Stock Return Predictability:
Is it There?*

Andrew Ang
Geert Bekaert
Columbia University and NBER

First Version: 4 March 2001
This Version: 14 August 2001

JEL Classification Codes: C12, C51, C52, E49, F30, G12
Keywords: present value model, predictability, international predictability, short rates, dividend yield, earnings yield, cashflow growth rates

*We thank Xiaoyan Zhang for help with data. We thank Kobi Boudoukh, John Campbell, John Cochrane, Charles Jones, Martin Lettau, Jun Liu, Sydney Ludvigson, Tano Santos, Bob Shiller, Ken Singleton, Meir Statman, Jeff Wurgler and seminar participants at Columbia University, the NY Federal Reserve and Stanford University for helpful comments. Andrew Ang thanks the Chazen Institute at Columbia University for financial support. Geert Bekaert thanks the NSF for financial support. Address: Columbia Business School, 3022 Broadway, Uris Hall, New York NY 10027. Andrew Ang: aa610@columbia.edu, Geert Bekaert: gb241@columbia.edu. The latest version of this paper can be downloaded from http://www.columbia.edu/∼aa610

Abstract

We ask whether stock returns in France, Germany, Japan, the UK and the US are predictable by three instruments: the dividend yield, the earnings yield and the short rate. The predictability regression is suggested by a present value model with earnings growth, payout ratios and the short rate as state variables. We find the short rate to be the only robust short-run predictor of excess returns, and find little evidence of excess return predictability by earnings or dividend yields across all countries. There is no evidence of long-horizon return predictability once we account for finite sample influence. Cross-country predictability is stronger than predictability using local instruments. Finally, dividend and earnings yields predict future cashflow growth rates both in the US and in other countries.

1 Introduction

A large body of empirical work has accumulated documenting excess stock return predictability. Among the most popular predictors are the nominal interest rate and the dividend yield.[1] The dividend yield appears to be the most popular stock return predictor used in applied work, but recently Lamont (1998) argues that the earnings yield, has independent forecasting power for excess stock returns in addition to the dividend yield.

The debate on what drives the predictability continues. It may reflect irrational investor behavior and hence be exploitable in trading strategies (Cutler, Poterba and Summers (1989)); it may reflect time-varying risk premiums (Kandel and Stambaugh (1990)), Campbell and Cochrane (1999), and Bekaert and Grenadier (2000)); or it may simply not be present in the data. This last possibility gains credibility considering the long list of authors criticizing the statistical methodologies in the predictability literature. The coefficients on the predictor variables are biased, since these variables are typically persistent, endogenous regressors correlated with returns innovations (Stambaugh (1999)). The standard focus on long-horizon regressions is problematic from a number of different perspectives. The distributions of the R^2 (Kirby (1997)) and the t-statistics on the coefficients, (Richardson and Stock (1989), Richardson and Smith (1991) and Hodrick (1992)) in long-horizon regressions are severely shifted to the right, leading to over-rejection of the no-predictability null. Researchers often forget to properly interpret various tests over different horizons by providing joint tests (Richardson (1993)). Finally, the possibility of decades of data mining clouds any inference regarding predictability for US stock returns (Lo and MacKinlay (1990) and Foster, Smith and Whaley (1997)).

In this paper, we re-examine the case for the predictability of short and long-horizon stock returns. We start by proposing a simple price earnings model, in which the variation in the price-earnings ratio and expected returns on equities is driven by three stochastic state variables, the payout ratio, earnings growth and the short rate. In this model, the earnings yield, the dividend yield and the short rate jointly capture any potential predictability, motivating a simple multivariate regression of excess stock returns over various horizons on these three variables as the main predictability regression. When certain parameter restrictions are met, the model has a constant expected excess return variant, in which the expected gross return on equity equals a

[1] For predictability of excess stock returns by the nominal interest rate see, among others, Fama and Schwert (1977), Campbell (1987), Breen, Glosten and Jagannathan (1989), Shiller and Beltratti (1992), Lee (1992) and Glosten, Jagannathan and Runkle (1993). Among those examining the predictive power of the dividend yield on excess stock returns are Fama and French (1988), Campbell and Shiller (1988a and b), Ferson (1989), Goetzmann and Jorion (1993, 1995), Hodrick (1992), Stambaugh (1999), Goyal and Welch (1999) and Valkanov (2001).

constant multiple of the short rate.

Given the considerable statistical challenges in establishing predictability, we precede our analysis of the data with a Monte Carlo analysis under the null of no predictability using the present value model. As Goetzmann and Jorion (1993, 1995) point out, Monte Carlo analysis with standard linear models ignores the fact that yield variables involve the inverse of price, an endogenous variable, which is also present in the denominator of the return on the left hand side. They conduct bootstrap exercises that impose this constraint, but their bootstrap keeps dividends non-stochastic at their data levels and therefore ignores the cointegration relation between dividends and price levels that characterizes rational pricing. By simulating the constant expected return variant of our earnings model, we accommodate this endogeneity constraint in an entirely coherent way.[2] Our analysis incorporating earnings yields shows that Ordinary Least Squares (OLS) or Hansen-Hodrick (1980) standard errors lead to severe over-rejections of the null hypothesis of no predictability at long horizons. The finite sample properties of the long-horizon regression t-statistics improve dramatically by removing the moving average structure in the error terms induced by summing returns over long horizons, as suggested by Richardson and Smith (1991), Hodrick (1992) and Boudoukh and Richardson (1993). For brevity, we refer to these alternative standard errors as Hodrick (1992) standard errors.

Armed with well-behaved t-statistics, we establish that the predictability evidence for US returns is surprisingly weak. In fact, the only variable retaining significance is the short rate, and it is only significant at short horizons. To mitigate data snooping concerns, we investigate analogous predictability regressions for four other countries, France, Germany, Japan and the UK. Interestingly, we find that the predictability coefficients are not robust across countries in sign or magnitude, except for the short rate effect. When we pool the regression across countries, the short rate remains the only significant predictor of excess stock returns.

Finally, we also investigate a number of cross-country predictability regressions, examining whether any predictors have predictive power across countries. Unlike Bekaert and Hodrick (1992) and Ferson and Harvey (1993), we only find evidence of strong predictability when we pool across countries. With cross-sectional information from international data we find that US instruments are strong predictors of foreign equity returns, unlike local instruments. The local short rate effect is subsumed by the predictive power of the US short rate. We also confirm and extend Bekaert and Hodrick (1992)'s finding that yield variables have predictive power for excess returns in the foreign exchange market. We conclude that the current predictability

[2] Bollerslev and Hodrick (1996) provide a detailed Monte Carlo analysis in the context of a present value model with constant and time-varying expected return variants which also imposes this constraint, but their solution to the present value model is only approximately true.

debate focuses on the wrong horizon (long-run instead of short-run), the wrong instruments (yield variables instead of interest rates) and the wrong setting (US segmented market instead of a globally integrated market).

In a rational no-bubble model, price-dividend and price-earnings ratios reflect the expected value (of a non-linear function) of future cashflow growth rates and discount rates. Since the variation of price-dividend or price-earnings ratios can be attributed to the variation of future cashflow growth, future discount rates or both, our results have two possible interpretations. First, it is possible that yield variables do predict returns but that linear regressions have little power to detect this predictability given the noisiness in returns. Second, it is possible that a large fraction of the variation in dividend and earnings yields reflects the predictability of future cashflow growth. Whereas previous results in the literature appear inconsistent with this possibility (for example, Cochrane (1992)), we find clear evidence of cashflow predictability by yield instruments.

The remainder of the paper is organized as follows. Section 2 sets out the empirical framework, including the present value earnings model and the predictability regressions. Section 3 describes the econometric estimation, the Monte Carlo analysis and describes the data. Section 4 considers the predictability in US returns, whereas Section 5 investigates and compares predictability in all 5 countries. Section 6 investigates return predictability across countries. Section 7 investigates cashflow predictability. Section 8 concludes and offers an interpretation of our results.

2 Theoretical and Econometric Framework

2.1 Data Description

Our data set consists of equity total return (price plus dividend) indices from Morgan Stanley Capital International (MSCI) for the US, Japan, UK, Germany and France. The short-term interest rates we use are 1 month EURO rates from Datastream. The sample period is from February 1975 to December 1999 for the US, UK, France and Germany and from January 1978 to December 1999 for Japan. MSCI provide dividend and earnings yields which use dividend and earnings summed over the past 12 months. Although this is restrictive, monthly earnings and dividend levels are impossible to use because they are dominated by seasonal components, given that most firms have a December-end calendar year. Hence, monthly de-seasonalized earnings and dividends are unobservable.

Table (1) reports summary statistics of returns and instruments. The historical log equity premium is around 7.5% in the Anglo-Saxon countries, 6.9% in France and Germany and only 3.6% in Japan. Excess return volatility is over 18% in all countries, except for the US where it is only 15%. The highest correlation of excess returns is between Germany and France, at 59.4%, and the lowest are between Japan and the other countries, which are around 35%. Hence, there is fairly large cross-sectional variation in excess returns. All three instruments - the log dividend yields dy^{12}, log earnings yield ey^{12} and short rates r are highly persistent. We use the superscript ^{12} on the dividend and earnings yields to indicate that these are constructed using dividends and earnings summed up over the last 12 months.

As Bekaert and Hodrick (1992) discuss, MSCI report price (capital appreciation) and total returns (including income). The total return is an estimate constructed from annualized dividends (summed over the previous twelve months). For some countries there is a discrepancy between the MSCI dividend yield dy^{12} and the implied annualized dividend yield constructed from the price and total return indices due to a differential tax treatment across the two series.[3] We checked our results by re-constructing the total return using the MSCI dividend yield dy^{12} with the price return and found them to be unchanged.

In the bottom panel of Table (1) we report summary statistics for a much longer sample of US excess returns and instruments using a quarterly frequency from 1935 to 1999. The stock returns are total returns (including reinvested dividends) on the S&P Composite Index obtained from Ibbotson Associates, and the dividend and earnings yields are from the *Security Price Index Record*, published by Standard & Poor's Statistical Service. The S&P yields are also constructed using dividends and earnings summed over the past year. The short rates are yields on 3 month T-bills. This dataset is similar to that used by Lamont (1998), but over a longer sample. Summary statistics are similar to the monthly MSCI data set, except the mean of the short rate is lower (4.1% versus 7.7%).

2.2 A Simple Present Value Model

Modern predictability regressions consider the predictability of excess stock returns, the return on equity over and above the return on a nominally risk-free security of the same holding period, which is known one period in advance. Since there is substantial time-variation in interest rates, and it is likely that expected stock returns vary with the interest rate, the hypothesis of interest is the constancy of the conditional equity premium, not the constancy of expected stock returns.

[3] For a discussion on how taxes are treated see MSCI Methodology and Index Policy, 1999. The MSCI dividend yield dy^{12} series and the implied annualized dividend yield from the price and total return indices are identical for the US and Japan, but differ for the UK, France and Germany.

Building present value models that imply constant excess stock returns, but allow time-varying interest rates is a non-trivial matter. Most of the recent work on present value models with time-varying discount rates builds on Campbell and Shiller (1988b) who, by linearizing returns around steady state log price dividend ratios, obtain a tractable linear present value model in which it is straightforward to impose the constancy of expected excess returns while allowing for variation in interest rates. More recently, the term structure models in the affine class (see Duffie and Kan (1996)) have been applied to stock pricing to yield tractable pricing equations in many settings without linearization (see Ang and Liu (2001) and Bekaert and Grenadier (2000)). We deviate from this literature by presenting a model for price earnings ratios.

Stock returns from time t to $t+1$ can be decomposed as:

$$Y_{t+1} \equiv \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{EA_{t+1}}{EA_t} \frac{PE_{t+1} + PO_{t+1}}{PE_t}$$

where P_t is the stock price at time t, D_{t+1} is the dividend paid at time $t+1$, $PO_t = D_t/EA_t$ is the payout ratio of dividends D_t to earnings EA_t, and PE_t is the price-earnings ratio. Defining g_t as log growth in earnings $g_t = \log(EA_t/EA_{t-1})$ and po_t as the log payout ratio $po_t = \log(PO_t)$, we have:

$$Y_{t+1} = \exp(g_{t+1}) \cdot \frac{PE_{t+1} + \exp(po_{t+1})}{PE_t} \tag{1}$$

In this pricing framework, we have decomposed dividend growth into earnings growth and a payout ratio. From a theoretical perspective, dividend growth should suffice to price stocks, but our decomposition may yield more accurate pricing formulas in finite samples. First, in finite samples using dividends may be problematic, since they are often manipulated, smoothed, or set to zero, making them poor indicators of the true value-relevant cashflows in the future. It is no surprise that in the real world analysts almost entirely focus on earnings growth. Second, the decomposition simply increases the information set for prediction, and includes a model that features only dividend growth as a special case. Finally, even a simple autoregressive model for payout ratios and earnings growth implies a more intricate ARMA process for dividend growth that would be difficult to estimate from dividend data alone (see Bansal and Yaron (2000) for evidence of an MA component in US dividend growth).

The model has three state variables, the short rate r_t, log earnings growth g_t and the log payout ratio po_t. Denote $X_t = (r_t\, g_t\, po_t)'$ which we assume to follow a first-order Vector Autoregression:

$$X_t = \mu + AX_{t-1} + \epsilon_t \tag{2}$$

where $\epsilon_t \sim IID\ N(0, \Sigma)$. To price equity, we use the Dividend Discount Model:

$$P_t = \mathrm{E}_t \left[\sum_{i=1}^{\infty} \Pi_{t+i} D_{t+i} \right], \tag{3}$$

where Π_{t+i} is the stochastic discount factor applying to payoffs at time $t+i$. To ensure the absence of arbitrage we model the one-period log pricing kernel m_{t+1}, such that:

$$m_{t+1} = -r_t - \frac{1}{2}\gamma' \Sigma \gamma + \gamma' \epsilon_{t+1}, \tag{4}$$

where γ is a 3×1 vector containing the prices of risk and the discount factor can be written as:

$$\Pi_{t+i} = \exp\left(\sum_{j=1}^{i} m_{t+j} \right).$$

We also impose conditions on the parameters $\theta = [\mu', \text{vec}(A)', \text{vech}(\Sigma)', \gamma']'$ so that the transversality condition

$$\lim_{i \to \infty} \Pi_{t+i} P_{t+i} = 0$$

is satisfied, ruling out bubbles.

Proposition 2.1 *In this economy the price-earnings ratio is given by:*

$$PE_t \equiv \frac{P_t}{EA_t} = \sum_{i=1}^{\infty} \exp(a(i) + b(i)' X_t) \tag{5}$$

where $a(i)$ and $b(i)$ are given by the recursive relations:

$$\begin{aligned} a(i+1) &= a(i) + (e_2 + b(i))'\mu + (e_2 + b(i))'\Sigma\gamma + \frac{1}{2}(e_2 + b(i))'\Sigma(e_2 + b(i)) \\ b(i+1) &= -e_1 + A'(e_2 + b(i)) \end{aligned} \tag{6}$$

with starting values:

$$\begin{aligned} a(1) &= (e_2 + e_3)'(\mu + \Sigma\gamma) + \frac{1}{2}(e_2 + e_3)'\Sigma(e_2 + e_3) \\ b(1) &= -e_1 + A'(e_2 + e_3) \end{aligned} \tag{7}$$

where e_i is a 3×1 vector of zeros with a 1 in the ith place.

The $b(i+1)$ term reveals that an increase in the short rate decreases the price earnings ratio, unless it simultaneously predicts higher earnings growth in the future with a feedback coefficient larger than 1 ($A_{21} > 1$, where subscripts denote matrix elements). Similarly, if earnings growth shows positive persistence, higher earnings growth leads, ceteris paribus, to higher price earnings ratios.

Since only dividend growth can be priced, we must link the price of risk of dividend growth to the price of risk of the payout ratio and earnings growth. Using the fact that log dividend growth g^d_{t+1} can be written as $g^d_{t+1} = \Delta po_{t+1} + g_{t+1}$, this is accomplished by setting:

$$\text{cov}_t(m_{t+1}, g^d_{t+1}) = \text{cov}_t(m_{t+1}, \Delta po_{t+1} + g_{t+1}).$$

Observation 2.1 *For the three-factor $X_t = (r_t \, g_t \, po_t)'$ system to yield the same pricing relation as a two-factor model using $(r_t \, g^d_t)'$ the prices of risk of dividend growth, earnings growth and log payout ratio (γ_d, γ_e and γ_{po} respectively) must satisfy:*

$$\gamma_d = \gamma_e = \gamma_{po} \tag{8}$$

Intuitively, both an increase in earnings growth or an increase in the payout ratio increase dividend growth by the same amount. Hence the price of risk ought to be the same for earnings growth and log payout. We can then impose the constraint $\gamma_e = \gamma_{po} = \bar{\gamma}$.

There is a large class of models in this system where appropriate parameter restrictions ensure that the expected excess simple return and volatility is a constant multiple of the gross short rate:

Corollary 2.1 *Denote the companion matrix $A = (A_{ij})$ and the conditional covariance matrix $\Sigma = (\Sigma_{ij})$. Let the vector $b(i) = b = (b_1 \, b_2 \, b_3)'$ in equation (6) be constant for all i and be given by:*

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} A_{21} + A_{31} - 1 \\ A_{22} + A_{32} \\ A_{32} + A_{33} \end{pmatrix}. \tag{9}$$

Then if the following restrictions are satisfied:

$$\begin{aligned} b'\Sigma\gamma &= e_3'\Sigma\gamma \\ A_{11} &= -\frac{A_{21}(A_{22} + A_{32}) + A_{31}(A_{23} + A_{33} - 1)}{A_{21} + A_{31} - 1} \\ A_{12} &= -\frac{A_{22}(A_{22} + A_{32}) + A_{32}(A_{23} + A_{33} - 1)}{A_{21} + A_{31} - 1} \\ A_{13} &= -\frac{A_{23}(A_{22} + A_{32}) + A_{33}(A_{23} + A_{33} - 1)}{A_{21} + A_{31} - 1} \\ (e_2 + e_3)'\Sigma\gamma &< -\left(\frac{1}{2}(e_2 + b)'\Sigma(e_2 + b) + (e_2 + b)'\mu\right), \end{aligned} \tag{10}$$

the conditional expected simple risk premium is a multiple of the gross short rate and given by:

$$\mathrm{E}_t[Y_{t+1} - \exp(r_t)] = (c-1) \cdot \exp(r_t) \tag{11}$$

where $c = \exp(-(e_2+e_3)'\Sigma\gamma)$*. The unconditional expected simple return is given by:*

$$\mathrm{E}[Y_{t+1} - \exp(r_t)] = (c-1) \cdot \exp(\bar{\mu}_r + \frac{1}{2}\bar{\sigma}_r^2) \tag{12}$$

where $\bar{\mu}_r$ *and* $\bar{\sigma}_r^2$ *are the unconditional mean and variance of* r_t *respectively.*

The first restriction in equation (10) links the prices of risk γ, the covariance Σ and the pricing vector b. It is trivially satisfied if $\gamma = 0$, in which case $c = 1$ and there is no risk premium. If $\gamma \neq 0$ and $b \neq e_3$, then this condition imposes a restriction on the relative prices of risk of the interest rate γ_r with the price of risk of dividend growth $\bar{\gamma}$, where $\gamma = (\gamma_r \, \bar{\gamma} \, \bar{\gamma})'$:

$$\frac{\gamma}{\bar{\gamma}_r} = -\frac{b_1\Sigma_{11} + b_2\Sigma_{21} + (b_3-1)\Sigma_{31}}{b_1\Sigma_{12} + b_2\Sigma_{22} + (b_3-1)\Sigma_{32} + b_1\Sigma_{13} + b_2\Sigma_{23} + (b_3-1)\Sigma_{33}}.$$

The second to fourth restrictions in equation (10) impose non-linear restrictions on A, so that the first row of A, the dynamics of the short rate, is linked to the second and third rows of A, the dynamics of earnings growth and payout. Normally, when interest rates move away from their unconditional mean the resulting change in discount rates and prices would induce predictable components in returns. The restriction on the companion form engineers an opposite cashflow effect that neutralizes the price change induced by the change in the interest rate. The final restriction in equation (10) results from imposing transversality. This is satisifed if γ is sufficiently negative.

In Corollary 2.1 the individual components of the equity return in equation (1), like earnings growth and payout, may be predictable. The restrictions of equation (10) ensure that the expected capital gain and dividend income, in excess of the risk-free return, move in such a way to exactly off-set each other so that the simple expected excess return is a multiple of the nominal rate. Hence, a regression of $Y_{t+1} - \exp(r_t)$ on the nominal rate would actually yield a positive coefficient equal to $c-1$. However, the scaled expected return, $\mathrm{E}_t[Y_{t+1}/\exp(r_t)]$ is constant and equal to c. The constant c is a function of the correlation between dividend growth innovations (the sum of the earnings growth and payout ratio innovations) with the pricing kernel. The predictability regressions typically run in the literature do not correspond to any of these two concepts, since they use log returns, $\tilde{y}_{t+1} \equiv \log(Y_{t+1}) - r_t$. It is straightforward to show that up to second order terms, the expected log risk premium is constant in this homoskedastic model.

Corollary 2.1 nests several important cases. For example, if $b = (-1, 1, 1)'$ and $A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{11} & A_{12} & A_{13} \\ -A_{11} & 1-A_{12} & 1-A_{13} \end{pmatrix}$, so that the short rate and earnings growth have the same conditional expected mean, then the model is similar to Goldstein and Zapatero (1996) except with a stochastic payout ratio. Another important special case is if $b = e_3$ with $A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ 1-A_{21} & -A_{22} & 1-A_{23} \end{pmatrix}$. Under this economy, conditional expected dividend growth is equal to a constant $\mu_g + \mu_{po}$ plus the current short rate ($\mathrm{E}_t[g^d_{t+1}] = \mathrm{E}_t[g_{t+1}] + \mathrm{E}_t[po_{t+1}] - po_t = \mu_g + \mu_{po} + r_t$). We select this particular simple model case as the null model for Monte Carlo purposes. Here, the true monthly de-seasonalized (but empirically unobserved) price-dividend ratio is constant but the monthly price-earnings ratio moves through time. However, the observed dividend and earnings yields constructed using dividends or earnings summed over the past 12 months are stochastic.

2.3 Predictability Regressions

The main regression we consider is:

$$\tilde{y}_{t+k} = \alpha + \beta_k' z_t + \epsilon_{t+k,k} \tag{13}$$

where

$$\tilde{y}_{t+k} = (12/k)((y_{t+1} - r_t) + \cdots + (y_{t+k} - r_{t+k-1}))$$

is the annualized k-month excess return for the aggregate stock market, and $y_{t+1} - r_t$ is the excess 1 month return from time t to $t+1$. All returns are continuously compounded. Our Present Value Model implies that $\mathrm{E}_t(\tilde{y}^i_{t+k})$ is constant and hence β^k_i is zero for all k. The error term $\epsilon_{t+k,k}$ follows a $MA(k-1)$ process under the null of no predictability because of overlapping observations. The instruments z_t consist of the log dividend yield dy^{12}_t, the log earnings yield ey^{12}_t, and continuously compounded monthly short rate r_t.

The log payout ratio po^{12}_t is linearly related to the dividend yield and earnings yield $po^{12}_t = dy^{12}_t - ey^{12}_t$. The three predictive instruments are endogenous instruments in our Present Value Model. However, they should capture the predictability present under the null of the present value model, because there is a one-to-one (non-linear) mapping between earnings yield, dividend yield and the short rate and our three state variables. One reason variables such as dividend yields may predict future returns more generally is the presence of price in the denominator. On the one hand, the presence of price on both sides of the regression may worsen small sample biases in the regressions (Goetzmann and Jorion (1993)). On the other hand, since price reflects all information about future expected returns and cashflow growth rates, its presence may capture genuine predictability. If the Present Value Model we present is truth, price would not be necessary to capture time-variation in expected returns, and all information should be captured by the three state variables, or transformations of them.

In a globally integrated world, predictability is likely also to extend across borders. In Section 6, following Bekaert and Hodrick (1992), we consider cross-country regressions of the form:

$$\tilde{y}^i_{t+k}\ (e^i_{t+k}) = \alpha_i + \beta_i' z_t + u^i_{t+k} \tag{14}$$

where $\tilde{y}^i_{t+k}$ are k-period annualized excess equity returns in local currency for country i, and e^i_{t+k} are k-period annualized exchange rate returns USD per foreign currency for foreign country i. The instruments in z_t we consider are log dividend yields for the US, log earnings yields for the US, and one-month risk-free rates for the US, and the foreign country counterparts of these variables. Note that as we use continuously compounded returns $\tilde{y}^{usd}_{t+1} = \tilde{y}^i_{t+1} + e^i_{t+1}$.

The regressions in equations (13) and (14) can be estimated by OLS. We consider three estimators of the standard errors. First, OLS standard errors are appropriate if there is no serial correlation of the error term and the error terms are homoskedastic. We use them as a benchmark even when $k > 1$, in which case they likely underestimate the true sampling error. Second, to account for the overlap in the residuals for $k > 1$ and to capture potential heteroskedasticity in returns, we use a heteroskedastic extension of Hansen and Hodrick (1980) standard errors. Using GMM the parameters $\theta = (\alpha^i\,(\beta^i_k)')'$ in equation (13) have an asymptotic distribution (see Hodrick (1992)) $\sqrt{T}(\hat{\theta} - \theta) \overset{a}{\sim} N(0, \Omega)$ where $\Omega = Z_0^{-1} S_0 Z_0^{-1}$, $Z_0 = \mathrm{E}(x_t x_t')$, $x_t = (1 z_t')'$ and S_0 is estimated by:

$$\hat{S}_0 = C(0) + \sum_{j=1}^{k-1} [C(j) + C(j)'] \tag{15}$$

where

$$C(j) = \frac{1}{T} \sum_{t=j+1}^{T} (w_{t+k} w'_{t+k-j})$$

and $w_{t+k} = \epsilon_{t+k,k} x_t$. This estimator of S_0 is not guaranteed to be positive semi-definite. If it is not, we use a Newey-West (1987) estimate of S_0 with k lags. We refer to these standard errors as Robust Hansen-Hodrick (1980) standard errors.

Finally, we report what we call Hodrick (1992) standard errors. This estimator exploits covariance stationarity to remove the overlapping nature of the error terms in the standard error computation. Instead of summing $\epsilon_{t+k,k}$ into the future to obtain an estimate of S_0, Hodrick (1992) sums $x_t x'_{t-j}$ into the past:

$$\hat{S}_0 = \frac{1}{T} \sum_{t=k}^{T} wk_t wk_t' \tag{16}$$

where

$$wk_t = \epsilon_{t+1,1}\left(\sum_{i=0}^{k-1} x_{t-i}\right).$$

In our Monte Carlo analysis we run a horse race between these three estimators. We find Hodrick standard errors to be far superior, and most of our results exclusively focus on t-statistics computed with the Hodrick standard errors. Readers not interested in the details of the Monte Carlo analysis can skip Section 3, although we feel that it contains some important results.

Apart from running univariate regressions, we are mindful of Richardson's (1993) critique of predictability tests testing for only one particular horizon k and we provide a number of joint tests across horizons. To test if the predictability coefficients are statistically significant across n horizons $k_1 \ldots k_n$ we set up the simultaneous equations:

$$\begin{aligned} \tilde{y}_{t+k_1} &= \alpha_{k_1} + \beta'_{k_1} z_t + u_{t+k_1} \\ &\vdots \\ \tilde{y}_{t+k_n} &= \alpha_{k_n} + \beta'_{k_n} z_t + u_{t+k_n} \end{aligned} \tag{17}$$

Denote the vector of coefficients $\beta = (\alpha_{k_1}\beta'_{k_1} \ldots \alpha_{k_n}\beta'_{k_n})'$. In practice, an estimate $\hat{\beta}$ of β is obtained by performing OLS on each equation. Appendix D details the construction of joint tests across horizons accomodating Hodrick standard errors.

When we consider predictability in multiple countries, we also provide joint tests of no predictability across countries and we estimate pooled coefficients across countries. Such a pooled estimation mitigates the data mining problem plaguing US data and increases efficiency and power under the null of no predictability. Here we estimate the system:

$$\tilde{y}^i_{t+k} = \alpha_i + \beta'_i z^i_t + u^i_{t+k} \tag{18}$$

for $i = 1 \ldots N$ countries, subject to the restriction $\beta_i = \bar{\beta}\ \forall i$, but imposing no restrictions on α_i across countries. We take i = US, UK, France, Germany, Japan. The econometrics underlying the pooled estimation is detailed in Appendix E.

3 Finite Sample Properties of Various Estimators

We calibrate our Present Value Model of the price-earnings ratio and use it as the DGP for a Monte Carlo analysis. We solely use US data for the Monte Carlo analyses, but the results are so clear-cut that there is little reason to suspect they would not extend to DGP's calibrated using data from other countries. All Monte Carlo experiments use 5,000 replications.

To perform the Monte Carlo analysis we must estimate the parameters (μ, A, Σ, γ). We proceed in two steps. First, we estimate the VAR parameters (μ, A, Σ). Second, we calibrate $\gamma = (\gamma_r \, \bar{\gamma} \, \bar{\gamma})'$ to fit the observed equity premium in the data. We set the price of interest rate risk to zero, ($\gamma_r = 0$) so that γ is a scalar. Appendix F fully describes the calibration procedure and the results. The estimation is complicated by the fact that our data uses earnings and dividends summed up over the past year but our model requires (unobserved) monthly earnings and dividends.

The null model matches the equity premium and almost matches the volatility of excess returns. Among the more interesting parameter estimates is the effect of the short rate on future earnings growth, which is significantly negative. The model does a poor job of matching dividend yield variability; the implied standard deviation of log dividend yields is 0.0861 compared to 0.3915 in the data. The model does a better job in matching earnings yield variability, but only matches a half of the observed variability (the implied (empirical) standard deviation of log earnings yields is 0.2056 (0.4059)).[4]

With the fully calibrated model we simulate 5000 samples of 299 observations to examine the empirical distribution of the various test statistics. We construct dividend yields dy^{12} and earnings yields ey^{12} using summed dividends and earnings over the past year as in the actual data:

$$\begin{aligned}\frac{D_t^{12}}{P_t} &= \frac{D_t + D_{t-1} + \cdots + D_{t-11}}{P_t} \\ \frac{EA_t^{12}}{P_t} &= \frac{EA_t + EA_{t-1} + \cdots + EA_{t-11}}{P_t}\end{aligned} \tag{19}$$

The earnings yield can be computed by:

$$\frac{EA_t^{12}}{P_t} = \frac{EA_t}{P_t} + \frac{EA_{t-1}}{P_{t-1}}\frac{P_{t-1}}{P_t} + \cdots + \frac{EA_{t-11}}{P_{t-11}}\frac{P_{t-11}}{P_t}, \tag{20}$$

where $PE_t = P_t/EA_t$ can be evaluated using Proposition 2.1, and

$$\frac{P_{t-i}}{P_t} = \frac{P_{t-i}}{EA_{t-i}}\frac{EA_{t-i}}{EA_t}\frac{EA_t}{P_t} = \frac{PE_{t-i}}{PE_t}\left[\exp(g_{t-i+1} + \cdots + g_t)\right]^{-1}, \tag{21}$$

allowing us to evaluate each term in equation (20). The dividend yield can be written as:

$$\begin{aligned}\frac{D_t^{12}}{P_t} &= \frac{D_t}{EA_t}\frac{EA_t}{P_t} + \frac{D_{t-1}}{EA_{t-1}}\frac{EA_{t-1}}{P_{t-1}}\frac{P_{t-1}}{P_t} + \cdots + \frac{D_{t-11}}{EA_{t-11}}\frac{EA_{t-11}}{P_{t-11}}\frac{P_{t-11}}{P_t} \\ &= \frac{PO_t}{PE_t} + \frac{PO_{t-1}}{PE_{t-1}}\frac{P_{t-1}}{P_t} + \cdots + \frac{PO_{t-11}}{PE_{t-11}}\frac{P_{t-11}}{P_t},\end{aligned} \tag{22}$$

[4] Naturally, the model would be rejected imposing restrictions from the levels of price-dividend or price-earnings ratios. For example, see Cliff (2001).

where $PO_t = \exp(po_t)$ and P_{t-i}/P_t can be evaluated using equation (21). The dividend and earnings yields used in the regressions are $dy_t^{12} = \log(D_t^{12}/P_t)$ and $ey_t^{12} = \log(EA_t^{12}/P_t)$ respectively.

In Table (2) we report the mean and standard deviation of the empirical distribution for the coefficients in the various regressions we consider. First, in the dividend regression, we find substantial upward bias. This is consistent with the bias one would find in a standard linear return, dividend yield system (see Stambaugh (1999)) and derives from the negative correlation between return and dividend yield innovations and the autocorrelation of dividend yields. Second, in the earnings regression we find a downward bias. Why might this be the case? In our price-earnings model with constant expected returns, the variation in returns is dominated by variation in earnings growth rates. The price-earnings ratio is not constant but dependent on the payout ratio. This implies that the earnings yield, even when summed over 12 periods, is likely to be primarily negatively correlated with the payout ratio, but in the DGP earnings growth rates and payout ratios are highly negatively correlated, so that returns and earnings yield innovations end up being positively correlated. This is not true for dividend yields, since the log dividend yield can be written as the sum of the log payout ratio (negatively correlated with the earnings growth rate and hence with return innovations) and the log earnings yield, and the first effect dominates. Note that the small sample coefficient in the dividend yield regression has the same positive sign that the dividend yield predictability literature finds (Fama and French (1988) and Hodrick (1992)), while the negative small sample coefficient on the earnings yield is what Lamont (1998) finds.

Third, in the bivariate regressions, the univariate biases are accentuated, with the bias on the dividend yield coefficient reaching 0.19 for $k = 1$. The biases decrease slightly with the horizon. Lamont (1998) finds positive signs on the dividend yield and negative signs on the earnings yield in a bivariate regression. Our biases are exactly the same sign as his result. Finally, in the trivariate regression, the biases for the earnings and dividend yield coefficients become larger still in absolute magnitude (0.24 for the dividend yield coefficient; -0.11 for the earnings yield coefficient). However, the bias on the short rate coefficient is negligible.

In Table (3) we examine the size properties of significance tests. We report empirical sizes for tests of size 10% and of size 5%, but we focus our discussion on the 5% tests. At $k = 1$, the univariate regressions display negligible size distortions, but for the bivariate and trivariate regressions, all tests slightly over-reject at asymptotic critical values and the empirical sizes exceed the nominal sizes. The worst distortion occurs for the dividend yield coefficient in the trivariate regression, with the nominal size being 8.6% for the OLS estimator, and 8.9% for the robust Hansen-Hodrick and Hodrick estimators (which coincide for $k = 1$). For longer horizons, the performance of the OLS estimator rapidly deteriorates with the empirical size exceeding 54% for a 5% test in the dividend regression and exceeding 72% in all other regressions at $k = 60$. Accounting for the overlap in the error terms using Hansen-Hodrick standard errors improves the small sample performance but not enough to yield a reliable test. The empirical size for a 5% test at $k = 60$ is at least 39.1% in the case of the dividend regression.

Table (3) shows that Hodrick standard errors are far superior to OLS and Hansen-Hodrick standard errors. The OLS and Hansen-Hodrick standard errors are too small; the former ignore the serial correlation in the error terms, whereas the latter under-estimate it, because the autocorrelation estimates are downward biased. For $k = 12$, there is negligible size distortion for the univariate regressions, whereas for the bivariate and trivariate regressions, the size distortion is at most 4.9% for the dividend yield coefficient in the trivariate regression (a 9.9% empirical size). Note that the Hodrick test now also over-rejects. For $k = 60$, the size distortions actually become smaller, with the worst size distortion occurring again for the dividend yield coefficient in the trivariate regression (a 7.9% empirical size for the 5% test). For the earnings yield regression, the Hodrick test is conservative with an empirical size of 3.6%. In summary, the Hodrick standard errors display overall far superior small sample properties, and using the asymptotic p-values is unlikely to dramatically affect statistical inference. For that reason, we report the Hodrick standard errors for all of our empirical tests.

4 Predictability in US Excess Stock Returns

Panel A of Table (4) contains our main results regarding the predictability of US excess stock returns. To allow comparison with Lamont (1998)'s article, we report univariate and bivariate yield regressions in addition to our main trivariate specification. We report t-statistics in parentheses based on Hodrick standard errors. For the full sample and looking over three horizons, the yield regressions produce no significant coefficients, although the dividend yield is a significant predictor of of excess stock returns in the bivariate regression at the 10% level, However, its sign is negative, not positive! Lamont, on the other hand, finds positive coefficients for both yield variables in univariate regressions, but a positive coefficient on dividend yields and a negative coefficient on the earnings yield in the bivariate regression. His main point is that the predictive power of the dividend yield stems from the role of dividends in capturing the permanent component of prices, whereas the negative coefficient on the earnings yield is due to earnings being a good measure of business conditions, which captures counter-cyclical risk aversion. According to Lamont, there is information in earnings and dividends over and above price. Unfortunately, our results appear inconsistent with this story. Only for $k = 60$ do the regression coefficients have the sign predicted by Lamont's analysis, but for long horizons Lamont finds that only price mattered. In the trivariate regression, the same sign pattern appears for the yield variables: negative on dividends and positive on earnings which reverses for $k = 60$.

Note that accounting for biases would not help recover the Lamont pattern. In the univariate regressions, the bias-corrected dividend yield coefficient would be even more negative since the bias is positive, and the negative bias in the earnings yield regression is too small to reverse the sign of the coefficients. For the bivariate and trivariate regressions, accounting for biases would make the pattern we observe (and which is opposite to what Lamont finds) even more striking.

Figure (1) shows the pattern over different horizons more clearly for both the univariate and bivariate regressions (the inclusion of the interest rate does not change the pattern very much). In both univariate regressions, the coefficients on the dividend or earnings yield turn increasingly more negative until around the 30 month horizon after which they start to increase without becoming positive. In the bivariate regression, we clearly see how the "Lamont-pattern" of positive dividend yield coefficients and negative earnings coefficients requires setting k equal to 50 months or higher. The right hand side of the three panels shows the corresponding t-statistics for OLS, Robust Hansen-Hodrick and Hodrick standard errors. Given the general lack of statistical significance, it is clearly pointless to try and interpret the sign changes. Interestingly, if we had relied on the OLS or Robust Hansen-Hodrick standard errors, we would have concluded there was significant predictive power in the dividend yield variable, which again drives home the importance of the simulation experiments in Section 3.

Goyal and Welch (1999) point out that the dividend yield predictability is not robust to the addition of the last decade of high returns coinciding with low dividend yields and is actually highly unstable in general. Therefore, Panel A of Table (4) reports results for shorter subsamples eliminating either the last 5 or the last 8 years in the sample. Whereas the sign of the coefficients now better corresponds to what Lamont (1998) finds, statistical significance is still lacking. Lettau and Ludvigson (2001) also find that the Lamont-pattern does not hold with late 1990's data with a quarterly sampling frequency.

The results described so far are pretty bleak if finding predictability were the objective of this study. However, there is yet another regressor in our fundamental regression, the short rate.[5] Whatever the sample period we use, the short rate enters significantly at the 99% level in the $k = 1$ regression and its impact on the equity premium is remarkably robust in terms

[5] If a detrended short rate is used instead of the level of the short rate, following Campbell (1991) and Hodrick (1992), the results are unaltered. If the regression is also augmented by dummy variables from Oct 1979 to Oct 1982 to account for the monetary targeting period, the short rate coefficient remains significant.

of magnitude across the various sample periods. A 1% increase in the annualized short rate decreases the equity premium by about 3.7%. The predictive power of the short rate dissipates quickly for longer horizons. The coefficient slowly becomes less negative and eventually even slightly positive.

We also conduct joint tests across regressors. Joint tests across the dy^{12} and ey^{12} regressors in the Lamont regression yield p-values of 0.2294, 0.3760 and 0.1788 for horizons $k = 1$, 12 and 60 respectively. Joint tests across dy^{12}, ey^{12} and r in the trivariate regression yield p-values of 0.0139, 0.3723 and 0.0985 for horizons $k = 1, 12, 60$. In this regression for $k = 1$, the dividend yield and earnings yield are also individually significant at the 5% level.

One problem with our results is that our sample is shorter than Lamont's. Panel B of Table (4) reports trivariate regression results using quarterly S&P data from 1935 or 1952 to the present. For the longest sample, we find no significant predictability at all. The coefficients on the yield variables are typically small but positive and have large standard errors. The short rate coefficient remains negative and its t-statistic is much higher than the t-statistics for the yield variables, especially at short horizons. The lack of significance for the short rate variable may be due to the fact that interest rate data are hard to interpret before the Treasury Accord, as short term interest rates were pegged by the Federal Reserve during the 1930's and 1940's. Once we restrict attention to post-1952 data, the short rate becomes highly significant with the expected negative sign at both the 1-month and 12 month horizon. The yield variables coefficients remain insignificant and the sign pattern we uncovered before (negative dividend yield coefficients and positive earnings yield coefficients) resurfaces. Table (4) clearly demonstrates that our results are not an artifact from focusing on post-1970 data.

Finally, following the lead of Richardson (1993), Table (5) reports tests of predictability over three horizons, $k = 1$, 12 and 60, simultaneously. The table also lists results for four other countries, which we discuss in the following section. If we now focus on the first column, the US results, we see that there is only strong evidence for yield predictability, if we consider the joint predictability of earnings yields and dividend yields in the trivariate specification. Despite the short-lived nature of the predictable patterns, the short rate still significantly (at the 5% level) predicts excess returns at the one month, 12 month and 5 year horizons.

5 Predictability of Excess Stock Returns in Five Countries

Our previous results suggest that the predictability patterns formerly found in US data appear not to be robust to the addition of the last few years of the 1990's and that statistical significance only occurs when the short rate is used as a predictor. It is conceivable that the lack of predictive power is simply a small sample phenomenon, due to the very special nature of the last decade for the US stock market. It is equally probable that the previous results of strong predictability and interesting predictability patterns are a statistical fluke. International evidence should help us sort out these two interpretations of the data. If we cannot confirm the previous predictability patterns for any countries, and if yield variables do not appear to predict stock returns in other counties, it seems likely that data mining and other statistical problems have led researchers in the US astray regarding the predictive power of the yield variables. We can also increase or decrease our confidence in the short rate as a robust predictor of equity returns using international stock return data. Finally, pooled estimation across countries can lead to powerful evidence on the robustness and magnitude of predictability patterns across countries.

We begin by summarizing the patterns in univariate and bivariate yield regressions. Figure (2) displays the dividend yield coefficients and their t-statistics using Hodrick standard errors The UK coefficient pattern is strikingly similar to that of the US but, as in the US, not a single t-statistic reaches higher than 2.00 and it is not surprising that the joint tests for $k = 1$, 12 and 60 in Table (5) fail to reject the null of no predictability. The coefficient patterns in the three other countries are more akin to these prevalent in the US in earlier samples, in that the coefficients increase with horizon, but in both France and Germany they start out negative. In Japan, the individual coefficient for dy^{12} is borderline significant in the univariate regression and the joint test across horizons in Table (5) rejects at the 10% level.

In Figure (3), we repeat the same graphs for the earnings yield regression. The earnings yield predictability coefficients again are similar for the US and the UK, following a reverse J-pattern. For France and Germany they have a U-shaped pattern, whereas they increase monotonically with the horizon for Japan. Individual significance again only occurs for certain horizons in Japan, but nowhere else. Joint tests in Table (5) fail to reject the null of no predictability in France and Germany but there is a borderline rejection in Japan (at the 5.7% level) and a rejection at the 5% level in the UK. Inspection of the UK coefficients in Figure (3) reveals that the $k = 1$, 12, 60 choice very fortunately avoids the lowest spot in the t-statistics curve and might somewhat overstate the true predictability. We conclude that the univariate yield predictability patterns are not terribly robust across countries and not very significant statistically.

Do we observe the Lamont pattern of positive dividend yield and negative earnings yield coefficients in international data? Figure (4) simply shows the coefficient patterns. Again (not reported), the individual t-statistics using Hodrick standard errors barely ever reach significance. Two countries, Japan and France, show a pattern somewhat reminiscent of the Lamont findings, with negative coefficients for the earnings variable and positive ones for the dividend yield at short horizons. In France, the coefficients further diverge as the horizon lengthens, whereas in Japan they converge. The UK pattern of the coefficients is similar to that in France but the initial earnings yield coefficient is actually slightly positive. In Germany, the sign pattern and its evolution over horizons matches the one we found for the US, with positive (negative) earnings yield (dividend yield) coefficients eventually switching sign, but the switch occurs much earlier than in the US. Joint tests across horizons reveal no significant rejections of the null of no predictability, except in the case of Japan, where earnings and dividend yields jointly predict stock returns at the 5% level (see Table (5)). However, this may be due to a multicollinearity problem in the regression for Japan (see below).

The coefficient patterns for the yield variables that we observe for the bivariate regressions qualitatively persist for the trivariate regressions (except for Japan) and hence we do not show them in a figure. However, the t-statistics are generally somewhat larger, resulting in joint rejections of the null of no predictive power for the yield variables in three countries, the US, the UK and Japan at the 1% level. So the yield variables appear to have some predictive power when considered together and over three horizons simultaneously. However, the pattern in coefficients we see is very different across countries and differs from what Lamont found, except in Japan.[6]

Figure (5) displays the coefficient patterns for the annualized short rate and its associated t-statistics in the trivariate regression. Strikingly, this coefficient pattern is much more robust across countries and a similar shape for the coefficient patterns appear for univariate regressions (not reported). For all countries, the one-month coefficient is negative between -3.73 for the US and -0.97 for Japan. For 4 of the 5 countries, the coefficient increases monotonically with horizon, leveling off at around 0.55 for the US, France, and Germany and at slightly less than zero for the UK (-.55). In Japan, the coefficient never reaches zero, but the horizon dependence is not monotonically increasing. The t-statistics are generally larger in absolute magnitude for short horizons with the exception again being Japan. At the one-month horizon the short rate coefficients are statistically significant only for the US and UK. When we pool across horizons in Table (5), only the US short rate retains significant predictive power at the 5% level, which is not surprising given the pattern of the coefficients and t-statistics in Figure (5). What is surprising is the very high p-value for Japan. Inspection of the graph reveals than the joint test happens to select two k's out of only a small set of k's that yield coefficients close to zero.

[6] For Japan, we observe a Lamont pattern at short horizons (less than 15 months) but at horizons longer than 15 months the earnings yield coefficients become positive. The dividend yield coefficients are less than the earnings yield coefficients between horizons 20 and 40 months. At long horizons (40-60 months) both dividend and earnings yields are positive, with the dividend yield coefficients greater than the earnings yield coefficients. This figure is available upon request.

We conclude that the international evidence on predictability does not support Lamont's findings regarding the predictability of earnings and dividend yields. There is only weak evidence in favor of it in three countries, and the international data do not reveal a consistent, interpretable data pattern. However, the short rate robustly predicts excess stock returns, but its effect is limited to short forecasting horizons (mostly one month). The short rate coefficient is not significantly different from zero for all countries. Table (5) reports joint tests of predictability for the three coefficients. The null of no predictability is rejected at the 10% level in Germany, and at the 5% level in the US and UK. It is also rejected at the 1% level in Japan which is surprising given the lack of significance of the individual coefficients. However, this is mainly due to multicolinearity in the regressions, and a near-singular covariance matrix of the regressor coefficients.

One way to come to more clear-cut conclusions regarding the magnitudes and significance of the coefficients is to pool the estimation across countries. Under the null of no predictability, the pooled estimation should enhance efficiency considerably given that the correlation of returns across countries is not very high. Unfortunately, we have to exclude Japan, because of the considerable difference in data coverage both in terms of sample size and in terms of variables (we use levels of earnings yields rather than log earnings yields, because of the presence of negative earnings in the Japanese series). Table (6) reports pooled predictability coefficients, t-statistics and joint tests. We also report a test of the over-identifying restrictions, described in Appendix E. For all of our specifications, this test fails to reject the restrictions imposed by equal coefficients across countries. In the Lamont regressions with only the yield variables, we fail to find statistical significance for the yield variables, both when tested individually and jointly. The coefficients get closer to significant levels at longer horizons. In terms of sign at $k = 1$, the dominant pattern appears to be negative dividend yield coefficients and positive earnings yield coefficients, a pattern opposite to that found by Lamont (1998).

In the trivariate system, both the dividend yield and earnings yield coefficients are positive for $k = 1$, but the earnings yield coefficient turns negative at longer horizons. Consequently, the Lamont pattern again fails to show. Moreover, none of the coefficients is individually significant, although the t-statistics increase with horizon. A joint test on the yield variables also fails to reject the null of zero coefficients. On the other hand, the short rate coefficient is -1.7334 at the one month horizon and significant at the 5% level. The coefficient increases to -0.34 at $k = 60$ but loses statistical significance. Joint tests across the three variables fail to reject the null of no predictability for all three horizons at the 5% level, but the test rejects at the 10% level for $k = 1$. We conclude that the only robust and significant predictor of excess stock returns in 5 countries appears to be the short rate.

6 Cross-Country Predictability

The previous discussion implicitly considered our 5 countries to be segmented markets and did not allow the possibility of cross-country influences. However, in an integrated market, global discount rates should price equity returns on all markets, and we may find common components in the predictable components of returns. To investigate this, we extend our trivariate regression to a 6 variable regression, looking at pairs of countries. This set-up is an extension of the regressions run by Bekaert and Hodrick (1992), who regress equity and foreign exchange returns on the dividend yields in two countries and the forward premium. Since the forward premium is the interest differential through Covered Interest Parity, our regression frees up an implicit constraint on the interest rate coefficients in the Bekaert and Hodrick regressions. Like Bekaert and Hodrick, we also investigate the predictability of foreign exchange returns. Our main contribution here is to examine the added role earnings yields may play.

Table (7) reports the main results for the predictability of equity excess returns by local and US instruments. A number of striking results emerge. First, there is not a single significant coefficient for the 12 and 60-month horizons, confirming once again that predictability is not a long-horizon phenomenon. Second, the only significant coefficients we report are US dividend yields significantly predicting returns in France at the 5% level and in Germany at the 1% level. Hence, the strongest predictability pattern is a cross-country effect. The sign of the coefficient is negative, as it is in the other countries. Moreover, the US earnings yield consistently carries a positive sign but fails to reach statistical significance. The local yield variables mostly have positive but insignificant coefficients. Third, the local short rate is no longer significant and increases in value, relative to its value in the domestic trivariate regression we studied before. For France and Germany, it even becomes positive. However, the US short rate enters with a negative sign in every regression and the magnitude of the coefficient is large, varying between -2.00 in Japan to -4.17 in Germany. Although the coefficients are never significant at the 5% level, the t-statistics are all 1.00 or larger.

These coefficient patterns suggest that cross-country predictability may be stronger than domestic predictability. This may be the case in an integrated world where perhaps shocks affecting global discount rates are best reflected in the instruments of the dominant stock market, the US. To examine this further, Table (8) reports p-values from a series of joint tests of predictability. The first set, labeled with US as ”Base,” concerns the regressions of Table (7). The base predictability column contains a test of the null of zero coefficients for the base instrument, in this case, the US instruments. The local predictability column reports p-values for tests of predictability using only the local instruments. This column is likely to confirm the results of the trivariate regressions we reported earlier. The US instruments jointly only significantly predict German excess returns. In the other sub-panels, we make other countries the base country. For example, in the second set we look at excess returns in the US, France, Germany and Japan and our predictor instruments comprise local and UK instruments.

We find a number of interesting significant predictability patterns in Table (8). First, no foreign country instruments predict US returns. However, in the presence of foreign instruments which have no predictive power, there is still significant predictive power of US domestic instruments, in particular the short rate, for US excess returns. Second, we find only two significant base country predictors: US instruments predict German returns, and German instruments predict Japanese returns. Finally, the local predictability results confirm the rejections we found for the US in Table (5), no matter which foreign instruments are included in the regression. For the UK, we no longer reject, which may reflect a loss of power due to the introduction of three new regressors. For Japan, we also no longer reject, which simply indicates that the inclusion of the foreign instruments resolved the singularity problem we faced with the regular estimation.

The results in Table (8) may be weak because of the inclusion of too many highly correlated regressors. Therefore, Table (9) reports predictability results using only US instruments, including a pooled estimation. The results are indeed stronger than what we reported in Table (7). First, long-horizon predictability remains rather weak to non-existent. Second, the yield variables retain their sign patterns (a reverse Lamont pattern) and are now significant in both France and Germany. The US short rate consistently has a negative sign and is now significant in the UK and Germany. Overall, the US instruments predict excess equity returns in these countries better than the local instruments! When we pool the estimation across countries, we find very strong results, with all three coefficients being significant at the 1% level. A 1% increase in the US dividend yield reduces the equity premium in other countries by about 50 basis points, an increase in the earnings yield increases international risk premiums by about 65 basis points, whereas an increase in the US short rate of 1% decreases the equity premium by almost 3.5%. It may be that US factors dominate global discount rates and that this leads to the observed pattern, but it seems hard to come up with an international asset-pricing model that would explain these predictability patterns.

Such an international model would also have to capture predictability patterns in exchange rate returns. We measure the exchange rate return for a US based investor, using the logarithmic exchange rate change (in dollars per foreign currency) plus the foreign-US interest rate differential. This return is the topic of the vast literature on the Unbiasedness Hypothesis in international finance. If no instruments predict this return, the interest differential or forward premium is an unbiased predictor of future exchange rate changes. Whereas earlier work finds very strong rejections of this hypothesis, recent tests yield weaker results (see Bekaert and Hodrick (2001)). In Table (10) we report regressions of foreign exchange returns on the US instruments and the instruments of the currency's country. Since France and Germany now share a common currency as of 1 January, 1999, and were included in the EMS during the 1990's we include only the US Dollar-Deutsch Mark exchange rate.

In the Unbiasedness literature it is customary to regress foreign exchange returns or exchange rate changes onto the forward premium or interest differential, an exercise which typically results in strong negative slope coefficients. In our framework, this would correspond to finding negative coefficients on the US interest rate and positive ones on the foreign interest rate. This pattern is only valid for the pound; in the other countries the US interest rate enters with a positive sign and is not statistically significantly different from zero. In the UK, both interest variables are significantly different from zero, whereas the only other significant interest rate coefficient is the Japanese interest rate for the yen equation. Perhaps surprisingly, some of the yield variables do seem to have predictive power for foreign exchange returns, but no clear pattern emerges. For example, the US earnings yield is only significant in the pound equation, and then only for $k = 12$, whereas the dividend yield is significant for the Deutsche Mark equation at both $k = 1$ and $k = 12$. Local instruments are also significant. For example, UK dividend yields predict pound foreign exchange returns, and German dividend yields predict mark foreign exchange rates. Perhaps such complex patterns are not surprising in a globally integrated world. An international pricing model typically requires the exchange rate change to be the difference of the two pricing kernels in the two countries, so that factors driving equity prices in both countries may affect exchange rates as well.

To obtain more powerful tests, we also conduct a number of joint tests in the bottom panel of Table (10) across horizons $k = 1$, 12 and 60 months. For all of our tests, we find strong rejections of the null of no predictability for the Japanese foreign exchange returns, but these are hard to interpret because of the collinearity problems that plague the covariance matrix in this case. Therefore we focus our discussion on the pound and mark returns. First, we contrast the predictive power of local versus US instruments. We fail to find significant rejections of the null of no predictability in both cases, but local instruments seem to have more predictive power than US instruments. Second, we contrast the predictive power of the interest rate instruments, with the predictive power of the yield variables.[7] Surprisingly, the yield variables are significant at the 5% level in the pound return regression and at the 1% level in the mark return regression, but the interest rate variables are not significant. Whereas the literature has typically focused on

[7] Bekaert and Hodrick (1992) and Bauer (2001) also predict foreign exchange returns with dividend yield variables.

interest rate instruments to predict returns, they do not appear strong predictors, relative to yield variables. Third, joint tests strongly reject the null of no predictability in both regressions. We conclude that for foreign exchange predictability returns, there is strong predictability by yield variables but not by interest rates, the instruments used in the standard literature.

7 Cashflow Predictability

The results in this paper overturn many of the conventional, well-accepted results regarding the predictive power of the dividend and earnings yield for stock returns. As is true in our present value model, the dividend yield is a particularly natural predictor for stock returns. The basic equation underlying any present value model with a transversality condition imposed is:

$$\frac{P_t}{D_t} = \mathrm{E}_t \left[\sum_{i=1}^{\infty} \exp \left(\sum_{j=1}^{i} (-\delta_{t+j} + g^d_{t+j}) \right) \right], \tag{23}$$

where δ is the discount rate and g^d is dividend growth. In our model, the latter variable can be decomposed into earnings growth and the change in the payout ratio. Since the price dividend ratio varies through time and is in fact quite variable, the infinite sum on the right hand side must show predictable time-variation as well. In particular, low price-dividend ratios today (or high dividend yields) imply that either discount rates in the future are high, future cashflow growth rates are low, or both. Cochrane (1992) provides a variance decomposition of the variability of the price-dividend ratio based on equation (23) and finds that most of the variation in the price dividend ratio comes from variation in discount rates. The underlying reason for this result is that he does not find much predictable variation in dividend growth rates.

In this section, we ask whether earnings and dividend growth rates are predictable by our instruments. If they are not, it must be the case that linear regressions simply have little information on the predictive power of dividend yields and earnings yields for stock returns and cash flow growth rates. If they are, it might be possible to reconcile our weak return predictability results with a world where price dividend ratios and price earnings ratios show substantial variation over time.

Table (11) repeats the regressions of Table (4) for the US, but replaces returns with either dividend growth rates or earnings growth rates. The earnings growth rates are computed as follows:

$$g_t^{12} = \log(ey_t^{12}/ey_{t-1}^{12} \times P_t/P_{t-1}), \tag{24}$$

with a similar expression replacing ey_t^{12} with dy_t^{12} for dividend growth rates.[8]

Let us first focus on the univariate regressions with the dividend or earnings yield as an instrument. For the sample using MSCI data, dividend yields fail to forecast future dividend growth. However, we find earnings growth to be predictable by the yield variables at long horizons. When we start the sample in 1952 and use quarterly S&P data, this predictability of earnings growth becomes stronger and more prevalent at the shorter horizons. The earnings yield now predicts dividend growth at short horizons. Overall, higher yields forecast lower earnings growth and higher dividend growth. The former result is consistent with the intuition in equation (23), where high yields may forecast future low cashflow growth, proxied by earnings growth. The latter result may be due to dividend growth being a poor proxy for cashflow growth.

For the trivariate regressions, the monthly MSCI data sample does not display strong predictability, except for the earnings yield predicting dividend growth at short horizons and the short rate predicting earnings growth at the 12 month horizon. Nevertheless, joint tests over the three variables reject the null of no predictability of dividend or earnings growth at the 1% level for $k = 12$, $k = 60$ and joint across all horizons. The quarterly S&P data replicate the coefficient patterns of the MSCI data, but with generally smaller standard errors, especially for the dividend growth regressions. Dividend and earnings yields now significantly predict dividend growth at all horizons. This time, higher dividend yields signal lower future dividend growth, and higher earnings yield signal negative dividend growth at longer horizons. The predictability of earnings growth is much less strong and the short rate does not significantly predict cash flows although the t-stats are often above 1.[9] The only case for which a joint test over the three variables does not reject is for earnings growth with a $k = 60$ horizon.

Table (12) repeats the exercise for the trivariate regressions but pools data over 4 countries (the US, the UK, France and Germany). Dividend growth appears predictable only by the two yield variables (with the same sign patterns as in Table (11)) for $k = 1$. A joint test across the three instruments rejects at the 5% level for $k = 1$. However, what is particularly striking is the strong predictability of earnings growth. High dividend yields signal high future earnings growth; high earnings yields signal low future earnings growth rates. It is conceivable that the first effect is a genuine price effect (higher prices in response to predicted rises in future earnings), whereas the second finding may reflect mean reversion in earnings growth. Joint tests of predictability reject at the 1% level for all three horizons. Moreover, the coefficient pattern is very close to the pattern uncovered for the US, making the results economically robust across

[8] Note that the one-period log growth rate of earnings represents $\exp(g_{t+1}^{12}) = (EA_{t+1} + \cdots + EA_{t-11})/(EA_t + \cdots + EA_{t-11})$.

[9] Nissim and Penman (2001) find significant predictability of the earnings of individual firms by short rates.

countries.

We conclude that yield variables do indeed predict future cash flows, but this predictability becomes visible only by focusing on earnings growth rates and international stock returns data, which have not been the focus of previous research in this area.

8 Conclusions

The predictable components in equity returns uncovered in empirical work over the last 20 years have had a dramatic effect on finance research. Theoretical research on equilibrium models uses the predictability evidence as a stylized fact to be matched. The partial equilibrium dynamic asset allocation literature investigates the impact of the predictability on hedging demands. Much of the focus has been on the predictive prowess of the dividend yield, especially at long horizons, but Lamont (1998) shows that earnings yields have independent predictive power. In this article, we pose the question whether this predictability is real. After carefully accounting for small sample properties of standard tests, our answer is surprising but important. We show that the standard predictability patterns are not statistically significant, and not robust across countries or sample periods. Moreover, there is no evidence of long-horizon predictability in any of the 5 countries we examine. In this sense, the predictability that has been the focus of most recent finance research is simply not there.

Nevertheless, we do find that stock returns are predictable, calling for a re-focus of the predictability debate in four directions. First, our results suggest that predictability is mainly a short-horizon, not a long-horizon, phenomenon. Second, the strongest predictability comes from the short rate and not from yield variables with price in the denominator. The result that the short rate predicts equity returns goes back to at least Fama and Schwert (1977), but somehow recent research has failed to address what might account for this predictability and has mostly focused on dividend yield predictability. Third, there are tantalizing cross-country predictability patterns that appear stronger than domestic predictability patterns. The emergence of a globally integrated capital market over the last 20 years should refocus research towards determinants of global discount rates. Finally, we demonstrate that dividend and earnings yields have predictive power for forecasting future cashflow growth rates. Hence, a potentially important source of variation in price-earnings and price-dividend ratios is the predictable component in earnings growth rates.

We hope that our results, in the short run, affect the asset allocation literature, which often has taken predictability of the dividend yield variable as given and generally ignores the variation and predictive power of the short rate on stock returns. In the longer run, we hope our results stimulate research on theoretical models that might explain the predictability patterns we demonstrate, particularly short rate predictability at short horizons and across countries.

Appendix

A Proof of Proposition 2.1

From the Dividend Discount Model we can write:

$$\begin{aligned}
\frac{P_t}{EA_t} &= \mathrm{E}_t\left[\sum_{i=1}^{\infty}\Pi_{t+i}\frac{EA_{t+i}}{EA_t}\frac{D_{t+i}}{EA_{t+i}}\right] \\
PE_t &= \mathrm{E}_t\left[\sum_{i=1}^{\infty}\exp\left(\sum_{j=1}^{i}m_{t+j}\right)\exp\left(\sum_{j=1}^{i}g_{t+j}\right)\exp(po_{t+i})\right] \\
&= \mathrm{E}_t\left[\sum_{i=1}^{\infty}\Pi_{t+i}G_{t+i}PO_{t+i}\right] \qquad \text{(A-1)}
\end{aligned}$$

where $G_{t+i} = \exp\left(\sum_{j=1}^{i} g_{t+j}\right)$ and $PO_{t+i} = \exp(po_{t+i})$.

We claim that:

$$\mathrm{E}_0[\Pi_i G_i PO_i] = \exp(a(i) + b(i)'X_0), \qquad \text{(A-2)}$$

which we show by induction.

The initial conditions are given by:

$$\begin{aligned}
\mathrm{E}_0[\Pi_1 G_1 PO_1] &= \mathrm{E}_0[\exp(-r_0 - \frac{1}{2}\gamma'\Sigma\gamma + \gamma'\epsilon_1)\exp(g_1)\exp(po_1)] \\
&= \exp(-e_1'X_0 - \frac{1}{2}\gamma'\Sigma\gamma)\mathrm{E}_0[\exp(\gamma'\epsilon_1 + (e_2+e_3)'(\mu + AX_0 + \epsilon_1)] \\
&= \exp(a(1) + b(1)'X_0) \qquad \text{(A-3)}
\end{aligned}$$

where

$$a(1) = (e_2+e_3)'(\mu + \Sigma\gamma) + \frac{1}{2}(e_2+e_3)'\Sigma(e_2+e_3)$$

and

$$b(1) = -e_1 + A'(e_2+e_3)$$

To prove the recursive relation, assume this relation holds for i. Then using iterative expectations:

$$\begin{aligned}
\mathrm{E}_0[\Pi_{i+1}G_{i+1}PO_{i+1}] &= \mathrm{E}_0\left[\exp(m_1+g_1)\cdot\mathrm{E}_1\left[\exp\left(\sum_{j=1}^{i}m_{1+j}+g_{1+j}\right)\exp(po_{i+1})\right]\right] \\
&= \mathrm{E}_0\left[\exp(-r_0 - \frac{1}{2}\gamma'\Sigma\gamma + \gamma'\epsilon_1)\exp(e_2'X_1)\exp(a(i)+b(i)'X_1)\right] \\
&= \exp(-e_1'X_0 - \frac{1}{2}\gamma'\Sigma\gamma + a(i))\mathrm{E}_0[\exp(\gamma'\epsilon_1 + (e_2+b(i))'(\mu + AX_0 + \epsilon_1)] \\
&= \exp(a(i+1) + b(i+1)'X_0) \qquad \text{(A-4)}
\end{aligned}$$

where

$$a(i+1) = a(i) + (e_2+b(i))'\mu + (e_2+b(i))'\Sigma\gamma + \frac{1}{2}(e_2+b(i))'\Sigma(e_2+b(i))$$

and

$$b(i+1) = -e_1 + A'(e_2+b(i)).$$

Hence the price-earnings ratio is given by:

$$\begin{aligned}
PE_t &= \sum_{i=1}^{\infty}\mathrm{E}_t\left[\Pi_{t+i}G_{t+i}PO_{t+i}\right] \\
&= \sum_{i=1}^{\infty}\exp(a(i) + b(i)'X_t) \qquad \text{(A-5)}
\end{aligned}$$

B Proof of Observation 2.1

We would like pricing kernel variability with earnings growth and log payout to be the same if were to use dividend growth. Denote $\xi_{t+1} = -\frac{1}{2}\gamma'\Sigma\gamma + \gamma'\epsilon_{t+1}$. In particular, we would like $\text{cov}_t(g^d_{t+1}, -\xi_{t+1})$ in a system with state variables $(r_t\ g^d_t)'$ as $\text{cov}_t(\Delta po_{t+1} + g_{t+1}, -\xi_{t+1})$ in our system with state variables $(r_t\ g_t\ po_t)'$. We note that

$$\text{cov}_t(\Delta po_{t+1} + g_{t+1}, -\xi_{t+1}) = -(e_2 + e_3)'\Sigma\gamma.$$

Denoting Σ^d as the 2×2 covariance matrix under the dividend growth system, and γ^d as the 2×1 vector of prices of risk under this system, then

$$\text{cov}_t(g^d_{t+1}, -\xi_{t+1}) = -e_2'\Sigma^d\gamma^d$$

Imposing equality between these two expressions we have the following relation:

$$\gamma_r(\sigma_{rg} + \sigma_{r,po}) + \gamma_g(\sigma^2_g + \sigma_{g,po}) + \gamma_{po}(\sigma_{g,po} + \sigma^2_{po}) = \gamma^d_r\sigma_{r,g^d} + \gamma^d_{g^d}\sigma^2_{g^d} \tag{B-1}$$

where $\gamma = (\gamma_r\ \gamma_g\ \gamma_{po})'$ from our three-factor earnings growth and log payout system and $\gamma^d = (\gamma^d_r\ \gamma^d_{g^d})'$ from a dividend growth system. Note that

$$\sigma_{r,g^d} = \text{cov}_t(r_{t+1}, \Delta po_{t+1} + g_{t+1}) = \sigma_{r,po} + \sigma_{r,g}$$

and

$$\sigma^2_d = \text{cov}_t(\Delta po_{t+1} + g_{t+1}, \Delta po_{t+1} + g_{t+1}) = \sigma^2_{po} + 2\sigma_{g,po} + \sigma^2_g$$

Hence we can re-write equation (B-1) as:

$$\gamma_g\sigma^2_g + (\gamma_g + \gamma_{po})\sigma_{g,po} + \gamma_{po}\sigma^2_{po} = \gamma^d_{g^d}(\sigma^2_g + \sigma_{g,po} + \sigma^2_{po}) \tag{B-2}$$

This relation is satisfied if $\gamma_g = \gamma_{po} = \gamma^d_{g^d}$.

C Proof of Corollary 2.1

Using equation (1) we can write:

$$\begin{aligned} \text{E}_t[Y_{t+1}] &= \frac{1}{PE_t}\left\{\text{E}_t[\exp((e_2+e_3)'X_{t+1}) + \text{E}_t\left[\sum_{i=1}^{\infty}\exp(a(i) + (b(i)+e_2)'X_{t+1})\right]\right\} \\ &\equiv \frac{V_t + Z_t}{PE_t} \end{aligned} \tag{C-1}$$

The term V_t is given by:

$$\begin{aligned} V_t &= \exp((e_2+e_3)'\mu + (e_2+e_3)'AX_t + \frac{1}{2}(e_2+e_3)'\Sigma(e_2+e_3)) \\ &= \exp(r_t)\exp(a(1) + b(1)'X_t)\exp(-(e_2+e_3)'\Sigma\gamma). \end{aligned} \tag{C-2}$$

The term Z_t is given by:

$$\begin{aligned} Z_t &= \sum_{i=1}^{\infty}\exp(a(i) + (b(i)+e_2)'\mu + \frac{1}{2}(b(i)+e_2)'\Sigma(b(i)+e_2) + (b(i)+e_2)'AX_t) \\ &= \exp(r_t)\sum_{i=1}^{\infty}\exp(a(i+1) + b(i+1)'X_t)\exp(-(e_2 + b(i)'\Sigma\gamma) \end{aligned} \tag{C-3}$$

Suppose that (i) $(e_2 + b(i))'\Sigma\gamma$ is not a function of i and $(e_2 + b(i))'\Sigma\gamma = (e_2+e_3)'\Sigma\gamma$, (ii) $b(1) = -e_1 + A'(e_2+e_3)$ and (iii) $b(i+1) = -e_1 + A'(e_2 + b(i))$. Then the expected simple total return is:

$$\begin{aligned} \text{E}_t[Y_{t+1}] &= \frac{1}{PE_t}\exp(-(e_2+e_3)'\Sigma\gamma)\exp(r_t)\left(\sum_{i=1}^{\infty}\exp(a(i) + b(i)'X_t\right) \\ &= c\cdot\exp(r_t) \end{aligned} \tag{C-4}$$

where $c = \exp(-(e_2 + e_3)'\Sigma\gamma)$. The simple risk premium is given by:

$$\mathrm{E}_t[Y_{t+1} - \exp(r_t)] = (c-1)\cdot\exp(r_t) \tag{C-5}$$

which is zero if $\gamma = 0$. The unconditional risk premium is:

$$\mathrm{E}[Y_{t+1} - \exp(r_t)] = (c-1)\cdot\exp(\bar{\mu}_r + \frac{1}{2}\bar{\sigma}_r^2) \tag{C-6}$$

where $\bar{\mu}_r$ and $\bar{\sigma}_r^2$ are the unconditional mean and variance of r_t respectively.

If $\gamma \neq 0$ then $b(i)$ must be constant for all i. The second condition, $b(1) = -e_1 + A'(e_2 + e_3)$, defines the $b(i) \equiv b$ vector in equation (9). The first condition $b(i)'\Sigma\gamma = e_3'\Sigma\gamma$, gives the first restriction in equation (10). The non-linear restrictions on A in equation (10) are derived from substituting the definition of b into the third condition $b = -e_1 + A'(e_2 + b)$.

The final restriction in equation (10) is a transversality condition. Under conditions $(i) - (iii)$ the PE_t can be written as:

$$\begin{aligned} PE_t &= \left(\sum_{i=1}^{\infty} \exp(a(i))\right)\exp(b'X_t) \\ &\equiv \bar{A}\exp(b'X_t). \end{aligned} \tag{C-7}$$

Examining the recursion for $a(i)$ in equation (6), we have:

$$a(i+1) = (e_2+e_3)'(\mu+\Sigma\gamma) + \frac{1}{2}(e_2+e_3)'\Sigma(e_2+e_3) + i\cdot[(e_2+b)'(\mu+\Sigma\gamma) + \frac{1}{2}(e_2+b)'\Sigma(e_2+b)].$$

Hence:

$$\begin{aligned} \bar{A} &= \sum_{i=1}^{\infty}\exp(a(i)) \\ &= \frac{\exp((e_2+e_3)'(\mu+\Sigma\gamma) + \frac{1}{2}(e_2+e_3)'\Sigma(e_2+e_3))}{1 - \exp((e_2+b)'(\mu+\Sigma\gamma) + \frac{1}{2}(e_2+b)'\Sigma(e_2+b))} \end{aligned} \tag{C-8}$$

The last restriction in equation (10) ensures this expression is well-defined.

D Testing Predictability Across Horizons

The moment conditions for the system in equation (17) are:

$$\mathrm{E}(h_{t+\bar{k}}) \equiv \mathrm{E}\begin{pmatrix} h_{t+k_1} \\ \vdots \\ h_{t+k_n}\end{pmatrix} = \mathrm{E}\begin{pmatrix} u_{t+k_1}x_t \\ \vdots \\ u_{t+k_n}x_t\end{pmatrix} = \mathrm{E}(u_{t+\bar{k}} \otimes x_t) = 0 \tag{D-1}$$

where $x_t = (1 z_t')'$, a $K \times 1$ vector and $u_{t+k} = (u_{t+k_1} \dots u_{t+k_n})'$.

From standard GMM $\sqrt{T}(\hat{\theta} - \theta) \overset{a}{\sim} N(0, \Omega)$ with $\Omega = Z_0^{-1}S_0Z_0^{-1}$, $Z_0 = (I_n \otimes \mathrm{E}(x_t x_t'))$ and

$$S_0 = \mathrm{E}(h_{t+\bar{k}}h'_{t+\bar{k}}) = \mathrm{E}\big((u_{t+\bar{k}}u_{t+\bar{k}}) \otimes (x_t x_t')\big) \tag{D-2}$$

The Hodrick (1992) estimate $\hat{S}_T^b$ of S_0 is given by:

$$\hat{S}_T^b = \frac{1}{T}W'W \tag{D-3}$$

where W is a $T \times Kn$ matrix $W = (W_{k_1} \dots W_{k_n})$ where W_k $T \times n$ is given by $W_k = (w'_{1+k}, \dots w'_{T+k})$, and w_{t+k}, $K \times 1$, is:

$$w_{t+k} = e_{t+1}\left(\sum_{i=0}^{k-1} x_{t-i}\right), \tag{D-4}$$

since under the null of no predictability the one-step ahead errors $e_{t+i} = u_{t+1}$ are uncorrelated and $u_{t+k} = e_{t+1} + \cdots + e_{t+k}$. Denoting $X = (x'_1, \ldots x'_T)$, $T \times K$, an estimate of Z_0^{-1} is given by:

$$\hat{Z}_T^{-1} = \frac{1}{T}(I_n \otimes (X'X)^{-1}) \tag{D-5}$$

To test the hypothesis $C\beta = 0$ we use the Newey (1985) χ^2 test:

$$(C\hat{\beta})'[C\hat{\Omega}C']^{-1}C\hat{\beta} \sim \chi^2_{\text{rank}(C)} \tag{D-6}$$

with $\hat{\Omega} = \hat{Z}_T^{-1}\hat{S}_T^b\hat{Z}_T^{-1}$.

E Testing Predictability Pooling Cross-Sectional Information

Let the dimension of z_t be $(K-1)$ so there will be a total of K regressors, including the constant terms α_i for each of N countries. In equation (18) denote the free parameters $\theta = (\alpha_1 ... \alpha_N \bar{\beta}')'$, and the unrestricted parameters stacked by each equation $\beta = (\alpha_1\beta'_1 \ldots \alpha_N\beta'_N)'$. We can estimate the system in equation (18) subject to the restriction that $C\beta = 0$, where C is a $NK \times (N-1)(K-1)$ matrix of the form:

$$C = \begin{pmatrix} \tilde{0} & I & \tilde{0} & -I & \tilde{0} & \ldots & & & \\ \tilde{0} & O & \tilde{0} & I & \tilde{0} & -I & \ldots & & \\ \vdots & & & & & & & & \\ \tilde{0} & O & \tilde{0} & \ldots & & & & \tilde{0} & -I \end{pmatrix} \tag{E-1}$$

where $\tilde{0}$ is a $(K-1) \times 1$ vector of zeros, O is a $(K-1) \times (K-1)$ matrix of zeros, and I is a $(K-1)$ rank identity matrix.

Denote

$$\begin{aligned} \tilde{y}_{t+k} &= (\tilde{y}^1_{t+k} \ldots \tilde{y}^N_{t+k})' && (N \times 1) \\ x^i_t &= (1 z^{i'}_t) && (K \times 1) \\ u_{t+k} &= (u^1_{t+k} \ldots u^N_{t+k})' && (N \times 1) \\ X_t &= \begin{pmatrix} x^1_t & & 0 \\ & \ddots & \\ 0 & & x^N_t \end{pmatrix} && (NK \times N). \end{aligned} \tag{E-2}$$

Then the system can be written as:

$$\tilde{y}_{t+k} = X'_t\beta + u_{t+k} \tag{E-3}$$

subject to $C\beta = 0$. To write in compact notation let $Y = (\tilde{y}'_{1+k} \ldots \tilde{y}'_{T+k})'$, $X = (X'_1 \ldots X_T)'$, $U = (u'_{1+k} \ldots u'_{T+k})'$. Then the compact system can be written as:

$$Y = X\beta + U \quad \text{subject to } C\beta = 0 \tag{E-4}$$

A consistent estimate $\hat{\beta}$ of β is given by:

$$\hat{\beta} = \beta^{ols} - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}C\beta^{ols} \tag{E-5}$$

with $\beta^{ols} = (X'X)^{-1}X'Y$. This gives us an estimate $\hat{\theta}$ of θ.

The moment conditions of the system in equation (E-3) are:

$$\mathrm{E}(h_{t+k}) = \mathrm{E}(X_t u_{t+k}) = 0 \tag{E-6}$$

By standard GMM $\hat{\theta}$ has distribution

$$\sqrt{T}(\hat{\theta} - \theta) \overset{a}{\sim} N(0, (D'_0 S_0^{-1} D_0)^{-1}) \tag{E-7}$$

with

$$D_0' = \mathrm{E}\left[\frac{\partial h_{t+k}}{\partial \theta'}\right] \quad \text{(E-8)}$$

and

$$S_0 = \mathrm{E}(h_{t+k}h_{t+k}') \quad \text{(E-9)}$$

The Hodrick (1992) estimate $\hat{S}_T^b$ of S_0 is given by:

$$\hat{S}_T^b = \frac{1}{T}\sum_{t=k}^{T} wk_t wk_t' \quad \text{(E-10)}$$

where wk_t $(NK \times 1)$ is

$$wk_t = \left(\sum_{i=0}^{k-1} X_{t-i}\right) e_{t+1}. \quad \text{(E-11)}$$

Under the null hypothesis of no predictability $u_{t+k} = e_{t+1} + \ldots e_{t+k}$ where e_{t+1} are the 1-step ahead serially uncorrelated errors. This is the SUR equivalent of the Hodrick (1992) estimate for univariate OLS regressions.

An estimate $\hat{D}_T$ of D_0 is given by:

$$\hat{D}_T' = \frac{1}{T}\sum_{t=0}^{T} \frac{\partial h_{t+k}}{\partial \theta'}, \quad \text{(E-12)}$$

$\theta = (\alpha_1 \ldots \alpha_N \bar{\beta}')$ with

$$-\frac{\partial h_{t+k}}{\partial \theta'} = \begin{bmatrix} 1 & z_t^{1'} & & & & 0 & \\ & & 1 & z_t^{2'} & & & \\ & & & & \ddots & & \\ & & 0 & & & 1 & z_t^{N'} \\ z_t^1 & z_t^1 z_t^{1'} & z_t^2 & z_t^2 z_t^{2'} & \ldots & z_t^N & z_t^N z_t^{N'} \end{bmatrix} \quad \text{(E-13)}$$

The estimate $\hat{\theta}$ has distribution

$$\sqrt{T}(\hat{\theta} - \theta) \overset{a}{\sim} N(0, [\hat{D}_T'(\hat{S}_T^b)^{-1}\hat{D}_T]^{-1}). \quad \text{(E-14)}$$

There are $(N + K - 1)$ free parameters in θ with NK moment conditions. This gives $NK - (N + K - 1)$ over-identifying restrictions. The Hansen (1982) χ^2 J-test of over-identifying restrictions is given by:

$$J = T(\bar{h}'(\hat{S}_T^b)^{-1}\bar{h}) \sim \chi^2(NK - (N + K - 1)) \quad \text{(E-15)}$$

with

$$\bar{h} = \frac{1}{T}\sum_{t=0}^{T} h_{t+k}. \quad \text{(E-16)}$$

F Calibrating the Constant Expected Return Present Value Model

We turn first to the VAR parameter estimation. We observe earnings growth rates g_t^{12} and log payout ratios po_t^{12}, but our Present Value Model requires monthly growth rates g_t and log payout ratios po_t. We construct g_t^{12} at a monthly frequency from earnings yields ey_t^{12} using $g_t^{12} = \log(ey_t^{12}/ey_{t-1}^{12} \times P_t/P_{t-1})$. If EA_t denotes monthly earnings, then g_t^{12} is related to g_t by:

$$\begin{aligned} g_t^{12} &= \log\left(\frac{EA_t + EA_{t-1} + \cdots + EA_{t-11}}{EA_{t-1} + EA_{t-2} + \cdots + EA_{t-12}}\right) \\ &= \log\left(\frac{EA_{t-11}(1 + e^{g_{t-10}} + \cdots + e^{(g_{t-10}+\cdots+g_t)})}{EA_{t-12}(1 + e^{g_{t-11}} + \cdots + e^{(g_{t-11}+\cdots+g_{t-1})})}\right) \\ &= g_{t-11} + \log(1 + e^{g_{t-10}} + \cdots + e^{(g_{t-10}+\cdots+g_t)}) \\ &\quad - \log(1 + e^{g_{t-11}} + \cdots + e^{(g_{t-11}+\cdots+g_{t-1})}) \end{aligned} \tag{F-1}$$

In addition, if D_t denotes monthly dividends then po_t^{12} is related to po_t by:

$$\begin{aligned} po_t^{12} &= \log\left(\frac{D_t + D_{t-1} + \cdots + D_{t-11}}{EA_t + EA_{t-1} + \cdots + EA_{t-11}}\right) \\ &= \log\left(\frac{EA_{t-11}[e^{po_{t-11}} + e^{po_{t-10}}e^{g_{t-10}} + \cdots + e^{po_t}e^{(g_{t-10}+g_{t-9}+\cdots+g_t)}]}{EA_{t-11}[1 + e^{g_{t-10}} + \cdots + e^{(g_{t-10}+\cdots+g_t)}]}\right) \\ &= \log(e^{po_{t-11}} + e^{(po_{t-10}+g_{t-10})} + \cdots + e^{(po_t+g_{t-10}+g_{t-9}+\cdots+g_t)}) \\ &\quad - \log(1 + e^{g_{t-10}} + \cdots + e^{(g_{t-10}+\cdots+g_t)}) \end{aligned} \tag{F-2}$$

Equations (F-1) and (F-2) show that the relation between monthly growth rates and payout ratios and their counterparts using earnings and dividends summed over the past year is highly non-linear. In particular, the use of summing past earnings and dividends over the past twelve months potentially induces very high autocorrelation up to 11 lags.

To estimate the VAR on X_t we use Simulated Method of Moments (SMM) (Duffie and Singleton (1993)). We use a two-step SMM procedure and impose a restricted companion form A where $A_{12} = A_{13} = 0$. The latter assumption is motivated by an analysis of a VAR on $(r_t\, g_t^{12}\, po_t^{12})$, in which we fail to reject that no variables Granger-cause interest rates. In the first step, we estimate the equation for r_t on US EURO 1 month rates since we have monthly data on interest rates. In the second step, holding the parameters for r_t fixed, we estimate the remaining parameters in A, μ and Σ using the first and second moments of g_t^{12} and po_t^{12}. We also use the moments in $\mathrm{E}[X_t^{12} X_{t-12}^{12'}]$ relating to g_t^{12} and po_t^{12}. The lag length is set at 12 since the first 11 lags are affected by the autocorrelation induced by the non-linear filters in equations (F-1) and (F-2). We compute the weighting matrix using the data, so we need not iterate on the weighting matrix.

Table (A-1) reports our results. We report two estimations, an Alternative Model which is exactly identified, and the Null Model which is estimated subject to the restrictions ensuring constant expected returns in Corollary 2.1. Focusing first on the Alternative Model, payout ratios are close to a random walk with no other significant feedback coefficients. Earnings growth on the other hand shows little persistence with the coefficient on past earnings growth barely significantly different from zero. However, high current payout ratios predict high future earnings growth, perhaps because they reflect permanent rather than transitory earnings. High short rates significantly reduce future expected earnings growth. A 1% increase in interest rates leads to a 25 basis point decrease in expected earnings growth.

When we estimate the covariance parameters of the Null Model on the moments of the Alternative Model, we obtain a singular covariance matrix, as the correlation of g_t and po_t approaches -1. Therefore we hold the covariance matrix Σ from the Alternative Model fixed, and estimate μ and a restricted companion matrix A using the first the and cross-moments of the Alternative Model estimation. Using a χ^2 test, we fail to reject the Null Model versus the Alternative Model with a p-value of 0.9252. The Null Model retains the important feedback from interest rates to earnings growth but makes the interest rate feedback to payout ratios much larger than before. The feedback from payout ratios to earnings growth rates remains intact as well, but the earnings growth rate is no longer persistent.

The estimation of the covariance matrix Σ has two important features. First, the conditional volatility of innovations to g_t is much more volatile than innovations to g_t^{12} (0.0647 versus 0.0233 from an unreported VAR on $(r_t\, g_t^{12}\, po_t^{12})$). The filter in equations (F-1) and (F-2) smooths out some of the volatility in earnings growth by summing earnings over the past twelve months. This implies that equity returns are more volatile using monthly earnings than summed annual earnings. Second, shocks to g_t and po_t are negatively correlated (-0.8496), which is true in the annual data as well. This might be due to the unusual smoothing that occurs in most corporate dividend policies. High temporary earnings are not paid out, so they decrease the payout ratio.

The constrained VAR ties down most of the parameters, but we still have to determine the price of risk for dividend growth $\bar{\gamma}$. We calibrate $\bar{\gamma}$ to the premium in the sample. The log risk premium is produced by simulation using 100,000 observations, while the simple risk premium is calculated using equation Corollary 2.1 (and checked by simulation). Setting $\bar{\gamma} = -4.655$, we match both the log and the simple risk premium. The annualized volatility of the log (simple) excess return corresponding to $\bar{\gamma} = -4.655$ is 0.1367 (0.1387), which is slightly below the annualized volatility in the sample 0.1477 (0.1472). The main source of equity return volatility in the model is the volatility of payout ratios and earnings growth rates, since the risk premium is constant. In fact, the implied estimate of the conditional volatility of dividend growth is $\sqrt{(e_2 + e_3)'\Sigma(e_2 + e_3)}$, which is 0.1368 annualized.

References

[1] Ang, A., and J. Liu, 2001, "A General Affine Earnings Valuation Model," forthcoming *Review of Accounting Studies*.

[2] Bansal, R., and A. Yaron, 2000, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," working paper, Duke University.

[3] Bauer, G., 2001, "The Foreign Exchange Risk Premium over the Long Run," working paper, University of Rochester.

[4] Bekaert, G., and S. Grenadier, 2000, "Stock and Bond Pricing in an Affine Economy," working paper, Columbia University.

[5] Bekaert, G., and R. J. Hodrick, 1992, "Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets," *Journal of Finance*, 47, 2, 467-509.

[6] Bekaert, G., and R. J. Hodrick, 2001, "Expectations Hypotheses Tests," *Journal of Finance*, 56, 4, 1357-1399.

[7] Bollerslev, T., and R. J. Hodrick, 1996, "Financial Market Efficiency Tests," in Pesaran, M. H., and M. R. Wickens, eds, 1996, *The Handbook of Applied Econometrics Vol 1, Macroeconomics*, Blackwell, 415-458.

[8] Boudoukh, J., and M. Richardson, 1993, "The Statistics of Long-Horizon Regressions," *Mathematical Finance*, 4, 2, 103-120.

[9] Breen, W., L. R. Glosten and R. Jagannathan, 1989, "Economic Significance of Predictable Variations in Stock Index Returns," *Journal of Finance*, 44, 5, 1177-89.

[10] Campbell, J. Y., 1987, "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373-399.

[11] Campbell, J. Y., 1991, "A Variance Decomposition for Stock Returns," *Economic Journal*, 101, 157-179.

[12] Campbell, J. Y., and J. H. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205-251.

[13] Campbell, J. Y., and R. J. Shiller, 1988a, "Stock Prices, Earnings and Expected Dividends," *Journal of Finance*, 43, 3, 661-676.

[14] Campbell, J. Y., and R. J. Shiller, 1988b, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 3, 195-228.

[15] Cliff, M. T., 2001, "Asset Pricing Tests under Price-Dividend Ratio Restrictions," working paper, Purdue University.

[16] Cochrane, J. H., 1992, "Explaining the Variance of Price-Dividend Ratios," *Review of Financial Studies*, 5, 243-280.

[17] Cutler, D. M., J. M. Poterba, and L. H. Summers, 1989, "What Moves Stock Prices?" *Journal of Portfolio Management*, 15, 3, 4-12.

[18] Duffie, D. and R. Kan, 1996, "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 4, 379-406.

[19] Duffie, D., and K. J. Singleton, 1993, "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61, 4, 929-952.

[20] Fama, E., and G. W. Schwert, 1977, "Asset Returns and Inflation," *Journal of Financial Economics*, 5, 115-146.

[21] Fama, E., and F. French, 1988, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3-26.

[22] Ferson, W. E., 1989, "Changes in Expected Security Returns, Risk and the Level of Interest Rates," *Journal of Finance*, 44, 5, 1191-1217.

[23] Ferson, W. E., and C. R. Harvey, 1993, "The Risk and Predictability of International Equity Returns," *Review of Financial Studies*, 6, 527-566.

[24] Foster, F. D., T. Smith, and R. E. Whaley, 1997, "Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R-Squared," *Journal of Finance*, 52, 2, 591-607.

[25] Glosten, L. R., R. Jagannathan, and D. E. Runkle, 1993, "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 5, 1779-1801.

[26] Goetzmann, W. N., and P. Jorion, 1993, "Testing the Predictive Power of Dividend Yields," *Journal of Finance*, 48, 2, 663-679.

[27] Goetzmann, W. N., and P. Jorion, 1995, "A Longer Look at Dividend Yields," *Journal of Business*, 68, 483-508.

[28] Goldstein, R., and F. Zapatero, 1996, "General Equilibrium with Constant Relative Risk Aversion and Vasicek Interest Rates," *Mathematical Finance*, 6, 3, 331-340.

[29] Goyal, A., and I. Welch, 1999, "The Myth of Predictability: Does the Dividend Yield Forecast the Equity Premium?" working paper, UCLA.

[30] Hansen, L., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

[31] Hansen, L., and R. Hodrick, 1980, "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829-853.

[32] Hodrick, R., 1992, "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5, 3, 357-386.

[33] Kandel, S., and R. F. Stambaugh, 1990, "Expectations and Volatility of Consumption and Asset Returns," *Review of Financial Studies*, 3, 2, 207-32.

[34] Kirby, C., 1997, "Measuring the Predictable Variation in Stock and Bond Returns," *Review of Financial Studies*, 10, 3, 579-630.

[35] Lamont, O., 1998, "Earnings and Expected Returns," *Journal of Finance*, 53, 5, 1563-1587.

[36] Lee, B. S., 1992, "Causal Relations Among Stock Returns, Interest Rates, Real Activity and Inflation," *Journal of Finance*, 48, 1591-1603.

[37] Lettau, M., and S. Ludvigson, 2001, "Consumption, Aggregate Wealth and Expected Stock Returns," *Journal of Finance*, 56, 3, 815-849.

[38] Lo, A., and A. C. MacKinlay, 1990, "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3, 431-468.

[39] Newey, W., 1985, "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, 29, 3, 229-56.

[40] Newey, W., and K. West, 1987, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.

[41] Nissim, D., and S. Penman, 2001, "An Empirical Analysis of the Effect of Changes in Interest Rates on Accounting Rates of Return, Growth and Equity Values," working paper, Columbia University.

[42] Patelis, A. D., 1997, "Stock Return Predictability and the Role of Monetary Policy," *Journal of Finance*, 52, 5, 1951-1972.

[43] Richardson, M., 1993, "Temporary Components of Stock Prices: A Skeptic's View," *Journal of Business and Economic Statistics*, 11, 2, 199-207.

[44] Richardson, M., and T. Smith, 1991, "Tests of Financial Models in the Presence of Overlapping Observations," *Review of Financial Studies*, 4, 2, 227-254.

[45] Richardson, M., and J. H. Stock, 1989, "Drawing Inferences from Statistics Based on Multiyear Asset Returns," *Journal of Financial Economics*, 25, 323-348.

[46] Shiller, R. J., and A. E. Beltratti, 1992, "Stock Prices and Bond Yields: Can Their Comovements Be Explained in Terms of Present Value Models?" *Journal of Monetary Economics*, 30, 25-46.

[47] Stambaugh, R. F., 1999, "Predictive Regressions," *Journal of Financial Economics*, 54, 375-421.

[48] Valkanov, R., 2001, "Long-Horizon Regressions: Theoretical Results and Applications to the Expected Returns/Dividend Yields and Fisher Effect Relations," forthcoming *Journal of Financial Economics*.

Table 1: Sample Moments of Returns and Instruments

Sample Moments of MSCI Monthly Data

	$\tilde{y}$	dy^{12}	ey^{12}	r
US				
mean	0.0766	-3.3632	-2.6271	0.0769
stdev	0.1477	0.4059	0.3915	0.0340
ρ	-0.0080	0.9819	0.9839	0.9699
UK				
mean	0.0765	-3.0780	-2.4601	0.1027
stdev	0.1848	0.2556	0.3608	0.0347
ρ	-0.0154	0.9601	0.9690	0.9521
France				
mean	0.0689	-3.2192	-2.8219	0.0948
stdev	0.2088	0.4085	0.6171	0.0473
ρ	0.0795	0.9803	0.9473	0.8464
Germany				
mean	0.0690	-3.3512	-2.7435	0.0576
stdev	0.1880	0.3324	0.4757	0.0245
ρ	0.0582	0.9794	0.9812	0.9806
Japan				
mean	0.0358	0.0106	0.0284	0.0454
stdev	0.1911	0.0051	0.0174	0.0302
ρ	0.0275	0.9808	0.9811	0.9678

Correlations of MSCI Monthly Excess Returns

	US	UK	France	Germany
UK	0.5790			
France	0.5050	0.5035		
Germany	0.4689	0.4620	0.5940	
Japan	0.3453	0.3669	0.3751	0.3383

Sample Moments of Quarterly US S&P Data

	$\tilde{y}$	dy^{12}	ey^{12}	r
mean	0.0821	-3.2485	-2.6109	0.0409
stdev	0.1671	0.3603	0.3603	0.0326
ρ	0.1218	0.9480	0.9478	0.9559

Summary statistics of excess returns and instruments. In the table $\tilde{y}$ represents excess returns, dy^{12} log dividend yields, ey^{12} log earnings yields, r short rates and earnings growth rates g^{12}. Excess returns and short rates are continuously compounded. Monthly equity returns are from MSCI and the monthly short rates are EURO 1 month rates. Quarterly equity returns (including re-invested dividends) for the US are for the S&P Composite Index obtained from Ibbotson Associates and quarterly short rates are US 3 month T-bill yields. The monthly (quarterly) excess return for a country is equal to the equity return in local currency less the EURO 1 month rate (3 month T-bill yield) for that country. Monthly dividend and earnings yields are from MSCI and quarterly dividend and earnings yields are from the *Security Price Index Record* published by Standard & Poor's Statistical Service. Both sources use the sum of dividends (earnings) over the past twelve months. The mean and standard deviation of $\tilde{y}$, r and g^{12} are obtained by multiplying the sample mean and standard deviation by 12 (4) and $\sqrt{12}$ (2) respectively for monthly (quarterly) returns. ρ denotes the sample autocorrelation. For monthly MSCI returns, the sample period is from Feb 1975 to Dec 1999 for the US, UK, France and Germany and from Jan 1978 to Dec 1999 for Japan. Earnings yields for Japan are negative during 1999, so we report levels, instead of logs, for dividend and earnings yields for Japan. For quarterly S&P returns for the US, the sample period is from Mar 1935 to Dec 1999.

Table 2: Small Sample Coefficients from the Null Model

Panel A: Dividend Regression Coefficient

horizon	$k = 1$	$k = 12$	$k = 60$
mean	0.0692	0.0547	0.0517
stdev	0.3216	0.2015	0.0879

Panel B: Earnings Regression Coefficient

horizon	$k = 1$	$k = 12$	$k = 60$
mean	-0.0331	-0.0304	-0.0243
stdev	0.1735	0.1548	0.1169

Panel C: Lamont Regression Coefficients

horizon	$k = 1$		$k = 12$		$k = 60$	
	dy^{12}	ey^{12}	dy^{12}	ey^{12}	dy^{12}	ey^{12}
mean	0.1877	-0.0913	0.1603	-0.0829	0.1175	-0.0604
stdev	0.3767	0.2207	0.2758	0.2108	0.1649	0.1607

Panel D: Trivariate Regression Coefficients

horizon	$k = 1$			$k = 12$			$k = 60$		
	dy^{12}	ey^{12}	r	dy^{12}	ey^{12}	r	dy^{12}	ey^{12}	r
mean	0.2411	-0.1093	0.0003	0.1983	-0.0954	-0.0010	0.1266	-0.0624	-0.0004
stdev	0.3932	0.2386	0.1357	0.2828	0.2228	0.1199	0.1649	0.1613	0.0784

The table reports the mean and standard deviation of the small sample distribution of the slope coefficients in the regression: $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return, z_t is the log dividend yield alone in Panel A, the log earnings yield alone in Panel B, the log dividend yield and log earnings yield together in Panel C, and the log dividend yield, log earnings yield, and the short rate in Panel D. The small sample distribution is based on 5000 replications of a sample size of 299 observations using the Constant Expected Return Null Model as the DGP.

Table 3: Size Properties of T-Statistics from the Constant Expected Return Null Model

Dividend Regression

	$k=1$		$k=12$		$k=60$	
Nominal size	0.100	0.050	0.100	0.050	0.100	0.050
OLS	0.105	0.052	0.471	0.402	0.611	0.541
Robust Hansen-Hodrick	0.108	0.055	0.210	0.144	0.470	0.391
Hodrick	0.108	0.055	0.106	0.052	0.144	0.076

Earnings Regression

	$k=1$		$k=12$		$k=60$	
Nominal size	0.100	0.050	0.100	0.050	0.100	0.050
OLS	0.106	0.052	0.601	0.539	0.786	0.747
Robust Hansen-Hodrick	0.111	0.055	0.190	0.124	0.475	0.403
Hodrick	0.111	0.055	0.096	0.044	0.079	0.036

Lamont Regression

	$k=1$		$k=12$		$k=60$	
Nominal size	0.100	0.050	0.100	0.050	0.100	0.050
Dividend coefficients only						
OLS	0.133	0.069	0.583	0.514	0.786	0.722
Robust Hansen-Hodrick	0.137	0.071	0.254	0.187	0.529	0.457
Hodrick	0.137	0.071	0.145	0.081	0.139	0.077
Earnings coefficients only						
OLS	0.136	0.075	0.661	0.600	0.835	0.800
Robust Hansen-Hodrick	0.138	0.078	0.223	0.152	0.497	0.426
Hodrick	0.138	0.078	0.127	0.066	0.098	0.051

Extended Lamont Regression

	$k=1$		$k=12$		$k=60$	
Dividend coefficients only						
Nominal size	0.100	0.050	0.100	0.050	0.100	0.050
Dividend coefficients only						
OLS	0.157	0.086	0.613	0.549	0.788	0.744
Robust Hansen-Hodrick	0.162	0.089	0.297	0.223	0.561	0.486
Hodrick	0.162	0.089	0.170	0.099	0.145	0.079
Earnings coefficients only						
OLS	0.146	0.084	0.668	0.610	0.826	0.801
Robust Hansen-Hodrick	0.149	0.086	0.249	0.175	0.524	0.448
Hodrick	0.149	0.086	0.131	0.075	0.091	0.047
Short Rate coefficients only						
OLS	0.123	0.064	0.643	0.578	0.794	0.761
Robust Hansen-Hodrick	0.127	0.068	0.226	0.152	0.486	0.407
Hodrick	0.127	0.068	0.119	0.060	0.089	0.048

The table lists nominal versus empirical size properties of the OLS, Robust Hansen-Hodrick and Hodrick t-statistics. We simulate 5000 samples of length 299 from the Constant Expected Return Null Model, calculate the t-statistics for each method and record the percentage of observations greater than the nominal critical values under the null hypothesis of no predictability.

Table 4: Predictability of US Excess Returns

Panel A Monthly MSCI Data

k	Univariate Regressions dy^{12} only	ey^{12} only	Lamont Regression dy^{12}	ey^{12}	Trivariate Regression dy^{12}	ey^{12}	r
			Full Sample 1975:02 - 1999:12				
1	-0.0820	-0.0415	-0.2994	0.2445	-0.3896	0.5747	-3.7341
	(-1.0571)	(-0.5301)	(-1.6837)	(1.3720)	(-2.1524)*	(2.5620)*	(-2.7539)**
12	-0.1063	-0.0673	-0.2581	0.1635	-0.2710	0.2433	-0.9604
	(-1.1315)	(-0.8086)	(-1.3550)	(1.0190)	(-1.4465)	(1.3738)	(-0.8302)
60	-0.0942	-0.0859	0.2280	-0.2394	0.1815	-0.2616	0.6659
	(-0.5249)	(-0.9866)	(0.2890)	(0.5094)	(0.2168)	(-0.5890)	(0.7111)
			Restricted Sample I 1975:02 - 1994:12				
1	0.0957	0.0552	0.2272	-0.0975	0.5494	0.0518	-4.5453
	(0.6615)	(0.5674)	(0.5705)	(-0.3725)	(1.3249)	(0.1939)	(-2.9723)**
12	0.0480	0.0217	0.1859	-0.1009	0.2837	-0.0598	-1.3252
	(0.3260)	(0.2324)	(0.4980)	(-0.4547)	(0.7318)	(-0.2634)	(-1.0870)
60	0.0415	0.0141	0.2574	-0.1917	0.1991	-0.2045	-0.6636
	(0.2662)	(0.0957)	(0.7205)	(-0.5266)	(0.6007)	(-0.5679)	(1.2746)
			Restricted Sample II 1975:02 - 1991:12				
1	0.1050	0.0760	0.1417	-0.0306	0.5471	0.0315	-4.5079
	(0.4763)	(0.4094)	(0.2553)	(-0.0633)	(0.9757)	(0.0656)	(-2.9744)**
12	0.0874	0.0700	0.0747	0.0111	0.1960	0.0281	-1.3268
	(0.3904)	(0.3559)	(0.1629)	(0.0256)	(0.4201)	(0.0654)	(-1.0924)
60	0.0873	0.0320	0.4615	-0.3306	0.3669	-0.3210	0.5557
	(0.3317)	(0.1363)	(1.0938)	(-0.7625)	(0.9128)	(-0.7834)	(0.5873)

Panel B Quarterly S&P Data

	k	dy^{12}	ey^{12}	r
	1	0.0191	0.0366	-1.3041
		(0.0977)	(0.2022)	(-1.5907)
Full Sample	12	-0.0142	0.1066	-0.9730
1935:qtr 1 - 1999:qtr 4		(-0.0990)	(0.8504)	(-1.3773)
	60	0.0187	0.0791	-0.5721
		(0.1435)	(0.9288)	(-0.9810)
	1	-0.1142	0.2256	-2.6650
		(-0.7250)	(1.4898)	(-3.4288)**
Post Treasury Accord Sample	12	-0.0655	0.1811	-1.8716
1952:qtr 1 - 1999:qtr 4		(-0.3902)	(1.2223)	(-2.4594)*
	60	0.1205	-0.0163	-0.6510
		(0.5002)	(-0.0976)	(-0.8995)

We estimate regressions of the form $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return, with instruments z_t, where k is in months. The variables in z_t include dy_t^{12}, log dividend yield using dividends summed over the past 12 months, ey_t^{12}, log earnings yield using earnings summed over the past 12 months, and r_t, the annualized risk-free rate (1 month EURO rate in Panel A and the 3 month T-bill rate in Panel B). In Panel A, we use monthly MSCI data. In Panel B, we use quarterly S&P data. T-statistics are in parentheses and calculated using Hodrick standard errors. T-statistics significant at 95% (99%) are denoted with * (**).

Table 5: Excess Return Predictability in 5 Countries

	US	UK	France	Germany	Japan†
			Univariate Regressions		
dy^{12}	0.8556	0.5014	0.6207	0.5821	0.0818
ey^{12}	0.5869	0.0404*	0.8732	0.8803	0.0572
			Lamont Regression - dy^{12} and ey^{12}		
dy^{12} only	0.6145	0.6129	0.5464	0.1394	0.9496
ey^{12} only	0.3723	0.8116	0.7030	0.1881	0.7433
Joint dy^{12} and ey^{12}	0.1495	0.1502	0.8271	0.2453	0.0101*
			Trivariate Regression - dy^{12}, ey^{12}, r		
dy^{12} only	0.7016	0.5474	0.8857	0.3523	0.9691
ey^{12} only	0.0849	0.2583	0.6519	0.2966	0.7915
Joint dy^{12} and ey^{12}	0.0097**	0.0111*	0.7897	0.4082	0.0010**
r only	0.0399*	0.1063	0.3802	0.1279	0.9318
Joint dy^{12}, ey^{12}, r	0.0105*	0.0312*	0.7260	0.0984	0.0018**

The table lists p-values of joint tests across horizons $k = 1, 12, 60$. P-values less than 5% (1%) are asterixed * (**). dy^{12} denotes the log dividend yield, ey^{12} the log earnings yield and r the short rate. Japan's regressions (marked by †) are performed in levels, not logs. The data set is monthly MSCI data, from 1975:02-1999:12 for the US, UK, France and Germany, and from 1978:01-1999:12 for Japan.

Table 6: Pooled-Country Excess Return Regressions

Lamont System dy^{12} and ey^{12}

	$k=1$	$k=12$	$k=60$
dy^{12}	-0.0234	0.0352	0.1084
	(-0.2803)	(0.3594)	(1.2303)
ey^{12}	0.0046	-0.0590	-0.0593
	(0.0859)	(-1.1053)	(-1.4923)
χ^2 Test p-values			
J-test	0.2605	0.5343	0.1667
$dy, ey = 0$	0.9592	0.5167	0.2253

Trivariate System dy^{12}, ey^{12}, r

	$k=1$	$k=12$	$k=60$
dy^{12}	0.0719	0.0726	0.0860
	(0.7884)	(0.7119)	(1.0057)
ey^{12}	0.0236	-0.0521	-0.0613
	(0.4626)	(-1.0446)	(-1.7079)
r	-1.7334	-0.6554	-0.3357
	(-2.5247)*	(-1.1852)	(0.8094)
χ^2 p-values			
J-test	0.1001	0.3960	0.1397
$dy, ey, r = 0$	0.0851	0.4849	0.1451
$dy, ey = 0$	0.4951	0.5653	0.1613
$r = 0$	0.0116*	0.2359	0.4183

The predictability regression $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return with instruments z_t, is estimated jointly across the US, UK, France and Germany, constraining the coefficients to be the same across countries (see Appendix E for details). Japan is excluded from estimation because of a shorter sample and negative earnings yields during 1999. The constants in the regressions are allowed to differ across countries (and are not reported in the table). We report the Lamont system where $z_t = (dy^{12}\; ey^{12})$ and a trivariate system with $z_t = (dy^{12}\; ey^{12}\; r_t)$, where dy^{12} and ey^{12} denote the log dividend and earnings yields respectively, and r is the annualized one-month short rate. T-statistics of the predictability coefficients are in parentheses and are based on Hodrick (1992) standard errors. In the χ^2 tests, the J-test is a test of the over-identifying restrictions and the other tests refer to joint test of coefficients equalling zero. T-statistics significant at levels greater than than 95% are asterixed *. The data set is monthly MSCI data, from 1975:02-1999:12

Table 7: Cross-Country Predictability of Excess Returns

k	US dy^{12}	US ey^{12}	US r	Local dy^{12}	Local ey^{12}	Local r
Local Market is UK						
1	-0.4806	0.1215	-3.2367	0.5282	0.5427	-2.0319
	(-1.1352)	(0.2201)	(-1.5434)	(1.0346)	(1.0204)	(-1.1247)
12	-0.3486	0.2964	-0.4831	0.4699	-0.1061	-0.8179
	(-1.1527)	(0.8534)	(-0.2761)	(1.2299)	(-0.4139)	(-0.6138)
60	0.0467	-0.0024	0.6602	0.3559	-0.1419	-0.6771
	(0.0532)	(-0.0052)	(0.4802)	(1.1660)	(-1.2958)	(-0.4557)
Local Market is FR						
1	-0.6061	0.6967	-3.0557	0.0512	-0.0188	0.1898
	(-1.8558)*	(1.7115)	(-1.1329)	(0.2383)	(-0.2786)	(0.1090)
12	-0.3639	0.3448	-1.3904	0.0991	-0.0622	0.3813
	(-1.1643)	(1.2332)	(-0.7624)	(0.5041)	(-1.0004)	(0.4261)
60	0.0527	-0.1538	0.6449	0.1095	-0.0292	0.3064
	(0.0330)	(-0.2238)	(0.3249)	(0.4004)	(-0.3712)	(0.7332)
Local Market is GER						
1	-1.0831	0.8229	-4.1665	0.6142	0.0012	0.0589
	(-2.6604)**	(1.8206)	(-1.7929)	(1.4097)	(0.0081)	(0.0205)
12	-0.6014	0.4281	-1.4525	0.5002	-0.1343	-1.1546
	(-1.5260)	(1.3244)	(-0.9198)	(1.2788)	(-1.0086)	(-0.4912)
60	0.1639	-0.1531	0.1716	0.2981	-0.1729	0.8250
	(0.1153)	(-0.2313)	(0.0982)	(0.8964)	(-1.7115)	(0.5810)
Local Market is JAP†						
1	-6.9948	3.6414	-1.9965	0.1794	0.0011	-0.4109
	(-0.5427)	(0.6959)	(-0.9995)	(1.0262)	(0.0145)	(-0.2348)
12	7.0783	-1.3938	-1.5191	0.1658	-0.0166	-0.4580
	(0.5893)	(-0.2835)	(-0.8545)	(0.8256)	(-0.2091)	(-0.3125)
60	16.8843	-5.0788	-0.1481	0.0568	0.0173	0.0583
	(0.4863)	(-0.6249)	(-0.1294)	(0.2003)	(0.4359)	(0.0589)

We regress local excess equity returns $\tilde{y}_{t+k}$ of horizon k onto US instruments and local instruments. Hodrick standard errors are used to calculate t-statistics given in parantheses. T-statistics significant at the 95% (99%) level are asterixed * (**). The regressions for Japan are run in levels for dy^{12} and ey^{12} (not logs).

Table 8: Cross-Country Predictability of Excess Returns χ^2 Tests

Base	Local	Base Pred	Local Pred
US	UK	0.1639	0.1761
	FR	0.3016	0.9773
	GR	0.0498*	0.2441
	JP†	0.7870	0.5249
UK	US	0.8656	0.0213*
	FR	0.5197	0.9311
	GR	0.6385	0.6695
	JP†	0.4137	0.4154
FR	US	0.5299	0.0306*
	UK	0.3826	0.1797
	GR	0.9929	0.6612
	JP†	0.6355	0.1822
GR	US	0.4925	0.0491*
	UK	0.2993	0.1587
	FR	0.4666	0.9912
	JP†	0.0250*	0.6342
JP†	US	0.5905	0.0013**
	UK	0.4989	0.1619
	FR	0.9183	0.2911
	GR	0.9322	0.6202

We regress local excess equity returns on base country instruments and local instruments (dy^{12}, ey^{12} and annualized short rates r). For example, the top entry regresses UK equity excess returns on US instruments and UK instruments; the second entry regresses French equity excess returns on US instruments and French instruments. We report χ^2 p-value tests for the hypothesis of base country predictability ("Base Pred") and local country predictability ("Local Pred") at a one-month horizon. P-values less than 5% are asterixed *, those less than 1% are asterixed **. Regressions for Japan (indicated by †) are run in levels, not logs.

Table 9: US Predictability of Foreign Excess Returns

	Coefficients			χ^2 Predictability Tests		
k	US dy^{12}	US ey^{12}	US r	dy^{12}, ey^{12}	r	dy^{12}, ey^{12}, r
UK on US instruments						
1	-0.2954	0.6406	-4.3467	0.1107	0.0354*	0.1867
	(-1.2797)	(1.8791)	(-2.1040)*			
12	-0.1080	0.2140	-1.1907	0.6400	0.4651	0.8233
	(-0.5372)	(0.8931)	(-0.7305)			
60	0.2872	-0.1720	0.2219	0.8947	0.8036	0.8584
	(0.4010)	(-0.4460)	(0.2487)			
FR on US instruments						
1	-0.5716	0.7090	-3.0522	0.0679	0.1789	0.1179
	(-2.2632)*	(2.1677)*	(-1.3422)			
12	-0.3033	0.3647	-1.4670	0.4244	0.4004	0.5711
	(-1.1262)	(1.2971)	(-0.8410)			
60	0.3506	-0.2421	0.3816	0.8133	0.7547	0.8258
	(0.3277)	(-0.4238)	(0.3124)			
GR on US instruments						
1	-0.6258	0.7185	-2.5147	0.0279*	0.0708	0.0308*
	(-2.4711)*	(2.6343)**	(-1.8071)*			
12	-0.3903	0.4392	-1.1823	0.1366	0.2607	0.1989
	(-1.4450)	(1.9760)*	(-1.1247)			
60	0.3340	-0.3268	1.2753	0.1431	0.3402	0.0910
	(0.2691)	(-0.4968)	(0.9537)			
JP on US instruments						
1	-0.2984	0.5406	-2.5062	0.2503	0.1667	0.4201
	(-1.0893)	(1.5631)	(-1.3828)			
12	0.0469	0.2005	-1.8474	0.3647	0.2339	0.5681
	(0.1673)	(0.6744)	(-1.1903)			
60	0.9097	-0.4119	-0.1968	0.3525	0.8254	0.3183
	(0.7357)	(-0.5853)	(-0.2207)			
Pooled Estimation (excluding JP) on US instruments						
1	-0.4706	0.6607	-3.4119	0.0061**	0.0032**	0.0036**
	(-2.8419)**	(3.1855)**	(-2.9440)**			
12	-0.2682	0.3153	-1.2001	0.1341	0.1683	0.1690
	(-1.6076)	(2.0033)*	(-1.3776)			
60	0.2883	-0.2506	0.6361	0.3242	0.4708	0.3839
	(0.3771)	(-0.6151)	(0.7212)			

We regress UK, FR, GR and JP excess equity returns $\tilde{y}_{t+k}$ for horizon k on US instruments only (US dy^{12}, ey^{12} and annualized short rates r). The last panel presents the estimations of predictability coefficients constraining the coefficients to be the same across countries, excluding Japan. The constants in the regressions are allowed to differ across countries (and are not reported in the table). T-statistics of the predictability coefficients are in parenthesis and are calculated using Hodrick (1992) standard errors. In the χ^2 tests, we perform a joint test of coefficients equalling zero and report p-values. For the joint estimation the χ^2 p-values for a J-test of over-identification are 0.2876, 0.0940 and 0.6197 for $k = 1$, 12 and 60 respectively. P-values less than 5% (1%) are asterixed * (**).

Table 10: Cross-Country Predictability of Exchange Rate Returns

k	US dy^{12}	US ey^{12}	US r	Local dy^{12}	Local ey^{12}	Local r
US-UK Exchange Rate Return						
1	-0.1618	0.0846	-2.3076	-0.5544	0.0014	1.9074
	(0.8279)	(0.2819)	(-2.2235)*	(-2.3974)*	(0.0061)	(1.7577)*
12	-0.1189	0.5339	-2.1985	-0.2221	-0.3607	0.9876
	(-0.6261)	(1.9995)*	(-2.5520)*	(-1.0779)	(-2.0673)*	(1.2725)
60	-0.0202	0.0551	-0.2927	0.0408	-0.1638	0.0085
	(-0.0463)	(0.2400)	(-0.3894)	(0.2360)	(-2.4828)*	(0.0105)
US-GR Exchange Rate Return						
1	0.6462	-0.1047	0.0707	-0.8342	0.0790	-0.0619
	(3.2389)**	(-0.4852)	(0.0514)	(-3.8443)**	(0.9796)	(-0.0385)
12	0.5111	0.0105	-0.8178	-0.7206	0.1011	0.0671
	(2.5128)*	(0.0607)	(-0.8983)	(-3.5886)**	(1.2660)	(0.0505)
60	0.2615	-0.1337	0.4888	-0.2570	0.0459	-0.0697
	(0.5061)	(-0.5518)	(0.6322)	(-1.9243)	(0.9329)	(-0.1090)
US-JP† Exchange Rate Return						
1	4.7153	-6.0483	0.0012	0.2085	-0.0666	3.7250
	(0.5470)	(-1.6499)	(0.0007)	(1.5573)	(-1.2474)	(-2.2651)*
12	-0.4377	-1.7852	-1.4859	-0.0136	0.0256	1.6005
	(-0.0519)	(-0.6045)	(-1.1226)	(-1.1096)	(0.4801)	(1.5244)
60	11.5440	-3.6162	0.2162	-0.1345	0.0227	0.1839
	(0.4527)	(-0.6076)	(0.2755)	(-0.7003)	(0.9389)	(0.2619)

χ^2 Tests Joint Across Horizons $k = 1$, 12, 60

	US	Local	Short Rates	Yields	All Variables
US-UK	0.1876	0.0615	0.2869	0.0465*	0.0000**
US-GR	0.2982	0.1449	0.5286	0.0051**	0.0000**
US-JP	0.0162	0.0002**	0.0049**	0.0047**	0.0000**

We regress exchange rate returns $\tilde{e}_{t+k}$ (expressed in dollars per foreign currency) onto US instruments and local (foreign country) instruments. We investigate horizons $k = 1$, 12 and 60. Hodrick standard errors are used to calculate t-statistics given in parentheses. T-statistics significant at the 95% (99%) level are asterixed * (**). The regressions for Japan are run in levels for dy^{12} and ey^{12} (not logs). The bottom panel reports p-values of χ^2 tests joint across horizons $k = 1$, 12, 60 for exchange rate predictability. The column labeled "US" tests joint predictability of US dividend yields dy^{12}, earnings yields ey^{12} and US short rates r. The column labeled "Local" tests joint predictability of foreign country dy^{12}, ey^{12} and r. The column labeled "Short Rate" tests for joint US and foreign country r predictability. The column labeled "Yields" tests for joint US and foreign country dy^{12} and ey^{12} predictability. Finally, the column labeled "All Variables" tests joint predictability of all US and foreign country instruments. P-values less than 5% (1%) are asterixed * (**).

Table 11: Predictability of US Cashflow Growth

Panel A Monthly MSCI Data 1975:02 - 1999:12

	Univariate Regressions		Trivariate Regression			Joint χ^2 tests	
k	dy^{12} only	ey^{12} only	dy^{12}	ey^{12}	r	k	p-value
Dividend Growth							
1	0.0187	0.0325	-0.0773	0.1316	-0.3899	1	0.0929
	(0.8584)	(1.4840)	(-1.5712)	(2.0094)*	(-0.6604)		
12	0.0121	0.0252	-0.0786	0.1696	-1.0409	12	0.0046**
	(0.4627)	(1.0222)	(01.5810)	(3.3896)**	(-2.4520)		
60	-0.0068	0.0010	-0.0486	0.0768	-0.5373	60	0.0000**
	(-0.1341)	(0.0386)	(-0.1917)	(0.5684)	(-1.7697)		
						1,12,60	0.0000**
Earnings Growth							
1	-0.0496	-0.0524	-0.0201	0.0001	-0.5350	1	0.3272
	(-1.7145)	(-1.4867)	(-0.1876)	(0.0008)	(-0.9218)		
12	-0.0695	-0.1019	0.1274	-0.1029	-1.5681	12	0.0020**
	(-1.9793)*	(-2.4530)*	(1.2425)	(-0.8554)	(-3.1758)**		
60	-0.1198	-0.1022	0.2395	-0.2105	-0.6542	60	0.0029**
	(-1.6343)	(-2.2915)*	(0.9584)	(-1.5339)	(-1.8304)		
						1,12,60	0.0000**

Panel B Quarterly S&P Data 1952:qtr 1 - 1999:qtr 4

	Univariate Regressions		Trivariate Regression			Joint χ^2 tests	
k	dy^{12} only	ey^{12} only	dy^{12}	ey^{12}	r	k	p-value
Dividend Growth							
1	0.0218	0.0510	-0.1773	0.2123	-0.4448	1	0.0000**
	(1.0589)	(2.8944)**	(-5.5105)**	(5.4212)**	(-1.6237)		
12	0.0113	0.0362	-0.1517	0.1692	-0.3666	12	0.0008**
	(0.5129)	(2.0304)*	(-5.1372)**	(4.8437)**	(-1.3685)		
60	0.0007	0.0125	-0.0798	0.0721	-0.0825	60	0.0001**
	(0.0584)	(1.4616)	(-2.5831)*	(3.1690)**	(-0.5354)		
						1,12,60	0.0000**
Earnings Growth							
1	-0.0973	-0.0721	-0.2083	0.1229	-0.7354	1	0.0055**
	(-2.9309)**	(-2.3000)*	(-2.3237)*	(1.3630)	(-1.6867)		
12	-0.0860	-0.0978	-0.0148	-0.0652	-0.7793	12	0.0027**
	(-2.4085)*	(-3.0500)**	(-0.1618)	(-0.7598)	(-1.8278)		
60	-0.0590	-0.0760	0.1552	-0.1908	0.1241	60	0.3693
	(-1.4615)	(-2.6924)**	(1.3119)	(-2.1677)*	(0.2921)		
						1,12,60	0.0364*

We run the predictability regression $\tilde{g}^{12}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{g}^{12}_{t+k} = 12/k(g^{12}_{t+1} + \cdots + g^{12}_{t+k})$ is the cumulated and annualized k-period ahead growth rate of earnings (or dividends), with instruments z_t and k is in months. One-period log growth rates of dividends are constructed similarly. The variables in z_t include dy^{12}_t, log dividend yield using dividends summed over the past 12 months, ey^{12}_t, log earnings yield using earnings summed over the past 12 months, and r_t, the annualized risk-free rate (1 month EURO rate in Panel A and the 3 month T-bill rate in Panel B). In Panel A, we use monthly MSCI data. In Panel B, we use quarterly S&P data. The last column lists p-values of χ^2 tests for joint significance of all three instruments in the trivariate regression for horizons k. T-statistics are in parentheses and calculated using Hodrick standard errors. T-statistics significant at 95% (99%) are denoted with * (**). P-values less than 5% (1%) are also marked with * (**)

Table 12: Pooled-Country Predictability of Cashflow Growth Rates

	Dividend Growth				Earnings Growth			
k	dy^{12}	ey^{12}	r	χ^2 p-val	dy^{12}	ey^{12}	r	χ^2 p-val
1	-0.0849	0.0544	0.1228	0.0155*	0.1505	-0.3664	-0.1929	0.0000**
	(-3.0797)**	(2.6216)**	(0.5291)		(2.1300)*	(-4.3137)**	(-0.4478)	
12	-0.0379	0.0228	-0.3567	0.0920	0.4218	-0.5148	-2.2115	0.0000**
	(-1.3022)	(1.2352)	(-1.7638)		(6.1882)**	(-7.1127)**	(-6.1564)**	
60	0.0152	-0.0099	-0.1418	0.6610	0.2337	-0.2297	0.0823	0.0000**
	(1.0057)	(-1.7650)	(-0.9202)		(3.0799)**	(-5.3331)**	(0.3360)	

The predictability regression $\tilde{g}^{12}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$, where $\tilde{g}^{12}_{t+k} = 12/k(g^{12}_{t+1} + \cdots + g^{12}_{t+k})$ is the cumulated and annualized k-period ahead growth rate of earnings (or dividends), with instruments z_t and k is in months, is estimated jointly across the US, UK, France and Germany, constraining the coefficients to be the same across countries (see Appendix E for details). Japan is excluded from estimation because of a shorter sample and negative earnings yields during 1999. The constants in the regressions are allowed to differ across countries (and are not reported in the table). T-statistics of the predictability coefficients are in parentheses and are based on Hodrick (1992) standard errors. The χ^2 test is a joint test that all the β coefficients equal zero. T-statistics significant at 95% (99%) are denoted with * (**). P-values less than 5% (1%) are denoted with * (**).

Table A-1: Null Model VAR Estimation

Constant and Companion Form

	const	Companion form A		
Null Model				
	μ	r_{t-1}	g_{t-1}	po_{t-1}
r_t	0.0002	0.9711	0.0000	0.0000
	(0.0001)	(0.0140)		
g_t	0.0216	-0.2894	-0.0078	0.0079
	(0.0012)			
po_t	-0.0222	1.2894	0.0078	0.9921
	(0.0012)	(0.0900)	(0.0105)	(0.0010)
Alternative Present Value Model				
	μ	r_{t-1}	g_{t-1}	po_{t-1}
r_t	0.0002	0.9711	0.0000	0.0000
	(0.0001)	(0.0140)		
g_t	0.0125	-0.2452	0.2739	0.0096
	(0.0037)	(0.0597)	(0.1669)	(0.0036)
po_t	-0.0104	0.0584	0.0204	0.9867
	(0.0024)	(0.9690)	(0.2237)	(0.0044)

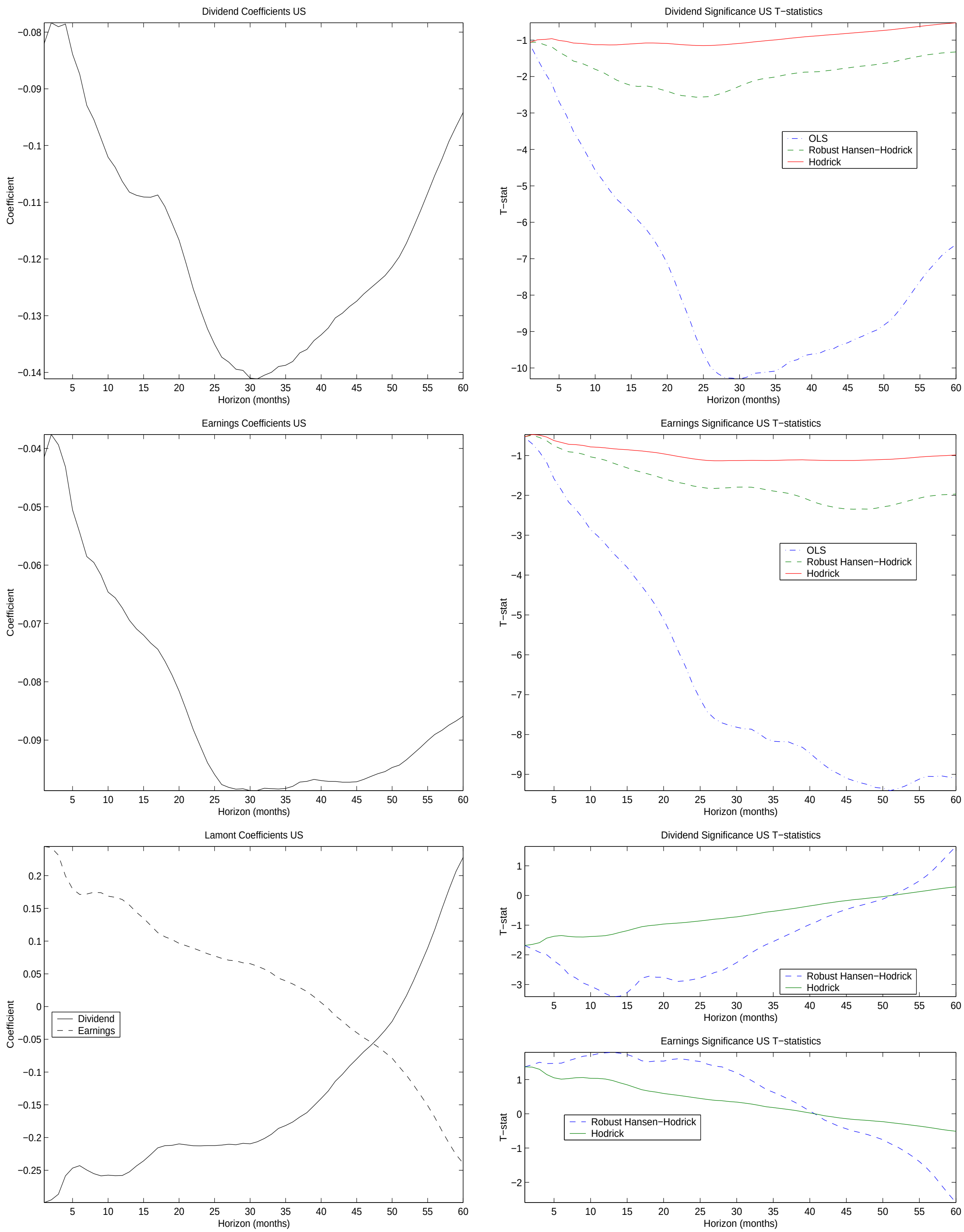
Conditional Volatilities and Correlations

	Volatility	Conditional Correlations		
	σ	r_t	g_t	po_t
r_t	0.0007	1.0000		
	(0.0000)			
g_t	0.0647	0.1038	1.0000	
	(0.0137)	(0.0318)		
po_t	0.0351	-0.1758	-0.8496	1.0000
	(0.0129)	(0.4035)	(0.1874)	

We estimate the monthly VAR $X_t = \mu + AX_{t-1} + \epsilon_t$, $\epsilon_t \sim N(0, \Sigma)$ of $X_t = (r_t \, g_t \, po_t)'$, where r_t is the continuously-compounded risk-free rate, g_t is monthly (unobserved) earnings growth and po_t is the monthly (unobserved) log payout ratio. We set $A_{12} = A_{13} = 0$, where subscripts denote matrix elements (row and column). Estimation of the Alternative VAR proceeds in two steps. First, we estimate the equation for r_t on US EURO 1 month rates. Second, holding these parameters as fixed, we estimate the remaining parameters in μ, A, and Σ using Simulated Method of Moments. We match the first and second moments of MSCI data on log earnings growth and log payout ratio which use summed earnings and dividends over the past year in their construction. We also use the cross-moment of current (annual) growth and (annual) payout with lagged one-year growth and payout. The Alternative Model is exactly identified. To estimate the Null Model we hold fixed the covariance matrix Σ from the Alternative Model. Then using first and cross-moments we estimate μ and A. The Null Model's VAR is estimated subject to the restrictions:

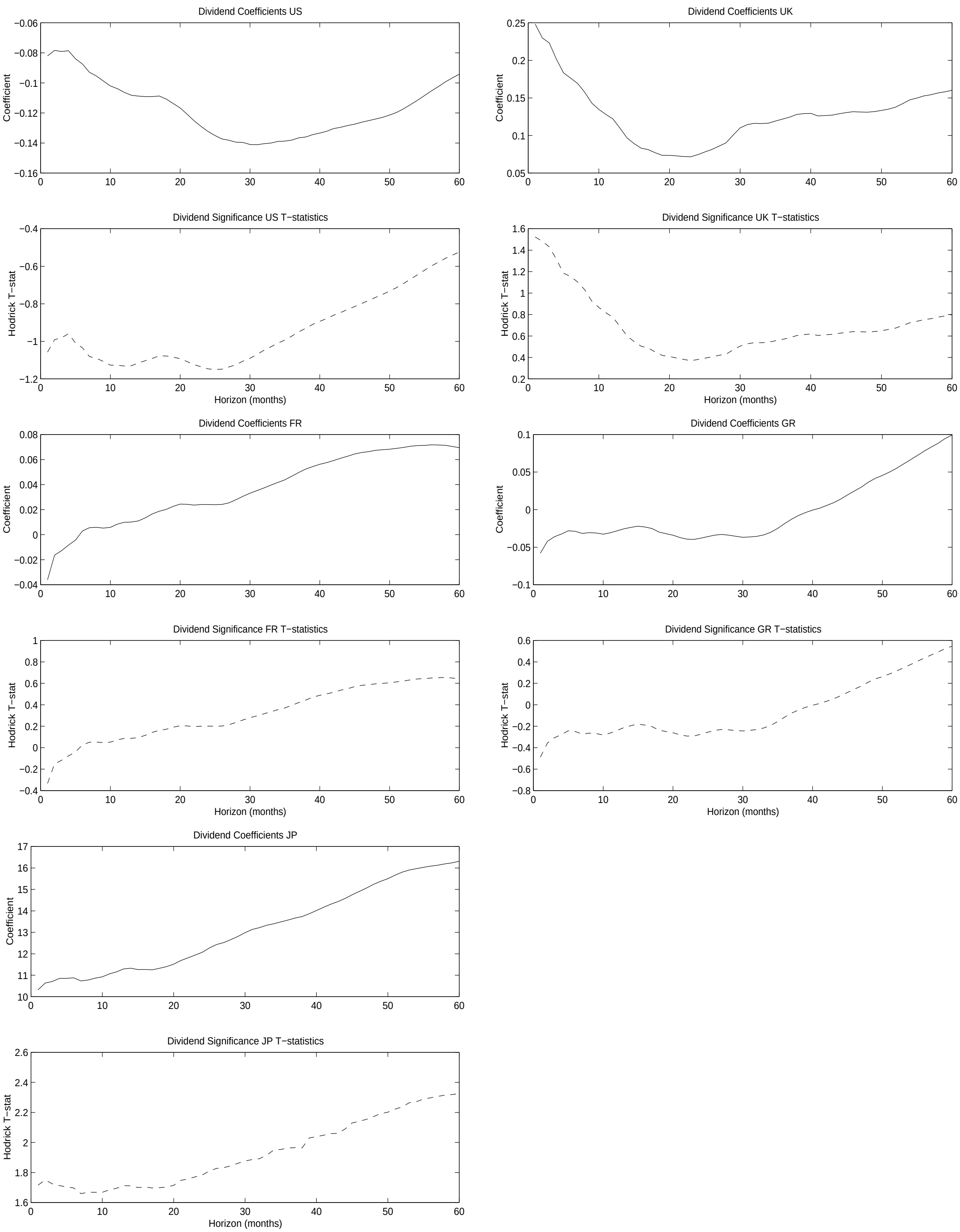
$$A = \begin{pmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & A_{23} \\ 1 - A_{21} & -A_{22} & 1 - A_{23} \end{pmatrix}.$$

The Null Model is over-identified. A χ^2 test of the over-identify restriction yields a statistic of 0.4710, which has a p-value of 0.9252.



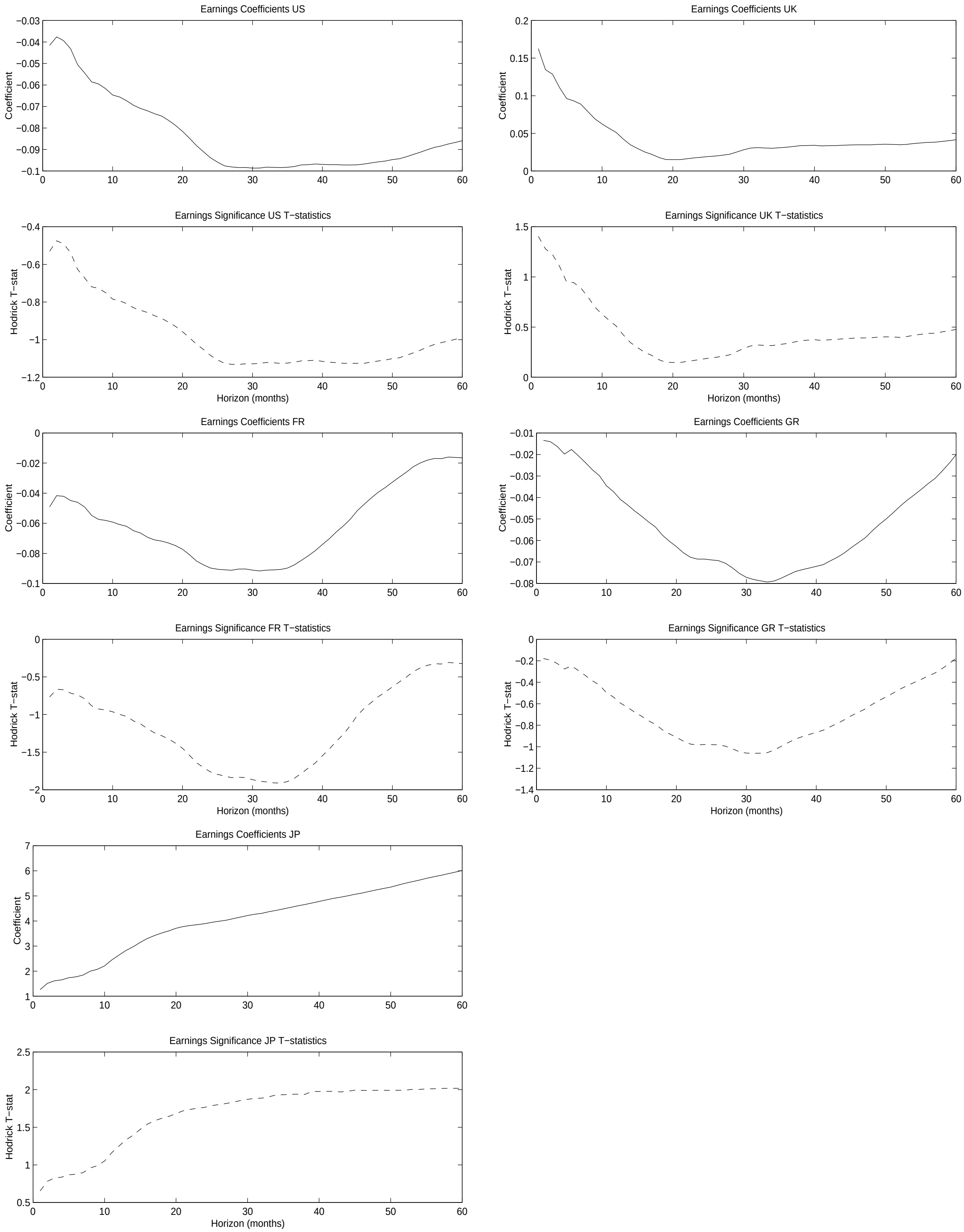
The left column shows coefficients β in the regression $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return, with instruments $z_t = dy_t^{12}$ (log dividend yields) in the top row, $z_t = ey_t^{12}$ (log earnings yields) in the middle row and $z_t = (dy_t^{12}, ey_t^{12})'$ in the bottom row. The right column shows t-statistics from these regressions. Horizons k are on the x-axis.

Figure 1: Coefficients and Standard Errors from US Regressions



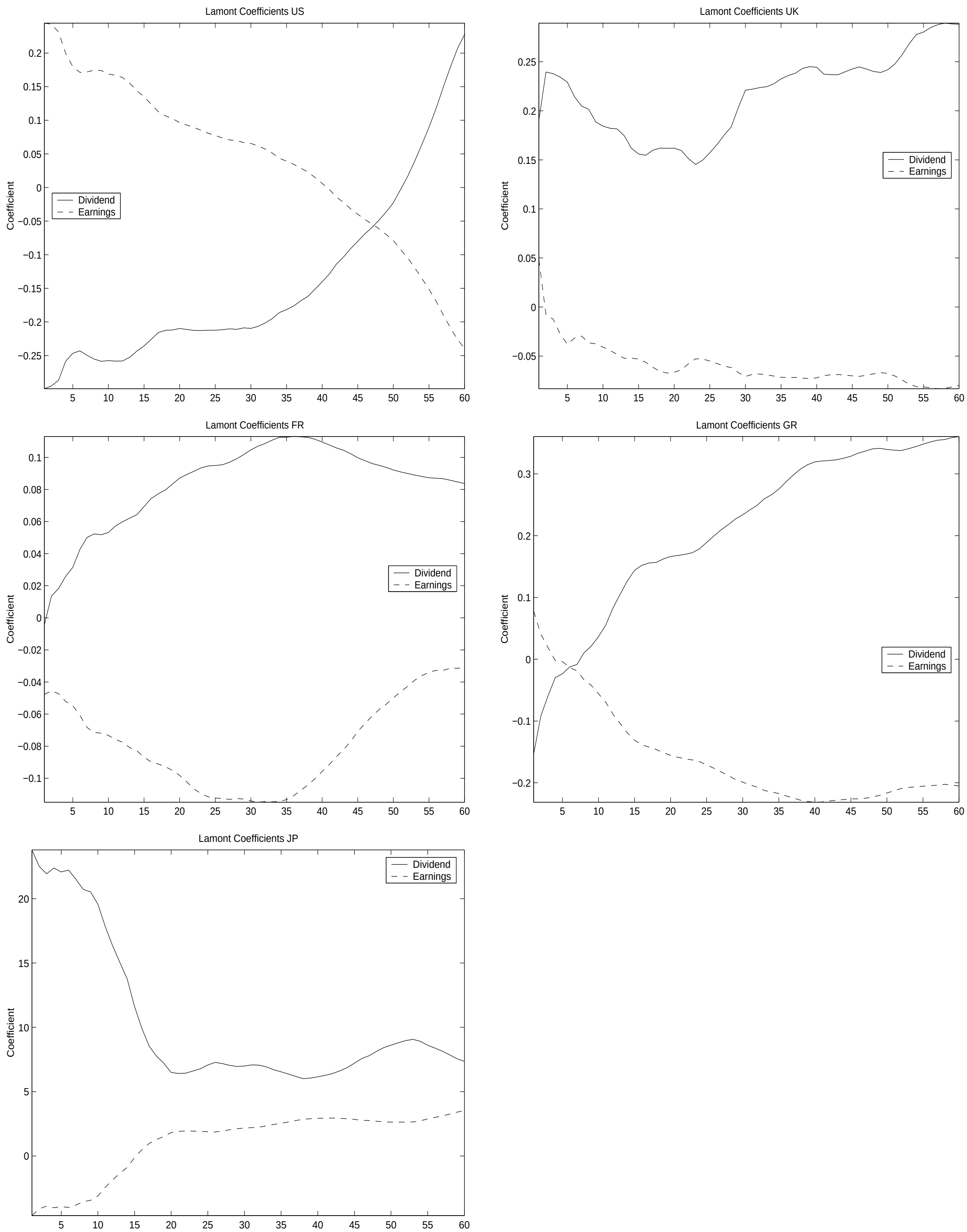
Dividend coefficients and t-statistics calculed using Hodrick standard errors for the regression $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return, with instruments $z_t = dy_t^{12}$. Horizons k are on the x-axis. For Japan, the dividend yield is in levels but for all other countries dy_t^{12} represents the log dividend yield.

Figure 2: Dividend Regressions in 5 Countries



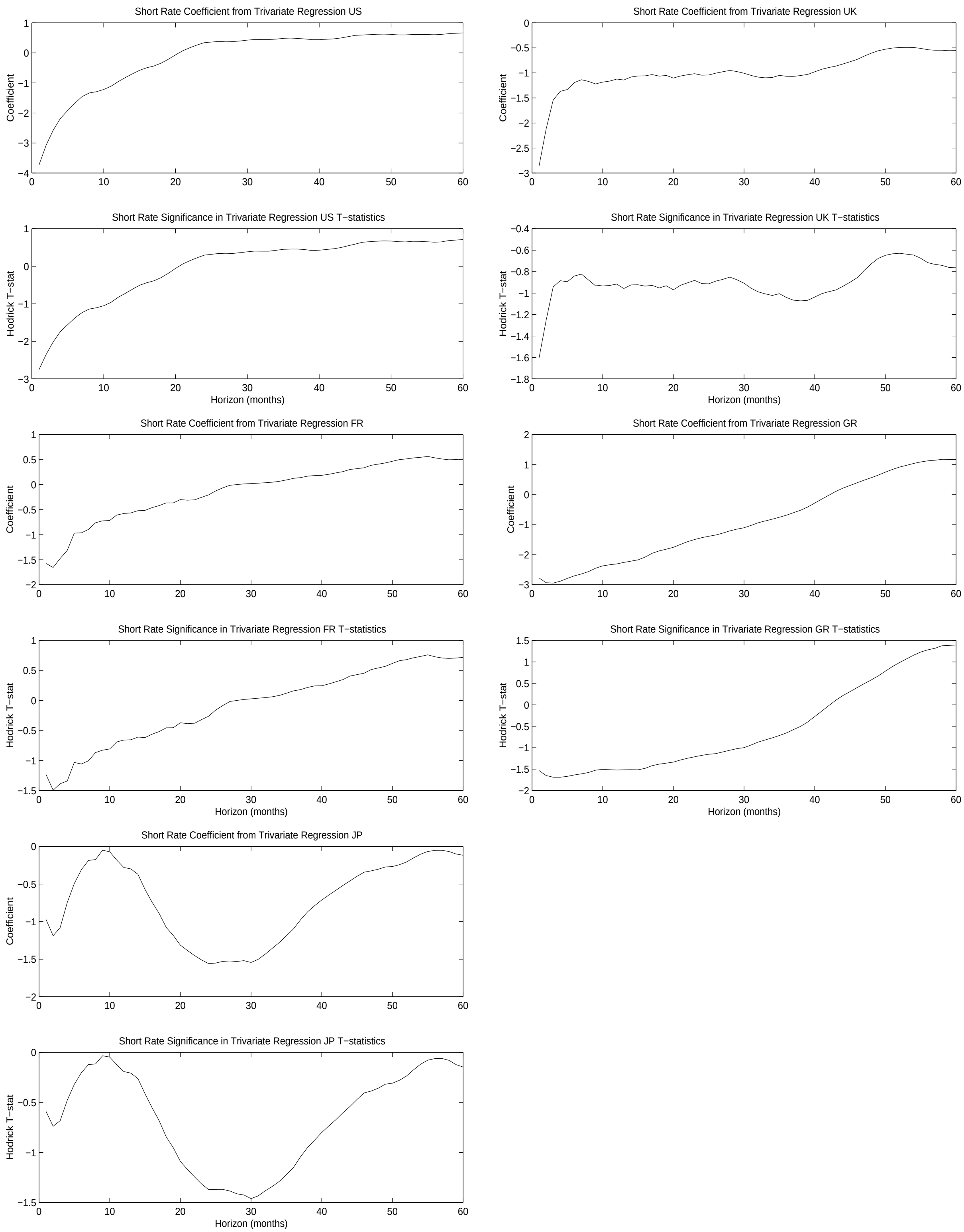
Earnings coefficients and t-statistics calculated using Hodrick standard errors for the regression $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return, with instruments $z_t = ey_t^{12}$. Horizons k are on the x-axis. For Japan, the earnings yield is in levels but for all other countries ey_t^{12} represents the log earnings yield.

Figure 3: Earnings Regressions in 5 Countries



Dividend and earnings coefficients β from the regression $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return, with instruments $z_t = (dy_t^{12}\; ey_t^{12})'$. Horizons k are on the x-axis. For Japan dividend and earnings yields in levels are used, all other countries use log dividend and earnings yields.

Figure 4: Lamont Regressions Coefficients in 5 Countries



Short rate coefficients and t-statistics constructed using Hodrick standard errors from the trivariate regression $\tilde{y}_{t+k} = \alpha + z_t'\beta + \epsilon_{t+k,k}$ where $\tilde{y}_{t+k} = 12/k(y_{t+1} + \cdots + y_{t+k})$ is the cumulated and annualized k-period ahead return, with instruments $z_t = (dy_t^{12}\, ey_t^{12}\, r_t)'$. The short rate r_t is annualized. We report only the coefficient on r_t. Horizons k are on the x-axis. For Japan dividend yields dy_t^{12} and earnings yields ey_t^{12} are in levels, for the other countries these are log dividend and earning yields respectively.

Figure 5: Short Rate Coefficients in Trivariate Regressions in 5 Countries