A cointegration approach to the lead-lag effect among size-sorted equity portfolios

Angelos Kanas
Department of Economics
University of Crete
74100 Rethymnon
Crete, Greece

Tel: 0030 831 77427
E-mail: a-kanas@econ.soc.uoc.gr

and

Georgios P. Kouretas *
Department of Economics
University of Crete
74100 Rethymnon
Crete, Greece

Tel: 0030 831 77412
E-mail: kouretas@econ.soc.uoc.gr

* : Correspondence author

September 2001

An earlier version of this paper was presented at the Conference of International Federation of Operation Research Societies (IFORS) on 'New Trends in Banking Management', Athens 1-3 April 2001, the International Conference on 'The Econometrics of Financial Markets', Delphi 22-25 May 2001, the 8th International Conference of the Multinational Financial Society, Lake Garda 23-27 June 2001 and the 10th Conference of the European Financial Management Association, Lugano 27-30 June 2001 and thanks are due to conference participants for numerous helpful comments. This paper has also benefited from comments by seminar participants at Athens University of Economics and Business, European University Institute, University of Crete and University of Cyprus. We would also like to thank without implicating Richard Baillie, George Constantinides, Laurence Copeland, Antonis Demos, Dimitrios Georgoutsos, John Goddard, Soren Johansen, Katarina Juselius, Eric Renault, Enrique Sentana, Nickolaos Travlos and Elias Tzavalis for helpful comments on an earlier draft of this paper.

I. Introduction

In their seminal work, Lo and MacKinlay (1990) document a lead-lag relation between weekly returns of size-sorted portfolios for the US market. Using the cross-autocorrelation test statistic, they demonstrate that returns of portfolios consisting of large-capitalisation stocks ('large-firm' portfolios) lead (i.e. are positively cross-autocorrelated with lagged) returns of portfolios consisting of small-capitalisation stocks ('small-firm' portfolios), but not vice-versa.[1] This relation indicates a complex mechanism of information transmission between small- and large-firms portfolio returns [Merton (1987), Badrinath, Kale and Noe (1995)]. Specifically, Lo and MacKinlay (1990) argue that this relation may be evidence of a lagged adjustment of small-firm portfolio prices, namely that information shocks are transmitted first to large and then to small firms.[2] An important implication that emerges from Lo and MacKinlay's findings refers to the short-run predictability of portfolio returns, namely that returns of large-firm portfolios can be used to reliably predict returns of small-firm portfolios in the short-run.[3]

This paper contributes to the literature on the lead-lag effect in several ways. We develop a formal framework illustrating how the lead-lag effect in returns is compatible with cointegration between the contemporaneous price of the small-firm portfolio and the lagged price of the large-firm portfolio. We show that a lead-lag effect in returns is a necessary (but not sufficient) condition for cointegration between the contemporaneous small-firm portfolio price and the lagged large-firm portfolio price. Cointegration between the current small-firm portfolio price and the lagged large-firm portfolio price can be interpreted as evidence of a *long-run* lead-lag relation among prices of size-sorted portfolios. Thus, we seek to extend Lo and MacKinlay's (1990)

---

[1] Previous research on stock return predictability includes Conrad and Kaul (1988, 1989), Conrad, Hameed and Niden (1994), and Lo and MacKinlay (1988).

[2] Badrinath, Kale and Noe (1995) have argued that a lead-lag effect may be related to the level of institutional ownership of firms. Another factor which is highly correlated with firm size and is consistent with lagged information transmission between large and small firms is the information set-up cost [Merton (1987)].

short-run approach, based on cross-correlations of returns over relatively short (weekly) return horizons, to the long-run as investors might have long holding periods.[4] Another important feature of cointegration in this context is that it carries important implications for the short-run predictability of the small-firm portfolio returns, namely that we can employ an error correction model to obtain more accurate short-run predictions. We employ the Autoregressive Distributed Lag (ARDL) approach, recently advanced by Pesaran and Shin (1998), to estimate the long-run parameters of the cointegrating relation between small- and large-firm portfolio prices, and also obtain the error correction model for predicting small-firm portfolio returns. The ARDL-based estimators of the long-run coefficients have the advantage of being super-consistent, and valid inferences on long-term parameters can be made using standard normal asymptotic theory. Furthermore, recent Monte Carlo evidence by Gerrard and Godfrey (1998) has indicated that diagnostic tests of the specification of the ECM can be sensitive to the method used to estimate the long-run coefficients that yield the ECM. These authors have concluded that the ARDL approach should be preferred in estimating the long-run coefficients of the cointegrating relation.

A new data-set for the UK equity market is used, containing three sets of equity portfolios, with each set consisting of ten portfolios. The first set contains equal-number-divided and value-weighted size-sorted (decile) portfolios. The portfolios contain approximately equal numbers of stocks, and within a portfolio returns are value-weighted. The second set contains equal-number-divided and equally weighted size-sorted portfolios. The portfolios contain approximately equal numbers of stocks, and within a portfolio returns are equally weighted.[5] The main feature of these two sets is that, in each set, portfolios have different capitalisation sizes, i.e. they are size-sorted. We therefore have two alternative weighting schemes for constructing size-

---

[3] Studies which have addressed the issue of short-term stock returns predictability based on their past history include Conrad and Kaul (1988, 1989), Chan (1988), Jegadeesh (1990), Lehmann (1990), Jegadeesh and Titman (1993, 1995), Levich and Thomas (1993), and Lo and MacKinlay (1990).

[4] Kasa (1992) argues that the correlation of stock returns over short run holding periods may be misleading for investors who have long holding periods. Moreover, Gallagher (1995) contends that correlation is a 'static' test applicable to short-run holding horizons.

sorted portfolios, namely a scheme yielding equally weighted and a scheme yielding value-weighted portfolios. The third set contains equal-value-divided and value-weighted portfolios. The portfolios are approximately equal in terms of the aggregated value of their stocks, and within a portfolio returns are value-weighted. Thus, in contrast to the two first sets, this set contains portfolios of approximately equal capitalisation sizes. The portfolios in all three sets are rebalanced at the end of each year. We wish to compare the test results for these three portfolio sets, and draw conclusions as to whether a long-run lead-lag effect arises in all three sets or not. If the lead-lag effect exists in the first two sets and not in the third then one could conclude that the lead-lag effect is driven by the capitalisation size. We also address the question of whether the existence of a lead-lag effect is determined by the particular weighting scheme, namely an equally-weighting and a value-weighting scheme, employed in the construction of the first two sets of portfolios.

Evidence of cointegration between contemporaneous small-firm portfolio prices and lagged large-firm portfolio prices is found only for size-sorted portfolios and not for equal-size portfolios, thereby indicating the importance of size in driving a long-run lead-lag effect. This result echoes the findings of Banz (1981) and Fama and French (1992) regarding the role of size in explaining asset returns. For size-sorted portfolios, the large-firm portfolio price appears to be the 'long-run forcing variable' for the explanation of the small-firm portfolio prices. It is important to note that, small-firm portfolio prices cannot be treated as 'long-run forcing variables' for the explanation of large-firm portfolio prices. These results suggest a long run lead-lag relation between the small- and large-firm portfolio prices, according to which small-firm portfolio returns lag large-firm portfolio returns and not vice-versa. This result holds for both sets of size-sorted portfolios, thereby indicating that the weighting scheme does not affect the existence of the lead-lag effect. We next compare out-of-sample forecasts of small-firm portfolio returns obtained from an ARDL-based ECM against an alternative model, namely a model

[5] Chen et al. (1986) also, used this weighting scheme.

without the error correction term. Our results indicate that the forecasts from the ECM models outperform the forecasts from the other two models, on the basis of the Root Mean Square Error (RMSE) criterion. Thus, we document that cointegration between the price of small-firm portfolios and the lagged price of large-firm portfolios may be utilised to improve on forecasts for small-firm portfolio returns. Overall, our results should be of interest to technical analysts, institutional investors and portfolio managers who seek to identify profitable portfolio strategies on the basis of past returns, as well as to 'producers' of asset-pricing models, who seek to identify relevant variables capable of explaining asset returns.

The structure of the paper is as follows. Section 2 develops a framework illustrating the relevance of cointegration in testing for a long-run lead-lag relation, and discusses the ARDL approach to cointegration. Section 3 describes the data set used in this study. Section 4 discusses the empirical findings. Finally section 5 concludes.

II. Cointegration and prices of size-sorted portfolios

A. Cointegration and lead-lag effect

Bossaerts (1988) conducted an early study of cointegration and asset prices.[6] Bossaerts assumed a Lucas-type, one-good pure exchange economy with a representative consumer, where dividends are all consumed, and showed that cointegration between contemporaneous asset prices may arise because of the approximate separation properties of this economy with a risk-averse representative consumer. In the present paper, we focus on Lo and MacKinlay's (1990) approach, and discuss how cointegration between the *lagged* large-firm portfolio price and the contemporaneous small-firm portfolio price is compatible with the lead-lag effect.

[6] Other studies include those by Granger (1986), who equated tests of cointegration between asset prices as tests for market efficiency, Campbell and Shiller (1988), who discuss the links between cointegration and

Lo and MacKinlay (1990) argue that the lead effect from large-firm portfolio returns to small-firm portfolio returns may be the result of information affecting first the prices of large market value securities and then the prices of small market value securities. In this section, we illustrate that lagged price adjustment to common factor shocks is compatible with cointegration between the lagged large-firm portfolio price and the contemporaneous small-firm portfolio price. Consider the returns of a large-firm portfolio, denoted by R_{Lt}, and the returns of a small-firm portfolio, denoted by R_{St}. Lo and MacKinlay (1990) employ a single factor and assume that lagged factor shocks affect the current returns of small-firm portfolios; the smaller the market value of a security, the longer the lag in the price adjustment. For convenience, we assume here that information shocks may affect small-firm portfolio returns only up to one lag, with the importance of the lagged shock to portfolio's returns declining as the market value of the portfolio increases. According to Lo and MacKinlay's model, the returns for the large-firm portfolio are given by:

$$R_{Lt} = \mu_L + \beta_{1L} f_t + \varepsilon_{Lt} \tag{1}$$

and the returns for a small-firm portfolio with a lagged adjustment to information shocks are:

$$R_{St} = \mu_{Si} + \beta_{1S} f_t + \beta_{2S} f_{t-1} + \varepsilon_{St} \tag{2}$$

where R_{it} is the return for security i at time t ($i = S, L$), f_t is a white noise factor shock at time t, ε_t is a stationary idiosyncratic error term. To illustrate the role of the lagged adjustment of the small-firm portfolio price in entailing cointegration between small- and large-firm portfolio prices, let us for the moment assume that no lagged price adjustment occurs and thus, the returns of a small-firm portfolio are given by (2)', namely:

$$R_{St} = \mu_{Si} + \beta_{1S} f_t + \varepsilon_{St} \tag{2'}$$

both dividend valuation models and term structure of interest rates, and Brenner and Kroner (1992), who discuss the relation between cointegration and no arbitrage pricing.

Summing to get prices, we obtain equations (3) and (4) which give the price for the large-firm portfolio and the price of the small-firm portfolio without lagged price adjustment respectively:

$$p_{Lt} = \mu_L t + \beta_{1L} F_t + \sum_{j=0}^{t} \varepsilon_{L,t-j} \qquad (3)$$

$$p_{St} = \mu_S t + \beta_{1S} F_t + \sum_{j=0}^{t} \varepsilon_{S,t-j} \qquad (4)$$

where F_t is the sum of factor shocks from time 0 to (t). Solving equation (3) for F_t and substituting F_t into (3) yields:

$$p_{st} = at + bp_{Lt} + v_t \qquad (5)$$

where $a = [\mu_s - (\beta_{1S} / \beta_{1L})\mu_L]$, $b = (\beta_{1S} / \beta_{1L})$, and $v_t = \sum_{j=0}^{t} \varepsilon_{S,t-j} - b\sum_{j=0}^{t} \varepsilon_{L,t-j}$.

In the absence of a lagged price adjustment, equation (5) implies a regression between the contemporaneous prices of the two portfolios, with an idiosyncratic nonstationary error term v_t. This non-systematic error term implies that the residuals of regression (5) will be nonstationary and thus there will be no cointegration between the two portfolio prices.

Assume now that there is a lagged adjustment in the price of the small-firm portfolio given by equation (2). Summing to get the price of this portfolio yields equation (6) :

$$p_{St} = \mu_S t + [\beta_{1S} + \beta_{2S}]F_{t-1} + \beta_{1S} f_t + \sum_{j=0}^{t} \varepsilon_{S,t-j} \qquad (6)$$

where F_{t-1} is the sum of factor shocks from time 0 to (t-1). Next, consider the price of the large-firm portfolio with 1 lag, given by equation (7):

$$p_{L,t-1} = \mu_L (t-1) + \beta_{1L} F_{t-1} + \sum_{j=0}^{t-1} \varepsilon_{L,t-j} \qquad (7)$$

Solving for F_{t-1} yields:

$$F_{t-1} = \frac{1}{\beta_{1L}} [p_{L,t-1} - \mu_L t + \mu_L - \sum_{j=0}^{t-1} \varepsilon_{L,t-j}] \qquad (8)$$

Substituting (8) into (6) yields

$$p_{S,t} = \gamma\mu_L + (\mu_S - \gamma\mu_L)t + \gamma p_{L,t-1} + \beta_{1S} f_t + w_t \quad \Rightarrow$$

$$p_{S,t} = \gamma\mu_L + (\mu_S - \gamma\mu_L)t + \gamma p_{L,t-1} + e_t \qquad (9)$$

where $\gamma = [(\beta_{1S} + \beta_{2S}) / \beta_{1L}]$, w_t is an idiosycratic error term, $w_t = \sum_{j=0}^{t} \varepsilon_{S,t-j} - \gamma \sum_{j=0}^{t-1} \varepsilon_{L,t-j}$, and

$e_t = \beta_{1S} f_t + \sum_{j=0}^{t} \varepsilon_{S,t-j} - \gamma \sum_{j=0}^{t-1} \varepsilon_{L,t-j}$.

As shown in equation (9), a lead-lag effect from the large-portfolio returns to small-portfolio returns (i.e. a lagged adjustment of the price of the small-firm portfolio) implies a regression of the small-firm portfolio price on the *lagged* price of the large-firm portfolio, a trend, and the white noise common factor f_t. Under a lagged small-portfolio price adjustment, the existence of cointegration (i.e. the stationarity of the residuals e_t) depends on the relative importance of the white noise common factor f_t and the idiosyncratic nonstationary term w_t. In other words, lagged price adjustment entails including a white noise factor f_t in equation (9), which introduces a stationary component in the residuals of equation (9). Without lagged price adjustment, f_t does not enter equation (9) and therefore, there is no such component in the residuals (and thus, there is no tendency towards cointegration). Consequently, Lo and MacKinlay's (1990) model has two conflicting time-series features: the lagged small-firm portfolio price adjustment to information shock, which entails a cointegrating relation between the small-firm portfolio price and the lagged large-firm portfolio price, and an idiosyncratic error term which obscures this link. If this idiosyncratic error term is not 'corrected', it will cause the prices of small- and large-firm portfolios to diverge. Thus, a lead-lag effect in portfolio returns may entail cointegration in portfolio prices if the idiosyncratic term is 'sufficiently' small, and may not entail cointegration if it is 'sufficiently' large. If cointegration is indeed found, then we could conclude that this idiosyncratic term may be sufficiently small, and that there is a lead-lag effect in the long-run.[7] Evidence of cointegration in portfolio prices can be interpreted as a long-run lead-lag effect.

In the empirical analysis, we test for cointegration between the prices of size-sorted portfolios using the well-known Phillips and Hansen (PH) (1990) Fully Modified-OLS procedure. This method has the advantage of being valid under a wide range of different distributional assumptions regarding the error terms (Moore and Copeland, 1995). We consider pairs of different size portfolios, and test for bivariate (pairwise) cointegration in the prices of a 'small'- and a 'large'-firm portfolio in order to be consistent with Lo and MacKinlay's (1990) approach of considering two portfolios at a time. Subject to establishing cointegration, we proceed to estimating the coefficients of the cointegrating vector, i.e. the long-run coefficients, and error correction models using the recently developed ARDL approach.

B. The Auroregressive Distributed Lag (ARDL) Cointegration Approach

In this section, we outline the ARDL approach to estimating a cointegrating relation, and discuss its advantages compared to other approaches. According to Pesaran and Shin (1998), the general ARDL(p,q) model is given by the following equation:

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \phi_t y_{t-i} + \beta' x_t + \sum_{i=0}^{q-1} \beta_i^{*'} \Delta x_{t-i} + u_t \tag{10}$$

$$\Delta x_t = P_1 \Delta x_{t-1} + P_2 \Delta x_{t-2} + \ldots + P_s \Delta x_{t-s} + \varepsilon_t$$

where x_t is the k-dimensional I(1) 'forcing' variables which are not cointegrated among themselves, u_t and ε_t are serially uncorrelated disturbances with zero means and constant variance-covariances, and P_i are k x k coefficient matrices such that the vector autoregressive process in Δx_t is stable. In the case of bivariate cointegration, we set $k = 1$. By setting $k = 1$, we

[7] Bossaerts (1988) considers a class of economies that also generate cointegrated asset prices. However, these economies are Lucas-type one-good pure exchange economies in which all dividend payments are assumed to be consumed. In our paper, we depart from Bossaerts assumption of a Lucas-type economy, and assume in the empirical analysis that dividends are reinvested as opposed to consumed.

avoid the problem of cointegration among the 'forcing' variables x_t, and are consistent with Lo and MacKinlay's (1990) pairwise approach to the lead-lag relation. In the above formulation, $\alpha_0, \alpha_1, \beta, \beta_1^*, ...\beta_{q-1}^*, \phi = (\phi_1, ...\phi_p)$ are the short-run parameters which are important in estimating the long-run coefficients defined by the ratios $\delta = \alpha_1 / \phi(1)$, and $\theta = \beta / \phi(1)$, where $\phi(1) = 1 - \sum_{i=1}^{p} \phi_i$. The ARDL approach also assumes that there is a stable long-run relation between the two variables, y and x. In the case where u_t and ε_t are correlated, the above ARDL specification is augmented with an adequate number of lagged changes in the regressors. The degree of augmentation required depends on whether q>s+1 or not. The augmented model is given by:

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \phi_t y_{t-i} + \beta' x_t + \sum_{i=0}^{m-1} \pi_i^* \Delta x_{t-i} + n_t \qquad (11)$$

where $m = max(q, s+1)$, $\pi_i = \beta_i^* - P_i' d$, d is a 1 x 1 vector containing the contemporaneous correlation between u_t and ε_t. Thus, the ARDL approach requires inserting enough lags of the 'forcing' variables in order to endogenise y_t. By doing this, the problem of endogenous regressors and serial autocorrelation can be simultaneously corrected for (Pesaran and Shin, 1998, page 16). The order of the distributed lag function on y_t and the forcing variable x_t are selected using the Akaike Information Criterion (AIC) or the Schwartz Bayesian Criterion (SC).[8] Setting the maximum orders of p and q equal to 12 (for monthly data), we compare the maximised values of the log-likelihood functions of the $(m+1)^{k+1} = (m+1)^2$ different ARDL models. We select the final model by finding those p and q which maximise the above mentioned selection criteria. Once the model has been selected, it is estimated using OLS to obtain the short-run parameters. Next, we estimate the long-run coefficients of the cointegrating relation $y_t = a + \delta t + \theta x_t + v_t$ by

[8] Monte Carlo evidence by Pesaran and Smith (1998) indicates that the Schwartz Bayesian Criterion tends to be preferred to the AIC.

$$\hat{a} = [\hat{a}_0 /(1 - \hat{\phi}_1 - ... - \hat{\phi}_p)], \tag{12a}$$

$$\hat{\delta} = [\hat{a}_1 /(1 - \hat{\phi}_1 - ... - \hat{\phi}_p)], \tag{12b}$$

$$\hat{\theta} = [(\hat{\beta}_0 + \hat{\beta}_1 + ... \hat{\beta}_q) /(1 - \hat{\phi}_1 - ... - \hat{\phi}_p)] \tag{12c}$$

As shown by (12a)-(12c), the long-run coefficients ($\hat{a}, \hat{\delta}, \hat{\theta}$, as calculated from equations 12a, 12b, and 12c) reflect the short-run parameters, namely the coefficients of the lags of the dependent and independent variables in the ARDL(*p,q*) model ($\hat{a}_0, \hat{a}_1, \hat{\delta}_0, \hat{\delta}_1, \hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_{p,} \hat{\phi}_1 ..., \hat{\phi}_p$ in equation 11). Thus, the long-run coefficients capture the effects of the lagged variables in the cointegrating relation. Pesaran and Shin (1998) show that these ARDL-based estimators of the long-run coefficients are super-consistent and, more importantly, valid inferences on these estimators can be made using standard normal asymptotic theory. The standard errors of the estimates of the long-run coefficients can in principle be obtained by the so-called 'delta' method.[9]

The use of the ARDL estimation procedure is directly comparable to the semi-parametric, Phillips-Hansen Fully Modified-OLS approach to estimation of cointegrating relations. Monte Carlo evidence by Pesaran and Shin (1998) indicates that the ARDL approach using the SC selection criterion and combined with the delta method of computing standard errors generally dominates the PH approach, especially with regard to the size-power performance of the tests on the long-run parameters. Moreover, the ARDL estimators are 'substantially less biased' than the PH estimators. Finally, recent Monte Carlo evidence by Gerrard and Godfrey (1998) indicates that the ARDL approach is preferable to other methods of estimating the long-run coefficients of the cointegrating relation, in the light of the sensitivity of the diagnostic tests of the specification of the ECM to alternative estimation methods.

---

[9] Another approach to obtaining standard errors is Bewely's (1979) regression approach. As Pesaran and Shin (1998) argue, both methods yield identical results and a choice between them is only a matter of computational convenience. In the present paper, we used the delta method.

III. The data-set

A data-set for the UK stock market is constructed to test for cointegration in the prices of size-sorted equity portfolios. We use the London Business School Share Price Database (LSPD) to obtain monthly stock return data covering the period from January 1955 to December 1994. The LSPD tapes contain the returns series for approximately 6000 companies, comprising several different samples: (i) a complete coverage of stocks after 1974 of every UK listed stock, and (ii) a fully comprehensive coverage of stocks over the 1955-1974 period based on a random sample of one-third of existing issues and new flotations. The data consists of 480 observations of monthly total returns, including reinvested dividends and capitalisation changes where appropriate, i.e. each return is calculated as $\log_e[(P_t + D_t)/P_{t-1}]$, where P_t is the stock price at time *t*, and D_t is the dividend at time *t*. We effectively assume that all dividends are reinvested, in line with Lo and MacKinlay (1990) who employed *total* returns to explain autocorrelation and cross-correlation patterns. We next use these cum-dividend returns to compute stock prices.

The stock price series computed from the LSPD cum-dividend returns series are used to create a data-set of UK stock market data, namely three sets of ten portfolios in each set. The first set comprising equal-number-divided and value-weighted size portfolios is referred to as NV portfolios. The second set comprising equal-number-divided and equally-weighted size portfolios is referred to as NE portfolios. The third set comprising equal-value-divided and value-weighted size portfolios is referred to as VV portfolios. Note that Lo and MacKinlay's (1990) results were based on equally-weighted size-sorted portfolios. Under each scheme, 'Portfolio 1' contains the smallest firms, 'Portfolio 2' the next smallest, and so on up to 'Portfolio 10' which contains the largest firms.

As both the cointegration tests and the ARDL approach require that the variables involved be I(1), we test for a unit root in each of the ten portfolios price series in each set. We employ the Kwiatkowski et al. (1992) (KPSS) test.[10] The null hypothesis of this test is that the series is stationary against the alternative hypothesis of nonstationarity. The 5% critical value for the test with trend is 0.146. Results are reported in Table 1. As shown in this Table, the test statistic for all portfolio prices is higher than 0.146, thereby rejecting the null of stationarity. In contrast, the test statistic for all returns series is lower than the critical value, thereby failing to reject the null of stationarity. Therefore, we conclude that all ten portfolios' prices are I(1).[11] Thus, we proceed to test for cointegration and, subject to establishing cointegration, we estimate the long-run coefficients and the error correction models using the ARDL approach.

IV. Empirical findings

A. Cointegration results

We showed in Section 2 that the lead-lag effect in returns is compatible with cointegration between the contemporaneous price of small-firm portfolios and the lagged price of large-firm portfolios. In this section, we test for pairwise cointegration between the current price of a small-firm portfolio and the lagged price of a large-firm portfolio for all possible combinations of portfolio pairs in each portfolio set. All cointegration tests refer to the period from January 1955 to December 1993, leaving the period of the last year (1994) for out-of-sample forecasting. The well-known Phillips and Hansen (1990) cointegration test was employed. The

[10] The KPSS unit root test is based on the assumption that a time series y_t is the sum of a deterministic trend t, a random walk r_t and a stationary error ε_t:

$$y_t = \xi t + r_t + \varepsilon_t.$$

The random walk is $r_t = r_{t-1} + u_t$ where u_t are iid $(0, \sigma_u^2)$. In this framework, for the null hypothesis that y_t is trend stationary to be true, the variance of the random walk component, σ_u^2, should be equal to zero. Testing of the null hypothesis that y_t is stationary around a level, is carried out by omitting the time trend. The test statistic is defined as

$$n = T^{-2}\sum_{t=1}^{T} S_t^2 \; / \; s^2(l)$$

where T is the sample size, S_t is the sum of the residuals when the series is regressed on an intercept and a time trend, and $s^2(l)$ is a consistent non-parametric estimate of the long-run variance of the error term. Critical values for the KPSS test, n, without a trend (n_μ) or with a trend (n_τ) are found in Kwiatkowski et al. (1992). We calculate the KPSS test statistics using a number of 8 lags, l, in the estimation of the long-run variance of residuals, on the basis of the Kwiatkowski et al. (1992, page 174) criterion of choosing l at the value at which the test statistic settles down.

null hypothesis is no cointegration. The null is rejected if the computed test statistic is smaller than -20.4935, the 5% critical value (Phillips and Ouliaris, 1990, Table Ib). The results are reported in Table 2, Panels A, B, and C for the NE, NV, and VV portfolios respectively.[12] All cointegration tests are based on the contemporaneous price of Portfolio *i* and the lagged price of Portfolio *j*, *i, j* = 1…10, *j* > *i*. To illustrate, consider Panel A. The PH test statistic for the null of no cointegration between the lagged price of Portfolio 2 (*j* = 2) and the current price of Portfolio 1 (*i* = 1) is –0.63. Similarly, the test statistic for the null of no cointegration between the lagged price of Portfolio 10 (*j* = 10) and the current price of Portfolio 1 (*i* = 1) is –24.20. As shown in Panel A (NE Portfolios), there is evidence of cointegration between Portfolio 1 (the smallest-firm Portfolio) and Portfolios 10, 9, 8 and 7; between Portfolio 2 and Portfolios 9, 8 and 7; between Portfolio 3 and Portfolios 4, 5, 6,7,8,9, 10, between Portfolios 4 and 5; Portfolios 6 and 9; and Portfolios 8 and 9. Results from Panel B (NV Portfolios) are similar to those for the NE Portfolios. Specifically the current price of Portfolio 1 is cointegrated with the lagged price of Portfolios 7, 8, 9, and 10; the current price of portfolio 2 with the lagged price of Portfolios 8, 9, and 10; and the current price of Portfolio 3 with the lagged price of Portfolios 4, 5, …10. For NV Portfolios, there is no evidence of cointegration between Portfolios 4 and 5; 6 and 9; and 8 and 9. Despite these slight differences, the overall results for the NE and NV portfolios are quite similar and indicate that lagged large-firm portfolio prices cointegrate with current prices of small-firm portfolios. The similarity of the results for NE and NV portfolios indicates that the finding of cointegration is not dependent on the method of portfolio construction (i.e. value-weighting vs equally-weighting schemes).

[11] We also applied the augmented Dickey Fuller (ADF) test, and found similar results. These results are not reported here, but are available upon request.
[12] To save space, we only report the trace test statistics. The maximal eigenvalue statistics yield similar results, and are available upon request.

We next turn to the cointegration results for the VV portfolios which represent value-weighted portfolios. These results are reported in Panel C of Table 2. As shown in this Panel, there is hardly any evidence of cointegration between portfolio prices. The null hypothesis of no cointegration can be rejected only in two cases, namely for the pairs of Portfolios 6 and 7, and Portfolios 8 and 9. Compared with the results from Panels A and B, we could argue that the existence of cointegration between portfolios is driven by differences in portfolio capitalisation size. VV portfolios represent equal aggregated value of stocks the number of which varies considerably among the ten portfolios. Thus, Portfolio 1, although composed of relatively small-firms, represents equal aggregated market value to that of Portfolio 10, which contains a different (smaller) number of large-firms. Lack of cointegration between prices of equal-size portfolios is compatible with no lagged information transmission adjustment, shown in equations (1), (2)', (3), (4) and (5). Finally, this finding is also in line with the lagged price adjustment hypothesis of Lo and MacKinlay (1990)[13].

The next step of our analysis is the estimation of the coefficients of the cointegrating vector for the cases where cointegration was found in Table 2. These coefficients are for the cointegrating relation $y_t = a + \delta t + \theta x_t + v_t$, and not for the $y_t = a' + \delta' t + \theta' x_{t-1} + v_t$, because, according to the ARDL approach, one needs to capture the long-run coefficients of the cointegrating vextor (namely, *α,δ, θ*), and not the short-run coefficients (namely, *α', δ', θ'*). The long-run coefficients incorporate any effects of the short-run parameters of the lagged effects of the variables in the ARDL (*p,q*) model of equation (11), as they are calculated using formulae (12a) –(12c) which are repeated here for convenience.[14]

$$\hat{a} = [\hat{a}_0 / (1 - \hat{\phi}_1 - ... - \hat{\phi}_p)],$$

[13] By construction, VV portfolios are not as well diversified as NE and NV portfolios. As the level of portfolio diversification is reflected upon the variance and not the price of the portfolios, and given that we test for cointegration among the prices of different portfolios, one would expect that the relatively low degree of diversification of VV portfolios does not influence the cointegration results.

$$\hat{\delta} = [\hat{a}_1 / (1 - \hat{\phi}_1 - ... - \hat{\phi}_p)],$$

$$\hat{\theta} = [(\hat{\beta}_0 + \hat{\beta}_1 + ... \hat{\beta}_q) / (1 - \hat{\phi}_1 - ... - \hat{\phi}_p)]$$

Panels A and B of Table 3 report the results of the estimated long-run coefficients $\hat{\theta}$ for the NE and NV portfolios respectively.[15] The empty cells are for the pairs of portfolios for which no cointegration was found in Table 2. The upper diagonal for each Panel reports the long-run coefficients $\hat{\theta}$ for cointegration relations in which the 'large'-firm portfolio is the independent (right-hand-side) variable in the cointegrating relation. The lower diagonal reports the coefficients $\hat{\theta}$ for cointegrating relations in which the 'small'-firm portfolio is the independent variable. Asymptotic t-statistics are also reported underneath the estimated coefficient. Finally, the order of the ARDL(*p,q*) model, upon which the estimated long-run coefficients were based, is also reported for each portfolio pair[16]. For instance in Panel A, the long-run coefficient in the cointegrating relation where the independent variable is Portfolio 10 and the dependent variable is Portfolio 1 is 1.80, with an asymptotic t-statistic of 5.33. The order of the ARDL model on the basis of which this long-run coefficient was derived is ARDL(3,2). As shown in Panel A, the long-run coefficient of the large-firm portfolio is statistically significant in every cointegrating relation between a large-firm and a small-firm portfolio. Therefore, the large-firm portfolio prices are long-run 'forcing' variables for small-firm portfolio prices in all cases where cointegration was found. This finding is compatible with Lo and MacKinlay's (1990) arguments that the lead-lag effect is from large- to small-firm portfolios. We next examine whether the small-firm portfolio prices can also be regarded as long-run forcing variables for the large-firm portfolio

[14] The estimated short-run parameters of the ARDL(p,q) model are not reported here, as they are not of direct relevance, and are available upon request.

[15] To save space, we do not report the other long-run coefficients, namely coefficients $\hat{\alpha}$, and $\hat{\delta}$. These are available on request.

[16] The order of the estimated ARDL models was based on the BIC, hence the difference in the order across the different ARDL models. We also estimated the models using the AIC. In some cases, the AIC-based order was different from that under the BIC, but the results are qualitatively similar.

prices, by placing the small-firm portfolio price on the right-hand-side of the cointegrating relation and the large-firm portfolio as the dependent variable. The results of the estimated long-run coefficients $\hat{\theta}$ are now reported in the lower diagonal part of Panel A. As shown in the lower diagonal, when Portfolio 1 is the independent variable, the long-run estimated coefficient is not statistically significant. This implies that the price of Portfolio 1 is not a long-run forcing variable for the prices of Portfolios 7, 8, 9, and 10, whereas, as found above, each of these Portfolios is a long-run forcing variable for Portfolio 1. The same conclusion applies to Portfolio 2. For Portfolio 3, we find that the long-run coefficient is not statistically significant if the dependent variable is Portfolio 10, 9, 8 or 7, which are relatively large-firm portfolios. As mentioned above, however, each of these Portfolios is a long-run forcing variable for Portfolio 3. Therefore, we can argue that for the largest-firm portfolios (7, 8, 9, and 10), portfolios 1, 2, and 3 are not long-run forcing variables, whereas the largest-firm portfolios are long-run forcing variables of the smallest-firms portfolios. In the cointegrating relations where the dependent variable is Portfolio 6, 5, or 4 and the independent is Portfolio 3, the long-run coefficient is statistically significant. This may be due to the fact that the size difference between these portfolios is relatively small[17]. Similar results hold for the NV portfolios. Larger-firm Portfolios 10, 9, 8, and 7, and Portfolios 10, 9, and 8 are long-run forcing variables of smaller-firm Portfolios 1 and 2, respectively, whereas Portfolios 1 and 2 are not long-run forcing variables of the larger-firm portfolios. This is in line with Lo and MacKinlay's (1990) finding that there is only a lead effect and no lag effect from small- to large-firm portfolio prices.

[17] This result may be related to the finding of cointegration between Portfolio 3, and Portfolio 4, Portfolio 5, and Portfolio 6 in Tables 2A and 2B.

B. Out-of-sample predictions of small-firm portfolio returns: Error correction models based on ARDL approach

For the cases where cointegration was found, and after estimating the long-run coefficients of the cointegrating vector, we proceed to obtain the error correction representation of the selected ARDL(*p,q*) model for portfolio returns. This model is used to obtain out-of-sample forecasts for the returns of small-firm portfolios for the period January 1994 – December 1994. The error correction model is the selected ARDL(*p,q*) model expanded by the error correction term, which is calculated using the ARDL-based long-run coefficients estimated during the 1955-1993 period[18]. The error correction representation of the selected ARDL(p,q) model is given by equation (13):

$$\Delta y_t = \delta_0 + \delta_1 t + \sum_{i=1}^{p} d_t \Delta y_{t-i} + \sum_{i=1}^{m} g_i \, \Delta x_{t-i} + \gamma ect_{t-1} + n'_t \qquad (13)$$

where $m = max(q, s+1)$, Δy_t is the portfolio returns, and ect_{t-1} is the lagged error correction term estimated using the ARDL approach. The coefficient of the ect_{t-1}, γ, is expected to be negative and statistically significant. The error correction models are estimated for the period January 1955 to December 1993. The results from estimating the error correction models are reported in Table 4, Panels A and B for the NE and NV portfolios respectively. As shown in both Panels, in all models the lagged error correction term enters with a negative sign and is statistically significant, thereby justifying the relevance of error correction models in out-of-sample forecasting.

We next proceed to obtain out-of-sample forecasts for the returns of portfolios for the period January 1994 to December 1994, using the estimated coefficients of the error correction models. For comparison purposes, we also obtain out-of-sample forecasts from same order ARDL(*p,q*) competing models which do not include the error correction term. The latter models can be regarded as missing the long-run forcing effect of the large-firm portfolios on the returns

[18] The order of the estimated error correction models is based on the BIC, hence the difference in the order across the different models.

of the small-firm portfolios documented in the previous section. The accuracy of the forecasts of the two competing models is measured using the Root Mean Squared Error (RMSE). The RMSEs of the foreacsts from the two competing models are reported in Table 5, Panels A and B for the NE and NV portfolios respectively. As shown in both Panels, the RMSE of the model with the error correction term is always smaller than the RMSE from the model without the error correction term. To test the statistical significance of the difference of the two RMSEs, we employ the nonparametric exact finite-sample Wilcoxon's signed-rank test, recommended by Diebold and Mariano (1995). The H_0 is that the mean-squared-errors of the error correction model is equal to that of its rival model, i.e. the RMSEs of the two models are the same. The last columns of Table 5, Panels A and B, show that the null hypothesis is rejected at the 5% significance level in each case, which implies that the RMSE of the error correction model is significantly smaller than that of the model without the error correction term. Consequently, the out-of-sample forecasts of small-firm portfolio returns from the models incorporating the long-run forcing effect of the large-firm portfolios are statistically more accurate than the forecasts from the models which do not include this long-run forcing effect.

E. Conclusions

Lo and MacKinlay's (1990) finding of a lead-lag effect in the returns of size-sorted portfolios was attributed to lagged information transmission to small-firm portfolio returns. We have developed a formal framework which illustrates how lagged information transmission may entail cointegration between the current price of small-firm portfolios and the lagged price of large-firm portfolios. If there were no lagged information transmission, then cointegration would not arise. We have shown that Lo and MacKinlay's (1990) lead-lag effect is a necessary condition for cointegration between the lagged price of large-firm portfolios and the contemporaneous price of the small-firm portfolios. We tested for cointegration between the current price of a small-firm portfolio and the lagged price of a large-firm portfolio using UK stock market data. Two sets of size-sorted portfolios and a set of equal-size UK equity portfolios have been constructed. One set of size –sorted portfolios comprises equally-weighted portfolios, while the other set comprises value-weighted portfolios.

The results from the cointegration tests indicated that for the two sets of size-sorted portfolios, there is substantial evidence of cointegration. Furthermore, using the recently advanced ARDL approach to cointegration, we conclude that large-firm portfolio prices are long-run forcing variables for small-firm portfolio prices and not vice versa. This result is in line with Lo and MacKinlay's (1990) finding that there is a lead effect from the large- to the small-firm portfolio returns and not vice-versa. For equal-size portfolios, we fail to find evidence of cointegration. This finding is not incompatible with Lo and MacKinlay's (1990) results, as the lead-lag effect should arise for portfolios which have different market capitalisations. Thus, we conclude that the capitalisation size drives the long-run lead-lag effect between small- and large-firm portfolio prices. This lead-lag effect is not affected by the weighting scheme (i.e. value-weighting vs equal-weighting) used to construct size-sorted portfolios.

For the portfolios for which cointegration was found, we estimated error correction models using the ARDL approach, and obtained out-of-sample forecasts of the returns of small-firm portfolios. An alternative model without the error correction term was also considered for comparing the accuracy of these forecasts. On the basis of the RMSE statistic, we found that the error correction models have significantly superior forecasting performance, thereby highlighting the relevance of cointegration between the lagged large-firm portfolio price and the current small-firm portfolio price in predicting small-firm portfolio returns. These results are of interest to technical analysts, portfolio managers and 'producers' of asset pricing models in seeking to identify relevant variables which explain asset returns and for forecasting.

References

Badrinath, S. G., J. R. Kale, and T. H. Noe. "Of shepherds, sheep, and the cross correlations in equity returns." *Review of Financial Studies*, 8 (1995),.

Banz, R. W. "The relationship between return and market value of common stocks." *Journal of Financial Economics,* 9 (1981), 3-18.

Bewley, R. " The direct estimation of the equilibrium response in a linear dynamic model." *Economics Letters*, 3 (1979), 357-361.

Bossaerts, P. "Common stationary components of asset prices." *Journal of Economic Dynamics and Control*, 12 (1988), 347-364.

Brenner, R. J., and K. F. Kroner. "Arbitrage and cointegration." Working Paper, University of Arizona (1992).

Campbell, J.Y., and R. J. Shiller. "Interpreting cointegrated models." *Journal of Economic Dynamics and Control*, 12 (1988), 505-522.

Chan, K. C. "On contrarian investment strategy." *Journal of Business*, 61 (1988), 147-163.

Chen, N. F., R. Roll, and S.A. Ross. "Economic forces and the stock-market." *Journal of Business*, 59 (1986), 3, 383-403.

Cheung Y. W., and L. K. Ng. "A causality-in-variance test and its implications to financial market prices." *Journal of Econometrics*, 72 (1996), 33-48..

Cohen, K., S. Maler, R. Scwartz, and D. Whitcomb. *The microstructure of securities markets.* Prentice-Hall, Englewood Cliffs, N.J. (1986).

Conrad, J., M.N. Gultekin, and G. Kaul. "Asymmetric predictability of conditional variances." *Review of Financial Studies*, 4 (1991), 597-622.

Conrad, J., A. Hameed, and C. Niden. "Volume and autocovariances in short-horizon individual security returns." *Journal of Finance*, 49 (1994), 1305-1329.

Conrad, J., and G. Kaul. "An anatomy of trading strategies." *Review of Financial Studies*, 11 (1998), 489-519.

Conrad, J., and G. Kaul. "Time-variation in expected returns." *Journal of Business*, 61 (1988), 409-425.

Conrad, J., and G. Kaul. "Mean-reversion in short-horizon expected returns." *Review of Financial Studies*, 2 (1989), 225-240.

Diebold, F. X., and R S Mariano. "Comparing predictive accuracy." *Journal of Business and Economic Statistics*, 13 (1995), 253-63.

Gallagher, L. "Interdependencies among the Irish, British and German stock markets." *The Economic and Social Review*, 26 (1995), 131-147.

Fama, E. F., and K. R. French. "The cross-section of expected stock returns." *Journal of Finance*, XLVII (1992), 427-465.

Gerrard, W. J., and L. G. Godfrey. "Diagnostic tests for single-equation error correction and autoregressive distributed lag models." *Manchester School*, 66 (1998), 222-237.

Granger, C. W. J. "Developments in the study of cointegrated economic variables." *Oxford Bulletin of Economics and Statistics*, 48 (1986), 213-228.

Hamilton, J. D. *Times Series Analysis*. Princeton University Press, Princeton (1994).

Jegadeesh, N., and S. Titman. "Returns to buying winners and selling losers: Implications for stock market efficiency." *Journal of Finance*, 48 (1993), 65-91.

Jegadeesh, N. "Evidence of predictable behaviour of security returns." *Journal of Finance*, 45 (1990), 881-898.

Jegadeesh, N., and S. Titman. "Overreaction, delayed reaction, and contrarian profits." *Review of Financial Studies*, 8 (1995), 973-993.

Kasa, K. "Common stochastic trends in international stock markets." *Journal of Monetary Economics*, 29 (1992), 95-124.

Levich, R., and L. Thomas. "The significance of technical trading-rule profits in the foreign exchange market: a bootstrap approach." *Journal of International Money and Finance*, 12 (1993).,451-474.

Lehmann, B. "Fads, martingales, and market efficiency." *Quarterly Journal of Economics*, 105 (1990), 1-28.

Lo. A., and C. MacKinlay. "Stock market prices do not follow random walks: evidence from a simple specification test." *Review of Financial Studies*, 1 (1988), 41-66.

Lo. A., and C. MacKinlay. "When are contrarian profits due to stock market overreaction?." *Review of Financial Studies*, 3 (1990), 175-206.

Merton, R. C. "A simple model of capital market equilibrium with incomplete information." *Journal of Finance*, 42 (1987), 483-510.

Moore M. J., and L. S. Copeland. "A comparison of Johansen and Phillips-Hansen cointegration tests of forward market efficiency: Baillie and Bollerslev revisited." *Economics Letters*, 47 (1995), 131-135.

Poterba, J., and L. Summers. "Mean reversion in stock prices: evidence and implications." *Journal of Financial Economics*, 22 (1988), 25-79.

Pesaran, M. H., and Y. Shin. "An autoregressive distributed lag modelling approach to cointegration analysis." in *Econometrics and Economic Theory in the 20th century*, The Ragnar Frinch Centennial Symposium, Chapter 11, S. Strom, ed. Cambridge: Cambridge University Press (1998).

Phillips, P.C.B., and B. Hansen. "Statistical inference in instrumental variables regression with I(1) processes." *Review of Economic Studies*, 57 (1990), 99-125.

Phillips, P. C. B., and S. Ouliaris. "Asymptotic properties of residual based tests for cointegration." *Econometrica*, 58 (1990), 165-193.

Stiglitz, J.E. "Using tax policy to curb speculative trading." *Journal of Financial Services*, 3 (1989), 101-115.

Table 1: KPSS stationarity tests

	NE portfolios		NV portfolios		VV portfolios	
	Prices	**Returns**	**Prices**	**Returns**	**Prices**	**Returns**
Portfolio 1	1.28 *	0.06	1.35 *	0.07	0.64 *	0.05
Portfolio 2	0.61 *	0.07	0.57 *	0.07	1.02 *	0.04
Portfolio 3	0.70 *	0.07	0.67 *	0.07	1.24 *	0.02
Portfolio 4	0.49 *	0.06	0.46 *	0.06	1.25 *	0.04
Portfolio 5	0.49 *	0.06	0.50 *	0.06	1.37 *	0.03
Portfolio 6	0.53 *	0.05	0.54 *	0.05	1.70 *	0.02
Portfolio 7	0.77 *	0.05	0.76 *	0.04	1.70 *	0.03
Portfolio 8	0.90 *	0.04	0.94 *	0.04	1.21 *	0.02
Portfolio 9	0.99 *	0.04	0.99 *	0.04	1.74 *	0.02
Portfolio 10	1.51 *	0.03	1.67 *	0.03	1.58 *	0.05

Notes:

1. The KPSS stationarity test statistic is defined as

$$n = T^{-2}\sum_{t=1}^{T}S^{2}_{t} \ / \ s^{2}(l)$$

where T is the sample size, S_t is the sum of the residuals when the series is regressed on an intercept and a time trend, and $s^2(l)$ is a consistent non-parametric estimate of the long-run variance of the error term. The null hypothesis is stationarity against the alternative of nonstationarity. We choose a number of 8 lags, l, in the estimation of the long-run variance of residuals, on the basis of the Kwiatkowski et al. (1992, page 174) criterion of choosing l at the value at which the test statistic settles down.

1. The 5% critical value of the KPSS test with a trend is 0.146, and is obtained from Kwiatkowski et al. (1992). If the test statistic is higher than 0.146, then reject the null of stationarity.

3. * indicates that the null of stationarity is rejected.

.

Table 2 . Phillips-Hansen tests for cointegration: 1955-1993

Cointegrating relation: $P_{i,t} = b_0 + b_1 P_{j,t-1}$, i,j = 1, …10, j>i.

Panel A: NE portfolios.

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5	Portfolio 6	Portfolio 7	Portfolio 8	Portfolio 9	Portfolio 10
Portfolio 1		-0.63	-4.09	-3.83	-5.40	-8.37	**-20.67 ***	**-21.15 ***	**-21.61 ***	**-24.20 ***
Portfolio 2			-9.50	-11.80	-13.29	-13.19	-14.95	**-21.20 ***	**-20.71 ***	-12.46
Portfolio 3				**-20.88 ***	**-20.99 ***	**-20.73 ***	**-21.88 ***	**-23.06 ***	**-25.24 ***	**-24.36 ***
Portfolio 4					**-33.57 ***	-16.84	-14.11	-13.24	-17.49	-8.77
Portfolio 5						-14.27	-11.51	-9.33	-13.81	-6.96
Portfolio 6							-12.11	-10.13	**-23.31 ***	-7.19
Portfolio 7								-8.22	-13.00	-6.49
Portfolio 8									**-20.60 ***	-7.84
Portfolio 9										-1.50
Portfolio 10										

Notes:

1. The Table reports the modified augmented Dickey-Fuller Z(a) test statistics calculated on the residuals of the corresponding cointegration regression estimated using the Phillips-Hansen fully modified ordinary least squares method.
2. The null hypothesis is of no cointegration. The 5% critical value is –20.4935 [Phillips and Ouliaris (1990), Table Ib].
3. * denotes that the null is rejected at the 5% level. Bolded test statistics indicate portfolio pairs for which there is cointegration.

Table 2 (continued)

Panel B: NV portfolios

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5	Portfolio 6	Portfolio 7	Portfolio 8	Portfolio 9	Portfolio 10
Portfolio 1		-0.43	-5.03	-4.24	-3.61	-7.77	**-20.87 ***	**-21.23 ***	**-26.30 ***	**-22.40 ***
Portfolio 2	——		-9.47	-12.01	-13.65	-12.89	-13.87	**-22.12 ***	**-21.51 ***	**-21.06 ***
Portfolio 3	——			**-20.61 ***	**-20.51 ***	**-20.53 ***	**-21.76 ***	**-22.87 ***	**-23.98 ***	-14.23
Portfolio 4	——				-13.23	-14.11	-14.01	-14.25	-16.81	-8.76
Portfolio 5	——					-14.29	-11.01	-9.78	-13.65	-6.90
Portfolio 6	——						-14.11	-10.25	-19.31	-7.96
Portfolio 7	——							-8.37	-13.31	-6.65
Portfolio 8	——								-20.00	-7.92
Portfolio 9	——									-1.30
Portfolio 10	——									

Notes:

1. See notes in Table 2, Panel A.

Table 2 (continued)

Panel C: VV Portfolios

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5	Portfolio 6	Portfolio 7	Portfolio 8	Portfolio 9	Portfolio 10
Portfolio 1		-2.73	-4.90	-3.99	-6.32	-8.07	-10.65	-9.98	-12.01	-14.30
Portfolio 2			-12.91	-10.76	-12.98	-11.98	-16.95	-11.54	-12.91	-15.91
Portfolio 3				-10.09	-14.54	-15.32	-12.14	-13.98	-16.98	-12.90
Portfolio 4					-12.45	-15.98	-12.21	-10.87	-17.91	-18.07
Portfolio 5						-11.23	-19.67	-9.65	-13.00	-6.41
Portfolio 6							**-22.11 ***	-10.55	-12.25	-7.91
Portfolio 7								-10.22	-11.64	-15.90
Portfolio 8									**-20.90 ***	-17.40
Portfolio 9										-7.40
Portfolio 10										

Notes:

1. See notes in Table 2, Panel A.

Table 3. Estimation of the long-run coefficients of the cointegrating relations using the ARDL approach:1955-1993

Panel A: NE Portfolios

Independent Variable in regression ____________ Dependent Variable In Regression	Port folio 1	Port folio 2	Port folio 3	Port folio 4	Port folio 5	Port folio 6	Port folio 7	Port folio 8	Port Folio 9	Port folio 10
Portfolio 1		____	____	____	____	____	**2.14 * (2.18) ARDL: (3,1)**	**1.97 * (4.24) ARDL: (3,1)**	**2.17 * (3.77) ARDL: (3,2)**	**1.80 * (5.33) ARDL: (3,2)**
Portfolio 2	____		____	____	____	____	____	**1.40* (4.20) ARDL: (3,1)**	**1.46 * (4.82) ARDL: (3,1)**	____
Portfolio 3	____			**1.28* (11.70) ARDL: (3,1)**	**1.45* (9.08) ARDL: (1,3)**	**1.30* (10.56) ARDL: (3,1)**	**1.27* (11.26) ARDL: (3,1)**	**1.35 * (8.70) ARDL: (3,1)**	**1.43 * (8.72) ARDL: (3,1)**	**1.45 * (3.25) ARDL: (3,2)**
Portfolio 4	____	____	**0.70 * (8.80) ARDL: (1,1)**		**1.04 * (42.35) ARDL: (2,1)**	____	____	____	____	____
Portfolio 5	____	____	**0.64 * (8.50) ARDL: (1,1)**	**0.93 * (40.5) ARDL: (1,1)**		____	____	____	____	____
Portfolio 6	____	____	**0.68 * (7.40) ARDL: (1,3)**	____	____		____	____	**1.27 * (5.37) ARDL: (3,1)**	____
Portfolio 7	0.15 (0.67) ARDL: (2,2)	____	0.67 (1.50) ARDL: (3,2)	____	____	____		____	**1.11 * (27.03) ARDL: (2,1)**	____
Portfolio 8	0.25 (1.75) ARDL: (2,2)	0.48 (1.82) ARDL: (2,2)	0.58 (1.81) ARDL: (3,2)	____	____	____	____		____	____
Portfolio 9	0.20 (1.26) ARDL: (1,2)	0.45 (1.90) ARDL: (2,2)	0.52 (1.32) ARDL: (1,3)	____	____	**0.55 * (2.90) ARDL: (1,3)**	____	**0.90 * (23.80) ARDL: (1,2)**		____
Portfolio 10	0.27 (1.72) ARDL: (2,1)	____	0.21 (0.60) ARDL: (2,2)	____	____	____	____	____	____	

Notes:

1. The reported coefficients are the long-run coefficients of the cointegration vectors for the pairs of portfolios for which cointegration is found in Table 2. In each cell, the order of the selected ARDL model is also reported. The selected model is based on the Schwartz Bayesian Criterion.
2. Asymptotic t-statistics are reported in parentheses underneath the corresponding long-run coefficient.
3. Statistical inference on the long run coefficients is based on standard normal asymptotic theory.
4. * denotes statistical significance at the 5% level of significance.

Table 3 (continued)

Panel B: NV portfolios

Independent Variable / Dependent Variable	**Port folio 1**	**Port folio 2**	**Port folio 3**	**Port Folio 4**	**Port Folio 5**	**Port folio 6**	**Port folio 7**	**Port folio 8**	**Port folio 9**	**Port folio 10**
Portfolio 1		——	——	——	——	——	**2.35 * (2.34) ARDL: (3,1)**	**1.97 * (5.05) ARDL: (3,1)**	**2.21 * (4.64) ARDL: (3,2)**	**1.84 * (5.82) ARDL: (3,2)**
Portfolio 2	——		——	——	——	——	——	**1.27 * (4.23) ARDL: (3,1)**	**1.46 * (4.92) ARDL: (3,1)**	**1.32 * (2.15) ARDL: (2,1)**
Portfolio 3	——	——		**1.28 * (11.82) ARDL: (3,1)**	**1.42 * (9.5) ARDL: (1,3)**	**1.26 * (10.72) ARDL: (3,1)**	**1.28 * (11.08) ARDL: (3,1)**	**1.35 * (8.27) ARDL: (3,1)**	**1.45 * (8.55) ARDL: (3,1)**	——
Portfolio 4	——	——	**0.71 * (9.11) ARDL: (1,1)**		——	——	——	——	——	——
Portfolio 5	——	——	**0.61 * (5.81) ARDL: (3,1)**	——		——	——	——	——	——
Portfolio 6	——	——	**0.69 * (7.31) ARDL: (3,1)**	——	——		——	——	——	——
Portfolio 7	0.18 (0.97) ARDL: (2,2)	——	**0.67 * (6.41) ARDL: (3,2)**	——	——	——		——	——	——
Portfolio 8	0.31 (1.89) ARDL: (1,2)	0.49 (1.89) ARDL: (2,2)	**0.57 * (4.31) ARDL: (3,2)**	——	——	——	——		——	——
Portfolio 9	0.25 (1.89) ARDL: (1,2)	0.43 (1.86) ARDL: (2,2)	0.50 (1.86) ARDL: (1,3)	——	——	——	——	——		——
Portfolio 10	0.28 (1.79) ARDL: (1,2)	-0.03 (-0.06) ARDL: (1,2)	——	——	——	——	——	——	——	

Notes:
See notes of Table 3, Panel A.

Table 4
Error Correction Models based on the ARDL approach: 1955-1993

Panel A: NE portfolios

Model			Cons tant	Trend	ECT_{t-1}	Coefficient of		Coefficient of	
Dependent Variable (Δy_t)	Independent Variable(Δx_t)	Model order				Δy_{t-1}	Δy_{t-2}	Δx_{t-1}	Δx_{t-2}
Portfolio 1	Portfolio 10	(2,1)	0.34 * (7.47)	0.11 (1.04)	-0.02 * (-3.28)	0.34 * (7.47)	0.12 * (2.54)	0.11* (3.55)	—
	Portfolio 9	(2,1)	-0.34 (-0.2)	-0.20 (-0.20)	-0.012 * (-2.46)	0.32 * (6.99)	0.12 * (3.05)	0.10* (3.20)	—
	Portfolio 8	(2,0)	0.24 (0.05)	0.27 (0.28)	-0.014 * (-2.66)	0.37 * (10.05)	0.09 * (2.53)	—	—
	Portfolio 7	(2,0)	0.002 (0.51)	0.57 (0.63)	-0.009 * (-1.96)	0.33 * (9.53)	0.097 * (2.79)	—	—
Portfolio 2	Portfolio 9	(2,0)	0.99 (0.25)	-0.71 (-0.9)	-0.016 * (-2.71)	0.33 * (10.95)	0.077 * (2.60)	—	—
	Portfolio 8	(2,0)	0.003 (0.80)	0.11 (0.17)	-0.013 * (-2.20)	0.27 * (9.88)	0.08 * (2.92)	—	—
	Portfolio 7	(2,0)	0.003 (0.85)	0.26 (0.41)	-0.013 * (-2.11)	0.22 * (8.68)	0.068 * (2.70)	—	—
Portfolio 3	Portfolio 10	(2,1)	0.005 (1.16)	-0.26 (-0.36)	-0.009 * (-1.96)	0.33 * (9.53)	0.097 * (2.79)	0.90* (3.10)	—
	Portfolio 9	(2,0)	0.99 (0.25)	-0.71 (-0.9)	-0.012* (-1.96)	0.22 * (4.97)	0.14 * (4.24)	0.08* (2.62)	—
	Portfolio 8	(2,0)	0.002 (0.50)	-0.38 (-0.72)	-0.026 * (-3.24)	0.18 * (7.77)	0.10 * (4.35)	—	—
	Portfolio 7	(2,0)	0.001 (0.46)	0.15 (0.03)	-0.031* (-3.55)	0.12 * (5.54)	0.084 * (3.85)	—	—
	Portfolio 6	(2,0)	-0.002 (-0.4)	-0.28 (-0.65)	-0.026 * (-3.12)	0.10 * (5.04)	0.71 * (3.74)	—	—
	Portfolio 5	(0,2)	-0.005 (-1.8)	-0.11* (-2.77)	-0.022* (-3.02)	—	—	0.05* (2.98)	0.05* (3.45)
	Portfolio 4	(2,0)	-0.003 (-1.5)	-0.11 (-2.64)	-0.026 * (-2.98)	0.03 (1.53)	0.54 * (3.26)	—	—
Portfolio 4	Portfolio 5	(1,0)	-0.003 (1.30)	0.001 (0.80)	-0.082* (-4.82)	0.04 * (3.03)	—	—	—
Portfolio 6	Portfolio 9	(2,0)	-0.002 (-0.5)	-0.83 (-1.92)	-0.015 * (-2.11)	0.17 * (10.5)	0.06 * (3.33)	—	—
Portfolio 8	Portfolio 9	(1,0)	-0.04* (-2.3)	-0.9* (-3.27)	-0.05 * (-3.90)	0.08* (8.16)	—	—	—

Table 4 (continued)

Panel B: NV Portfolios

Model: Dependent Variable (Δy_t)	Model: Independent Variable(Δx_t)	Model order	Constant	Trend	ECT_{t-1}	Coefficient of Δy_{t-1}	Coefficient of Δy_{t-2}	Coefficient of Δx_{t-1}	Coefficient of Δx_{t-2}
Portfolio 1	Portfolio 10	(2,1)	0.003 (0.60)	0.62 (0.66)	-0.02 * (-3.34)	0.33 * (7.27)	0.11 * (2.80)	0.10* (3.07)	—
	Portfolio 9	(2,1)	-0.04 (-0.9)	-0.73 (-0.82)	-0.015 * (-2.76)	0.30 * (6.51)	0.11 * (3.08)	0.09* (2.62)	—
	Portfolio 8	(2,0)	-0.01 (-0.4)	-0.18 (-0.21)	-0.015 * (-2.86)	0.32 * (9.45)	0.09 * (2.83)	—	—
	Portfolio 7	(2,0)	0.15 (0.03)	-0.33 (-0.43)	-0.008 * (-1.96)	0.29 * (8.83)	0.10 * (3.13)	—	—
Portfolio 2	Portfolio 9	(2,0)	0.49 (0.12)	-0.89 (-0.12)	-0.02 * (-2.78)	0.33 * (10.85)	0.079 * (2.63)	—	—
	Portfolio 8	(2,0)	0.003 (0.90)	0.13 (0.19)	-0.014 * (-2.20)	0.27 * (9.83)	0.081 * (2.72)	—	—
	Portfolio 7	(2,0)	0.002 (0.77)	0.28 (0.43)	-0.013 * (-2.13)	0.22 * (8.60)	0.068 * (2.69)	—	—
Portfolio 3	Portfolio 10	(2,1)	0.006 (1.40)	-0.12 (-0.16)	-0.009 * (-1.96)	0.23 * (5.17)	0.13 * (3.87)	0.094* (2.93)	—
	Portfolio 9	(2,0)	-0.01 (-0.3)	-0.09 (-1.54)	-0.028* (-3.71)	0.25 * (9.35)	0.095 * (3.98)	—	—
	Portfolio 8	(2,0)	0.002 (0.77)	-0.41 (-0.78)	-0.024 * (-3.13)	0.19 * (7.87)	0.10 * (4.44)	—	—
	Portfolio 7	(2,0)	0.001 (0.36)	-0.22 (-0.04)	-0.03* (-3.50)	0.12 * (5.50)	0.084 * (3.95)	—	—
	Portfolio 6	(2,0)	-0.67 (-0.3)	-0.23 (-0.55)	-0.025 * (-3.11)	0.10 * (5.03)	0.71 * (3.80)	—	—
	Portfolio 5	(0,2)	-0.004 (-1.8)	-0.11* (-2.76)	-0.024* (-3.06)	—	—	0.05* (3.02)	0.05* (3.55)
	Portfolio 4	(2,0)	-0.004 (-1.7)	-0.11 (-2.84)	-0.026 * (-3.05)	0.02 (1.43)	0.53 * (3.26)	—	—
Portfolio 4	Portfolio 5	(1,0)	0.18 (0.09)	0.91* (2.21)	-0.078* (-4.66)	0.04 * (3.21)	—	—	—
Portfolio 6	Portfolio 9	(2,0)	-0.002 (-0.7)	-0.92* (-2.05)	-0.016 * (-2.26)	0.18 * (10.1)	0.06 * (3.53)	—	—
Portfolio 8	Portfolio 9	(1,0)	-0.06* (-2.3)	-0.11* (-3.63)	-0.06 * (-4.33)	0.08* (8.04)	—	—	—

Table 5: Out-of-sample forecasting of 'small-firm' portfolio returns: January 1994 - December 1994

Panel A: NE Portfolios

Model		RMSE from model with the ECT	RMSE from model without the ECT	Wilcoxon's signed rank test of statistically significant difference of the two RMSEs. H_0: the two RMSEs are equal H_1: the two RMSEs are not equal
Dependent Variable	**Independent Variable**			
Portfolio 1	Portfolio 10	0.048	0.32	3.061 [0.00]
	Portfolio 9	0.041	0.65	4.612 [0.00]
	Portfolio 8	0.042	0.62	4.423 [0.00]
	Portfolio 7	0.040	0.65	4.324 [0.00]
Portfolio 2	Portfolio 9	0.020	0.14	3.123 [0.00]
	Portfolio 8	0.020	0.17	3.452 [0.00]
	Portfolio 7	0.017	0.13	3.062 [0.00]
Portfolio 3	Portfolio 10	0.031	0.32	4.001 [0.00]
	Portfolio 9	0.019	0.07	2.790 [0.00]
	Portfolio 8	0.018	0.10	2.589 [0.00]
	Portfolio 7	0.017	0.06	2.470 [0.00]
	Portfolio 6	0.016	0.011	1.001 [0.49]
	Portfolio 5	0.014	0.087	2.567 [0.00]
	Portfolio 4	0.014	0.100	2.698 [0.00]
Portfolio 4	Portfolio 5	0.013	0.021	1.021 [0.49]
Portfolio 6	Portfolio 9	0.011	0.080	2.542 [0.00]
Portfolio 8	Portfolio 9	0.008	0.034	1.000 [0.49]

Table 5 (continued)

Panel B: NV Portfolios

Model		RMSE from model with the ECT	RMSE from model without the ECT	Wilcoxon's signed rank test of statistically significant difference of the two RMSEs. H_0: the two RMSEs are equal H_1: the two RMSEs are not equal
Dependent Variable	Independent Variable			
Portfolio 1	Portfolio 10	0.050	0.321	2.864 [0.00]
	Portfolio 9	0.042	0.672	4.424 [0.00]
	Portfolio 8	0.042	0.625	4.123 [0.00]
	Portfolio 7	0.040	0.675	4.465 [0.00]
Portfolio 2	Portfolio 10	0.033	0.371	3.523 [0.00]
	Portfolio 9	0.021	0.142	2.652 [0.00]
	Portfolio 8	0.020	0.181	2.426 [0.00]
Portfolio 3	Portfolio 10	0.032	0.331	3.501 [0.00]
	Portfolio 9	0.019	0.094	2.790 [0.00]
	Portfolio 8	0.018	0.129	2.209 [0.02]
	Portfolio 7	0.018	0.081	2.170 [0.03]
	Portfolio 6	0.016	0.018	0.801 [0.45]
	Portfolio 5	0.013	0.077	1.567 [0.10]
	Portfolio 4	0.014	0.086	1.998 [0.04]