Dividend Yields for Forecasting Stock Market Returns

A Cointegration Analysis Based on the ARDL Approach

by

ANSGAR BELKE* and THORSTEN POLLEIT**

*University of Hohenheim, Stuttgart,
**Barclays Capital and Business School of Finance & Management, Frankfurt

2 September 2004

Abstract

This paper empirically assesses the ability of dividend yields to predict future stock returns in Germany under the assumption of the Efficient Market Hypothesis under Rational Expectations (EMH-RE). Since the order of integration of regressors are not exactly known, a new bounds procedure, namely an autoregressive distributed lag (ARDL) model, is applied to test for cointegrating relationships among future stock returns and today's dividend yield. This procedure is efficient for small samples and capable of dealing with the controversial issue of exogeneity of the dividend yield. Additionally, ARDL and error-correction models are estimated for (future) stock returns and the dividend yield on consistent estimates and standard normal asymptotic theory. Short-run and long-run impacts of the dividend yield on future stock returns in Germany are identified only if: (a) stock market returns are measured by the annualised one-month dividend growth;(b) if a very specific lag structure is imposed. Hence, it cannot be claimed as a rule, that dividend yields are generally useful for forecasting stock market returns.

*JEL Classification*s: C22, G12, G14.

*Keyword*s: Asset Prices, Autoregressive Distributed Lag Models, Cointegration, Dividend Yields, Long-Run Relationships, Stock Returns.

Contact:

Prof. Dr. Ansgar Belke
University of Hohenheim
Department of Economics
Chair of International Economics (520E)
70593 Stuttgart
Germany
Tel.: ++49(0)711-459-3246
Fax: ++49(0)711-459-3815
E-mail: belke@uni-hohenheim.de

Dr. Thorsten Polleit
Barclays Capital and
Honorary Professor at the
Business School of Finance & Management
Frankfurt
Germany
Tel.: ++49(0)69-56-04-32-29
Fax: ++49(0)69-7161-15 99
E-mail: thorsten.polleit@barclayscapital.de

1. Introduction

The forecasting power of the dividend yield (i.e. the ratio between dividend payments and the stock price) on future stock market returns is a hypothesis that has a long tradition among practitioners and academics (for example, Dow (1920), Ball (1978). The theoretical and empirical literature offers evidence that expected stock returns are predictable. However, the predictable component of stock market returns, or equivalently the variation through time of expected returns, is a relatively small fraction of return variances (Fama and Schwert (1977), Fama (1981), Keim and Stambaugh (1986) and French, Schwert and Stambaugh (1987)). Another interesting finding is that the power of the dividend yield to forecast future stock returns, measured by the simple coefficient of determination, increases with the time horizon under review (Fama and French (1988)). Fama and French (1988) offer two explanations: (i) that high autocorrelation causes the variance of unexpected returns to grow faster than the return horizon, and (ii) the growth of the variance of unexpected returns with the return horizon is attenuated by a discount rate effect: shocks to expected returns generate opposite shocks to current prices.

In view of these findings, several issues have remained unresolved from our point of view. First, by construction, the use of dividend yields and future stock returns in the framework of the efficient market hypothesis and rational expectations (EMH-RE) generally leads to *small sample sets*. Second, in linear regression analysis there is only a limited number of degrees of freedom due to the use of moving averages in calculating stock return measures. However, the distribution of the test statistics is only known for larger sample sets. As a consequence, there is often *no clear information on the integration and cointegration properties* of the data. Thus, whether variables should be introduced in differenced or levels form is questionable. The latest studies in this field (Fama and French (1988) and Domanski and Kremer (1998)) have not sufficiently addressed this issue.

A procedure that avoids these difficulties and appears to be eminently suitable for the problem at hand is that proposed by Pesaran, Shin and Smith (1996) and Pesaran and Shin (1999). First, it is *as efficient as possible* in the case of small samples. Second, it is capable of dealing with the controversial issue of *(lack of) exogeneity* of the dividend yield. Also, it has the additional advantage of yielding *consistent estimates of the long-run coefficients* that are asymptotically normal, *irrespective of whether the underlying regressors are I(0) or I(1)* and of the extent of cointegration. This is a key property since a second objection raised in the empirical finance literature is that it is not clear whether stock market performance measures (such as holding period return, dividend growth and holding period returns minus dividend growth) are stationary (I(0)) or integrated of order one (I(1)) within the specific sample chosen. However, economic reasoning would suggest that stock market returns should be stationary, i.e. stock market returns should not "outperform" (world) output growth on a sustained basis. A third objection against the usual procedures in assessing the impact of dividend yield on asset prices, is that these procedures do not allow one to distinguish clearly between long-run and short-run relationships, as they estimate VARs only in differences or only in levels. The procedure used in this paper will also allow *the correct dynamic structure* to be obtained.

In this paper we apply the procedure proposed by Pesaran, Shin and Smith (1996) on *monthly* data for the German stock market in the period August 1974 to September 2003. The approach used here involves *two steps*. As a *first step* we test the null hypothesis that there exists no long-run relationship between the levels of the variables under consideration using the bounds procedure by Pesaran, Shin and Smith (1998). In the spirit of their study, we suggest moving to the next stage only if the null hypothesis is rejected. The test is the standard Wald or F-statistic for testing the significance of the lagged levels of the variables in a first difference ARDL regression (with a non-standard distribution under the null). If the result is significant, we take the *second step* and estimate the *long-run coefficients* and the corresponding *error-correction models* using the ARDL procedure described in Pesaran and Shin (1999). The use of error-correction models has only recently become popular in analyzing the impacts of dividend yield on asset prices. However, it has not yet been applied to the relation between dividend yields and future stock returns in Germany.

In the first part of the paper, the bounds-testing procedure proposed by Pesaran, Shin and Smith (1996) is applied to the estimation of the impact of dividend yields on future German stock market returns for 3 to 48 months. The existence of a long-run relationship between future stock returns and the dividend yield is examined. In the second part of the paper, the respective long-run relationship and the respective short-term dynamics are estimated. Some new econometric techniques proposed by Pesaran and Shin (1999) are applied to improve on some of the critical points of earlier studies on the role of dividend yield variables in forecasting future stock returns.

2. Dividend Yield Impacts Future Stock Prices?

It is economically reasonable to think of a stock's fundamental value as the sum of a firm's discounted expected future cash flow. The discount rate used can be interpreted as the required (expected) rate of return that attracts investors to hold the asset in their portfolios. In an information efficient market, a stock's market price should then equal its fundamental value as calculated by all or the marginal investor depending on whether expectations are assumed to be homogenous or not. Applied to the stock market, this general valuation approach leads to the dividend discount model. In line with Campbell, Lo and Shiller (1997, pp. 260-2) the approximation formula for the continuously compounded one-period return on stocks is:[1]

(1) $$h_{t+1} = k + \rho p_{t+1} + (1-\rho)d_{t+1} - p_t$$

where h_{t+1} = approximate continuously compounded one-period return on stocks over the holding period t+1. p_t = log of stock price at the end of t; d_{t+1} = log of dividend paid out before the end of period t+1; $\rho \equiv 1/(1 + \exp(\overline{d-p}))$, where $\overline{d-p}$ = average of log of dividend yield; and $k = -\log(\rho) - (1-\rho)\log(1/\rho - 1)$.

Equation (1) shows a log-linear relation between stock prices, returns and dividends. It is a first-order linear difference equation in the stock price. Solving forward and imposing the terminal condition $\lim_{j\to\infty} \rho^j p_{t+j} = 0$ yields:

(2) $$p_t = \frac{k}{1-\rho} + \sum_{j=0}^{\infty} \rho^j \left[(1-\rho)d_{t+1+j} - h_{t+1+j}\right].$$

Equation (2) is an *ex post*-identity, which says that today's stock price is high if future dividends are high and/or future returns are low. Applying the conditional expectation operator $E_t x_{t+1} = E[x_{t+1} \mid \Omega_t]$, where Ω_t = market-wide information set available at the end of period t, and the law of iterative expectations, equation (2) be changed to an ex ante relation:

[1] See Cuthbertson et al. (1997), pp. 1005.

$$(3)\qquad p_t = \frac{k}{1-\rho} + \sum_{j=0}^{\infty} \rho^j \left[(1-\rho) E_t d_{t+1+j} - E_t h_{t+1+j} \right].$$

Assuming homogenous expectations and instantaneous market clearing, the log stock price always equals its single fundamental value, which is the specifically weighted, infinite sum of expected log dividends discounted by principally time-varying expected equilibrium returns. Combined with RE, equation (3) represents the rational valuation formula (RVF).

The log-linear approximation framework has two advantages. First, it allows a linear and thus simple analysis of the stock price behaviour. Second, it conforms with the empirically plausible assumption that dividends and stock return follow log-linear stochastic processes. For empirical analyses, equation (3) can be rearranged such that the log dividend yield (or log dividend–price ratio) is singled out as the left-hand variable:

$$(4)\qquad d_t - p_t = \frac{k}{1-\rho} + \sum_{j=0}^{\infty} \rho^j \left[-E_t \Delta d_{t+1+j} + E_t h_{t+1+j} \right].$$

The current dividend yield should predict future returns if the discount rates used by forward-looking investors actually depend on expected holding period returns for subsequent periods, and if these expectations do not deviate systematically, and too much, from realised returns.[2]

The log-linear relation between prices, dividends and returns provides an accounting framework, which provides an economic interpretation of the relationship between the dividend yields and future stock market return measures (Campbell, 1991; Cuthbertson, 1996). High prices must eventually be followed by high future dividends, low future returns, or some combination of the two. If investor expectations are consistent with this interpretation, high prices must be associated with high expected future dividends, low expected future returns, or some combination of the two. Similarly, high returns must be associated with upward revisions in expected future dividends, downward revisions in expected future returns, or some combination of the two.

3. Testing for the Existence of Long-Run Relations

The test for the existence of long-run relations between stock market returns and dividend yields was conducted for the German stock market for the period August 1974 to September 2003. We used monthly data provided by Datastream and calculated three alternative future stock market return measures (dependent variables): (i) annualised one-month continuously compounded stock returns (*h*), (ii) annualised one-month dividend growth rates in percent (Δd) and (iii) and the difference between the two ($h - \Delta d$).[3] The measures were calculated over holding periods of 1, 3, 12, 24, 36 and 48 months. These performance measures were regressed on the independent variable, that is the dividend yield (*dp*), after we have reassured that there is no problem of “reverse causation”, i.e. that the dividend yield really is the ‘forcing variable’.

2 See Domanski and Kremer (1998), p. 26. Note that the term $E_t h_{t+1+j}$ is the equivalent to the expected future discount rate of the RVF, a finding which will be explained in the following.

3 Regressions for dividend and profit growth are subject to the omitted variables problem because, in that case, expected stock returns introduce noise. To circumvent this problem the differences between *h* and $h - \Delta d$ were also calculated.

Figure 1 shows three scatter plots for the variables over a time horizon of 12-months. It shows cross-plots of three measures of stock returns against the dividend yield, respectively. The charts suggest that the positive (negative) relationship between the *dp* dividend yield and *h* and $h–\Delta d$ holds for the German stock market. Also, as indicated by theoretical considerations outlined earlier, the relation between *dp* and Δd is negative. However, what matters for our empirical analysis, is that the overall relationships in the charts show a clear positive or negative relation - rather than a vertical or horizontal one. Figure 2 shows the variables under review over time.

(Figure 1 about here)

(Figure 2 about here)

As a *first step*, we estimated the long-run relations between various stock market performance measures (measured over holding periods (*K*) ranging from one month to four years) and the dividend yield. The results in Table 1 represent baseline estimations, which will serve as benchmarks against which the results gained from the autoregressive distributed lag procedure will be evaluated later on in this paper.[4] As can be seen, the R-squared systematically increases with the forecast horizon. The same is valid for the Newey-West adjusted empirical realisations of the t-values for the dependent variable $h–\Delta d$. However, in the cases of $x = h$ and $x = \Delta d$ the t-values reach their maximum after three and two years, respectively. The slope coefficients reveal a positive sign with $x = h$ and $x = h–\Delta d$ and a negative one in the case of $x = \Delta d$. They tend to reach their maximum in absolute values after 3 months and decrease afterwards.

(Table 1 about here)

Although there are serious doubts about the statistical reliability of long-horizon regressions, these results seem to suggest that future stock returns, and especially future dividend growth, might contain predictable components that are reflected in the current dividend yield. On a purely statistical basis, the finding that the ability of the dividend yield to forecast future stock returns increases with the return horizon is widely attributed to the central fact that it is a rather persistent variable (Cochrane, 2001, pp. 391; Hodrick, 1992). Economically, the finding might indicate that market agents can forecast medium- and-long-term prospects of the economy much easier that short-term fluctuations. A relatively stable monetary framework, that is, for instance, a stable reaction and objective function of the central bank and relatively few serious financial market shocks might be held responsible for this outcome. Finally, it should be noted that the predictability of future stock returns does not contradict the efficient market hypothesis, which postulates only that abnormal returns are unpredictable, not that actual returns are unpredictable.

The rather low realisations of DW-statistics indicate serial autocorrelation in the residuals. We corrected for serial correlation and potential heteroskedasticity by using alternative t-statistics proposed by Newey and West (1987) to compensate the data-overlap for the forecasts beyond one month (this leads to serial correlation of the error terms, even under the null hypothesis of no stock return predictability through the dividend yield[5]). By this, we also cope with the need to use asymptotic theory to generate standard errors. This need emerges from the fact that the

[4] We also experimented with different truncation lags, but the results did not change materially.
[5] In this case, errors are correlated with the K-1 previous error terms.

dividend yield as the regressor is a predetermined value and is not exogenous (Campbell, Lo, and Shiller, 1997, pp. 334-336).

As in Domanski and Kremer (1998), we have dispensed with testing the order of integration of the dividend yield and the stock return at this stage of analysis. By doing so, one might interpret the results as providing preliminary evidence that future stock returns, and especially future dividend growth, contain predictable components with are reflected in the current dividend yield. However, it cannot be ruled out that the variables under consideration represent non-stationary series. If this is the case, cointegration theory prevents inferences to be made from the t-values of the coefficient estimates. Therefore, it is of interest to analyse if these results hold up robust when using the approach of Pesaran, Shin and Smith (1996) and Pesaran and Shin (1999), respectively.

3.1. Testing for Cointegration: The Pesaran, Shin and Smith ARDL Approach

3.1.1. Theoretical Background

As mentioned above, an important problem inherent in the residual-based tests and in some system-based tests for cointegration is the precondition that it must known with certainty that the underlying regressors in the model are I(1). However, given the low power of unit root tests, there will always remain a *certain degree of uncertainty with respect to the order of integration* of the underlying variables. For this reason, we now make use of the approach proposed by Pesaran, Shin and Smith (1996) to test for the existence of a linear long-run relationship, when the orders of integration of the underlying regressors are not known with certainty. The test is the *standard Wald or F statistic* for testing the significance of the lagged levels of the variables in a first-difference regression. The involved regression is an error-correction form of an autoregressive distributed lag (ARDL) model in the variables of interest.

More specifically, in the case of an unrestricted ECM, regressions of y on a vector of x's, the procedure first requires estimating the following model derived by Pesaran, Shin and Smith (1996, pp. 2 ff.):

$$\Delta y_t = a_{0y} + a_{1y} \cdot t + \phi y_{t-1} + \delta_1 x_{1,t-1} + \delta_2 x_{2,t-1} + \ldots + \delta_k x_{k,t-1} +$$

(8)

$$\sum_{i=1}^{p-1} \psi_i \Delta y_{1,t-i} + \sum_{i=0}^{q_1-1} \varphi_{1i} \Delta x_{1,t-i} + \sum_{i=0}^{q_2-1} \varphi_{i2} \Delta x_{2,t-i} + \ldots + \sum_{i=0}^{q_k-1} \varphi_{ik} \Delta x_{k,t-i} + \xi_{ty}$$

with ϕ and δ as the long-run multipliers, Ψ and φ as short-run dynamic coefficients, (p,q) as the order of the underlying ARDL-model (p refers to y, q refers to x), t as a deterministic time trend, k as the number of 'forcing variables', and ξ uncorrelated with the Δx_t and the lagged values of x_t and y_t.

As a second step, one has to compute the usual F-statistic for testing the joint significance of $\phi = \delta_1 = \delta_2 = ... = \delta_k = 0$. However, the asymptotic distributions of the *standard Wald or F statistic* for testing the significance of the lagged levels of the variables are *non-standard* under the null hypothesis that no long-run relationship exists between the levels of the included variables. Pesaran, Shin and Smith (1996) provide *two sets of asymptotic critical values*; one set assuming that all the regressors are I(1); and another set assuming that they are all I(0). These two sets of critical values provide a band covering all possible classifications of the regressors into I(0), I(1), or even mutually cointegrated.

A third step is required in order to use the appropriate bounds testing procedure. The test proposed by Pesaran, Shin and Smith (1996) is consistent with this. For a sequence of local alternatives, it has a non-central χ^2-distribution asymptotically. This is valid irrespective of whether the underlying regressors are I(0), I(1) or mutually cointegrated. The recommended procedure based on the F-statistic is as follows. The F-statistic computed in the second step is compared with the upper and lower 90, 95 or 99 percent critical value bounds (F_U and F_L). As a result, three cases can emerge. If $F > F_U$, one has to reject $\phi = \delta_1 = \delta_2 = ... = \delta_k = 0$ and conclude that there is a long-term relationship between y and the vector of x's. However, if $F < F_L$, one cannot reject either $\phi = \delta_1 = \delta_2 = ... = \delta_k = 0$ or the hypothesis that a long-run relationship does not exist. Finally, if $F_L < F < F_{U,}$ the inference has to be regarded as inconclusive. The order of integration of the underlying variables has to be investigated more deeply.

In order to select the so-called ‘forcing variables’, the above procedure should be *repeated* for ARDL regressions of *each* element of the vector of x's on the remaining relevant variables (including y). For example, in the case of k = 2, the repetition should concern the ARDL regressions of x_{1t} on (y_t, x_{2t}) and x_{2t} on (y_t, x_{1t}). If the linear relationship between the relevant variables is not 'spurious' can no longer be rejected, one can estimate coefficients of the long-run relationship by means of the ARDL-procedure. This estimation procedure is discussed in section 4.

3.1.2. Application to German Stock Market Data

Since the choice of the orders of the lagged differenced variables in the unrestricted ECM specification can have a significant effect on the test results, models in the log of stock market returns and the logs of the other mentioned stock market relevant variables are estimated for the orders p = q = 1, 4, 12. Finally, in the absence of *a priori* information about the direction of the long-run relationship between *h*, *Δd* or *h–Δd* and the other stock market variables, we estimate unrestricted ECM regressions of *h* (y) on the vector of stock market variables (x) as well as the reverse regressions of x on y. More specifically, in the case of the unrestricted ECM regressions of y on x, we re-estimate model (1) using monthly observations over a maximum sample from 1974(8) to 2003(9). In view of the monthly nature of observations we set the maximum orders to 12, (i.e. we estimate eq. (1) for the order of $p = q_1 = q_2 = 12$ over the same sample period 1974(8) to 2003(9)). It is important to note already at this early stage of investigation that we have to choose p and q *quite liberally* in order to endogenise the log of stock market returns (detailed proofs can be found in Pesaran and Shin (1999) and Pesaran, Shin and Smith (1996)). In addition, due to the seasonality in the data, first differences in the variables at order 12 are used.[6]

Like in any long-horizon analyses, we are aware of risks that some events such as, for instance, the German reunification, the introduction of the euro area on 1 January 1999, the international financial market crisis 1997-98 and, more recently, the international stock market crash around 2000-01, might have dramatically changed the pricing action in stock markets. We decided to rely more on estimates, which take German reunification explicitly into ac-

6 Unfortunately, this option is not available under Microfit 4.0 for the ARDL-estimation in the second part of this paper. Thus, seasonality has to be captured by implementing a sufficient number of lags of the endogenous variable on the RHS of the ARDL-equation.

However, the inclusion of deterministic seasonal dummies leaves inference concerning the relevant test-statistic unaffected. See Davidson and McKinnon (1993), p. 705. If we include seasonal dummies the results do not change substantially.

count by means of a point dummy D901. This dummy implies a *permanent* change in the relation between the stock market return and the other stock market relevant variables. We distinguish between *three different definitions of stock returns* (cases $x = h$, $x = \Delta d$, and $x = h–\Delta d$). Our models are structured as follows:

- *Model 1*: the holding period return, *h*, the dividend yield, *dp*, and a constant are included in the long-term relation.
- *Model 2*: the dividend growth, Δd, the dividend yield, *dp*, and a constant are included in the long-term.
- *Model 3*: the holding period return minus dividend growth, $h–\Delta d$, the dividend yield, *dp*, and a constant included in the long-term.

These specifications allow the dividend yield to slow down the adjustment to a new stock market equilibrium in the wake of a shock.[7] The three models represent the core implication derived above, namely that in the long run, the dividend yield is in long-term equilibrium with the average stock market return. Thus, the modelling approach is strictly *guided by theory*. The following estimations, like all other computations in this paper, have been carried out using the program Microfit 4.0 (see Pesaran and Pesaran (1997)). We now let the data tell us which of the above models case fits the German stock market data best. Tables 1a to 1c display the empirical realisations of the F-statistics for testing the existence of a long-run relationship between stock market return measures and the dividend yield. In all cases, the underlying equations pass the usual diagnostic tests for serial correlation of the residuals, for functional form misspecification and for non-normal and/or heteroskedastic disturbances.

The 90, 95 and 99 percent lower and upper critical values bounds of the F-test statistic that are dependent on the number of regressors and dependent on whether a *linear trend* is included or not, are originally given in Table B in Pesaran, Shin and Smith (1996) and usefully summarized in Pesaran and Pesaran (1997) (see Annex C, Statistical Tables, Table F). The critical value bounds for the application without trend are given in the middle panel of this Table F at the 90 percent level by 4.042 to 4.788, at the 95 percent level by 4.934 to 5.764 and at the 99 percent level by 7.057 to 7.815. For the application with a linear trend the respective upper bound critical values can be found in the lower panel of Table F: 5.649 to 6.335 (at the 90 percent level), 6.606 to 7.423 (at the 95 percent level) and 9,063 to 9.786 (at the 99 percent level). We took the upper bound critical values from these intervals and tabulate them in Tables 1a to 1c as the relevant conservative benchmarks to check the significance of the cointegration relationships. We also experimented with the inclusion of a dummy which approximated the international stock market turbulences and took the value 1 as from 2000(1) and 0 otherwise. We finally decided to put it into the test equation including a deterministic trend in order to grasp *inter alia*, the U-turn shape of the dividend yield curve for Germany with the trough in January 2000.

According to the empirical F-values in Tables 2a to 2c, we find that the null hypothesis of no long-run relationship in the case of unrestricted ECM regressions of the log of stock returns on the dividend yield and other open economy stock market variables is *rejected* in 18 cases at $\alpha = 0,05$ and in most of the cases even at the 1 percent level. 10 of these cases emerge if a

---

[7] In principle, a more sophisticated specification our hypothesis could have made the impact of dividend yield dependent on the sign of the error-correction term (negative, if the latter is positive and vice versa) via e.g. the sign function. However, this way of modelling is certainly beyond the scope of this paper.

deterministic trend is excluded. However, the null hypothesis of no cointegration tends not to be rejected if the moving average of the relevant variables is below 12 months (the only exception is *h*1 with trend) or if it is higher than 24 months (except $\Delta d36$ and $\Delta d48$ without trend).

(Table 2a. about here)

(Table 2b. about here)

(Table 2c. about here)

Overall, these results suggest *strong evidence in favour of the existence of a long-run relationship* between the (future) stock market return and the dividend yield and the constant, at least if the relevant variables are moving averages over 12 or 24 months. But in view of the high levels of cross-sectional and temporal aggregation, it is not possible to know *a priori* whether the dividend yield is the 'long-run forcing' variable for the average future stock market return performance. Therefore, we considered all possible regressions and substitute the *change* in the stock market return measures as the dependent variable in equation (8) by the *change* in the dividend yield, in order to test whether this relationship is *spurious* in the sense that we do not capture the 'correct direction of causation'. For instance, we have to ensure that the future stock market return is not among the forcing variables. The results of the reversed test equations are displayed in the final column of Tables 2a to 2c. In the case of $x = \Delta d$ and for a wide range of moving averages (12 to 48 months), we find that the *direction* of this relation is most likely to be *from the dividend yield to future stock market returns,* so that the variable *dp* can be considered as the 'long-run forcing' variable for the explanation of the variable Δd. As a consequence, in this case the parameters of the long-run relationship can now be estimated using the ARDL procedure discussed in Pesaran and Shin (1999). However, in the cases of $x = h$ and $x = h - \Delta d$ where the variables are 12-month moving averages, our bounds procedure reveals that the dividend yield and the stock returns are 'forcing variables' for each other (i.e. that there seems to be a two-way causation between them). However, in the cases of $x = h$ and $x = h - \Delta d$ where the variables are 24-month averages, future stock returns even appear to be the forcing variable' for the dividend yield. Therefore, in the following section *[??can you specify which section??]*, we will concentrate on the case $x = \Delta d$.

However, before this is done, some complementary tests for cointegration on the basis of models 1 and 2 in an earlier version of the paper should be conducted. When using cointegration analysis in the Johansen-framework (Johansen (1991, 1995)), we would first need to establish that all the underlying variables are I(1). However, such pre-testing results may adversely affect the test results based on cointegration techniques (Cavanaugh et al. (1995), Pesaran (1997)). This insight already motivated us to use the Pesaran, Shin and Smith (1996) approach and not to display the results here. In general, the results of these traditional cointegration exercises suggest cointegration relationships. However, even more important in the light of the current debate in the literature on the sources of dividend yield (reaction functions), cointegration is indicated if exogeneity is imposed (solely) on the dividend yield variable.

In addition, the *estimation of the long-run coefficients* and the associated *error-correction models* for the German stock market has been interpreted as an important completion of the analysis by Domanski and Kremers (1998, pp. 29), who limit their analysis of the impact of dividend yield on stock markets to a battery of estimations of single equations, in levels based on monthly data. As a result, we explicitly take into account the existence of a *long-term* stock market relationship and the *short-term deviations* from it as a driving force of short-term movements in future stock returns. By this, we allow the dividend yield to have short-term *and* long-term (and by this again, additional short-term) impacts on the future stock return.

4. Applying the ARDL-Approach to Cointegration Analysis

4.1. Theoretical Background

4.1.1. Estimating Long-Term Coefficients

We start with the problem of estimation and hypothesis testing in the context of the following ARDL(p,q)-model:

$$(9) \qquad y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=0}^{q} \beta_i x_{t-i} + u_t,$$

with serially uncorrelated residuals u_t. We will make use of the possibility to model an additional vector of deterministic variables such as dummies, or some exogenous variables with fixed lags.[8] In addition, we assume the existence of a long-run relationship between the levels of y_t and x_t in the light of the above testing procedures. In order to handle the above model correctly, we have to distinguish between *two cases*:

case 1) x_t and u_t are uncorrelated,

case 2) x_t and u_t are correlated, and x_t is characterised by the following finite order AR(s) process:

$$(10) \qquad x_t = \sum_{i=1}^{s} \rho_i x_{t-i} + v_t .$$

This AR-process could also consist of lagged values of Δy_t, but not of the levels or lagged values of y_t (Pesaran and Shin (1999), p. 14).

Under case 1, one can end up with the long-run relationship between y and x by first estimating the parameters of the ARDL model by OLS and then estimating the parameters of the cointegrating relationship $y_t = \alpha + \delta t + \theta x_t + v_t$ by:

$$(11) \qquad \hat{\alpha} = \frac{\hat{\alpha}_0}{1 - \hat{\phi}_1 \ldots - \hat{\phi}_p}, \; \hat{\delta} = \frac{\hat{\alpha}_1}{1 - \hat{\phi}_1 \ldots - \hat{\phi}_p}, \; \hat{\theta} = \frac{\hat{\beta}_0 + \hat{\beta}_1 + \ldots + \hat{\beta}_q}{1 - \hat{\phi}_1 - \hat{\phi}_2 \ldots - \hat{\phi}_p}.$$

The standard errors of these estimates can in principle be obtained by the so-called 'delta-method' in the usual fashion (Pesaran and Pesaran (1997), p. 404; Pesaran and Shin (1999), pp. 16 f.; and Serfling (1980)).[9] However, we prefer computational convenience to calculate their asymptotic standard errors using the so-called 'Bewley's Regression' (Bewley (1979)). Both procedures lead to numerically identical estimates of standard errors of the estimates (Bårdsen (1989)). We estimate the regression:

$$(12) \qquad y_t = \alpha + \delta t + \theta x_t + \sum_{i=0}^{p-1} d_i \Delta y_{t-i} + \sum_{i=0}^{q-1} c_i \Delta x_{t-i} + v_t$$

---

8 We leave this vector out in our representation of equation (9) for reasons of simplicity.

9 The relevant option is option 5 in the 'Post Regression Menu'. This allows one to estimate non-linear functions of the parameters in one's regression model.

by the instrumental variable method referring to 1, t, x_t, Δx_t, ..., Δx_{t-q+1}, y_{t-1}, y_{t-2}, ..., and y_{t-p} as the relevant instruments. Taking the above choice of instruments into account, it again becomes obvious how important it is to *choose p and q quite liberally*. The final result consists of instrumental variable estimates of α, δ and θ.

Under case 2 (x_t and u_t in eq. (9) are correlated) the ARDL procedure and the use of OLS is still valid in the end. However, it requires the estimation of an *augmented* version of the original model, if s (the order of the x_t process in eq. (10)) is larger than q. Let m be equivalent to max (s,q). In this case, the appropriate ARDL model to be used for estimation of the cointegrating regression deviates from eq. (9) and results as follows:

$$(13) \qquad y_t = \alpha_0 + \alpha_1 \cdot t + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=0}^{m} \beta_i x_{t-i} + u_t .$$

In other words, the ARDL approach necessitates putting in *enough lags* of the 'forcing variables' in order to endogenise y_t (here: the German stock returns), before estimation and inference are carried out. By this, one can simultaneously correct for the problem of endogenous regressors (our case 2) and for residual autocorrelation (Pesaran and Shin (1999), p. 16). The estimation of the *long-run* cointegrating parameters can now be calculated in the same way as for case 1. The main reason for the presence of additional lagged changes in y_t and the lagged changes in x_t (which are introduced to deal with the residual serial correlation problem) as additional I(0) variables is they do not affect the asymptotic properties of the OLS estimates of the long-run coefficients (Pesaran and Shin (1999), pp. 14 ff.).[10] However, there exist two decisive differences between the ARDL models depicted by eqs. (9) and (13). These differences refer, first, to the order of lagged Δx_t's and, second, to the interpretation of their coefficients.

In summary, we make use of two important facts resulting from appropriate augmentation of the order of the ARDL-model. First, the OLS estimators of the *short-run* parameters are $\sqrt{T}$-consistent with the asymptotically singular covariance matrix. Second, the ARDL-based estimators of the *long-run* coefficients are super-consistent. Thus, valid inferences on the long-term parameters can be made using standard normal asymptotic theory (Pesaran and Shin 1999). We prefer this approach since it has the additional advantage of yielding consistent estimates of the long-run coefficients that are asymptotically normal, irrespective of whether the underlying regressors are I(0) or I(1), (Pesaran and Shin (1999), p. 17). However, what is most important to us is that the ARDL procedure is valid even if there is some doubt about the unit-root properties of some of the variables y and x (as e.g. the dividend yields and the stock returns). Following Pesaran and Shin (1999), in the case where x_t and u_t are uncorrelated, the ARDL procedure (in contrast to other procedures often proposed in the literature for estimation of cointegrating relations) works irrespective of whether x and y are I(1) or are near I(1) processes.

When estimating the long-run relationship, one of the most important issues is the *choice of the order of the distributed lag function* on y_t and the 'forcing variables' x_t for the unrestricted ECM model. One possibility would be to carry out the *two-step* ARDL estimation approach advanced by Pesaran and Shin (1999), in which the lag orders p and q are selected at first by the *Akaike (AIC)* or the *Schwarz information criteria (SIC)*, (or as a substitute by the *Hannan-Quinn criterion* (HQC, Hannan and Quinn (1979)) or by Theil's (1971) *R-Bar Squared crite-*

---

[10] However, the estimation of the short-run effects still requires an explicit modelling of the contemporaneous dependence between u_t and v_t. Cf. Pesaran, Shin (1998), p. 2 and 15.

rion).[11] The excellent Monte Carlo results *for small samples* gained by Pesaran and Shin (1999) compared with the Fully-Modified OLS estimation procedure by Phillips and Hansen (1990) speak strongly in favour of this two-step estimation procedure.

Setting the maximum orders for p and the q's to 12 (with an eye on the monthly nature of our data), we compare the maximised values of the log-likelihood functions of the different $(m+1)^{k+1}$ (with m: maximum lag and k: number of 'forcing variables') ARDL models. Most importantly, all the models have to be estimated based on the same sample period, namely (m+1, m+2, ... , n) (Belke (2000)). We select the final model by finding the values of p and q that maximise the absolute values of the above mentioned selection criteria. Then the selected model is estimated by the OLS procedure described above. These estimates will be referred to as AIC-ARDL and SIC-ARDL in this paper.

Before applying the two-step procedure to our German stock market data, the model selection criteria are explained briefly. If one denotes the maximised values of the log-likelihood functions of different ARDL (p,q) models by LL (p,q), then

(14) AIC (p,q) = LL(p,q) - (p+q+1+k), and

(15) SIC (p,q) = LL(p,q) - 1/2*(p+q+1+k)*logN,

with N = sample size and k = number of exogenous/deterministic regressors in the ARDL model.[12] The AIC-ARDL and the SIC-ARDL estimates perform very similar in small samples as shown by Pesaran and Shin (1999). However, the SIC reveals an even slightly better performance in most of the experiments. This may be a reflection of the fact that the SIC, like the HQC, are consistent in the sense that for sufficiently large samples they choose the correct model (Belke (2000)). However, the assumption is that the true model is among the models under consideration. Consistency is not valid for the AIC or the R-Bar Squared criterion. However, this does not necessarily mean a disadvantage for our purposes, because one can never be sure that the 'true' model does in fact belong to the models actually investigated. Moreover, the SIC tends to select a more parsimonious specification than, for example, the AIC. However, the fact that a quite liberal specification might be very important for the reasons described above, again speaks in favour of the AIC.[13]

4.1.2. Estimating the Coefficients of the Error-Correction Model

The derivation of the error-correction model from the ARDL equation (9) can best be understood by rewriting this equation in vector representation:

$$(16)\quad \phi(L,p)y_t = \sum_{i=1}^{k}\beta_i(L,q_i)x_{it} + \delta' w_t + u_t,$$

with

$$\phi(L,p) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$$

---

[11] However, one drawback in practical work is that one has to set the maximum lag orders p and q *a priori* although the 'true' orders of the ARDL (p,m) model are not known *a priori*. Cf. Pesaran, Shin (1998), pp. 3 and 16.

[12] In the case of the ARDL model shown in eq. (9) with an intercept and a deterministic trend, k amounts to 2.

[13] For a further description and discussion of the model selection criteria cf. Lütkepohl (1991), section 4.3, and Pesaran and Pesaran (1997), pp. 352 ff.

$$\beta_i(L,q_i) = 1 - \beta_{i1}L - \beta_{i2}L^2 - ... - \beta_{iq_i}L^{q_i}, \text{ for } i = 1, 2, .., k.$$

L represents the usual lag operator ($Ly_t = y_{t-1}$). The variable w_t represents a s x 1 vector of deterministic variables such as an intercept, seasonal dummies, time trends, or exogenous variables with fixed lags. In section 4.1.1., it was shown in a bit more generalised fashion (see eq. (11)), that the long-run response of y_t to a unit change in x_{it} can be estimated as:

$$(17) \quad \hat{\theta}_i = \frac{\hat{\beta}_{i0} + \hat{\beta}_{i1} + ... + \hat{\beta}_{i\hat{q}_i}}{1 - \hat{\phi}_1 - \hat{\phi}_2 ... - \hat{\phi}_{\hat{p}}}.$$

Similarly, the long-run response of y_t to a unit change in the deterministic exogenous variable w_t with fixed lags is given by a transformation of the OLS estimate of δ in eq. (16) for the chosen ARDL model:

$$(18) \quad \hat{\psi} = \frac{\hat{\delta}(\hat{p}, \hat{q}_1, \hat{q}_2, ..., \hat{q}_k)}{1 - \hat{\phi}_1 - \hat{\phi}_2 ... - \hat{\phi}_{\hat{p}}}.$$

If one substitutes

$$y_{t-s} = y_{t-1} - \sum_{j=1}^{s-1} \Delta y_{t-j} \quad (s = 1, 2, ... , p) \text{ and}$$

$$x_{i,t-s} = x_{i,t-1} - \sum_{j=1}^{s-1} \Delta x_{i,t-j} \quad (s = 1, 2, ... , q_i)$$

into (16) and then rearranges the different terms appropriately, one gets the error-correction model corresponding to the above ARDL equation:

$$(19) \quad \Delta y_t = -\phi(1,\hat{p})EC_{t-1} + \sum_{i=1}^{k} \beta_{i0}\Delta x_{it} + \delta' \Delta w_t - \sum_{j=1}^{\hat{p}-1} \phi_j^* \Delta y_{t-j} - \sum_{i=1}^{k} \sum_{j=1}^{\hat{q}_i - 1} \beta_{ij}^* \Delta x_{i,t-j} + u_t,$$

with the error-correction term:

$$(20) \quad EC_t = y_t - \sum_{i=1}^{k} \hat{\theta}_i x_{it} - \hat{\psi}' w_t.$$

The error-correction parameter that measures the quantitative importance of the error-correction term, is represented by:

$$(21) \quad -\phi(1,\hat{p}) = -(1 - \hat{\phi}_1 - \hat{\phi}_2 - ... - \hat{\phi}_{\hat{p}}).$$

The other coefficients ϕ_j^* and β_{ij}^* refer to the short-term dynamics of the model under consideration. They are in the case of the lagged differences of the dependent variable determined by:

$$(22) \quad \begin{aligned} \hat{\phi}_1^* &= \phi_{\hat{p}} + \phi_{\hat{p}-1} + ... + \phi_3 + \phi_2 \\ \hat{\phi}_2^* &= \phi_{\hat{p}} + \phi_{\hat{p}-1} + ... + \phi_3 \\ & \\ \hat{\phi}_{\hat{p}-1}^* &= \phi_{\hat{p}} \end{aligned}$$

and in the case of the changes in the 'long-run forcing' variables respectively by

$$\beta^*_{i1} = \beta_{i,\hat{q}_i} + \beta_{i,\hat{q}_i-1} + \ldots + \beta_{i,3} + \beta_{i,2}$$
$$\beta^*_{i2} = \beta_{i,\hat{q}_i} + \beta_{i,\hat{q}_i-1} + \ldots + \beta_{i,3} \qquad (23)$$
$$\vdots$$
$$\beta^*_{i,\hat{q}_i-1} = \beta_{i,\hat{q}_i}$$

The remaining cointegration parameters $\hat{\theta}_i$ and $\hat{\psi}$ are estimated using the equations (17) and (18).

Based on the above relations, we finally derive the estimated parameters of the ECM described in eq. (19) from the estimated coefficients of the underlying ARDL model. The estimated standard errors of these estimates allow for non-zero co-variances between the estimates of the short-run and the long-run coefficients.

4.2. Application to German Stock Market Data

The estimation of the long run parameters and the associated error-correction model for the unrestricted ECM regression of the stock market returns, cases $x = h$, $x = \Delta d$, and $x = h\text{–}\Delta d$ (which we abbreviate in the following as h, d, or hd), on the dividend yield dp is now carried out using the *two-step ARDL estimation approach* proposed by Pesaran and Shin (1999).

4.2.1. Estimating the Orders of the Distributed Lag Functions

As emphasised already, the most important issue is the *choice of the order of the distributed lag function* on y_t and the 'forcing variables' x_t for the unrestricted ECM model when estimating the long-run relationship. We prefer to carry out the two-step ARDL estimation approach by Pesaran and Shin (1999) and apply it to our model 2 (x=d, without trend), where firstly the lag orders p and q are selected by the *Akaike* or the *Schwarz information criteria*, the *Hannan-Quinn* or the *R-Bar Squared criterion*. The selected model has been estimated by the OLS procedure. Setting the maximum orders for p and the q's to 12 (since we use monthly data), we compare the maximised values of the log-likelihood functions of the $(m+1)^{k+1}$ (with m: maximum lag and k: number of 'forcing variables') different ARDL models. Table 3 shows the selected lag order and the corresponding maximising empirical values of the model selection criteria, AIC and SIC (the values of the other two criteria are available on request), for each variants of the model (MA = 12, 24, 36, 48 months). The sequence of the lag orders (p, q_1, q_2 ...) always corresponds to the sequence of the variables in both models. Both selection criteria point at Model 2 (MA 12 months) without trend, as the best fitting model.

(Table 3 about here)

4.2.2. Estimating Long-Run Relationships

The estimation results for the long-run relationship between German stock market returns and different stock market variables are displayed in Tables 4a to 5b. Applying the ARDL approach, we only focus on specifications without a deterministic trend in the cointegration vector.[14] The values in brackets represent the standard errors of the parameter estimates. The associated estimated error correction regressions are obtained later.

(Table 4a. about here)

(Table 4b. about here)

(Table 5a. about here)

(Table 5b. about here)

The long-run coefficients based on the selected ARDL models estimated over the maximum period 1974.8 to 2003.9 are listed in Tables 4a to 5b. The results throughout the Tables 4a to 5b show that the *long-run elasticity* of dividend growth with respect to the *dividend yield* is negative/positive, which is in line with theoretical reasoning.[15] The specifications according to the SIC-, and AIC- model selection criteria yield very *similar point estimates*. However, the lag order specifications differ dependent on the choice of the number of months in the moving average specification. In addition, the estimated standard errors vary depending on the specific model selection criterion and on the order of the selected ARDL model.

4.2.3. Estimating Final Error-Correction Models and Model Selection

After determining the lag order and the long-run coefficients for each ARDL model, we can derive the estimates for the error correction models (as explained in section 4.1.2). One further issue that needs to be addressed before the best specification can be selected is: what criterion should one make the final selection? We started with the four possible criteria (introduced in section 4.1.1.), but made our final choice based on two of them, namely on the Akaike - and the Schwarz information criterion, AIC and SIC.

In order to select the best performing ARDL-model, the *significance* of the resulting *ECM-parameters* or, alternatively in cases of identical samples, the *empirical values of the two information criteria* are compared. The advantage of the AIC lies in its property to generally lead to a higher order of ARDL model than the SIC. This tendency leads in turn, to smaller estimated standard errors and a higher chance of white-noise property of the residuals.[16] However, the SIC is again chosen as the alternative to the AIC because it asymptotically determines the true model under certain preconditions (Belke (2000)). Table 2 shows the empirical realisations of both information criteria. These values are already maximised in the sense that they refer to ARDL-models whose orders have already been selected by the respective information criterion. As already stated, we selected the model displayed in Table 3a - (Model 2, $x = d$, without trend, MA = 12 months).

In Table 5, the values in brackets are the t-values of the error-correction parameter estimates based on the selected model. Taking a closer look at the estimated error correction parameter (estimated error-correction model, Table 6), the main result is that the error correction coeffi-

---

[14] Recall that cointegration was established by our cointegration tests in section 3.1 preponderantly if a deterministic trend is excluded.

[15] As it is well-known from cointegration theory, we should not draw any inference from the t-values of the coefficient estimates. However, for instance Domanski and Kremer (1998) clearly violate this key guideline when interpreting the estimation results of their table 1 on pp. 30.

[16] It has already been mentioned that a less parsimonious specification is preferred on theoretical grounds.

cient is highly significant as compared with the usual t-distribution.[17] The estimated error-correction parameter has the correct negative sign. Its size, estimated at a magnitude of around -0.06 to -0.12, suggests a moderate speed of convergence to equilibrium. The most conservative critical t-values (leading to the lowest chance of rejection of the non-cointegration hypothesis) for the ECM parameter estimates can be taken from Banerjee, Dolado and Mestre (1992), Appendix Table 4. For the selected model we choose the critical value for one exogenous regressor, ECM with a constant and no deterministic trend and around 300 observations ($\alpha = 0.05$), as falling between a range from 3.27 (100 obs.) to 3.23 (500 obs.). Even in this extreme case, two of the three estimated error-correction parameters are significant at $\alpha = 0.05$.

(Table 6 about here)

At first glance, the R-squared appear to be rather low and corresponds with values observed by Domanski and Kremer (1998). However, this pattern is not exceptional for an ECM modelled for financial market variables. The models *fit very well* on average, explaining almost 7 percent of the variations in future stock market returns (*changes* in the (logs of) h, Δd, or $h-\Delta d$). This is even valid when the fit is measured by the R-Bar-Squared. In all cases, the underlying ARDL equations also pass the diagnostic tests for the serial correlation of residuals, for functional form misspecification and for non-normal and homoskedastic disturbances. Beyond the highly significant ECM parameter, some but not all of the estimated coefficients of the selected ECMs are also significant (the reported standard errors allow for the sampling variations in the estimated long-run coefficients) and are *of a similar magnitude across the different specifications* selected by the two criteria.[18]

Tables 7 and 8 contain the detailed results for the selected error-correction model, giving some intuition on the order of magnitude of the detected impact of dividend yield on stock market returns. The dividend yield is in both selected cases (ARDL (1,0) and ARDL (3,5)) significant and reveals the correct negative sign.

(Table 7 about here)

(Table 8 about here)

Overall, the results which support short- and long-term impacts of the dividend yield on future German stock returns appear to be supported from another angle: on the basis of a fully specified stock market model, of monthly data (which seem to be appropriate to capture the short-term dynamics), of an econometric procedure whose reliability is not dependent on the order of integration of the included variables and which additionally takes into account deviations from equilibrium long-term relationships between stock market variables as 'driving forces' of the short-term dynamics in future German stock returns. As outlined earlier, the coefficient of dp is positive in the case of the dependent variables h and $h-\Delta d$, and negative if Δd is the dependent variable, as suggested by theoretical reasoning. However, it has to be kept

---

17 Under the assumption that the vector of cointegrating parameters is given the distribution of the t-statistics can be approximated in many cases by the standard normal distribution. This would also legitimise the use of the student-t-distribution for a judgment on the significance of the error-correction parameter. See Banerjee et al. (1993), pp. 230 ff., and Kremers, Ericsson and Dolado (1992), pp. 328 ff.

18 Our ARDL procedure does not allow to skip the seemingly insignificant variables, since they contribute to the fit according to the empirical realisations of the information criteria.

in mind that significant error-correction parameter estimates could be gained only for a small share of possible specifications.

5. Dynamic Forecasts for the Growth Rate of German Stock Returns Based on an Assessment of the Future Course of Dividend Yield

The selected error-correction models can also be used in forecasting the growth rate of German stock returns, conditional on the dividend yield variables including a constant but no trend. We dynamically forecast the growth rate of German stock returns over the period 2001M10 to 2002M9 (12-months in-sample forecasts). For this purpose, our estimation will dispense with the last year in the sample and we now have to re-estimate the specification given in Table 7 for the shorter sample. We will start with a forecast based on our preferred model 2 without trend selected by the SIC criterion and displayed in Table 4a ($x = d$ and MA = 12 months). As Table 9a reveals, the root mean squares of forecast error (approx. 8.14 per cent per month) does not compare favourably with the value of the same criterion computed over the estimation period which is less than half its value. Moreover, the model fails to forecast the extent of the future stock return growth in each forecasted month. Nevertheless, the model does not fail to forecast the sign except in two cases (i.e., March and April 2002). The plot of the dynamic forecast for the growth rate of future German stock returns in Figure 1a reveals that the forecasts are very close to the actual values for the estimation period, but not for the forecast period. All in all, the mixed results of the forecast do not totally corroborate the choice of this model.

Using the same model, Table 9b displays the dynamic in-sample 3-month ahead forecasts of German stock returns based on the dividend yield. The root mean squares of forecast errors of around 3.53 percent per quarter is, in contrast to the 12-month ahead forecast, comparable to the value of the same criterion computed over the estimation period. Additionally, the model does not miss the sign of future stock return growth in each forecasted month. Our main impression is corroborated by the plot of the dynamic forecasts for the growth rate of German stock returns in Figures 1a and 1b. Summarising, the results of the forecasts corroborate our choice of our model chosen by the SIC only if forecasts are made over a short horizon.

(Table 9a. about here)

(Table 9b. about here)

(Figure 3a. about here)

(Figure 3b. about here)

Finally we also enacted the same exercise for our other final model selected by the AIC criterion (Model 2 with $x = \Delta d$ and MA = 12 months). Here, we also differentiated between a 12-month and a 3-month in-sample forecast. The results are displayed in Tables 10a and 10b and in Figures 4a and 4b, respectively. Although the lag structure of the model chosen by the AIC is different, the results display a pattern similar to those for the model chosen by the SIC criterion.

(Table 10a. about here)

(Table 10b. about here)

(Figure 4a. about here)

(Figure 4b. about here)

6. Conclusions and Implications for the Debate on the Impacts of Dividend Yield on Asset Prices

For a few specifications, we find that the dividend yield has a statistically significant positive impact on the future stock returns in Germany: "low" stock prices relative to dividends forecast higher subsequent returns. In these cases, and in line with previous findings and theoretical considerations, we find that the power of dividend yields to forecast future stock expected returns increases with the return horizon. In the first part of the paper we conclude that the relationship between dividend yield and the future stock returns is *one-way from the first to the latter* if stock market returns are measured by the annualised one-month dividend growth rates in percent. Hence, (only) in this case the dividend yield variable can best be characterised as a so-called "forcing variable" of future stock returns. For other measures of the dividend yield used by us, we either find either a significant co-movement with causality going into both directions or no cointegration, depending on the lag structure.

Our results based on the ARDL approach corroborate findings by Domanski and Kremer (1998), who are able to detect a significant positive relationship between the magnitude of future stock returns and the level of the dividend yields in Germany. As indicated by the significant positive impact of the dividend yield in the I(0) part and the I(1) part of our estimated error-correction models, we find that even *short-run* increases in the dividend yield could have a temporary impact on future stock returns (i.e., the annualised one-month dividend growth) in addition to *permanent* ones. The latter finding had already been theoretically suggested by earlier studies of Fama and French (1988), Campbell, Lo and Shiller (1997) and Domanski and Kremer (1998). However, significant error-correction parameter estimates could be gained only for a small share of all possible specifications. Moreover, it proved to be extremely difficult to identify an empirical model with good forecast properties, at least for the longer term, i.e. 12 months. Therefore, it is not conclusive at this stage of analysis that dividend yields are generally useful for forecasting stock market returns.

We realise that the results are preliminary, not least because the questions posed in this paper have not been tackled based on the highly suitable autoregressive distributed lag approach à la Pesaran in the literature so far. However, the limited number of observations is no reason to be overly cautious any more. The procedure used in this article is robust with respect to small samples and the uncertainty of the order of integration of the included variables. Our approach was applied only to Germany, since replicating it for many others like the US would simply have taken too much space. We leave this task for future research. Moreover, empirical work could follow in the sense that it could try to exploit the progress in economic theory by imposing it as a restriction on the empirical models in order to exactly identify long-run relations (Pesaran, 1997). However, because of scarcity of knowledge in this area, there will still be a long way to go.

References

Banerjee, A., Dolado, J.J., Mestre, R. (1992): On Some Simple Tests for Cointegration, The Cost of Simplicity, mimeo, Institute of Economics, Aarhus/Denmark, later published as: Error-Correction Mechanism Tests for Cointegration in a Single-equation Framework, in: Journal of Time Series Analysis, Vol. 19, pp. 267-283, 1998.

Banerjee, A., et al.*[??should list names??]* (1993): Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data, Oxford.

Bårdsen, G. (1989): The Estimation of Long-Run Coefficients from Error Correction Models, in: Oxford Bulletin of Economics and Statistics, Vol. 51, pp. 345-350.

Belke, A. (1996): Testing for Unit Roots in West German and U.S. Unemployment Rates: Do 'Great Crashes' Cause Trend Breaks?, in: Applied Economics Quarterly, Vol. 42, pp. 327-360.

Belke, A. (2000): Political Business Cycles in the German Labour Market? Empirical Tests in the Light of the Lucas Critique, in: Public Choice, Vol. 104, pp. 225-283.

Bewley, R. (1979): The Direct Estimation of the Equilibrium Response in a Linear Economic Model, in: Economics Letters, Vol. 55, pp. 251-276.

Campbell, J.Y., Lo, A.W., Shiller, R.J. (1997): The Econometrics of Financial Markets, Princeton University Press, Princeton, NY.

Cavanaugh, C.L., Elliott, G., Stock, J.H. (1995): Inference in Models with Nearly Integrated Regressors, in: Econometric Theory, Vol. 11, pp. 1131-1147.

Cochrane, J.H. (2001): Asset Pricing, Princeton University Press, Princeton and Oxford.

Cuthbertson, K. (1996): Quantitative Financial Economics, Stocks, Bonds and Foreign Exchange, Chichester et al., Wiley.

Cuthbertson K., S. Hayes, D. Nitzsche (1997): The Behavior of UK Stock Prices and Returns: Is the Market Efficient?, in: Economic Journal, 107, July, 986-1008.

Davidson, R., McKinnon, J.G. (1993): Estimation and Inference in Econometrics, Oxford.

Domanski, D., and Kremer, M, (1998): What Do Asset Price Movements in Germany Tell Policy Makers, in: The Role of Asset Prices in the Formulation of Monetary Policy, Bank for International Settlement, Conference Papers, Basle, Vol. 5, pp. 24-44.

Engle, R.F., Granger, C.W.J. (1987): Cointegration and Error-Correction Representation: Estimation and Testing, in: Econometrica, Vol. 55, pp. 251-276.

Fama, E.F., Schwert, G.W. (1977): Asset Returns and Inflation, in: Journal of Financial Economicss, 5, pp. 115-146.

Fama, E.F. (1981): Stock Returns, Real Activity, Inflation, and Money, in: American Economics Review, 71, pp. 545-565.

Fama, E.F. and French, K.R. (1988): Dividend Yields and Expected Stock Returns, in: Journal of Financial Economics, 22, p. 3-25.

French, K.R., Schwert, G.W. and Stambaugh, R.F. (1987): Expected Stock Returns and Volatility, in: Journal of Financial Economics, 19, pp. 3-29.

Hamilton, J.D., Jordá, O. (2000): A Model for the Federal Funds Rate Target, NBER Working Paper 7847, National Bureau of Economic Research, Cambridge/MA.

Hannan, E.J., Quinn, B.G. (1979): The Determination of the Order of an Autoregression, in: Journal of the Royal Statistical Society, Vol. 41, pp. 190-195.

Hendry, D.F., Pagan, A.R., Sargan, J.D. (1984): Dynamic Specification, in: Griliches, Z., Intrilligator, M.D. (eds.), Handbook of Econometrics, Vol. II, Elsevier, Amsterdam.

Hodrick, R. (1992): Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, in: Review of Financial Studies, 5, pp. 357-386.

Johansen, S. (1991): Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models, in: Econometrica, Vol. 55, S. 251-276.

Johansen, S. (1995): Likelihood Based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press, Oxford.

Keim, D.B., Stambaugh, R.F. (1986): Predicting Returns in the Stock and Bond Markets, in: Journal of Financial Economics, 17, pp. 357-390.

Kremers, J.J.M., Ericsson, N.E., Dolado, J.J. (1992): The Power of Cointegration Tests, in: Oxford Bulletin of Economics and Statistics, Vol. 54, pp. 325-348.

Lütkepohl, H. (1991): Introduction to Multiple Time Series Analysis, Springer, New York.

Newey, W., West, K. (1987): A Simple, Positive-definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, in: Econometrica, Vol. 55, pp. 703-708.

Pesaran, M.H. (1997): The Role of Economic Theory in Modelling the Long-Run, in: Economic Journal, Vol. 107, pp. 178-191.

Pesaran, M.H., Pesaran, B. (1997): Working with Microfit 4.0 - Interactive Econometric Analysis, Oxford University Press.

Pesaran, M.H., Shin, Y. (1999): An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis, in: Strom, S., and P. Diamond (eds.), Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, chapter 11, Cambridge, Cambridge University Press., Manuscript Version 1998.

Pesaran, M.H., Shin, Y., Smith, R.J. (1996): Testing for the Existence of a Long-Run Relationship, DAE Working Paper No. 9622, Department of Applied Economics, University of Cambridge. Later published as:

Pesaran, M.H., Y. Shin, and R. J. Smith, (2001): Bounds Testing Approaches to the Analysis of Level Relationships, in: Hendry, D.F., and M.H. Pesaran (eds.), Journal of Applied Econometrics, Special issue in honour of J D Sargan - Studies in Empirical Macroeconometrics, Vol. 16, pp. 289-326.

Phillips, P.C.B., Hansen, B. (1990): Statistical Inference in Instrumental Variables Regression with I(1) Processes, in: Review of Economic Studies, Vol. 53, pp. 473-496.

Phillips, P.C.B., Ouliaris, S. (1990): Asymptotic Properties of Residual Based Tests for Cointegration, in: Econometrica, Vol. 58, pp. 165-193.

Serfling, R.J. (1980): Approximation Theorems of Mathematical Statistics, Wiley, New York.

Theil, H. (1971): Principles of Econometrics, Wiley, New York.

Data

All stock market data for Germany was taken from the Thomson Financials data base. The indices used cover around 80% of the stock market capitalization in Germany.

The following stock market return measures were calculated:

dp = natural logarithm of the dividend yield;

h = holding stock market returns (capital gains plus dividend returns, presented by the total stock market performance index), expressed as the annualised one-month continuously compounded stock return in percent;

Δd = dividend growth, expressed as the annualised one-month continuously compounded stock return in percent and

h–Δd = holding period return minus dividend growth.

In the text, a number behind a variables indicates the time horizon under review. For instance, *h*36 would indicate the holding period return over the coming 36-months. In the case of *dp*, a number would indicate the time horizon which is forecast by using the dividend yield.

Figure 1. – Stock returns and the dividend yield, scatter diagram

(a) Holding period against and dividend yield

(b) Dividend growth and dividend yield

(c) Holding period yield minus dividend growth against dividend yield

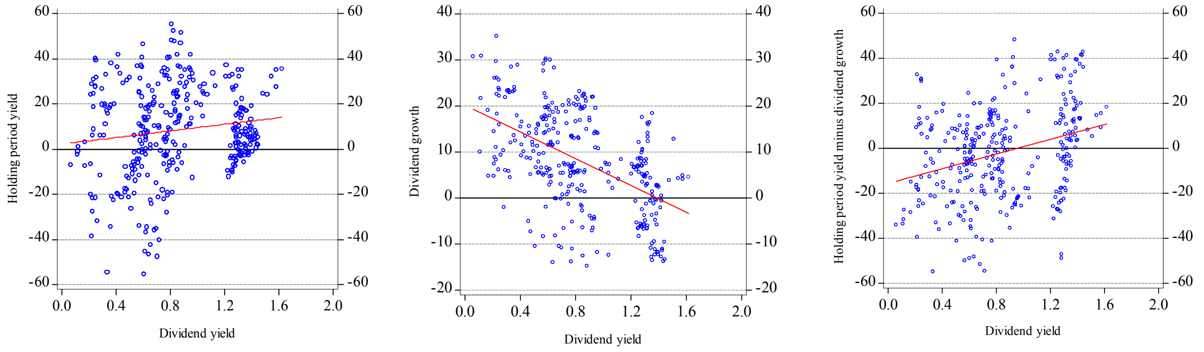
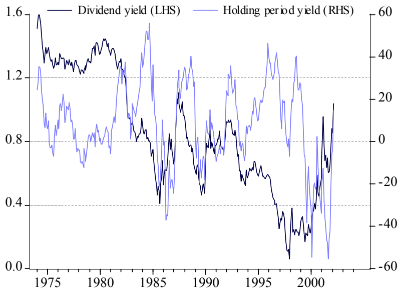
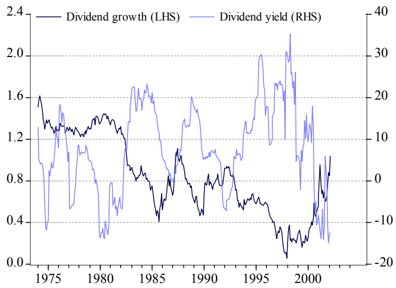


Figure 2. – Stock returns and the dividend yield over time, line diagram

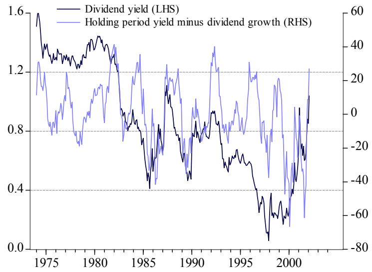
(a) Holding period against and dividend yield



(b) Dividend growth and dividend yield



(c) Holding period yield minus dividend growth against dividend yield



Data source: Thomson Financials; own calculations. – Time period: 1974.8 to 2003.9. Time horizon 12-month for all variables.

Figure 3a. – Dynamic in-sample 12-month ahead forecasts of the level of German stock returns based on the dividend yield (model 2 selected by SIC for $x = d$ and MA = 12 months)

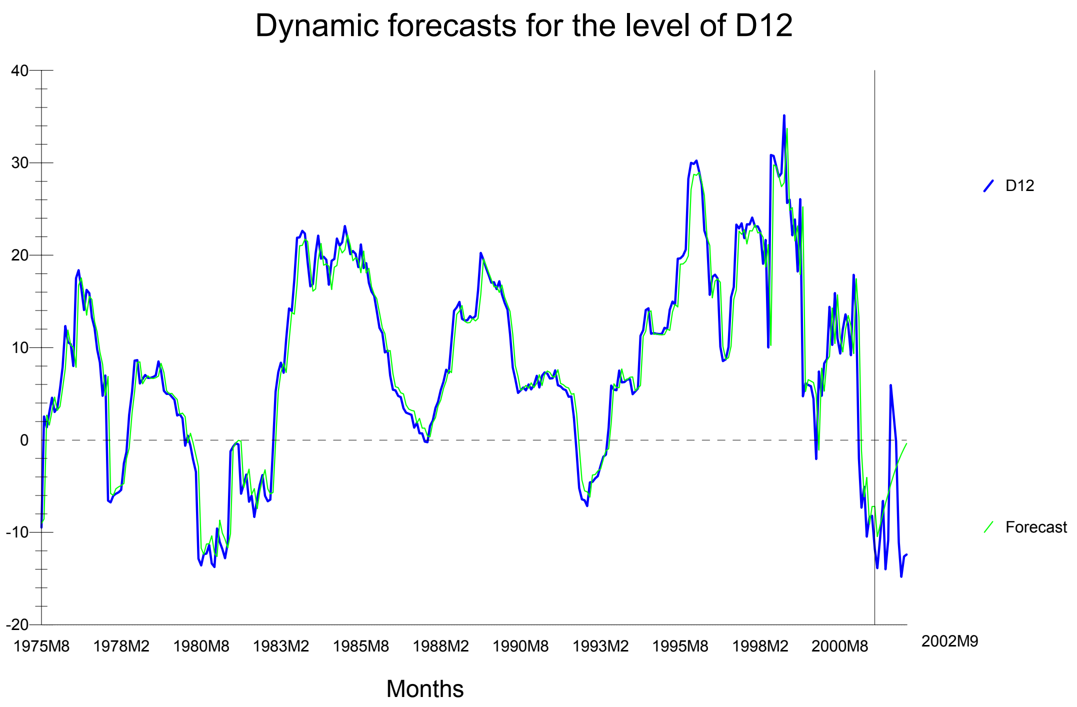


Figure 3b. – Dynamic in-sample 3-month ahead forecasts of the level of German stock returns based on the dividend yield (model 2 selected by SIC for $x = d$ and MA = 12 months)

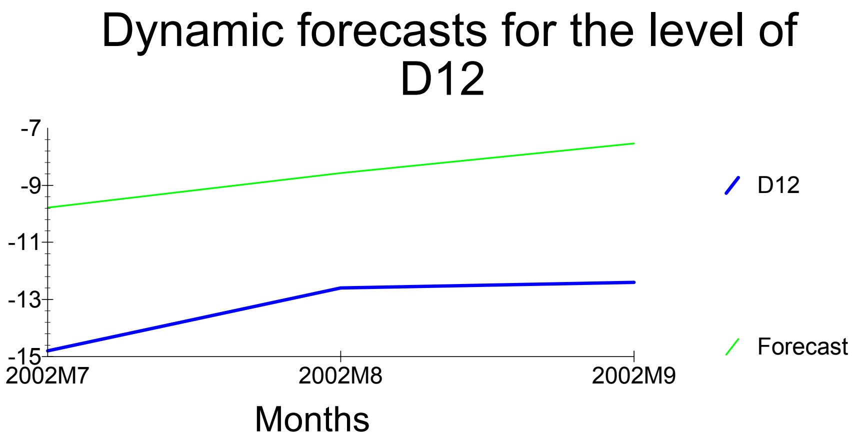


Figure 4a. – Dynamic 12-months ahead in-sample forecasts of the level of German stock returns based on the dividend yield (Model 2 selected by AIC for $x = \Delta d$ and MA = 12 months)

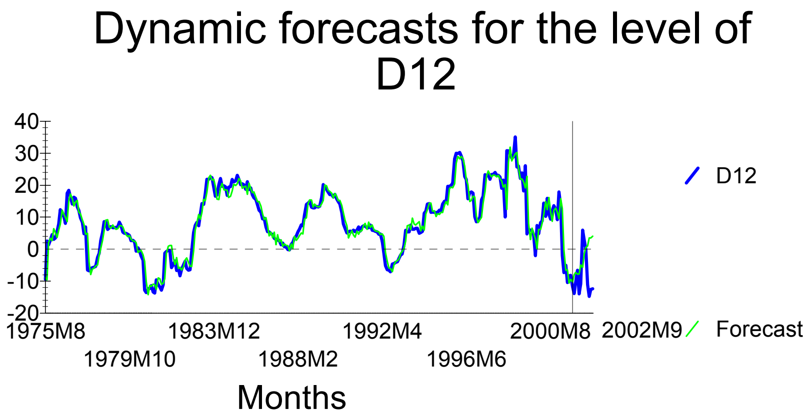


Figure 4b. – Dynamic 3-months ahead in-sample forecasts of the level of German stock returns based on the dividend yield (Model 2 selected by AIC for $x = d$ and MA = 12 months)

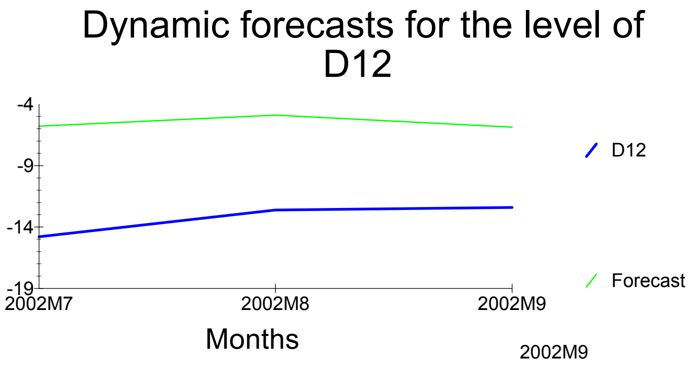


Table 1. – Long-horizon regressions of stock market measures on the log dividend yield and a constant for Germany

	x = h					
Forecast horizon K	1	3	12	24	36	48
$R^2(K)$	0.002	0.09	0.019	0.043	0.068	0.065
β(K)	7.943	9.556	7.361	7.704	7.512	5.786
t-value Newey West	0.833	1.100	1.280	1.622	1.763	1.660
	x = Δd					
$R^2(K)$	0.039	0.108	0.269	0.259	0.236	0.280
β(K)	-16.335	-16.470	-14.533	-11.388	-9.423	-9.293
t-value Newey West	-3.346	-3.752	-5.569	-4.848	-3.949	-4.437
	x = h–Δd					
$R^2(K)$	0.016	0.054	0.184	0.353	0.471	0.546
β(K)	24.279	26.030	21.894	19.092	16.935	15.080
t-value Newey West	2.370	2.869	4.123	5.748	6.342	7.010

Estimation period: August 1974 to September 2003, monthly data. *h* is the annualised one-month continuously compounded stock return in percent. Δd and Δp represent the annualised one-month continuously compounded dividend and profit growth rate, respectively. $\alpha(H)$ is the constant of the regression (not shown). $\beta(H)$ is the slope coefficient of the regression. Regression is estimated on the basis of OLS. $\varepsilon_{t+H,H}$ is the error term which is autocorrelated owing to data overlap for $H > 1$ under the null hypothesis of no predictability. Standard errors and t-values are corrected for serial correlation and heteroskedasticity in the equation using the New and West (1987), that is general covariance estimators that are consistent in the presence of both heteroskedasticity and autocorrelation of unknown form are used. The truncation lag, the parameter representing the number of auto-correlations used in evaluating the dynamics of the OLS residuals, has been chosen as 5.

Data source: Thomson Financials; own calculations.

Table 2a. – F-statistics for testing the existence of a long-run relationship between the stock market return and the dividend yield (model 1: $x = h$)

	Based on regressions with the change of stock returns d(h) as dependent variable		Based on regressions with the change of the dividend yield d(dp) as dependent variable	
MA-order of h	*Without trend*	*With trend*	*Without trend*	*With trend*
h1	0.375	**6.432**	0.042	0.086
h3	0.379	5.009	0.287	0.319
h12	**6.452**	**10.379**	**29.380**	**40.973**
h24	1.490	4.961	**12.587**	**13.130**
h36	1.446	1.930	4.727	5.282
h48	1.166	1.723	2.162	2.688
$F^C(0.1)$	*4.788*	*6.335*	*4.788*	*6.335*
$F^C(0.05)$	*5.764*	*7.423*	*5.764*	*7.423*
$F^C(0.01)$	*7.815*	*9.786*	*7.815*	*9.786*

Table 2b. – F-statistics for testing the existence of a long-run relationship between the stock market return and the dividend yield (model 2: $x = \Delta d$)

	Based on regressions with the change of stock returns d(Δd) as dependent variable		Based on regressions with the change of the dividend yield d(*dp*) as dependent variable	
MA-order of Δd	*Without trend*	*With trend*	*Without trend*	*With trend*
$\Delta d1$	0.345	0.255	0.058	0.067
$\Delta d3$	2.746	3.347	0.024	0.210
$\Delta d12$	**217.707**	**10.383**	0.142	0.610
$\Delta d24$	**39.919**	2.515	0.606	2.022
$\Delta d36$	**44.835**	3.400	0.638	5.160
$\Delta d48$	**48.312**	1.965	0.740	4.150
$W^C(0.1)$	*4.788*	*6.335*	*4.788*	*6.335*
$W^C(0.05)$	*5.764*	*7.423*	*5.764*	*7.423*
$W^C(0.01)$	*7.815*	*9.786*	*7.815*	*9.786*

Table 2c. – F-statistics for testing the existence of a long-run relationship between the stock market return and the dividend yield (model 3: $x=h-\Delta d$)

	Based on regressions with the change of stock returns d($h-\Delta d$) as dependent variable		Based on regressions with the change of the dividend yield d(dp) as dependent variable	
	Without trend	*With trend*	*Without trend*	*With trend*
$(h-\Delta d)1$	0.754	3.297	0.079	0.759
$(h-\Delta d)3$	1.269	2.950	0.033	0.324
$(h-\Delta d)12$	**30.585**	**30.983**	**18.206**	**20.318**
$(h-\Delta d)24$	1.112	2.606	**16.209**	**18.891**
$(h-\Delta d)36$	1.619	0.853	0.753	0.695
$(h-\Delta d)48$	0.620	0.383	0.101	0.070
$W^C(0.1)$	*4.788*	*6.335*	*4.788*	*6.335*
$W^C(0.05)$	*5.764*	*7.423*	*5.764*	*7.423*
$W^C(0.01)$	*7.815*	*9.786*	*7.815*	*9.786*

Notes: Maximum sample: 1974.8 to 2003.9. Lag orders: $p = q_1 = q_2 = 12$. We implemented a dummy which is coded as 1 from 2000(1) on, otherwise 0, into those regressions which also include a deterministic trend.

Table 3. – Empirical values of model selection criteria

ECM	*SIC-value of SIC - ARDL*	*AIC-value of AIC - ARDL*
Model 2 (MA 12 months)	-881.7076	-872.8644
without trend	ARDL (1,0)	ARDL (3,5)
Model 2 (MA 24 months)	-624.4295	-617.5638
without trend	ARDL (1,0)	ARDL (7,0)
Model 2 (MA 36 months)	-477.5989	-468.7692
without trend	ARDL (1,0)	ARDL (12,0)
Model 2 (MA 48 months)	-391.9962	-384.2233
without trend	ARDL (1,0)	ARDL (3,0)

Sample: For MA=12 months: 1975M8 to 2002M9. For MA=24 months: 1975M8 to 2001M9.
For MA=36 months: 1975M8 to 2000M9. For MA=48 months: 1975M8 to 1999M9.

Table 4a. – Estimated long-run coefficients using the ARDL approach
(model 2, $x = d$, without trend, MA = 12 months)

	SIC - ARDL (1,0)	*AIC - ARDL (3,5)*
Intercept	14.1207 (7.8346)	12.8791 (6.6367)
Dividend yield	-7.3630 (8.7523)	-6.4375 (7.3780)

Sample: 1975.8 to 2002.9. Values in brackets are the standard errors of the parameter estimates.

Table 4b. – Estimated long-run coefficients using the ARDL approach
(model 2, $x = d$, without trend, MA = 24 months)

	SIC - ARDL (1,0)	*AIC - ARDL (7,0)*
Intercept	281.5509 (1675.9)	-32.0534 (93.9957)
Dividend yield	-307.7547 (1884.1)	44.7870 (105.4863)

Sample: 1975.8 to 2001.9. Values in brackets are the standard errors of the parameter estimates.

Table 5a. – Estimated long-run coefficients using the ARDL approach
(model 2, $x = d$, without trend, MA = 36 months)

	SIC - ARDL (1,0)	*AIC - ARDL (10,0)*
Intercept	75.0618 (69.4743)	-.38877 (18.4145)
Dividend yield	-76.4702 (79.4155)	10.3068 (20.9638)

Sample: 1975.8 to 2000.9. Values in brackets are the standard errors of the parameter estimates.

Table 5b. – Estimated long-run coefficients using the ARDL approach
(Model 2, $x = d$, without trend, MA = 48 months)

	SIC - ARDL (1,0)	*AIC - ARDL (3,0)*
Intercept	53.6893 (28.2457)	83.8705 (112.1945)
Dividend yield	-50.3357 (31.3694)	-84.0636 (125.1270)

Sample: 1975.8 to 1999.9. Values in brackets are the standard errors of the parameter estimates.

Table 6. – Error correction parameter estimates

ECM	***ARDL (1,0)***	$\overline{R}^2$	***ARDL (3,5)***	$\overline{R}^2$	***ARDL (12,12)***	$\overline{R}^2$
Model Table 3a	-.064278 (-2.9427)	.022828	-.077150 (-3.3534)	.061555.	-.12009 (-3.9385)	.069402

Sample: 1975M8 to 2002M9. Model specifications as denoted in Table 3a; t-values of EC term in brackets.

Table 7. – Error correction representation of selected ARDL model 2 (ECM without trend): ARDL (1,0) model selected based on Schwarz Bayesian Criterion (SIC)

```
******************************************************************************
 Dependent variable is dD12
 326 observations used for estimation from 1975M8  to 2002M9
******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 dDP12                     -.47328             .63151            -.74944[.454]
 dINPT                      .90766             .66289             1.3692[.172]
 ecm(-1)                  -.064278            .021844            -2.9427[.003]
******************************************************************************
 List of additional temporary variables created:
 dD12 = D12-D12(-1)
 dDP12 = DP12-DP12(-1)
 dINPT = INPT-INPT(-1)
 ecm = D12 +   7.3630*DP12  -14.1207*INPT
******************************************************************************
 R-Squared                    .028841   R-Bar-Squared                  .022828
 S.E. of Regression            3.5384   F-stat.    F(  2, 323)    4.7962[.009]
 Mean of Dependent Variable -.0074529   S.D. of Dependent Variable      3.5795
 Residual Sum of Squares       4044.1   Equation Log-likelihood      -873.0273
 Akaike Info. Criterion     -876.0273   Schwarz Bayesian Criterion   -881.7076
 DW-statistic                  2.1331
******************************************************************************
 R-Squared and R-Bar-Squared measures refer to the dependent variable
 dD12 and in cases where the error correction model is highly
 restricted, these measures could become negative.
```

Table 8. – Error correction representation of selected ARDL model 2 (ECM without trend): ARDL (3,5) model selected based on Akaike Information Criterion (AIC)

```
****************************************************************************
 Dependent variable is dD12
 326 observations used for estimation from 1975M8  to 2002M9
****************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 dD121                    -.059012            .056414           -1.0461[.296]
 dD122                     .085184            .055962            1.5222[.129]
 dDP12                     -9.4090             3.6077           -2.6081[.010]
 dDP121                    -2.5191             3.6441           -.69128[.490]
 dDP122                    -1.8994             3.6461           -.52095[.603]
 dDP123                     .31673             3.6436           .086930[.931]
 dDP124                   -11.3108             3.6588           -3.0914[.002]
 dINPT                      .99362             .67690            1.4679[.143]
 ecm(-1)                  -.077150            .023007           -3.3534[.001]
****************************************************************************
 List of additional temporary variables created:
 dD12 = D12-D12(-1)
 dD121 = D12(-1)-D12(-2)
 dD122 = D12(-2)-D12(-3)
 dDP12 = DP12-DP12(-1)
 dDP121 = DP12(-1)-DP12(-2)
 dDP122 = DP12(-2)-DP12(-3)
 dDP123 = DP12(-3)-DP12(-4)
 dDP124 = DP12(-4)-DP12(-5)
 dINPT = INPT-INPT(-1)
 ecm = D12 +   6.4375*DP12  -12.8791*INPT
****************************************************************************
 R-Squared                    .087543   R-Bar-Squared                  .061555
 S.E. of Regression            3.4676   F-stat.    F(  8, 317)    3.7897[.000]
 Mean of Dependent Variable -.0074529   S.D. of Dependent Variable      3.5795
 Residual Sum of Squares       3799.7   Equation Log-likelihood      -862.8644
 Akaike Info. Criterion     -872.8644   Schwarz Bayesian Criterion   -891.7989
 DW-statistic                  2.0100
```

Table 9a. – Dynamic in-sample 12-month ahead forecasts of the level of German stock returns based on the dividend yield (model 2 selected by SIC for $x = d$ and MA = 12 months)

```
*******************************************************************************
 Based on 314 observations from 1975M8  to 2001M9 .
 ARDL(1,0) selected using Schwarz Bayesian Criterion.
 Dependent variable in the ARDL model is D12 included with a lag of 1.
 List of other regressors in the ARDL model:
 DP12            INPT
*******************************************************************************
```

Observation	Actual	Prediction	Error
2001M10	-13.8791	-10.4584	-3.4208
2001M11	-10.4747	-9.1794	-1.2952
2001M12	-6.6052	-7.9964	1.3912
2002M1	-13.9974	-6.8539	-7.1436
2002M2	-10.8478	-5.8113	-5.0365
2002M3	5.9525	-4.7927	10.7453
2002M4	2.8459	-3.8427	6.6885
2002M5	-.10917	-2.9708	2.8616
2002M6	-11.0711	-2.2081	-8.8629
2002M7	-14.8009	-1.5403	-13.2606
2002M8	-12.6037	-.90205	-11.7017
2002M9	-12.4071	-.38543	-12.0217

Summary Statistics for Residuals and Forecast Errors

	Estimation Period 1975M8 to 2001M9	Forecast Period 2001M10 to 2002M9
Mean	-.0000	-3.4214
Mean Absolute	1.9827	7.0358
Mean Sum Squares	11.2325	66.3285
Root Mean Sum Squares	3.3515	8.1442

Table 9b. – Dynamic in-sample 3-month ahead forecasts of the level of German stock returns based on the dividend yield (model 2 selected by SIC for *x* =*d* and MA = 12 months)

```
Dynamic forecasts for the level of D12
*******************************************************************************
 Based on 323 observations from 1975M8  to 2002M6 .
 ARDL(1,0).
 Dependent variable in the ARDL model is D12 included with a lag of 1.
 List of other regressors in the ARDL model:
 DP12            INPT
*******************************************************************************
 Observation             Actual             Prediction              Error
  2002M7                 -14.8009               -9.7841               -5.0168
  2002M8                 -12.6037               -8.5694               -4.0344
  2002M9                 -12.4071               -7.5338               -4.8733
*******************************************************************************


             Summary Statistics for Residuals and Forecast Errors
*******************************************************************************
                      Estimation Period            Forecast Period
                     1975M8  to 2002M6           2002M7  to 2002M9
*******************************************************************************
 Mean                            .0000                     -4.6415
 Mean Absolute                  2.0936                      4.6415
 Mean Sum Squares              12.4388                     21.7310
 Root Mean Sum Squares          3.5269                      4.6617
```

Table 10a. – Dynamic 12-months ahead in-sample forecasts of the level of German stock returns based on the dividend yield (Model 2 selected by AIC for $x = \Delta d$ and MA = 12 months)

```
*******************************************************************************
 Based on 314 observations from 1975M8  to 2001M9 .
 ARDL(6,7) selected using Akaike Information Criterion.
 Dependent variable in the ARDL model is D12 included with a lag of 6.
 List of other regressors in the ARDL model:
 DP12            DP12(-1)        DP12(-2)        DP12(-3)        DP12(-4)
 DP12(-5)        DP12(-6)        DP12(-7)        INPT
*******************************************************************************
 Observation             Actual             Prediction             Error
  2001M10               -13.8791               -8.0222               -5.8569
  2001M11               -10.4747               -7.2481               -3.2265
  2001M12                -6.6052               -7.8574                1.2522
  2002M1                -13.9974               -7.7671               -6.2303
  2002M2                -10.8478               -4.9877               -5.8601
  2002M3                  5.9525               -5.7072               11.6597
  2002M4                  2.8459               -1.4678                4.3137
  2002M5                 -.10917                .47107               -.58024
  2002M6                -11.0711                .99067              -12.0617
  2002M7                -14.8009                3.4534              -18.2544
  2002M8                -12.6037                3.5014              -16.1051
  2002M9                -12.4071                4.0621              -16.4692
*******************************************************************************

             Summary Statistics for Residuals and Forecast Errors
*******************************************************************************
                     Estimation Period             Forecast Period
                    1975M8  to 2001M9            2001M10 to 2002M9
*******************************************************************************
 Mean                         -.0000                        -5.6182
 Mean Absolute                 2.0586                         8.4892
 Mean Sum Squares             10.0881                       106.9708
 Root Mean Sum Squares         3.1762                        10.3427
*******************************************************************************
```

Table 10b. – Dynamic 3-months ahead in-sample forecasts of the level of German stock returns based on the dividend yield (model 2 selected by AIC for $x = d$ and MA = 12 months)

```
Dynamic forecasts for the level of D12
*******************************************************************************
 Based on 323 observations from 1975M8  to 2002M6 .
 ARDL(6,7).
 Dependent variable in the ARDL model is D12 included with a lag of 6.
 List of other regressors in the ARDL model:
 DP12            DP12(-1)        DP12(-2)        DP12(-3)        DP12(-4)
 DP12(-5)        DP12(-6)        DP12(-7)        INPT
*******************************************************************************
 Observation             Actual            Prediction             Error
  2002M7                -14.8009              -5.7937               -9.0072
  2002M8                -12.6037              -4.8943               -7.7095
  2002M9                -12.4071              -5.8713               -6.5358
*******************************************************************************


             Summary Statistics for Residuals and Forecast Errors
*******************************************************************************
                      Estimation Period            Forecast Period
                     1975M8  to 2002M6           2002M7  to 2002M9
*******************************************************************************
 Mean                      -.0000                      -7.7508
 Mean Absolute              2.1695                      7.7508
 Mean Sum Squares          11.3837                     61.0942
 Root Mean Sum Squares      3.3740                      7.8163
*******************************************************************************
```

Figure 4a. – Dynamic 12-months ahead in-sample forecasts of the level of German stock returns based on the dividend yield (Model 2 selected by AIC for $x = \Delta d$ and MA = 12 months)

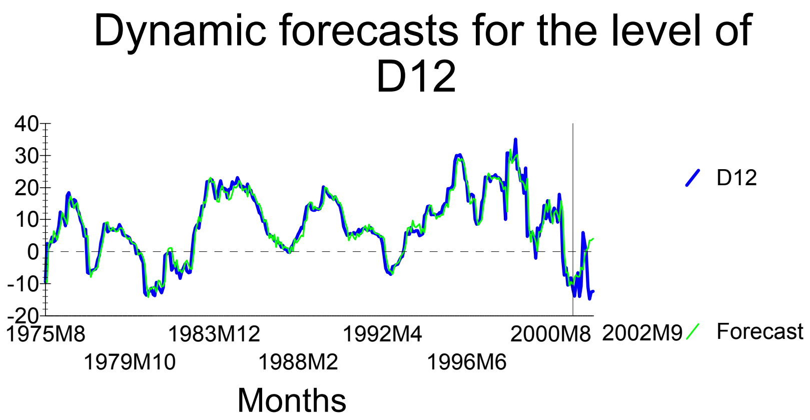


Figure 4b. – Dynamic 3-months ahead in-sample forecasts of the level of German stock returns based on the dividend yield (Model 2 selected by AIC for $x = d$ and MA = 12 months)