RENE CAPITELLI

The Dividend Puzzle: Some Empirical Evidence on the Swiss Stock Market

1. Introduction

The influence of dividend payments on share prices is an important but still highly controversial issue. Starting in the fifties, many authors have advocated high payout ratios believing that investors generally prefer liberal dividends to niggardly ones. The reason put forward was that dividends are cash in hand while capital gains are at best "in the bush" [1]. Consequently, investors were thought to bid up the price of high yield securities relative to low yield securities. However, in their now 1961 paper, MILLER and MODIGLIANI prove that in a world where taxes and transaction costs are absent, corporations' dividend policy, given their investment and financing decisions, does not influence the market value of a firm. But in reality it may still be the case that capital gains are preferred to dividends because capital gains receive preferential tax treatment. Under these circumstances, dividends should have a positive influence on expected stock returns. In Switzerland, for example, capital gains by "natural persons" are not taxed on the federal level. With few exceptions (e.g. Bern and Jura) this also applies at the cantonal level. By contrast, dividends and interest receipts are taxed as ordinary income whereas interest payments on debts can be deducted from taxable income.
Several empirical studies have been conducted in the past in order to shed light on the actual influence of dividends on equity returns. Unfortunately, the results are conflicting in that each of the above hypotheses has been confirmed at least once. Thus, there is not only an urgent need for further empirical clarification but for new data as well. We conduct the present study for the case of Switzerland. This allows us not only to test the above hypotheses on a new sample but in addition permits us to avoid some difficulties contained in previous studies. The fact that in Switzerland dividend payments are generally known with certainty at the beginning of an ex-dividend month avoids information biases normally induced by multiple definitions of expected dividend yield [2]. In addition, the estimation procedure applied is superior to those used in previous studies.
In section 2 we give a short description of the model we plan to test. Data related aspects are presented in section 3. Section 4 contains the empirical results. Conclusions are drawn at the end. A more technical presentation of the estimation procedures is given in the appendix.

2. Theoretical Arguments

The theoretical model to be tested is a special after-tax version of the capital asset pricing framework developed by LITZENBERGER and RAMASWAMY (1979). It includes wealth and income related constraints on borrowing along with personal taxes on a progressive scheme. Taxes are only levied upon dividends and interest payments but not upon capital gains. The model contains N risky assets (shares) and one risk-free security paying a constant rate R_F. Dividends on risky assets are paid at the end of the period and are assumed to be known with certainty at the beginning of the period. LITZENBERGER and RAMASWAMY (1979, eq. 15) and CAPITELLI (1989, eq. 3.1) then derive the following relation for equilibrium expected returns, which forms the basis of the tests reported subsequently [3]

$$E(R_i) - R_F = a + b\beta_i + c(d_i - R_F) \qquad (1)$$

where:

R_i Total pre-tax rate of return on security i, which is equal to the ratio of the value of the security at the end of the period plus dividends over its current value, less one

R_F Rate of return on the risk free asset

β_i Market risk of security i defined as $Cov(R_i, R_m)/Var(R_m)$; R_m = return on the market portfolio

d_i Dividend yield on security i, which is equal to the dividend amount per share divided by the current price per share

E Mathematical expectation operator

a,b,c Parameters to be estimated

According to equation (1) the expected excess return on any asset, $E(R_i)-R_F$, is a linear function of its systematic risk, β_i, and its excess dividend yield, d_i-R_F. In addition, the following points are worth noting: Firstly, if there is a wealth constraint on borrowing and if this constraint is binding for some individuals, the constant term labelled a in equation (1) should be positive (see CAPITELLI (1989), p. 72). Secondly, given risk averse individuals, the expected excess return on any asset should increase whenever systematic risk goes up (b > 0 in equation (1)).

And thirdly, the most interesting aspect for our purpose is represented by the parameter c, which reflects the influence of dividends on asset returns. c should be negative if, on average, investors prefer dividends to capital gains. The reason is that investors require smaller expected returns on shares paying high dividends under these circumstances. The opposite is true (c positive) if, on average, those investors dominate that prefer capital gains to dividends for tax reasons. However, firms may adjust their dividend policy to the wishes of different groups of investors, thereby influencing the value of the parameter c. The influence of dividends on expected stock returns would even vanish, c = 0, if firms have to obey no restrictions on their dividend policies and are therefore able to fully adjust to their dividend clientèle [4].

3. Data Related Aspects

Data on Swiss security rates of return (R_{it}) were calculated from the monthly stock price and dividend series supplied by Data Stream Inc., London [5]. In 1986, for example, these price and dividend tapes cover roughly 230 Swiss firms whose shares are traded on a regular basis at the Zurich Stock Exchange. Before the actual calculations were carried out, several corrections were made. First, the exact ex-dividend dates of the stocks in our sample could not be obtained from the above-mentioned tapes. For the period January 1974 to July 1986 we therefore looked up all ex-dividend dates of each stock in the official price notations of the Zurich Stock Exchange and the federal tax office, respectively. Second, a lot of errors in the dividend series due to uncontrolled use of raw data had to be corrected. Having done this we then constructed the monthly market return by weighting the return of each security on the tape mentioned above by its relative market value. To calculate the value weights we used the reported numbers of outstanding shares of each firm by August 1986 [6]. As a proxy for the risk-free rate, R_F, we used the 1-month Euro-SFr rate.

The dividend yield variable was proxied by $d_{it} = D_{it}/P_{it-1}$, if security i went ex-dividend in month t, and set equal to zero otherwise. D_{it} is taxable dividend per share. This construction actually assumes that investors know at the end of each month whether or not the subsequent month is an ex-dividend month and if it is, how much of a dividend security i will pay. Although this assumption is generally questionable for other countries (e.g. USA), it closely matches the Swiss reality. More specifically, in Switzerland dividends are normally paid only once a year. Approximately 1 - 1 1/2 months before the annual meeting of the stockholders takes place, the public learns from the daily newspapers whether or not a firm intends to pay a dividend. In addition, the proposed amount of the dividend is known. At the stockholder's meeting itself the proposed dividend is practically never changed. Thus, in the case of Switzerland, there is actually no doubt that investors know D_{it} with certainty at the beginning of an ex-dividend month. Another point we wish to make is also strongly related to this institutional aspect. As MILLER and SCHOLES (1982) point out, the results obtained in earlier studies on the relationship between excess returns and dividends are very sensitive to the definition of expected dividend yield. In particular, unwanted information effects are frequently introduced by the various short-run measures of expected dividend yield. However, given the fact that in Switzerland dividends are known in advance, biases introduced by announcement effects can safely be ruled out.

4. Empirical Results

The results for equation (1) are given in table 1. Just for comparisons, estimates of a so-called before-tax version are presented in the first row. The dividend variable is neglected in this case. According to the before-tax model of table 1, the estimated coefficient a is seen to be positive but not significantly different from zero at the 10%-level (two-sided test). The same is true for parameter b, which also lacks significance. Applying only a one-sided test, the critical t-value at the 10%-level is 1.28. Thus, the coefficient a would in this case be significant and the t-value of b would practically coincide with the critical t-value. However, the sign of both parameters is plausible. Consistent with investors risk aversion, b is positive. A positive value of a can in principle be explained by a binding margin constraint on borrowing or by some misspecification of the market porfolio (see ROLL, 1977). If we then turn to the after-tax version, we see that the above results change dramatically. According to

Table 1: Estimates of Equation (1).

	a	b	c
Before-tax model	0.00365 (1.51)	0.00490 (1.28)	
After-tax model	-0.00485 (-1.61)	0.00882 (1.83)	0.24840 (4.11)

Notes:
The estimation period extends from January 1977 to December 1985. t-values are given in parentheses below the estimated coefficients.

the second row of table 1, the sign of a is now reversed. But based on a two-sided test, a is still not significantly different from zero at the 10%-level. By contrast to the before-tax model, b is now significantly positive and therefore agrees better with the theoretical arguments. In addition, a rather strong dividend effect is shown to exist ($c > 0$), whose coefficient is significantly different from zero even at the 1%-level. The magnitude of the coefficient indicates that for every Swiss franc of taxable return investors require nearly 25 centimes of additional before-tax return.

However, a word of caution is necessary. It is possible that the positive coefficient on dividend yield is not a tax-effect and that this effect will be reversed in subsequent months. For example, dividends may be interpreted as a signal for a temporary decline in the number of profitable projects available to the respective firm. It is also possible that a stock's dividend yield is a proxy for the covariance of its return with assets not included in the value weighted market index used to calculate the betas. But if the calculated betas do not measure the whole market risk, or to put it differently, if the measured influence of the dividend yield is entirely due to omitted assets, this effect should be exactly the same in ex-dividend and non-ex-dividend months. In order to test whether there is a reversal or an equivalent re-enforcing effect in non-ex-dividend months, the following cross-sectional regressions are run

$$R_{it}-R_{Ft} = a + b\beta_{it} + c[\delta d^*_{it}-R_{Ft}] + f[(1-\delta)d^*_{it}] + e_{it} \quad (2)$$

The variable δ is a dummy which is set equal to unity in ex-dividend months and zero otherwise. The variable d^*_{it} is defined as D_{it}/P_{it-1}, where D_{it} is the last (including t and going back at most 18 months) dividend paid per share i. D_{it} is equal to zero if no dividend is paid during the current and the preceding 18 months. The variable $(1-\delta)d^*_{it}$ is thought to pick up the dividend influence in non-ex-dividend months whereas the variable $(\delta d^*_{it}-R_{Ft})$ measures the dividend influence in ex-dividend months. Note that $(\delta d^*_{it}-R_{FT})$ is identical to our previous variable $(d_{it}-R_{Ft})$. e denotes the error term in equation (2).
The results of these regressions are reported in table 2. As can be seen from table 2, the estimates for a, b and c are practically identical to those of table 1. In addition, the estimated coefficient for f is positive, thus indicating that there is no reversal effect. Moreover, if the dividend influence in table 1 is entirely attributable to omitted assets, this effect should be the same in ex-dividend and non-ex-dividend months. This obviously implies an approximately equal size of c and f, which is not the case. The estimated parameter f is not only much smaller than c but also lacks statistical significance. Thus we conclude that there is also no equivalent re-enforcing effect, implying that it is in fact reasonable to interpret the previously measured dividend effect of table 1 as a permanent tax effect.

Table 2: Estimates of Equation (2).

a	b	c	f
-0.00582 (-1.48)	0.00972 (1.92)	0.27462 (3.16)	0.0272 (0.42)

Notes:
The estimation period extends from January 1977 to December 1985. t-values are given in parentheses below the estimated coefficients.

Finally, we also present some evidence on a possible clientèle effect, which provides a reason to include a further parameter in our cross-sectional regressions. Holding beta constant, we no longer assume that expected before-tax returns are an increasing linear function of dividend yield. On the contrary, the relation is non-linear and implies that compensation for tax-penalty of dividends may vary over firms. This idea is due to ELTON and GRUBER (1970) whose evidence on the ex-dividend behavior of common stocks suggests that the coefficient on excess dividend yield may be a decreasing function of yield. Given restrictions on short-sales of risky assets, the theoretical explanation for this effect is that investors in low (high) tax brackets are those who pay the highest price for the high (low) dividend assets. Thus short-sale restrictions may basically result in a segmentation of security holdings according to investors' tax brackets [7]. We implement this approach by assuming that the coefficient c in equation (1) is a linear decreasing function of the i-th security's dividend yield. The hypothesized structure can therefore be estimated by running the following cross-sectional regressions

$$R_{it}-R_{Ft} = a+b\beta_{it}+g_0(d_{it}-R_{Ft})+g_1[d_{it}(d_{it}-R_{Ft})]+e_{it} \quad (3)$$

where:

$$c_i = g_0-g_1 d_i, \qquad g_0,g_1 > 0, \qquad i = 1, \dots , N_t$$

As the results in table 3 show, g_0 and g_1 have the expected sign and are both significantly different from zero at the 5%-level (two-sided test). Moreover, b is close to the value it had in table 1 and is still significant at the 10%-level. Thus we conclude that these results are entirely consistent with the proposed clientèle-model. The magnitude of g_1

Table 3: Estimates of Equation (3).

a	b	g_0	g_1
-0.00244	0.007953	0.7311	-12.3839
(-0.76)	(1.67)	(3.28)	(-2.13)

Notes:
The estimation period extends from January 1977 to December 1985. t-values are given in parentheses below the estimated coefficients.

suggests that for every percentage point in yield the implied tax rate for ex-dividend months declines by 0.124. For example, if the annual yield was 4%, the implied tax rate would be 73.1-12.384*4 = 23.5%.

5. Summary and Conclusions

This paper has concentrated on the empirical relation between dividends, share prices and taxes in Switzerland. The theoretical model we actually tested is basically an after-tax version of the capital asset pricing framework. It includes wealth and income related constraints on borrowing along with personal taxes on a progressive scheme. The income constraint tends to mitigate the effects of personal taxes on the equilibrium structure of share prices. The equilibrium structure implies, that in the absence of any supply adjustment in corporate dividends, the expected before-tax return on any asset is not only an increasing linear function of its systematic risk but also of its dividend yield. However, given a world where dividends and interest payments are taxed as ordinary income, unrestricted supply adjustments by value maximizing firms would induce an equilibrium price structure that is in no way different from the structure obtained in a world where taxes are absent. In order to shed some light on these conflicting aspects, a number of empirical investigations have been conducted in the past. Yet, their empirical results are contradictory to such an extend that further clarification seemed necessary and is the primary aim of the present paper. In addition, we tried to counter some problems which were apparent in the previous studies. In particular two points are worth noting: First, the Swiss stock market was chosen in order to avoid information biases normally induced by several definitions of expected dividend yield. Second, unlike prior studies that used arbitrary grouping or instrumental variables to correct for measurement errors, we used the sample estimates of the variance of observed betas to arrive at consistent estimates of the coefficients in the relations tested. Since our correction is directly implied by the distributional assumptions made, the consistency of our estimates is established on a more rigourous theoretical basis than in previous studies.

The main findings can be summarized as follows:

1. Expected pre-tax returns are an increasing linear function of market risk as measured by beta. Shares involving a higher beta therefore require higher compensation.
2. Investors on the Swiss stock market want to be compensated for the tax disadvantages suffered in connection with dividend payments. In ex-dividend months, the coefficient of dividend yield is significantly positive and just short of 0.25. Hence, for each taxable dividend franc, investors demand 25 centimes compensation in the form of an additional pre-tax return.
3. The measured dividend effect is not a temporary phenomenon, nor can it be triggered by a misspecification of our market portfolio. In particular, neither a reversal nor a re-enforcing effect can be observed in non-ex-dividend months. The measured dividend effect may thus best be explained by means of permanent tax compensation.
4. Tax compensation is not identical for all firms. On account of the tax disadvantage dividend payments entail, high-yielding stocks ought to be mainly held by investors from low tax brackets. Since such investors demand a correspondingly low tax compensation, the coeffi-

cient of dividend yield, i.e. the implicit tax rate, should drop as dividends rise. This phenomenon can be shown to exist in the case of Switzerland. The implicit tax rate falls by about 12.5% for every percentage point of additional dividend yield. For an annual dividend yield of 4% the implicit tax rate for investors operating in this segment of the Swiss stock market will for example average 24%.

Our results have a number of important practical implications. Through its effect on expected pre-tax return, dividend policy generally affects our choice of discount rate for investment projects. But if dividend policy does change the required return, then the investment decision is not independent of the financing decision. The discount rate in that case must recognize the payout rate, a fact that sharply contrasts with the theorem of Modigliani and Miller.

Our findings also imply that a company may lower its expected pre-tax return and thus increase its market value by pursuing an appropriate dividend policy. This brings us to the apparent conclusion that the dividend policy adhered to by Swiss companies is suboptimal. However, it must be noted that, in the case of asymmetric information, dividend payments may function as a signaling and sorting devise. Similar to insurance companies endeavoring to recognize good and bad risks on the basis of the retention amount an insured person opts for, shareholders may use dividend payments as an indicator for sorting profitable from less profitable firms. To ensure that this selection mechanism functions smoothly, firms must not be able to profit from false signals. True signals evidently have their price. Given asymmetric information, a sub-maximal market value might therefore only reflect those tax costs which, from a rational standpoint, ought to be accepted in order to prevent insider trading or false signals. This explanation still leaves one problem open. The costs of this signal, approximately 25 centimes per taxable dividend franc, appear very high. The question therefore arises as to whether or not a less expensive mechanism could be found for truely signaling to shareholders the information lead enjoyed by insiders. Should this be the case, then the dividend policy of Swiss companies really is less than optimal.

Appendix. Estimation Procedure

To keep the technical side within reasonable bounds, we will only sketch roughly the estimation procedure applied. For a more stringent description the reader should consult CAPITELLI (1989). Normally, a pooled cross-section and time-series approach is used within a three-step procedure to estimate the parameters a, b and c of the equilibrium relation (1). By applying time-series regressions, step one consists of estimating the market risk parameter beta for all points in time and all securities under investigation. In step two, the equilibrium relation (1) is then calculated for all test months by means of cross-sectional regressions. Since the betas estimated during step one are needed in step two, there is obviously a danger of measurement errors distorting our results. We will therefore discuss briefly how to counter this danger. Two correction techniques are commonly used. The first consists of grouping all securities into various portfolios on the basis of a specific criterion. Instead of the individual securities, the different portfolios are then examined. This procedure is based on the finding that individual measurement errors frequently offset each other in the aggregate. Such error-induced distortions can hence be weakened by grouping securities into portfolios. However, this procedure replaces one problem by another. Whereas the danger of estimation results being distorted by measurement errors can thus be reduced, one loses at the same time the information on the level of individual securities. The second correction technique, which is based on an instrumental variable approach, is also plagued by shortcomings. A key problem is the fact that very little is really known about the factors which affect the market risk of a security. Choosing suitable instruments is therefore inevitably an arbitrary process.

For these reasons, we propose a new method of error correction in this study. Instead of resorting to an arbitrary grouping or instrumental variable technique, the estimated variances of observed betas are used directly for adjustment purposes. The consistency of our estimator in step two, i.e. the property of estimated coefficients to converge in probability to their true values, is thus a direct consequence of the distributional assumptions made and can be derived under more general conditions than those encountered in earlier studies of dividend policy. Finally, cross-sectional estimates are pooled in step three. Of course, the whole procedure is very time-consuming. This is clearly illustrated by the fact that, in our random sampling, overall estimators are based on around 16,000 regressions. However, the resultant improvements more than compensate for the relatively complex programming and calculation work needed.

Estimates of each security's beta, β_{it}, and its associated standard error are obtained from time series OLS regressions of the security excess return on the market excess return for 35 months prior to t,

$$R_{i\gamma}-R_{F\gamma} = \delta_{it} + \beta_{it} + (R_{m\gamma}-R_{F\gamma}) + \varepsilon_{i\gamma} \qquad (A.1)$$

$$\gamma = t\text{-}35, \ldots\ldots, t\text{-}1$$

This regression is repeated for all securities on the Data Stream tapes from t=1 (January 1977) until t=T=108 (December 1985). January 1977 is chosen as a starting date because the available series begin in January 1974 and we want roughly three years of observations for estimating the β_{it}'s and their standard errors. To conduct the cross-sectional regression for a given month t, e.g.

$$R_{it}-R_{Ft} = a + b\beta_{it} + c(d_{it}-R_{Ft}) + e_{it} \qquad (A.2)$$

$$i = 1,\ldots,N_t$$

we replace β_{it} in (A.2) by its estimate from (A.1) and transform all the variables in (A.2) by the estimated standard error of β_{it}. This transformation is necessary to construct a consistent cross-sectional estimator (as shown by CAPITELLI 1989, pp. 84ff).

Cross-sectional regressions based on the transformed variables in (A.2) provide a sequence of estimates for a, b and c denoted by a_t, b_t and c_t respectively, where t extends from 1 to T. Each of these estimates is derived by using between 127 and 172 firms on a given cross-section. Given that the cross-sectional estimates are stationary and serially independent, the overall estimates of a, b and c are obtained by an arithmetic average over 108 periods as discussed in CAPITELLI (1989, pp. 80).

The econometric procedure described above applies equally well to the two factor Capital Asset Pricing Model, as in BLACK, JENSEN and SCHOLES (1972) or FAMA and MACBETH (1973). Just for comparison, we therefore estimated first a before-tax version of the model, which may be written as

$$R_{it}-R_{Ft} = a + b\beta_{it} + e_{it} \qquad (A.3)$$

$$i = 1,\ldots,N_t;\ t = 1,\ldots,T$$

Footnotes

[1] See for example GRAHAM and DODD (1951) and GORDON (1959).

[2] See MILLER and SCHOLES (1982).

[3] Specifically, the following assumptions are made: Markets are informationally efficient, transactions costs are zero and returns are normally distributed. Investors have homogeneous expectations and maximize Von Neumann type utility functions.

[4] BLACK and SCHOLES (1974) provide a detailed analysis of this issue.

[5] All series are corrected for capital-increases, stock splits etc. The monthly rate of return for security i, R_{it} is defined as $R_{it}=(P_{it}-P_{it-1}+D_{it})/P_{it-1}$ where P_{it} is the price of security i at the end of month t and D_{it} is the dividend of security i in month t. D_{it} is zero if t is a non-ex-dividend month.

[6] The information is taken from Bank Vontobel, Research Reports, August 1986, Zürich.

[7] See LITZENBERGER and RAMASWAMY (1980) for a detailed theoretical model.

References

BLACK, F., M. JENSEN and M. SCHOLES (1972):" The capital asset pricing model: Some empirical tests", in: M. Jensen, ed. studies in the theory of capital markets, pp. 79-121, Praeger, New York.

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, pp. 1-22.

CAPITELLI, R. (1989): "Dividenden, Steuern und Aktienpreise: Theorie und empirische Ergebnisse für die Schweiz", Verlag Haupt, Bern.

ELTON, E. and M. GRUBER (1970): "Marginal stockholder tax rates and the clientèle effect", Review of Economics and Statistics 52, pp. 68-74.

FAMA, E.F. and J.C. MACBETH (1973): "Risk, return and Equilibrium: Empirical tests", Journal of Political Economy 71, pp. 607-636.

GORDON, M.J. (1959): "Dividends, earnings and stock prices", Review of Economics and Statistics 41, pp. 99-105.

GRAHAM, B. and D.L. DODD (1951): "Security analysis: Principles and techniques", 3rd ed., McGraw-Hill, New York.

LITZENBERGER, R.H. and K. RAMASWAMY (1979): "The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence", Journal of Financial Economics 7, pp. 163-195.

LITZENBERGER, R.H. and K. RAMASWAMY (1980): "Dividends, short selling restrictions, tax-induced investor clientèles and market equilibrium", Journal of Finance 35, pp. 469-482.

MILLER, M. and M. MODIGLIANI (1961): "Dividend policy, growth and the valuation of shares", Journal of Business 4, pp. 411-433.

MILLER, M. and M. SCHOLES (1982): "Dividends and taxes: Some empirical evidence", Journal of Political Economy 90, pp. 1118-1141.

ROLL, R. (1977): "A critique of the asset pricing theory's tests", Journal of Financial Economics 4, pp. 129-176.