Testing for short-termism in the UK stock market

David Miles

The views expressed are those of the author and not necessarily those of the Bank of England. The author is grateful to Elroy Dimson, Bahram Pesaran, Steve Satchell, Ron Smith, Steven Hall and seminar participants from the Universities of Exeter and Warwick for many helpful comments on earlier drafts. He would also like to thank the Bulletin Group for editorial assistance in producing this paper. The author is also with Birkbeck College, University of London.

Issued by the Economics Division, Bank of England, London, ECC2R 8AH to which requests for individual copies should be addressed; envelopes should be marked for the attention of Bulletin Group. (*Telephone*: *071-601-4030.*)

ISBN 1 85730 011 4
ISSN 0142–6753

Contents

Abstract

This paper uses data on the stock market valuations of a large sample of UK companies to assess if that market displays short-termism. Tests are undertaken of whether discount rates, implicit in market valuations, applied to cash flows which accrue in the longer term are too high, both absolutely and relative to the rates applied to cash flows in the near term. We find *prima facie* evidence that these longer-term discount rates are too high, a result consistent with the existence of short-termism.

Testing for short-termism in the UK stock market

Introduction

The hypothesis that in the Anglo-Saxon economies the interaction of financial markets with managerial decision making results in a sub-optimal level of long-term investment has been discussed by numerous academics, politicians, industrialists and financiers. The claim that "short-termism" exists has been made, often stridently, by amongst others Lawson (1986), Greenspan (1990), Law (1986) and Cosh et al (1990). Here is former British Chancellor of the Exchequer Nigel Lawson:

> *"The big institutional investors nowadays increasingly react to short-term pressure on investment performance . . .* (and) *. . . are unwilling to countenance long term investment"* [quoted in Nickell and Wadhwani (1987)].

Counter claims have been made by Jensen (1986 and 1989), Healy et al, (1989), T.Boone Pickens (1986), Woolridge (1988) and, most recently, in an elegantly written piece by Marsh (1990). The counter attack has been robust. Pickens, for example, does not mince his words:

> *"Increasing acceptance of the short-term theory . . . has freed executives to scorn any shareholders they choose to identify as short-termers. Executives aim their contempt not only at the initiators of takeover attempts but at the arbitrageurs and institutional investors who frequently trade in and out of stocks".*

Theoretical models, generally based on the existence of asymmetric information between managers and owners of corporations, have been developed in which optimising agents rationally behave in a way which results in positive net present value projects with relatively long maturity systematically being rejected [see Webb (1991), Narayanan (1985), Schliefer and Vishny (1990), Stein (1989), Froot et al (1990)]. But empirical evidence relevant to the question of whether short-termism exists is noticeably thin on the ground. Nickell and Wadhwani (1987) undertook an empirical study of stock market valuations of UK companies and argued that their results showed that cash flows which accrue further in the future are discounted at sub-optimally high rates. This is the only empirical study which directly addresses the issue as to whether stock market valuations of UK corporations are "short term". But Marsh (1990) and Cooper and Marsh (1987) argue that the study is seriously flawed.

Evidence quoted more recently [eg Marsh (1990)] by opponents of the short-termism hypothesis—that stock market prices of companies frequently rise on announcements of new investment projects and that share prices of nearly all companies are several time current earnings per share—are not convincing. Perhaps in an efficient market share prices of UK stocks would rise even more than they do on investment announcements and P/E multiples be far higher.

The aim of this paper is to directly test for short-termism in the UK stock market by assessing whether the discount rates implicit in market valuations applied to cash flows which accrue in the longer term are too high, both absolutely and relative to the rates applied to cash flows in the near term. In the first part of the paper the link between excessively high discount rates for longer-term projects and popular conceptions of "short-termism" is analysed. In Section II some simple models, in each of which versions of the short-termism hypothesis is nested, are developed and econometric issues in the estimation of those models are addressed. Section III presents empirical results. Data sources are described in the Appendix.

Section I What is short-termism?

It is clear that all contributors to the debate agree that an implication of short-termism is that there is insufficient investment in projects which have relatively long maturities—some types of project with positive expected net present value when discounted appropriately both for risk and the pure rate of time preference of investors are systematically rejected. But there are several quite different hypotheses about how such a situation might persist. It is instructive to distinguish between theories which explain the phenomenon in terms of mis-pricing of companies by investors—ie mistaken perceptions by investors either of cash flows or of the level of risk involved in long-term projects—and those which do not rely on inappropriate valuations of projects by outsiders.

The simplest theory, perhaps story is a better word since generally no formal model is presented by its advocates, is that the stock market simply undervalues all long-term projects—either by systematically underestimating net cash flows for the longer term or by systematically applying inappropriately (to the risk involved) high discount rates to rational expectations of those cash flows. If this phenomenon exists it is not hard to tell a plausible story as to why long-term investment is suboptimally low so long as there is a link between managerial decisions on projects and stock market valuations. Such a link is, in turn, plausible—at least in the United Kingdom and United States—so long as the threat of hostile takeover, which I take to be against the interest of incumbent management, depends upon the current share price. The existence of remuneration packages for managers which depend upon stock market valuations (more common in the United States than the United Kingdom) reinforces the link between self-interested managers' decisions on project appraisal and the current share price.

In passing we should note that the *source* of the mis-allocation of resources here is not the existence of hostile takeovers *per se*, which have been the focus of much managerial criticism [see Cosh et al (1990) and Law (1986)], but the (unexplained) mis-pricing of corporate securities. Furthermore, it is a specific form of mis-pricing which causes the problem; empirical tests designed to discover *general* forms of mis-pricing which may be inconsistent with the efficient markets hypothesis do not usually throw much light on the validity of this version of the short-termism hypothesis. The massive literature on the excess volatility of stock prices [see Shiller (1990) for an excellent account of the literature and for many of the seminal contributions] tells us little about short-termism because what is being tested is departures, in either direction, from so called fundamental value. Share prices can be excessively volatile if they oscillate between dramatic under and over-valuations yet it is unclear whether this has any implication for the relative incentives of managers to undertake short-term as opposed to long-term investments. In short, much of the literature on stock market efficiency and excess volatility has not addressed the specific form of mis-pricing which would generate under-investment in projects with long pay-off periods.

I noted above that the simple story of stock market mis-pricing of long term projects lacked a theoretical underpinning. There are, however, several models with impeccable theoretical credentials which do generate under-investment in long term projects [see, for example, Webb (1991), Stein (1989), Schliefer and Vishny (1990), Narayanan (1985)]. Yet these models do not generally imply that stock market valuations are irrational [note the title of Stein's (1989) paper]. The flavour of this class of model is captured with a simple story.

Suppose managers of companies have better information about the profitability of potential investment projects than do outsiders. Suppose also that outsiders, not surprisingly, assess the efficiency of managers by studying the ex-post profitability of projects which the managers have initiated. Finally, assume that managers rationally expect that their period of office is limited—perhaps to no more than the next three to five years—either because they will feel like a change of job or else retire or be sacked due to some misfortune unrelated to the projects they are currently assessing. (Given the average period of office of managing directors of UK corporations this final assumption is surely not unacceptable.) One implication of this set of assumptions is that self-interested managers might systematically reject long-term projects with positive net present values in favour of less profitable projects which generate positive cash flows in the near term [see Narayanan (1985) for a formal statement]. The important thing about this fable is that the stock market rationally values all projects conditional on the information available to it. The problem is the asymmetry of information between owners and managers and the (perfectly rational) limited time horizon of the latter. The policy response to this form of market failure is unclear; tying managers to companies for longer periods certainly ameliorates the direct problem though it generates some fairly obvious other difficulties and in practice may be in the interests of neither owners nor managers. Reducing asymmetries of information—by improving the transparency of company accounts, by increasing the number of non-executive directors on boards, by encouraging more active monitoring by insititutional investors—may be beneficial. But in this type of model it is often in the interests of *both* good managers and owners to reduce information asymmetries so their continued existence suggests real costs to further improving the flow of information. (Stein's 1989 paper shows how the asymmetry of information leads to outcomes analogous to the prisoner's dilemma; everyone could be better off if the information asymmetries were removed. So if asymmetries of information persist in such a world it must be because they are hard to eliminate.)

Even though the class of models where short-termism is due to information asymmetries do not usually imply irrational valuation of companies by outsiders (indeed most of the theoretical work *assumes* rational expectations on the part of investors) there are empirical implications which are common to these models and the simpler story based on systematic, but irrational, market under-valuation of long-term cash flows. For example, the Schleifer and Vishny (1990) paper depends upon asymmetric information between managers and owners and has a small group of arbitrageurs who aim to take advantage of any mis-pricing of firms. They show that risk-averse managers will prefer short-term projects so long as arbitrageurs have limited funds and that mis-pricing of long-term projects by the market is only slowly removed. There is no irrationality in the model—merely different classes of potential investors with varying access to, and abilities to assess, information. Yet the result is that managers tend to avoid long-term projects unless they are especially profitable because, effectively, the risk premium attached to longer-term investment is high. Those long-term investments which are undertaken tend to be highly profitable to offset the disinclination of risk averse managers to undertake investments where the probability of mis-pricing is high due to the thinness of arbitrage for projects with longer time horizons. Thus, long-term projects which *are* undertaken are, *ex post*, highly profitable. A similar argument can be used to show that Narayanan's (1985) model of rational shareholders and managers implies that those longer-term investments that are undertaken are highly profitable, though undervalued by the stock market. This implication of many models of rational behaviour with asymmetric information is the same as that generated by a simpler (non-theoretical) set up simply based on the assumption of irrational undervaluation of longer-term cash flows by the market. This irrationality prompts self-interested managers only to undertake long-term projects if they do not significantly depress the share price—which implies that acceptable long-term projects need to be highly profitable.

In the next section a strategy is outlined for testing whether projects with long-term cash flows do need to be highly (and excessively) profitable to prevent share prices falling. I take this to be an implications of short-termism and one which implies that the discount rates applied by the market to longer-term cash flows are, other things equal, higher than those applied to cash flows accruing in the near term.

Section II Assessing discount rates applied to cash flows

We aim to use stock market and company accounts data on a large sample of firms to assess the discounts implicit in stock market valuations which are applied to (expected) future cash flows which accrue at different horizons. We first develop some simple relations between stock prices and cash flows in a world with no short-termism, then generalise the model to allow for myopia.

Denote the required, or expected, one period return on the shares of company j during period t by $E(R_j)_t$. (Expectations are formed at the end of t-1). Most models of equilibrium required returns—the CAPM [Sharpe (1963, 1964), Lintner (1965) and Mossin (1966)], the Arbitrage Pricing Theory [Ross (1976)] and the consumption CAPM [Lucas (1978)]—write the required return as the sum of a risk-free return and a risk premium:

$$E(R_j)_t = R_{f_t} + \pi_{jt} \tag{1}$$

Where: R_{f_t} is the riskless return that can be earned in period t (which is not company specific) and π_{jt} is the risk premium for period t (which is company specific).

Equilibrium models of risk (CAPM etc) are derived from optimisation problems where investors do not display short-termism; they do, of course, have a rate of pure time preference which is reflected by $Rf_t > 0$, but that is a different thing. If markets do not display short-termism, and we use a measure of risk premia based on a correctly specified equilibrium asset pricing model, we should find that actual returns only differ randomly from expected returns. Ignoring (for the moment) tax effects actual returns are

$$R_{jt} = \frac{P_{j,t+1} + D_{j,t+1} - P_{j,t}}{P_{j,t}} \tag{2}$$

Where P_{jt+1} is the share price of firm j at the beginning of period t+1; D_{jt+1} are dividends per share paid in period (t+1). We assume here that dividends are paid at the start of the period. Assuming an efficient market we can write

$$R_{j_t} = E(R_{j_t}) + \varepsilon_{jt} = \frac{E\,(P_{j,t+1} + D_{j,t+1})}{P_{jt}} - 1 + \varepsilon_{jt} \tag{3}$$

where ε_{jt} is a forecast error which is uncorrelated with $E(R_{jt})$. Using (1) in (3) and re-arranging:

$$P_{jt} = \frac{E(P_{j,t+1}) + E(D_{j,t+1})}{1 + R_{f_t} + \pi_{jt}} \tag{4}$$

Assuming future risk premia and risk-free rates are known and solving (4) forward yields:

$$P_{jt} = \sum_{i=1}^{N} \frac{E(D_{jt+i})}{\prod_{k=0}^{i-1}(1 + Rf_{t+k} + \pi_{jt+k})} + \frac{E(P_{j,t+N})}{\prod_{k=0}^{N-1}(1 + Rf_{t+k} + \pi_{jt+k})} \tag{5}$$

In (5) all expectations are dated t.

We now make two simplifications to (5).

First, we assume that the company specific risk premium, π_j, is expected to be constant across time periods. This is a strong assumption, but one which is not implausible. Asset pricing theories generally imply that risk premia depend upon fundamental characteristics of companies—covariability of returns with those of other companies (the CAPM); sensitivity of earnings to macro and micro shocks (the APT); covariability of returns with consumption (the consumption CAPM). It seems reasonable to think that investors would not predict significant changes in these fundamental characteristics to occur in the future.

In the case of the CAPM each company's risk premium is a function of its βeta. Since the average βeta of a sample of companies who comprise a high proportion of the market must be close to unity it is not possible for those βetas to systematically follow a trend; on average betas must be constant. An important point to note is that the assumption that company risk premia in equation (5) are expected to remain constant does not imply that cash flows which accrue far in the future are no more risky than those that accrue in the near term. Risk premia are measured at an annual rate and are compounded. If, for example, the risk premium for a company were 10% and the safe rate were zero the discount factor applied to expected cash flows accruing in ten years is $1/(1.1)^{10}$ while that for five year flows is $1/(1.1)^5$; expected cash flows on ten year projects would need to be over 60% higher than on five year projects to compensate for risk. In short the assumption of constant risk premia makes the perceived risk for different expected cash flows proportional to the time until those flows accrue. [Merton (1973) shows that proportionality of risk premium to the time horizon of a project follows when cash flows follow a diffusion process with the conditional variance of flows proportional to their maturity.]

Second, we assume that the appropriate nominal risk-free return (expressed at a per period rate) to apply (at t) to cash flows accruing at different horizons is equal to the redemption yield on risk-free bonds which mature at those horizons. We denote the period t yield to maturity on a bond maturing at $t+k$ by $r_{t,t+k}$. Note that $r_{t,t+k}$ is known with certainty at t. But the legitimacy of using a default-free and known nominal yield to proxy the "safe return" is dubious given obvious uncertainty over future price inflation. Within the framework of the CAPM Friend et al (1976) have developed a model of equilibrium risk premia with uncertain inflation and where the only known return available is fixed in nominal terms. We return to this important point below when the problems of applying one-period asset pricing models with no uncertainty over inflation are considered in greater detail. [For a formal treatment of the changes which need to be made to one-period asset pricing models to be consistent with valuations of returns over many periods see Fama (1970), Merton (1973) and Stapleton and Subrahymanyam (1977).]

Making these two assumptions allows us to simplify equation (5) substantially and to write:

$$P_{jt} = \sum_{i=1}^{N} \frac{E(D_{j,t+i})}{(1+r_{t,t+i}+\pi_j)^i} + \frac{E(P_{j,t+N})}{(1+r_{t,t+N}+\pi_j)^N} \tag{6}$$

Finally, we allow for the impact of taxes. In the UK dividends paid to shareholders have already been taxed at the company level; corporations pay Advance Corporation Tax (ACT) on dividends. Shareholders effectively reclaim the tax paid on their behalf by the corporation and then pay income tax at their marginal rate. Denoting the imputation rate by s and the marginal tax rate of a shareholder by m a dividend of D paid by a company is worth $D(1\text{-}m)/(1\text{-}s)$ post tax.

Shareholders differ with respect to their marginal income tax rate; hence a given per share dividend paid by a company is worth different amounts to different shareholders. There is no firmly grounded theory of how corporations with given future expected dividends are valued in such a world [see King (1977)]. We follow the rather *ad hoc* procedure of using the economy-wide weighted *average* of the post tax valuations of corporate dividend payments to assess the fundamental worth of the company. Denoting the weighted average of $(1\text{-}m)/(1\text{-}s)$ by $(1\text{-}\bar{m})/(1\text{-}\bar{s})$ we now have our fundamental (no-myopia) share price equation:

$$P_{jt} = \sum_{i=1}^{N} \frac{\frac{(1-\bar{m})}{(1-\bar{s})}.E(D_{j,t+i})}{(1+r_{t,t+i}+\pi_j)^i} + \frac{E(P_{j,t+N})}{(1+r_{t,t+N}+\pi_j)^N} \tag{7}$$

Note that we assume that the tax rates applied to dividends at various horizons is constant. This assumption is only valid if at time t shareholders expect the ratio of their marginal income tax rate to the basic rate to be constant.

Note also that we have implicitly assumed that the effective capital gains tax rate is negligible. Since 1981 personal shareholders have only been taxed on real capital gains. Furthermore, shareholders are only taxed on net realised gains. With a generous tax free allowance (£5,000 per annum) and the ability to control the timing of net real gains it is unlikely that persons face a significant rate of capital gains tax. [See King and Fullerton, (1984, page 35.)] Pension funds and (since 1980) authorised unit and investment trusts are exempt from capital gains tax. Insurance companies, in principle, pay capital gains though Kay and King (1990) argue that '. . . *because of the Byzantine complexity of . . . regulations, the amounts of tax paid in practice by life insurance companies are not very large*' (page 82).

Three versions of the short-termism hypothesis can be considered by generalising equation (7). The first version identifies short-termism with excessively high discounts applied to cash flows accruing in the more distant future. One way to capture this is to allow the rate of discount actually applied to cash flows in, say, year 5 to exceed $(1+r_{t,t+5}+\pi_j)^5$. Suppose the actual discount is twice as high—ie is a discount rate more appropriate in an efficient market (assuming an approximately flat yield curve) to a cash flow accruing 10 years in the future. That discount is $(1+r_{t,t+5}+\pi_j)^{2 \times 5}$.

If the degree of excess discounting is consistent for all future cash flows we could generalise (7) to:

$$P_{jt} = \sum_{i=1}^{N} \frac{\frac{(1-\bar{m})}{(1-\bar{s})} . E(D_{j,t+i})}{(1+r_{t,t+i}+\pi_j)^{b.(i)}} + \frac{E(P_{j,t+N})}{(1+r_{t,t+N}+\pi_j)^{b.(N)}} \qquad (8)$$

Where b reflects the degree of short-termism. If $b = 1$ (8) is equivalent to (7). If $b = 2$ cash flows in year 1 are discounted at a rate more appropriate to flows in year 2, cash flows in year 5 are discounted at a rate more appropriate to cash flows in year 10 etc. Thus cash flows which accrue further in the future are discounted by amounts which are excessive to a greater degree the further in the future they accrue. [Cash flows accruing six months in the future are discounted such that the degree of impatience, measured as the calendar time between the actual cash flow and the appropriate time for that particular discount rate, is $(b-1)$. 6 months; for a 5 year cash flow the degree of impatience is $(b-1)$. 60 months—10 times as great.]

The hypothesis to test so as to assess whether short-termism in this sense exists is $b > 1$.

A rather different version of the short-termism hypothesis focuses upon excessively pessimistic forecasts of cash flows associated with projects with long time horizons. If we interpret equation (7) as our fundamental, no short-termism, pricing equation then we are assuming that the expectations involved are rational. Suppose, instead, that expectations differ systematically from these in the following way: expectations of cash flows for year t + i are only $(x)^i$ as great as the rational expectation, where $x < 1$. *If* x = 0.9 then projected cash flows for six months ahead are 95% of their rational expectation, whilst expectations for year 5 are only 59% of their non-myopic value. Allowing for this form of short-termism equation (7) becomes:

$$P_{jt} = \sum_{i=1}^{N} \frac{\frac{(1-\overline{m})}{(1-\overline{s})}.E(D_{j,t+i}).x^i}{(1+r_{t,t+i}+\pi_j)^i} + \frac{E(P_{j,t+N}).x^N}{(1+r_{t,t+N}+\pi_j)^N} \tag{9}$$

The hypothesis to test so as to assess whether short-termism in this sense exists is $x < 1$.

Note that it may be hard to distinguish between the two forms of short-termism. If there is little variation across companies in risk premia (π_j) *and* if the yield curve is relatively flat we can write (**8**) as:

$$P_{jt} = \sum_{i=1}^{N} \frac{\frac{(1-\overline{m})}{(1-\overline{s})}.E(D_{j,t+i})}{(1+\overline{r}+\overline{\pi})^{b.(i)}} + \frac{E(P_{j,t+N})}{(1+\overline{r}+\overline{\pi})^{b.(N)}} + \text{Second Order Terms} \tag{10}$$

Where $\bar{r}$ is the average level of the yield curve between t + 1 and t + N and π is the average risk premium across companies.

Note that (**10**) can be written:

$$P_{jt} = \sum_{i=1}^{N} \frac{\frac{(1-\overline{m})}{(1-\overline{s})}.E(D_{j,t+i})[(1+\overline{r}+\overline{\pi})^{1-b}]^i}{(1+\overline{r}+\overline{\pi})^i} + \frac{E(P_{j,t+N})[(1+\overline{r}+\overline{\pi})^{1-b}]^N}{(1+\overline{r}+\overline{\pi})^N} \tag{11}$$

Which is identical in form to (9) with x = $(1+\overline{r}+\overline{\pi})^{1-b}$. Thus if $b > 1$ we would find $x < 1$ (assuming $\overline{r}+\overline{\pi} > 0$).

In the empirical section of the paper both tests for short-termism are undertaken.

A third set of tests is applied to see if markets discriminate in a cruder way between short-term and long-term cash flows. Suppose that cash flows accruing more than, say, five years in the future are discounted more heavily than cash flows accruing over the first five years. There are several ways this might be modelled. We could allow the discount rate on cash flows for years five, six, . . . to be higher by an amount a_o. Letting the last dividend in equation (7) be $D_{j,t+5}$ we could then write this hypothesis as:

$$P_{jt} = \sum_{i=1}^{5} \frac{\frac{(1-\overline{m})}{(1-\overline{s})} E(D_{j,t+i})}{(1+r_{t,t+i}+\pi_j)^i} + \frac{E(P_{j,t+5})}{(1+r_{t,t+5}+\pi_j+a_o)^5} \quad \textbf{(12)}$$

And test for $a_o > 0$.

Two alternatives are to estimate

$$P_{jt} = \sum_{i=1}^{5} \frac{\frac{(1-\overline{m})}{(1-\overline{s})} E(D_{j,t+i})}{(1+r_{t,t+i}+\pi_j)^i} + \frac{E(P_{j,t+5})}{(1+r_{t,t+i}+\pi_j)^{\alpha.5}} \quad \textbf{(13)}$$

and

$$P_{jt} = \sum_{i=1}^{5} \frac{\frac{(1-\overline{m})}{(1-\overline{s})} E(D_{j,t+i})}{(1+r_{t,t+i}+\pi_j)^i} + \frac{E(P_{j,t+5})\lambda}{(1+r_{t,t+i}+\pi_j)^5} \quad \textbf{(14)}$$

In **(13)** excess discounting of longer term cash flows implies $\alpha > 1$. In **(14)** excess discounting implies $\lambda < 1$. Below we report the results of estimating **(12)**, **(13)** and **(14)**.

Notice that all these tests for short-termism are conditional on a particular model of risk ie of π_j. This is inescapable. Any results from estimation of equations (8) or (9), that is any estimated values of x and b, can be "explained" in terms of a time-varying risk premia. But it is only a rather special form of time-varying risk that could account for the finding that x (b) were significantly less than (greater than) 1; namely risk premia that rose sharply through time. As noted above, theoretical asset pricing models imply that risk premia depend upon covariability of company returns with other variables—returns on other assets, consumption or fundamental macro shocks. It is hard to think of a plausible model where for companies *as a whole* there is a systematic tendency for the relevant covariability of returns to rise over time.

We attempt to model company specific risk premia in the following way. We take the standard CAPM, where π_j is equal to the βeta of company j multiplied by the expected excess return on a diversified portfolio over the risk free rate, and make two amendments. First, as noted above, uncertainty over inflation implies that a known nominal yield is an inappropriate proxy for the "safe rate". In an annex to this paper the CAPM with uncertain inflation, as developed by Friend et al (1976), is described. It is shown there that one result of allowing for uncertainty over price inflation is that in an equation for π_j the coefficient on the company j βeta may well be negative, a result which implies that the more highly correlated with the average returns on other risky assets are company j returns the lower is risk *relative* to an asset with fixed nominal returns. If a diversified portfolio of risky assets (eg equities) are a good hedge against inflation this result would not be surprising.

The second amendment to the standard CAPM draws on Merton's (1973) dynamic model of asset pricing. Merton shows that uncertainty over future interest rates—perhaps the most obvious example of a changing investment opportunity set over time and one which is sufficient to invalidate the simple one-period CAPM—has an important impact on risk premia. His equation (34) (p882) shows that the simple CAPM must be amended to take account of correlation between the returns on company j and changes in interest rates. Since that correlation will depend upon the level of debt held by company j we should expect debt gearing to have a positive impact on π_j.

Our empirical model of risk premia is:

$$\pi_j = a_1 \cdot \beta_j + a_2 Z_j$$

Where a_1 and a_2, are co-efficients; β_j is the measured βeta for company j and Z_j is company j's level of debt gearing. All variables are measured at time t (subscript omitted). If the simple CAPM is valid we should find the co-efficient a_1 is equal to the difference between the expected return on the market and the safe rate; a_2 will be insignificant. If uncertainty over inflation matters and equities are perceived to be a good inflation hedge a_1 may be negative. If uncertainty over future interest rates is important a_2 might be positive.

As a check against failure to allow for time varying risk premia we estimate our basic specifications with covariability of company returns (betas) measured over various time periods.

Econometric and data issues

Data on a sample of 477 UK non-financial firms was used to estimate equations (7)–(9) and equations (12)–(14). The sample was chosen from the EXSTAT data base of company accounts, which also includes data on the end year company share price. All non-financial firms which had reported over the period 1975–89 were selected. The sample accounts for around one half of the market value of all UK quoted companies. (Details on the sample of companies are provided in the data appendix.)

For each of the nine years from 1980 to 1988 a different set of cross section regressions was estimated corresponding to the various specifications of short-termism.

By estimating models (7), (8) and (9) for our 477 companies in each of nine years we gain two benefits. First, we are able to gauge whether the degree of short-termism in the stock market changed over the 1980s. Second, we avoid the problem of our results depending upon a particular year when share prices in general may have been depressed (the early 1980s) or unsustainably high (1986 and much of 1987).

To illustrate exactly how the models were estimated we will describe in detail the specifications for share prices in 1984, the middle year in our sample. For that year non-linear cross section regressions using each companies share price at end-1984 as the dependent variable were estimated. The explanatory variables included (rational) expectations of future dividends and the future share price at end-1989, and current dated variables relevant to assessing company risk (βeta, gearing). The three specifications corresponding to (7), (8) and (9) which were estimated (omitting for the moment tax factors) were as follows:

$$P_{j,84} = \frac{D_{j,85}}{(1+r_{85}+a_1\beta_j+a_2Z_j)} + \frac{D_{j,86}}{(1+r_{86}+a_1\beta_j+a_2Z_j)^2} + \frac{D_{j,87}}{(1+r_{87}+a_1\beta_j+a_2Z_j)^3}$$
$$+ \frac{D_{j,88}}{(1+r_{88}+a_1\beta_j+a_2Z_j)^4} + \frac{D_{j,89}}{(1+r_{89}+a_1\beta_j+a_2Z_j)^5} + \frac{P_{j,89}}{(1+r_{89}+a_1\beta_j+a_2Z_j)^5} \tag{15}$$

$$P_{j,84} = \frac{D_{j,85}}{(1+r_{85}+a_1\beta_j+a_2Z_j)^b} + \frac{D_{j,86}}{(1+r_{86}+a_1\beta_j+a_2Z_j)^{2b}} + \frac{D_{j,87}}{(1+r_{87}+a_1\beta_j+a_2Z_j)^{3b}}$$
$$+ \frac{D_{j,88}}{(1+r_{88}+a_1\beta_j+a_2Z_j)^{4b}} + \frac{D_{j,89}}{(1+r_{89}+a_1\beta_j+a_2Z_j)^{5b}} + \frac{P_{j,89}}{(1+r_{89}+a_1\beta_j+a_2Z_j)^{5b}} \tag{16}$$

$$P_{j,84} = \frac{(D_{j,85}).x}{(1+r_{85}+a_1\beta_j+a_2Z_j)} + \frac{(D_{j,86}).x^2}{(1+r_{86}+a_1\beta_j+a_2Z_j)^2} + \frac{(D_{j,87}).x^3}{(1+r_{87}+a_1\beta_j+a_2Z_j)^3}$$
$$+ \frac{(D_{j,88}).x^4}{(1+r_{88}+a_1\beta_j+a_2Z_j)^4} + \frac{(D_{j,89}).x^5}{(1+r_{89}+a_1\beta_j+a_2Z_j)^5} + \frac{(D_{j,89}).x^5}{(1+r_{89}+a_1\beta_j+a_2Z_j)^5} \tag{17}$$

Where:

$P_{j,84}$ is the share price of the *j*th firm at the company report date in 1984; (Note: companies reporting in 1985 before mid-April are counted as 1984 reporters.)

$D_{j,85}$ is the dividend per share paid in the financial year 1985 (etc);

$P_{j,89}$ is the share price at the company 1989 report date (Note: companies reporting up to mid-April 1990 are counted as 1989 reporters);

β_j is the company βeta;

Z_j is the ratio of debt outstanding to the market value of the firm (gearing);

r_{85}, r_{86}, r_{87} . . . are the average yields to maturity at end-1984 on all UK Government bonds with 1, 2, 3 . . . years to maturity; and

a_1, a_2, b and x are parameters to be estimated.

For years before and after 1984 an identical procedure was followed and the same parameters estimated. In each case the level of company gearing, the share price and the yields to maturity of UK government bonds were all measured at the appropriate dates. As the dependent variables in these regressions are prices measured at various years the number of dividend terms appearing on the right hand side varies across regressions as we use a fixed end point of 1989. In the final year used to estimate the specifications—1988—there is only one dividend term, D_{1989}; in the early years there are up to ten different cash flows.

We use the McCullum (1976) technique to account for the expectations terms in **(7)**, **(8)** and **(9)** [see also Wickens (1982)]. More specifically, we use the actual values of dividends per share paid in each year, and the actual share price at end-1989, and estimate **(15)**, **(16)** and **(17)** by non-linear, two stage least squares. [The non-linear 2SL2 option in LIMDEP, version 6.0, was used; see Amemiya (1985) pages 245–49, for details of the procedure.] For each company future dividends and the company end-period share price were instrumented using company-specific lagged share prices, lagged dividends per share and lagged earnings per share. Five lags of each variable were used.

McCullum (op cit) and Wickens (op cit) analyse the consistency of the instrumental variables estimator in time series models with expectations of future variables. Equations **(15)**, **(16)** and **(17)** are cross section regressions. The difference is important. We cannot assume, for example, that expectations of share prices in 1989 formed in 1984 or in 1986 are, on average, correct because unexpected macroeconomic factors arising between 1984 and 1989 or between 1986 and 1989 may have made overall share prices in 1989 higher (or lower) than could rationally have been expected five or three years earlier; in time series regressions the rational expectations hypothesis is invoked to justify the assumption that on average these forecast errors are close to zero. The consistency of the IV estimator in our cross section regressions nevertheless holds provided a constant is included in the specifications and that there is no correlation between the observable company specific variables at each point and the deviation of the 1989 share price (and future dividends) from expected values *once account is taken of the overall unexpected performance of the economy.* To analyse this condition more formally and to show the importance of a constant in our specifications consider a highly simplified version of the model. Assume that no dividends are paid and that appropriate discount rates (π_j's) are known. For the 1984 regressions equation (9) can then be written

$$P_{j,84} = \frac{E(P_{j,89}).x}{(1+\pi_j)^5} \tag{18}$$

where x reflects the extent of short-termism and is the only parameter to be estimated. The expected share prices, at 1984, are unobservables. We can write

$$E(P_{j,89}) = P_{j,89} - u_{j,89} \tag{19}$$

where $u_{j,89}$ is the unobservable, *ex-post* forecast error for company *j*'s share price. (18) and (19) imply

$$P_{j,84} = \frac{(P_{j,89}).x}{(1+\pi_j)^5} - \frac{(u_{j,89}).x}{(1+\pi_j)^5} \tag{20}$$

Define $\bar{U}_{89}$ as the average *ex-post* forecast error across all *n* companies.

$$ie \quad \sum_{j=1}^{n} \frac{(u_j, 89)}{n} = \bar{U}_{89}$$

Thus $\bar{U}_{89} = \bar{P}_{j,89} - \overline{E(P_{j,89})}$

We now have

$$P_{j,84} = \frac{(P_{j,89}).x}{(1+\pi_j)^5} - \frac{(\bar{U}_{89}).\ x}{(1+\pi_j)^5} - \frac{(u_{j,89} - \bar{U}_{89}).\ x}{(1+\pi_j)^5} \tag{21}$$

Estimation of (21) to yield consistent estimates of *x* is possible provided valid instruments are available, ie variables correlated with actual 1989 prices but independent of the company specific *excess* forecasting errors $(u_{j,89} - \bar{U}_{89})$.

We use lags of prices, dividends per share and earnings per share as instruments (all observable at the time the dependent variable in each regression is measured) which gives consistent parameter estimates provided the degree to which future variables deviate from expected values over and above the average degree of over (or under) stock market performance do not depend on past performance. The fact (if such it be) that overall stock market performance in 1989 exceeded expectations made in, for example, 1984 (ie $\bar{U}_{89} > 0$) will not in itself give downward biased estimates of x. It would, however, generate a negative estimate of the equation constant provided the true x were positive.

The equally weighted geometric mean return on all UK equities between 1984 and 1989 was 26.2% at an annual rate. The capitalisation weighted geometric mean was 20.7%. Inflation over this period averaged 5.5%. (I thank Elroy Dimson for these figures.) These actual returns almost certainly exceeded expectations. We should therefore expect a negative constant in each of our specifications for 1984. Consistency of estimates is preserved provided there is no tendency for firms which did even better than the average excess performance to have paid particularly low (or high) dividends pre 1984 or to have had unusually low (or high) prices or earnings.

Since **(15)**, **(16)** and **(17)** are cross section regressions using the 477 observations on price at each year as the dependent variable it is not possible to include company specific dummies to pick up omitted factors unique to each company. However, the estimated risk premia are specific to each company since they are a function of company βeta and company gearing. Three measures of βeta were tried:

(a) the βeta estimated from share data in the five years *prior* to the end of the 1984/85 financial year;

(b) the βeta estimated from share price data in the five years up to end of financial year 1989/90; and

(c) a weighted averaged of the two.

Because the theoretical basis underlying the modelling of stock prices where investors have different tax rates is weak, and because the measurement of average marginal tax rates is subject to error, we also estimated three different tax versions of each of **(15)**, **(16)** and **(17)** for each year. The first version ignores tax. The second version uses an estimate of the average of $(i - m)/(i - s)$ from the middle of our sample (1984). To construct this we estimate the average marginal income tax rate in (1984), using the methodology of King and Fullerton (1984) (see appendix). That estimated weighted average tax rate is approximately 0.20; the basic rate of tax in 1984 was 0.30 implying a central estimate of $(1 - \overline{m})/(1 - \overline{s})$ of around 1.15. The second versions of **(15)**, **(16)** and **(17)** use this scaling factor of 1.15 for future dividends.

The third version allows the tax factor to be estimated. Because the model is highly non-linear in the parameters (*and* is estimated by instrumental variables) it was not computationally possible to simultaneously estimate this tax scaling factor with the other parameters. Instead we estimated the model with scaling factors varying between 1.07 and 1.25 and report the version with the lowest minimand. [The minimand in non-linear 2SLS, which is defined in Table 1, is analogous to the residual sum of squares in OLS; because of the use of instruments, straightforward comparisons of log likelihood functions or of regression standard errors is illegitimate but comparison of the minimands is (see Gallant and Jorgenson (1979).] The range of tax factors over which the search was made imply a range of weighted marginal tax rates on dividend income from .25 down to .12. Given the share ownership of tax exempt institutions (pension funds) and insurance companies, who pay a rate of tax below the basic rate, this range of scaling factor almost certainly includes the true weighted average of marginal tax rates. Whether using a weighted average of tax rates is the correct way to allow for tax remains unclear—hence we also report the no-tax version of each specification.

Exactly the same methodology (the non-linear errors in the variable technique) was used to estimate equation **(12)**, **(13)** and **(14)** allowing for short-termism only to "kick in" after five years.

Section III Results

As a preliminary check on the properties of the estimation procedure we estimated a regression using the non-linear IV technique to assess the average risk premia for our sample. We used the mid-year of our sample (1984) and estimate α_o and π from the cross section regression:

$$P_{j,84} = \alpha_o + \frac{D_{j,85}}{(1+r_{85}+\pi)} + \frac{D_{j,86}}{(1+r_{86}+\pi)^2} + \frac{D_{j,87}}{(1+r_{87}+\pi)^3}$$

$$+ \frac{D_{j,88}}{(1+r_{88}+\pi)^4} + \frac{D_{j,89}}{(1+r_{89}+\pi)^5} + \frac{P_{j,89}}{(1+r_{89}+\pi)^5} \qquad \textbf{(22)}$$

The results are shown in the top panel of Table 1. The estimate of π is .078 implying an average risk premium of 7.8%. This figure is re-assuringly close to the mean annual excess return (over the gilts yield) on the all stocks index for the period 1919–89 of 7.7% [see Office of Water Services (1991), Volume 2, Table A3.8 and Spackman (1991)], suggesting that neither our sample nor the estimation procedure are generating an implausible figure for the average discount applied to future expected cash flows. α_0 is estimated to be –13.5, around 13% of the average 1984 price, suggesting that *ex-post* returns between 1984 and 1989, on average, exceeded expectations.

Having established that the sample of companies used and the (highly-nonlinear) IV estimation procedure adopted do not generate unusual or implausible estimates of the average equity premium we move on to allow risk premia to vary across companies and for discount factors to vary over different horizons.

We begin describing the results by, once again, focussing on 1984, the middle year from those used. The lower panel of Table 1 and Tables 2 and 3 show the results of estimating **(15)**, **(16)** and **(17)** where risk premia vary across companies. In general using the βeta estimated over the five years up to end of the financial year 1984 gave a slightly better fit than using the end-1989 βeta or the weighted average of the two. But in no case were the differences in fit or in the parameter estimates significant. Furthermore, in all cases where the null of no short-termism could be rejected using the 1984 βeta the null could be rejected with even greater confidence using each of the other measures.

In all specifications the estimated constant was negative and significant which we noted above is exactly what we would expect if actual dividends and the end period share prices were, on average, higher in the period 1985–89 than expected in 1984.

Table 1 reveals that estimation of the model under the restriction of no-shortism generates plausible parameter estimates. For the no-tax case the coefficient on βeta, which in a world where the simple CAPM is valid equals the expected excess return on a diversified portfolio over and above the risk-free rate, is positive (though low) and implies an expected excess return on equities over safe assets of only 1%.

Gearing consistently shows up as a significant determinant of company risk premia; the coefficient of 17.8 in the no-tax version implies that a company with average gearing (which for our sample is 57%) has a risk premia of 10 full percentage points higher than an otherwise similar company with no debt.

The version of the model where dividends are multiplied by our point estimate of the tax factor gives a somewhat better fit. The vale of the tax scaling factor which minimised the function was at the top end of our range of values—a finding consistent across all specifications—though the improvement in the fit of the model is negligible over the range of scaling factors between 1.15 and 1.25. Over this range estimated parameters are insensitive to the assumption made about marginal tax rates.

Correlation between our instruments and fitted residuals should be low if parameter estimates are not to be seriously biased. For all specifications estimated a regression of the equation residuals upon our 18 instruments had a R^2 of close to 14%. With a sample size approaching 500 a formal test for the validity of our instrument sets, based on TR^2 being distributed x^2_k under the null with k the number of overidentifying restrictions, suggests problems. ($TR^2 \simeq 68$ with x^2_{14} only 30, even at 1%.) But with a large sample a test such as this for *zero* correlation is almost certain to result in rejection. Of greater importance is the order of magnitude of bias for parameter estimates, which depends on the size of the correlation. With less than 15% of error variability explained by our 18 instruments the degree of bias is likely to be small. As a check on this we deleted all post 1981 variables from the instrument set. The R^2 between instruments and errors fell to below 0.09; but the value of the key parameters to assess short-termism were little changed. Indeed the extent of short-termism was estimated as being greater with the much reduced set of instruments.

Table 2 shows the results of estimating **(16)**. In all of the specifictions the estimate of b is in excess of 1. Comparing the no tax specifications and the central tax case specifications of Table 2 with those of Table 1 reveals a substantial reduction in the value of the minimands. A quasi likelihood ratio test constructed from the minimands [see Gallant and Jorgenson (1979)] rejected the null of no short-termism in both cases. For the no tax case the quasi likelihood ratio was 5.82; for the central case the likelihood ratio was 5.0; under the null the ratio $\sim x_1{}^2$ with a 5% critical value of 3.84. The null of no short-termism can easily be rejected at the 5% level.

Although the contribution of gearing to company specific risk premia remains significant, once allowance is made for "short-termism" the coefficient on βeta becomes negative. In a world with known inflation this would be implausible, though we noted above that Friend et al (*op cit*) have shown that low, and negative, coefficients on beta are just what one should expect if equities are a hedge against unexpected movements in the general level of prices.

Table 3 shows the results of allowing for a slightly different version of short-termism. We noted above that it was possible that specifications **(7)** and **(8)** might generate similar results—at least in terms of rejecting the null of no myopia. Table 3 appears to confirm this intution. In all cases the no myopia restriction ($x = 1$) is rejected at the 5% level. For the no-tax and central case specficiations the quasi likelihood ratio statistics are 6.33 and 5.56 respectively. Furthermore, the negative coefficient on βeta which appeared in Table 2 is also found consistently across various specifications of **(17)**. Whether the results in Table 3 add weight to the hypothesis that there exists short-termism in UK stock markets depends on whether the parameter estimates can really measure something different from estimates of b reported in Table 2. As noted above, estimating b or x amounts to the same thing *if* there is little variability in discount rates. But constructing the discount rates implicit in the parameter estimates reported in the first row of Table 2 showed that cross-company variability was great.

The mean discount rate was just over 13% with a standard deviation of 7.9%. Table 1 also reveals that allowing discount rates to vary across companies (lower panel regressions) very substantially increases the fit of the model relative to a constant premium specification (upper panel); the null hypothesis of no variation in risk premia can be overwhelmingly rejected.

Table 4 shows estimates of equations (**12**), (**13**) and (**14**). Here the null of no short-termism is rejected most strongly. In version (**12**) our central estimate of a_0, a parameter reflecting excess discounting of longer term cash flows, is .155 with a t statistic of 3.8. This implies that discount rates applied to longer term cash flows (expressed at an annual rate) are 15 full percentage points higher than discounts applied to short-term flows. Put another way, discount rates applied to longer term flows are about double the rates applied to shorter-term flows.

Parameter α from specification (**13**) is estimated at 2.03, some 3.1 standard errors in *excess* of 1. Parameter λ, specification (**14**), is estimated at .53, 5.4 standard errors *below* 1. These results from estimating (**13**) and (**14**) are consistent with the results from estimating (**12**)—longer term flows are discounted at rates twice as high as shorter term flows.

Table 5 shows the results of estimating specifications similar to equation (**16**) for each of the four years before, and each of the four years after, 1984. In each of these eight cross section regressions the same parameters are estimated though we would certainly expect the value of the constant to differ across years since the extent to which future dividends and share prices, on average, diverged from their expected values is likely to have varied substantially across the decade. We would also expect the coefficient b to vary if the degree of excess discounting of longer term cash flows (ie the extent of short-termism) changed over time. Variability in the coefficients describing company risk premia over default-free bonds (ie $\hat{a}_1$ and $\hat{a}_2$) would vary if the degree to which equities were perceived as a hedge against inflation varied. The table reports specifications where the central value of the tax parameter was used; in all cases the results were hardly changed if the tax factor were set to one (ie the tax system were assumed neutral).

In every year from 1983 the parameter b is estimated as being significantly greater than one. In the earliest years of the period estimation proved difficult; convergence was slow and parameter values were, in one case (1982) wildly implausible. Since the model becomes increasingly non-linear in the parameters as prices are measured further from the end point, convergence problems are to be expected.

Table 6 shows the results from estimating our alternative model of short-termism [equation (**17**)]. Here, in every period save one (1981) the key parameter measuring the degree of short-termism (x) is below one; in each period from 1983 the deviation from the no-short-termism value of $x = 1.0$ is statistically significant.

What emerges from these results is apparent evidence of short-termism throughout much of the decade. Despite the problems with the regressions from the early years in the decade there is also evidence that the degree of short-termism then was less than in the period from 1983; convergence problems were less serious with the specifications reported in Table 6 and it is evident from those results that x is close to unity in 1980–82 whilst nearer 0.90 in subsequent years.

The constants in each specification show considerable variability across years. As noted above this is to be expected. In periods where the market as a whole may have been over-valued (1986, 1987) we should expect a negative *ex-post* average forecast error for future dividends and share prices which should generate a *positive* constant. In periods where future dividends and stock prices turned out higher than could have been anticipated (1983, 1984) the constant is *negative*. Tables 5 and 6 reveal that the 1983 and 1984 constants were negative whilst those for 1986 and 1987 were signficantly positive. It is reassuring that the values of b and x, which measure the degree of short-termism, are far less volatile than are the intercepts, suggesting that the IV estimator of these parameters is robust to changes in the degree to which expectations of future cash flows are fulfilled.

One important feature of the results in Tables 5 and 6 is that there is no tendency for years in which expectations of future dividends and prices turned out to be the most pessimistic to generate parameter estimates implying the highest degrees of short-termism. The average *ex-post* growth in share prices (expressed at an annual rate) over the period 1980–89 for our sample of companies was around 17%. For the period 1981–89 the figure is almost 18% and for 1982–89 nearly 19%. The estimated degree of short-termism from the regressions for those three-years is small; the null of no short-termism could not be rejected at any reasonable level based on share price behaviour at the beginning of the 1980s. In contrast, the average annual ex-post growth in share prices between 1987 and 1989 for our sample was only around 8%, yet the extent of short-termism implicit in our estimated values of b and x for that year was substantial and close to being the greatest for the 1980s. These results are inconsistent with the argument that our evidence for short-termism merely shows that the stock market generated higher returns in the 1980s than could have been expected and that apparently high discounts on future cash flows simply reflect unexpectedly high subsequent returns. If that argument were true it would imply both that the instruments used were invalid and that parameter estimates should show short-termism to have been at its worst in the early 1980s and non-existent in 1986–87.

There is significant variability in the parameters describing the risk premia (a_1, a_2) both across time periods and across the two specifications estimated for each year. This may reflect undue weight attached to the prices of particular firms in particular years. As a check on this we re-estimated all the models reported in Tables 6 and 7 using a generalised method of moments estimator; this estimator is equivalent to the two stage non-linear least squares estimator used above applied to the weighted observations on company prices for each year where the weights reflect the inverse of the expected conditional variance of each company's price. Applying this estimator generated less volatile values of $\hat{a}_1$ and $\hat{a}_2$; but in all cases the coefficients on company βeta were negative and on gearing positive. In no case did the heteroscedasticity adjustment remove evidence of short-termism; for the years since 1983 all estimated values of b remained significantly in excess of 1 whilst all values of x were less than 1.

Some possible explanations:

The most straightforward explanation of the results is that stock market valuations of corporations are short term. Our estimates of b are often around 1.8 implying that discount rates applied to cash flows which accrue six months in the future are more appropriate to flows accruing in eleven months while cash flows which are not expected to come through for five years are discounted as if they did not accrue for nine years. Our central esitmate of x suggests a somewhat different version of short-termism, but one no less dramatic in its implications for the degree of myopia. Table 6 revealed values of x around 0.90; this implies that cash flows which accrue six months in the future are underestimated by 5% relative to rational expectations but cash flows which do not accrue for five years are systematically underestimated by almost 40% [ie $(1-x^5)$]. Put another way, projects with only a six month time horizon need, on average, to be 5% more profitable than is optimal if companies which undertake them are not to suffer a decline in stock market value; projects with five years to maturity, however, need to be around 40% more proftiable than is optimal. On even the loosest definition of what constitutes clear sight this counts as serious myopia.

Allowing for short-termism only to "kick in" after five years suggests an even greater degree of excess discounting of long term cash flows. The results reported in Table 4 imply that cash flows accruing more than five years in the future are discounted at twice the rate of shorter term flows.

But as noted earlier it is always possible to explain these, and any other, results as being consistent with market efficiency provided one is prepared to accept as plausible highly variable *and* increasing risk premia. Our results can be attributed to mismeasurement of risk premia, but such claims lack force unless the pattern of risk premia needed to explain estimates of b, x, α, a_0 and λ well away from unity are also consistent with a plausible degree of risk aversion and a rational assessment of the underlying risk of projects. An estimate of b of 2.0 in equation **(16)**, for example, is perfectly consistent with no short-termism provided one accepts that a discount rate of $(1+r+\hat{\pi})^2$, and not $(1+r+\hat{\pi})$, is consistent with plausible attitudes towards uncertainty and efficient assessment of risk. In the current context that would imply that twice the risk-free rate, the square of company βeta, the square of the nominal interest rate and twice the cross product of the two should appear in the appropriate discount rate. This seems hard to rationalise. Perhaps it is even more difficult to rationalise the apparent doubling of (per period) discount rates implied by the results of estimating equations (**12**), (**13**) and (**14**).

The method of measuring company-specific risk premia we followed allowed premia to respond to βeta—measured both using past and future covariances of returns with a diversified portfolio. In general we found risk premia responded negatively to βeta—a decidedly non-standard result but not wholly implausible in a world with uncertain future prices and where equity prices over the longer term may move in line with consumer prices. We also allowed gearing to influence risk premia. Dropping βeta from specifications led to a reduction in the fit of models, but the crucial parameters (b, x, α, a_0 and λ) continued to imply short-termism. Other factors which might be relevant in measuring risk—for example the covariance between a company's return and aggregate consumption—were excluded. Whether this could account for our results must be the subject of future studies.

As regards tax we went to some length to allow for the influence of the operation of the UK imputation sustem. Results on short-termism proved robust to quite major changes in the assumed tax parameters. Possible errors arising from other factors can be listed but are hard to quantify; not allowing for the exact timing of the cash flows by assuming that dividends are paid at the end of each report year (as we have done) certainly induces errors. But the bias here is likely to be small (see Nickell and Wadhwani, 1987) and goes in favour of rejecting short-termism (dividends are paid no later than the report date and so actual cash flows are always nearer than assumed, so discounts should be *smaller* than in the specifications where no myopia is allowed).

But despite attempts to assess the robustness of results to variations in the treatment of tax, risk and to how different versions of short-termism might work the range of assumptions we have needed to make to estimate any of the models remains substantial. Ultimately, any rejection of the null hypothesis of no short-termism is conditional on the acceptance of some auxiliary assumptions. What our results show is not that short-termism certainly exists, but that those who believe it does *not* exist do need to explain something about the operation of security markets which makes longer-term cash flows appear to be discounted at higher rates than shorter-term flows.

In short our results present a puzzle which is comparable to, though distinct from, that posed by Mehra and Prescott (1985). They argued that the return on equities in the United States was too high to be consistent with reasonable estimates of the volatility of share prices and plausible values for risk aversion.

Just as that puzzle has proved hard to resolve it may be difficult to square the findings reported here with the theory that market valuations of UK equities reflect rational predictions of future cash flows adjusted appropriately for risk.

Data appendix

Definitions:

βeta	Company specific βetas are taken from the London Business School's "Risk Measurement Service". βetas are calculated from five years of monthly data on company yields and on the FT all share index.
Price	Company share prices are quoted in numbers of pence and measured on the company report date.
Dividends	Measured as total dividends per share, in pence, paid in the account year.
Earnings	Post tax, net of interest payments, earnings per share.
$r_{84,T}$	The average yield to maturity at end-1984 on all UK government bonds maturing in year T; quoted at an annual rate.
$\bar{m}$	Weighted average of the marginal tax rates of persons, pension funds, insurance companies and non-financial companies. Weights are calculated using CSO figures on shareholdings at end-1984. Marginal tax rates for different institutions are from data kindly provided by Mervyn King.
$\bar{s}$	Imputation rate at end-1984 = 30%.

Sources for company specific data (excluding βeta) are the EXTEL tapes including balance sheet and profit loss accounts items from the published accounts. The sample was of 477 non-financial companies. Total profits of this sample accounted for approximately 45% of the total profits of all UK industrial and commercial companies in 1989; the stock market valuation was 50% of the value of all ICCS.

Annex

The CAPM with uncertain inflation:

Friend et al (1976) derive the equilibrium required returns on capital assets with uncertain inflation and when the return on the "safe" asset is fixed only in nominal terms. Ignoring tax and assuming all assets are marketable they derive the condition:

$$E(R_j) = r_f + \sigma_{j\pi} + \left[\frac{E(R_m) - r_f - \sigma_{m\pi}}{\sigma_m^2 - \frac{\sigma_{m\pi}}{\alpha}} \right] \left(\sigma_{jm} - \frac{\sigma_{j\pi}}{\alpha}\right) \qquad \textbf{(1a)}$$

where

$E(R_j)$	=	expected return on asset j (nominal)
r_f	=	nominal risk free rate of return
$\sigma_{j\pi}$	=	covariance between return on the jth asset and inflation
$E(R_m)$	=	expected return on market (nominal)
$\sigma_{m\pi}$	=	covariance between return on the market and inflation
σ_{jm}	=	covariance between return on the jth asset and the market
α	=	ratio of value of risky (in nominal terms) assets to the value of all assets,
σ_m^2	=	variance of return on market.

It seems plausible that holding the market portfolio is a far better hedge against unexpected inflation than holding a single typical share ie $\sigma_{m\pi}$ is far in excess of $\sigma_{j\pi}$. If we ignore the $\sigma_{j\pi}$ terms as being of second order importance we can write (1a) as:

$$E(R_j) = r_f + \left[\frac{\frac{[E(R_m) - r_f]}{\sigma_m^2} - \frac{\sigma_{m\pi}}{\sigma_m^2}}{1 - \frac{\sigma_{m\pi}}{\sigma_m^2 \alpha}} \right] \sigma_{jm} \qquad \textbf{(2a)}$$

$$\frac{E(^{R}m) - rf}{\sigma_m^2}$$ is the "market price of risk" as defined in the simple CAPM.

Denote this λ, a number generally estimated as between 1 and 4 [see Hall, Miles and Taylor (1989) and Merton (1980)].

$$\frac{\sigma_{m\pi}}{\sigma_m^2}$$ is the slope coefficient from a regression of inflation on stock market returns.

Denote this $\beta_{m\pi}$, a number almost certainly less than one and probably positive.

(2a) now becomes

$$E(^{R}j) = r_f + \left[\frac{\lambda - \beta m\pi}{1 - \frac{\beta_{m\pi}}{\alpha}} \right] . \sigma_{jm} \tag{3a}$$

Letting the term in square brackets be denoted ϕ we can write

$$E(^{R}j) = r_f + (\sigma^2{}_m.\phi).\ \beta_j \tag{4a}$$

where β_j is the standard CAPM βeta of security j. Thus the sign of the coefficient on simple βeta in the CAPM amended for uncertain inflation is the same as the sign of ϕ. Since λ is almost certainly greater than $\beta_{m\pi}$, the sign is negative (positive) if $\beta m\pi$, exceeds (is less than) α.

Approximately 1/3 of the marketable assets of UK households are equities. Taking this as a central estimate of a we see that a negative coefficient on βeta will arise if $\beta_{m\pi} > 1/3$.

Table 1

Estimate of Equation (22) **Estimating Average Risk Premia**
Dependent Variable $P_{j,84}$

[Estimates from applying non-linear two stage least squares to equation (22).]

		Strd Error	Log Likelihood	R^2	Minimand
α_o	-13.54				
	(2.30)	72.57	-2720.6	.39	475672
π	.0781				
	(7.55)				

Estimates of Equation (15) Dependent Variable $P_{j,84}$

Tax Scaling factor	$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	Strd Error	Log Likelihood	SK	KU	R^2	Minimand
1	.0099	17.80	-12.87	65.45	-2671.4	-1.20	11.89	.50	303188
(no tax effects)	(.90)	(3.0)	(2.5)						
1.15	.008	17.70	-13.33	64.47	-2664.1	-1.23	12.11	.52	287404
(central case)	(.70)	(6.0)	(2.67)						
1.25	.0109	17.73	-13.67	63.99	-2660.6	-1.19	11.89	.52	280831
(max likelihood)	(.95)	(6.0)	(2.74)						

Notes

SK = coefficient of skewnes of equation residuals.

KU = coefficient of kurtosis of equation residuals.

$\hat{a}_1$ = coefficient on βeta.

$\hat{a}_2$ = coefficient on gearing.

t statistics in parenthesis. Mean of dependent variable: 106.36
Standard deviation dependent variable: 92.796

number of observations: 477

Minimand is $e'Z(Z'Z)^{-1}Z'e$ where $e = P_{84} - f(a_1, a_2, \ldots \text{CNST}, X)$

and Z is a matrix of instrumental variables.
X are the actual values of explanatory variables, in the RHS of (15).
$f(\ )$ is the functional form represented by the RHS of (15).

Table 2

Estimates of Equation (16) Dependent Variable $P_{j,84}$

Estimates of Equation (15) Dependent Variable $P_{j,84}$

Tax Scaling factor	$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	$\hat{b}$	Strd Error	Log Likelihood	SK	KU	Minimand	R^2
1	-.042	8.95	-10.22	1.654	63.46	-2656.6	-1.30	13.70	279740	.53
(no tax effects)	(3.0)	(3.4)	(2.0)	(5.5)						
1.15	-0.40	8.78	-10.71	1.678	62.66	-2650.5	-1.23	13.35	267692	.54
(central case)	(2.9)	(3.5)	(2.1)	(5.6)						
1.25	-.039	8.67	-11.01	1.694	62.14	-2646.6	1.19	13.14	260322	.55
(max likelihood)	(2.8)	(3.5)	(2.2)	(5.7)						

Notes

$\hat{a}_1$ = coefficient on βeta

$\hat{a}_2$ = coefficient on gearing

Minimand as in Table 1

Table 3

Estimates of Equation (17)   Dependent Variable $P_{j, 84}$

Tax Scaling factor	$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	$\hat{x}$	Strd Error	Log Likelihood	SK	KU	Minimand	R^2
1	-.071	14.98	-10.25	.929	63.44	-2656.4	-1.29	13.66	277700	.53
(no tax effects)	(2.1)	(5.3)	(2.0)	(30.4)						
1.15	-.069	14.91	-10.74	.926	62.64	-2650.4	-1.24	13.34	265560	.54
(central case)	(2.1)	(5.4)	(2.1)	(30.6)						
1.25	-.069	14.88	-11.05	.925	62.12	-2646.4	-1.26	13.20	258130	.55
(max likelihood)	(2.1)	(5.4)	(2.2)	(30.7)						

Notes

$\hat{a}_1$ = coefficient on βeta.

$\hat{a}_2$ = coefficient on gearing.

Other notes as in Table 1

Table 4

Estimation of Equation (12) (Tax Factor Set at 1.15)

$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	$\hat{a}_0$	Strd Error	Log Likelihood	Minimand	R^2	SK	KU
-.117	14.44	-10.71	.155	59.48	-2625.7	238838	.59	-1.06	12.80
(3.6)	(5.4)	(2.3)	(3.8)						

Estimation of Equation (13) (Tax Factor Set at 1.15)

$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	$\hat{\alpha}$	Strd Error	Log Likelihood	Minimand	R^2	SK	KU
-.049	7.36	-10.61	2.03	60.01	-2629.9	236492	.58	-0.93	11.51
(4.7)	(3.8)	(2.3)	(6.3)						

Estimation of Equation (14) (Tax Factor set at 1.15)

$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	$\hat{\lambda}$	Strd Error	Log Likelihood	Minimand	R^2	SK	KU
-.100	13.12	-10.68	.529	59.62	-2626.8	238398	.59	-0.96	12.50
(3.7)	(5.2)	(2.3)	(6.1)						

Notes as in Table 1

Table 5

Estimates of Equation (16) for various years
Non Linear Two Stage Least Squares

	$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	$\hat{b}$	Strd Error	Log Lik	Minimand	R^2
1980	-0.001	2.96	-15.34	1.136	50.05	-2541	238558	.19
	(0.1)	(3.0)	(3.3)	(6.1)				
1981	-0.012	4.53	-21.96	0.950	53.80	-2576	209016	.21
	(0.5)	(3.7)	(4.4)	(6.0)				
1982*	-0.64	68.46	10.19	0.20	31.25	-2317	465999	.80
	(4.1)	(3.7)	(4.7)	(2.3)				
1983	-0.032	5.27	-10.35	1.886	54.62	-2583	172252	.54
	(2.7)	(4.0)	(2.3)	(6.7)				
1984(1)	-0.040	8.78	-10.71	1.678	62.66	-2650	267692	.54
	(2.9)	(3.5)	(2.1)	(5.6)				
1985	-0.039	7.98	-2.40	2.045	61.75	-2644	295932	.67
	(3.1)	(3.2)	(0.5)	(5.9)				
1986	-0.057	8.78	12.76	2.67	72.86	-2722	653298	.69
	(5.8)	(3.1)	(2.3)	(5.7)				
1987	-0.035	10.72	27.02	2.385	86.77	-2806	713908	.71
	(1.8)	(2.2)	(4.2)	(3.1)				
1988(2)	-0.017	1.792	9.149	1.680	57.79	-2612	166024	.89
	(0.3)	(0.5)	(2.21)	(2.5)				

Notes

* convergence problems

(1) This is the same regression as reported in Table 2.
(2) IN 1988 both remaining dividends and the terminal price are only one period ahead so the excess discounting parameter is applied only to the terminal price and not to D_j89.

$\hat{a}_1$ = Coefficient on βeta.
$\hat{a}_2$ = Coefficient on gearing.

In all cases five lags of earnings per share, dividends per share and prices were used as instruments for future dividends and the end period share price.

In all cases the number of observations is 477.

Table 6

Estimates of Equation (17) for various years [t statistics in parenthesis]

	$\hat{a}_1$	$\hat{a}_2$	$\hat{CNST}$	$\hat{x}$	Strd Error	Log Lik	Minimand	R^2
1980	0.0006	4.21	-7.02	.962	46.24	-2504	219028	.31
	(0.1)	(4.2)	(1.7)	(39.0)				
1981	0.015	4.48	-12.50	1.019	49.80	-2539	220400	.32
	(0.7)	(5.6)	(4.5)	(39.9)				
1982	-0.032	6.96	-20.84	0.954	49.89	-2539	146658	.48
	(1.3)	(6.4)	(4.6)	(44.4)				
1983	-0.059	10.04	-10.44	0.914	54.69	-2584	172398	.54
	(2.0)	(5.5)	(2.3)	(35.2)				
1984	-0.069	14.91	-10.74	0.926	62.64	-2650	265560	.54
	(2.1)	(5.4)	(2.1)	(30.6)				
1985	-0.084	16.95	-2.30	0.893	61.61	-2642	292116	.67
	(2.3)	(4.4)	(0.5)	(27.2)				
1986	-0.146	23.96	12.65	0.844	72.65	-2721	650086	.69
	(3.3)	(4.3)	(2.3)	(21.7)				
1987	-0.089	27.37	27.11	0.885	86.54	-2804	710300	.71
	(1.3)	(3.7)	(4.2)	(15.5)				
1988(1)	0.022	2.99	9.16	0.919	57.80	-2612	166200	.89
	(0.3)	(0.5)	(2.2)	(13.0)				

Notes

In all cases the number of observations is 477.

(1) In 1988 there are only two cash f ows remaining—D_{j89} and P_{j89} both of which are one year ahead; the excess discounting para eter x is applied only to P_{j89}.

References

Amemiya, T (1985): "Advanced Econometrics", Basil Blackwell.

Cooper, I and Marsh, P (1987): Letter to the Financial Times, 16 March.

Cosh, A Hughes, A and Singh, A (1990): "Takeovers, Short-termism and Finance-Industry Relations in the UK", Department of Applied Economics, Cambridge, May.

Fama, E (1970): "Multiperiod Consumption and Investment Decisions", *Amercian Economic Review*, 60, pages 163–74.

Friend, I, Lanskroner, Y and Losq, E (1976): "The Demand for Risky Assets and Uncertain Inflation", *Journal of Finance,* XXXI, no 5, pages 1287–97.

Froot, K Scharfstein, D and Stein, J (1990): "Herd on the Street: Informational Ineffeciencies in a Market with Short-term Speculation", NBER Working Paper No 3250.

Gallant, A and Jorgenson, D (1979): "Statistical Inference for a System of Simultaneous, Non-Linear, Implicit Equations in the Context of Instrumental Variables Estimation", *Journal of Econometrics,* Vol II, pages 275–303.

Greenspan, A (1990): Speech quoted in Marsh (1990).

Hall, S; Miles, D K and Taylor, M (1989): "Modelling Asset Prices with Time-Varying betas", *The Manchester School,* Vol LVII, pages 340–56.

Healy, Palepu and Rubak (1989): NBER Discussion Paper No 3348.

Jensen, M (1986): "The Takeover Controversey: Analysis and Evidence", *Midland Corporate Finance Journal,* IV, pages 6–32.

Jensen, M (1989): "Eclipse of the Public Corporation" *Harvard Business Review,* September–October, pages 61–74.

Kay, J and King, M (1990): "The UK Tax System", fifth edition, Oxford University Press, Oxford.

King, M (1977): "Public Policy and the Corporation", McMillan, London.

King, M and Fullerton, D (1984): "The Taxation of Income from Capital: A Comparative Study of the United States, United Kingdom, Sweden and West Germany" (eds). Chicago, Chicago University Press.

Law, W (1986): "A Corporation is more than its Common Stock", *Harvard Business Review,* May-June, pages 80–3.

Lawson, N (1986): Speech by the then Chancellor of the Exchequer made on 9 June [quoted in Nickell and Wadhwani, (1987)].

Lintner, J (1965): "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets", *Review of Economics and Statistics,* pages 13–37.

Lucas, R (1978): "Asset Prices in an Exchange Economy", *Econometrica.*

Marsh, P (1990): "Short Termism on Trial".

McCallum, B (1976): "Rational Expectations and the Natural Rate Hypothesis", *Journal of Econometrics.*

Mehra, R and Prescott, E (1985): "The Equity Premium: A Puzzle", *Journal of Monetary Economics,* 15, pages 145–61.

Merton, R (1973): "An Inter-temporal Capital Asset Pricing Model", *Econometrica,* 41, No 5, pages 867–88.

Merton, R (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation", *Journal of Financial Economics,* Vol 8, No 3, pages 323–61.

Mossin, J (1966): "Equilibrium in a Capital Asset Market", *Econometrica,* October, pages 768–83.

Narayanan (1985): "Observability and the Payback Criterion", *Journal of Business,* July.

Nickell, S and Wadhwai, S (1987): "Myopia, The Dividend Puzzle, and Share Prices", LSE Centre for Labour Economics Discussion Paper No 272.

Office of Water Services (1991): "Cost of Capital: A Consultative Report", July.

Pickens, T Boone (1986): "Professions of a Short-termer", *Harvard Business Review,* May–June, pages 75–9.

Ross, S (1976): "The Arbitrage Theory of Capital Asset Pricing", *Journal of Economic Theory,* 13, pages 341–60.

Sharpe (1963): "A Simplified Model for Portfolio Analysis", *Management Science,* January, pages 277–93.

Sharpe (1964): "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk", *Journal of Finance,* September, pages 425–42.

Schleifer, A and Vishny, R (1990): "Equilibrium Short Horizons of Investors and Firms", *American Economic Review,* Vol 80, 2, pages 148–53.

Shiller, R (1990): "Market Volatility", MIT Press, Cambridge, Mass.

Spackman (1991): "Discount Rates and Rates of Return in the Public Sector: Economic Issues", UK Treasury Working Paper No 58.

Stapleton, R and Subrahymanyam (1977): "Multi-Period Equilibrium Asset Pricing Models", *Econometrica,* 46.

Stein, J (1989): "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behaviour", *Quarterly Journal of Economics,* pages 655–69.

Webb, D (1991): "Asymmetric Information and the Trade-Off Between Cash Flow and Net Present Value", LSE Financial Markets Group Discussion Paper No 117.

Wickens, M (1982): "The Efficient Estimation of Econometric Models with Rational Expectations", *Review of Economic Studies,* 49, pages 55–68.

Woolridge, J (1988): "Competitive Decline and Corporate Restructuring: Is a Myopic Stock Market to Blame?", *Journal of Applied Corporate Finance,* pages 23–36.

Bank of England Working Paper Series

Publication date in italics

1	Real interest parity, dynamic convergence and the European Monetary System *(June 1992)*	Andrew G Haldane Mahmood Pradhan
2	Testing real interest parity in the European Monetary System *(July 1992)*	Andrew G Haldane Mahmood Pradhan
3	Output, productivity and externalities—the case of banking *(August 1992)*	R J Colwell E P Davis
4	Testing for short-termism in the UK stock market *(October 1992)*	David Miles

Design and desktop publishing by the
Bulletin Group, Economics Division

Printed by the
PRINT & COPY SHOP