.52	0.80	0.86	0.01	0.50	0.54	0.99	1	
(9)	0.35	0.52	0.43	0.27	0.60	0.73	0.58	0.61	1

Table 2: Persistence of Dividend Yields

This table summarizes the dividend yields for the value-weighted portfolios formed according to the lagged one-year dividend yields in the formation period, after one, three, and five years over the period from 1927 to 2004. The dividend yields are at an annual frequency corresponding to December values. The values are expressed in percent.

	Lagged and Future Average Dividend Yields in Percent			
	Year 0	Year 1	Year 3	Year 5
No Dividend Portfolio	0	0.72	1.59	2.04
Lowest Dividend Decile	1.46	1.91	2.15	2.41
Decile 2	2.49	2.85	3.02	3.19
Decile 3	3.14	3.30	3.53	3.64
Decile 4	3.69	3.87	3.92	4.00
Decile 5	4.23	4.26	4.24	4.19
Decile 6	4.76	4.66	4.64	4.50
Decile 7	5.31	5.16	4.89	4.74
Decile 8	6.00	5.66	5.26	5.14
Decile 9	6.94	6.13	5.70	5.53
Highest Dividend Decile	9.32	7.35	6.54	6.15

Table 3: Measures of Tax Burden

This table summarizes the moments of three variables capturing the tax burden on a cross-section of equity securities over the period from 1927 to 2004. The lagged dividend yield y^{div} is based on the dividends of portfolio k during the last 12 months. The effective tax yield $\hat{\kappa}$ is defined according to equation (1) and the effective tax rate $\hat{\tau}$ is defined as the ratio between the effective tax yield $\hat{\kappa}$ and the average return over the whole sample period. Three different portfolio formation criteria are used: (1) 30 portfolios based on six dividend yield groups and five size groups; (2) 11 portfolios based on one portfolio including all non-dividend paying stocks and decile portfolios including dividend paying stocks; and (3) two portfolios based on one portfolio including non-dividend paying stocks and one portfolio including dividend paying stocks. The values are at an annual frequency and correspond to December values.

Panel A: 30 Dividend Yield and Size Portfolios

		Mean	Standard Deviation	Minimum	Maximum	Correlation
(1)	Dividend Yield	3.97	3.14	0.00	21.48	1
(2)	Effective Tax Yield	1.79	1.16	0.29	7.76	0.85
(3)	Effective Tax Rate	15.45	10.04	2.46	66.94	0.85

Panel A: 30 Dividend Yield and Size Portfolios

		Mean	Standard Deviation	Minimum	Maximum	Correlation
(1)	Dividend Yield	4.30	3.05	0.00	18.05	1
(2)	Effective Tax Yield	1.90	1.16	0.29	8.72	0.83
(3)	Effective Tax Rate	16.37	10.00	2.46	7.53	0.83

Panel A: 30 Dividend Yield and Size Portfolios

		Mean	Standard Deviation	Minimum	Maximum	Correlation
(1)	Dividend Yield	2.20	2.44	0.00	8.42	1
(2)	Effective Tax Yield	1.23	0.87	0.29	3.71	0.88
(3)	Effective Tax Rate	10.62	7.52	2.46	32.01	0.88

Table 4: Raw and Abnormal Returns for Portfolios Sorted by Dividend Yield
This table summarizes the raw and the abnormal returns for portfolios formed according to the initial dividend yields in the month after the portfolio formation over the period from 1927 to 2004. The factor loadings are computed on a rolling basis using 60 months of return data. The returns are expressed in percent per month and standard errors are summarized in parentheses. The significance levels are abbreviated with asterisks: One, two, and three asterisks denote significance at the 10, 5, and 1 percent level, respectively.

	Monthly Value-Weighted Returns in Percent			
	Raw Return	CAPM	Fama-French	Carhart
No Dividend Portfolio	1.04***	−0.20*	−0.26***	−0.26***
	(0.29)	(0.11)	(0.07)	(0.07)
Lowest Dividend Decile	0.86***	−0.18**	−0.09	−0.06
	(0.23)	(0.09)	(0.09)	(0.09)
Decile 2	0.92***	−0.13*	−0.02	0.05
	(0.21)	(0.07)	(0.07)	(0.07)
Decile 3	0.94***	−0.06	0.05	0.09
	(0.19)	(0.06)	(0.06)	(0.06)
Decile 4	0.96***	0.01	0.10*	0.13**
	(0.18)	(0.06)	(0.06)	(0.06)
Decile 5	0.95***	0.09	0.08	0.13**
	(0.17)	(0.06)	(0.06)	(0.06)
Decile 6	0.99***	0.09	0.03	0.05
	(0.17)	(0.07)	(0.06)	(0.06)
Decile 7	1.10***	0.20***	0.08	0.14**
	(0.17)	(0.07)	(0.06)	(0.06)
Decile 8	1.17***	0.26***	0.11*	0.11*
	(0.18)	(0.07)	(0.06)	(0.06)
Decile 9	1.16***	0.30***	0.12*	0.10
	(0.17)	(0.08)	(0.07)	(0.07)
Highest Dividend Decile	1.35***	0.43***	0.17**	0.12
	(0.17)	(0.09)	(0.08)	(0.08)
Difference Highest Dividend	0.32*	0.64***	0.43***	0.38***
Minus No Dividend	(0.19)	(0.15)	(0.10)	(0.10)
Difference Highest Dividend	0.50***	0.62***	0.26**	0.18
Minus Lowest Dividend	(0.16)	(0.14)	(0.13)	(0.12)

Table 5: Tax Capitalization Regressions

This table summarizes the tax capitalization coefficient δ of the following regression: $\tilde{\alpha}_{k,t} = \gamma + \delta\hat{\kappa}_{k,t} + \epsilon_{k,t}$, where $\tilde{\alpha}_{k,t}$ is the abnormal return of portfolio k at time t and $\hat{\kappa}_{k,t}$ is the tax yield of portfolio k at time t using data over the period from 1927 to 2004. Excess returns are computed by subtracting the value-weighted market return from the portfolio return. Factor-adjusted returns are computed by subtracting the expected returns of the CAPM, the Fama-French, and the Carhart models from the portfolio return. The factor loadings are computed on a rolling basis using 60 months of return data. Three different portfolio formation criteria are used: (1) 30 portfolios based on six dividend yield groups and five size groups; (2) 11 portfolios based on one portfolio including all non-dividend paying stocks and decile portfolios including dividend paying stocks; and (3) two portfolios based on one portfolio including non-dividend paying stocks and one portfolio including dividend paying stocks. The standard errors take into account clustering by time period and are summarized in parentheses. The significance levels are abbreviated with asterisks: One, two, and three asterisks denote significance at the 10, 5, and 1 percent level, respectively.

Panel A: 30 Dividend Yield and Size Portfolios

	Excess Return	CAPM	Fama-French	Carhart
Tax Yield	1.11**	1.29***	1.75***	1.37***
	(0.51)	(0.44)	(0.29)	(0.25)
Constant	0.08	0.02	−0.24***	−0.11**
	(0.12)	(0.11)	(0.06)	(0.05)
R-Squared (in Percent)	0.08	0.14	0.54	0.34

Panel B: 11 Dividend Yield Portfolios

	Excess Return	CAPM	Fama-French	Carhart
Tax Yield	1.12***	1.65***	1.30***	0.98***
	(0.34)	(0.33)	(0.30)	(0.26)
Constant	−0.09	−0.20***	−0.18***	−0.11**
	(0.06)	(0.06)	(0.05)	(0.05)
R-Squared (in Percent)	0.17	0.47	0.38	0.22

Panel C: 2 Dividend Yield Portfolios

	Excess Return	CAPM	Fama-French	Carhart
Tax Yield	−0.42	1.44**	1.87***	1.92***
	(0.88)	(0.68)	(0.43)	(0.43)
Constant	0.10	−0.23*	−0.31***	−0.31***
	(0.15)	(0.12)	(0.08)	(0.08)
R-Squared (in Percent)	0.01	0.20	0.85	0.89

Table 6: Tax Capitalization Regressions: Subperiod Evidence

This table summarizes the tax capitalization coefficient δ of the following regression: $\tilde{\alpha}_{k,t} = \gamma + \delta \hat{\kappa}_{k,t} + \epsilon_{k,t}$, where $\tilde{\alpha}_{k,t}$ is the abnormal return of portfolio k at time t and $\hat{\kappa}_{k,t}$ is the tax yield of portfolio k at time t. Excess returns are computed by subtracting the value-weighted market return from the portfolio return. Factor-adjusted returns are computed by subtracting the expected returns of the CAPM, the Fama-French, and the Carhart models from the portfolio return. The factor loadings are computed on a rolling basis using 60 months of return data. The 30 stock portfolios are formed based on six dividend yield groups and five size groups. The standard errors take into account clustering by time period and are summarized in parentheses. The significance levels are abbreviated with asterisks: One, two, and three asterisks denote significance at the 10, 5, and 1 percent level, respectively.

	Excess Return	CAPM	Fama-French	Carhart
1932-1949	0.85	0.27	2.01***	1.25**
	(0.98)	(0.96)	(0.64)	(0.53)
1950-1959	0.54	1.62***	2.01***	1.55***
	(0.63)	(0.60)	(0.43)	(0.42)
1960-1969	0.67	1.97	3.51***	3.15***
	(1.90)	(1.46)	(0.85)	(0.80)
1970-1979	3.33*	3.29**	0.75	1.18*
	(1.69)	(1.30)	(0.65)	(0.62)
1980-1989	3.93**	5.23***	2.14***	2.73***
	(1.68)	(1.35)	(0.74)	(0.65)
1990-2004	1.90	2.23	2.55**	1.79
	(2.49)	(1.91)	(1.20)	(1.12)

Table 7: Tax Capitalization Regressions Using Different Measures of the Equity Tax Burden

This table summarizes the tax capitalization coefficient δ of the following regression: $\tilde{\alpha}_{k,t} = \gamma + \delta\hat{\theta}_{k,t} + \epsilon_{k,t}$, where $\tilde{\alpha}_{k,t}$ is the abnormal return of portfolio k at time t and $\hat{\theta}_{k,t}$ is one of 12 different measures of the tax burden of portfolio k at time t using data over the period from 1927 to 2004. Excess returns are computed by subtracting the value-weighted market return from the portfolio return. Factor-adjusted returns are computed by subtracting the expected returns of the CAPM, the Fama-French, and the Carhart models from the portfolio return. The factor loadings used to compute the abnormal returns are computed on a rolling basis using 60 months of return data. The 30 stock portfolios are formed based on six dividend yield groups and five size groups. The standard errors take into account clustering by time period and are summarized in parentheses. The significance levels are abbreviated with asterisks: One, two, and three asterisks denote significance at the 10, 5, and 1 percent level, respectively.

	Excess Return	CAPM	Fama French	Carhart
(1) Tax Yield	1.11**	1.29***	1.75***	1.37***
(Base Case)	(0.51)	(0.44)	(0.29)	(0.25)
(2) Tax Yield with Fitted	1.18*	1.20**	1.98***	1.54***
Dividend Yield	(0.69)	(0.60)	(0.36)	(0.31)
(3) Tax Yield Using Declared Lagged	1.58***	1.91***	2.26***	2.83***
Quarterly Dividend	(0.56)	(0.52)	(0.34)	(0.31)
(4) Statutory Tax Yield for	1.59*	1.72**	2.36***	1.92***
$100,000 Real Income	(0.83)	(0.73)	(0.47)	(0.42)
(5) Statutory Tax Yield for	0.97**	1.04**	1.49***	1.17***
$250,000 Real Income	(0.48)	(0.42)	(0.28)	(0.24)
(6) Statutory Tax Yield for	0.44*	0.49**	0.75***	0.56***
Maximum Real Income	(0.23)	(0.20)	(0.14)	(0.12)
(7) Tax Yield with No	1.02**	1.22***	1.56***	1.21***
Capital Gains Realizations	(0.44)	(0.37)	(0.25)	(0.22)
(8) Tax Yield with Full	1.09	0.88	2.02***	1.59***
Capital Gains Realizations	(0.80)	(0.76)	(0.47)	(0.41)
(9) Tax Yield Including Stocks	1.03**	1.09**	1.72***	1.28***
in Tax-Sheltered Accounts	(0.51)	(0.45)	(0.30)	(0.26)
(10) Tax Yield with Varying	1.25**	1.45***	1.75***	1.42***
Average Portfolio Returns	(0.49)	(0.41)	(0.28)	(0.24)
(11) Lagged Tax Yield	0.86*	1.17***	1.78***	1.38***
	(0.52)	(0.44)	(0.27)	(0.24)
(12) Lagged Dividend Yield	0.33*	0.63***	0.63***	0.50***
	(0.19)	(0.15)	(0.11)	(0.10)

Table 8: Tax Capitalization Regressions With Alternative Econometric Methods
This table summarizes the tax capitalization coefficients δ using different estimation methodologies over the period from 1927 to 2004. The 30 stock portfolios are formed based on six dividend yield groups and five market capitalization groups. Excess returns are computed by subtracting the value-weighted market return from the portfolio return. Factor-adjusted returns are computed by subtracting the expected returns of the CAPM, the Fama-French, and the Carhart models from the portfolio return. The standard errors are summarized in parentheses. The significance levels are abbreviated with asterisks: One, two, and three asterisks denote significance at the 10, 5, and 1 percent level, respectively.

	Excess Return	CAPM	Fama French	Carhart
(1) Two-Stage Panel Regression with	1.11^{**}	1.29^{***}	1.75^{***}	1.37^{***}
Clustering (Base Case)	(0.51)	(0.44)	(0.29)	(0.25)
(2) Two-Stage Panel Regression with	0.18	1.96^{***}	2.22^{***}	1.77^{***}
Time-Fixed Effects and Clustering by Portfolio	(0.83)	(0.72)	(0.37)	(0.50)
(3) Two-Stage Prais-Winsten Regression with	1.13^{**}	1.29^{***}	1.75^{***}	1.37^{***}
Clustering	(0.56)	(0.45)	(0.30)	(0.25)
(4) Two-Stage Prais-Winsten Regression with	1.12^{*}	1.29^{***}	1.75^{***}	1.37^{***}
Panel-Corrected Standard Errors	(0.59)	(0.48)	(0.27)	(0.26)
(5) Two-Stage Fama-MacBeth with	0.44	2.70^{***}	2.04^{***}	1.64^{***}
Newey-West Standard Errors	(1.13)	(0.94)	(0.48)	(0.46)
(6) Two-Stage Fama-MacBeth by Month	1.09^{**}	1.26^{***}	1.74^{***}	1.36^{***}
Averaged over the 12 Months	(0.45)	(0.38)	(0.20)	(0.23)
(7) Simultaneous Panel Regression with	1.11^{**}	1.53^{***}	1.15^{***}	1.01^{***}
Clustering	(0.51)	(0.42)	(0.23)	(0.21)
(8) Simultaneous Panel Regression with	1.11^{**}	2.24^{***}	2.06^{***}	1.57^{***}
Constant Factor Loadings and Clustering	(0.51)	(0.42)	(0.31)	(0.31)
(9) Simultaneous Seemingly Unrelated	1.61^{***}	1.95^{***}	1.06^{***}	0.83^{***}
Regression	(0.31)	(0.27)	(0.17)	(0.17)

Table 9: Tax Capitalization Regressions: Dividend Yield vs. Tax Effect

This table summarizes the coefficients of the following regression: $\tilde{\alpha}_{k,t} = \beta_0 + \beta_1 \tau_t^{div} \times y_{k,t}^{div} + \beta_2 y_{k,t}^{div} + \beta_3 \tau_t^{div} + \epsilon_{k,t}$, where $\tilde{\alpha}_{k,t}$ is the abnormal return of portfolio k at time t, τ_t^{div} is the average marginal tax rate on dividends at time t, and $y_{k,t}^{div}$ is the dividend yield of portfolio k over the prior 12 months over the period from 1927 to 2004. Excess returns are computed by subtracting the value-weighted market return from the portfolio return. Factor-adjusted returns are computed by subtracting the expected returns of the CAPM, the Fama-French, and the Carhart models from the portfolio return. The factor loadings used to compute the abnormal returns are computed on a rolling basis using 60 months of return data. The 30 stock portfolios are formed based on six dividend yield groups and five size groups. The standard errors for the regression take into account clustering by time period and are summarized in parentheses. The significance levels are abbreviated with asterisks: One, two, and three asterisks denote significance at the 10, 5, and 1 percent level, respectively.

	Excess Return	CAPM	Fama-French	Carhart
Dividend Tax Rate × Dividend Yield	1.05	2.08	3.43***	2.67**
	(2.00)	(1.82)	(1.29)	(1.21)
Dividend Yield	−0.05	−0.09	−0.64	−0.48
	(0.79)	(0.73)	(0.52)	(0.50)
Dividend Tax Rate	0.19	−2.05*	−1.39**	−1.23***
	(1.27)	(1.20)	(0.54)	(0.48)
Constant	0.07	0.75	0.33	0.38**
	(0.49)	(0.47)	(0.21)	(0.18)
R-Squared (in Percent)	0.09	0.37	0.65	0.43

Table 10: Tax Capitalization Regressions Including Size and Book-to-Market in the Second Stage

This table summarizes the coefficients of the following regression: $\tilde{\alpha}_{k,t} = \gamma + \delta_0 \hat{\kappa}_{k,t} + \delta_1 z_{k,t}^{logmktcap} + \delta_2 z_{k,t}^{beme} + \delta_3 z_{k,t}^{mom} + \delta_4 \hat{\kappa}_{k,t} \times z_{k,t}^{logmktcap} + \delta_5 \hat{\kappa}_{k,t} \times z_{k,t}^{beme} + \delta_6 \hat{\kappa}_{k,t} \times z_{k,t}^{mom} + \epsilon_{k,t}$, where $\tilde{\alpha}_{k,t}$ is the abnormal return of portfolio k at time t, $\hat{\kappa}_{k,t}$ is the tax yield of portfolio k at time t, and $z_{k,t}^{logmktcap}$, $z_{k,t}^{beme}$, and $z_{k,t}^{mom}$ are the standardized logarithm of the market capitalization, the standardized book-to-market ratio, and the standardized prior one-year return of the stocks included in the respective portfolios over the period from 1927 to 2004. The variables are standardized by subtracting the mean across all portfolios at time t and by dividing the difference by the standard deviation of the corresponding variables at time t. Factor-adjusted returns are computed by subtracting the expected returns of the Carhart model from the portfolio return. The factor loadings are computed on a rolling basis using 60 months of return data. The standard errors take into account clustering by time period and are summarized in parentheses. The significance levels are abbreviated with asterisks: One, two, and three asterisks denote significance at the 10, 5, and 1 percent level, respectively.

	Dependent Variable: Abnormal Carhart Return		
Tax Yield	1.37***	1.40***	1.21***
	(0.25)	(0.25)	(0.24)
Size		−0.02	0.13***
		(0.02)	(0.04)
Book-to-Market		0.08***	0.17***
		(0.03)	(0.05)
Momentum		−0.01	−0.05
		(0.02)	(0.04)
Tax Yield × Size			−0.87***
			(0.18)
Tax Yield × Book-to-Market			−0.53*
			(0.28)
Tax Yield × Momentum			0.33*
			(0.19)
Constant	−0.11**	−0.12**	−0.08*
	(0.05)	(0.05)	(0.05)
R-Squared (in Percent)	0.34	0.44	0.63