Disappearing Dividends:

Implications on the Dividend-Price ratio and the Predictive Regressions

Chang-Jin Kim

University of Washington and Korea University

and

Cheolbeom Park[1]

Korea University

August, 2010

Abstract

The conventional dividend-price ratio is highly persistent, and the literature reports mixed evidence on its role in predicting stock returns. In particular, its predictive power seems to be sensitive to the choice of the sample period. We argue that the decreasing number of firms with traditional dividend-payout policy is responsible for these results, and develop a model in which the long-run relationship between the dividends and stock price is time-varying. An adjusted dividend-price ratio is stationary with considerably less persistence than the conventional dividend-price ratio. The adjusted dividend-price ratio is also shown to have a stable and statistically significant in-sample predictive power on stock returns, regardless of the firm sizes. For small-sized firms, the predictive regression model that employs the adjusted dividend-price ratio as a regressor beats the random-walk model in terms of the out-sample predictability.

Key Words: Stock Return Predictability, Adjusted Dividend-price ratio, Disappearing Dividends, Time-Varying Cointegration Vector,

JEL classification: G12, C12, C22

[1] **Kim:** Department of Economics, Korea University, Anam-dong, Seongbuk-gu, Seoul, Korea 136-701 (TEL: +82-2-3290-2215. E-mail: cjkim@korea.ac.kr).

Park (corresponding author): Department of Economics, Korea University, Anam-dong, Seongbuk-gu, Seoul, Korea 136-701. (TEL: +82-2-3290-2203. E-mail: cbpark_kjs@korea.ac.kr).

1. Introduction

Since Campbell and Shiller (1988), a large body of research provides evidence that the dividend price ratio predicts future stock returns.[2] However, recent empirical studies report evidence of structural breaks or instability in the return forecasting models. For example, Goyal and Welch (2003, 2008) and Bossaerts and Hillion (1999) suggest that the return forecasting models seem unstable, as diagnosed by their poor out-of-sample predictions in the presence of strong in-sample prediction. Other researchers like Rapach and Wohar (2006) and Paye and Timmerman (2006) provide direct evidence of structural breaks in the predictive regressions by employing rigorous tests. All of these researchers document deterioration in the predictability patterns in the US stock returns in the second half of the 1990s.

Fama and French (2001) and Allen and Michaely (2003) document that changes in the dividend payout policies has mad firms less likely to pay dividends, resulting in a decreasing dividend-price ratio. For example, the proportion of firms paying cash dividends falls from 66.5% in 1978 to 20.8% in 1999, as documented by Fama and French (2001). Recently, Lettau and van Nieuwerburgh (2007), argue that a failure to take into account the shifts in the mean of the dividend-price ratio is responsible for the instability in the predictive regressions. By appropriately accounting for the mean shifts in the dividend-price ratio in the predictive regressions, they were able to reconcile the inconsistent results for the in-sample and out-of-sample predictions. In the meantime, Boudoukh et al (2007) and Robertson and Wright (2006) note that, while the dividend price ratio changed remarkably during the 1980's and 1990's, the total payout ratio defined as dividends plus share repurchases over price has remained relatively stable. By using the total payout ratio in the predictive regressions, they were able to recover statistically and economically significant predictability at both short and long horizons. According to them, the seemingly unstable predictive regressions and the recent decline in the predictive power may be due to mis-measurement in the regressor of the predictive regression.

In this paper, we investigate the implications of changes in the firms' payout policy on the long-run relationship between the dividend and the stock price and on the stock return

[2] The claim that stock returns are predictable is different from the argument that investment into the stock market can provide extra profits. Predictable stock returns can be viewed as an equilibrium phenomenon. For example, Campbell and Cochrane (1999) show that the time-varying risk premium can generate predictable stock returns. Cecchetti, Lam and Mark (2000) argue that stock return predictability may be attributable to distorted beliefs. Bansal and Yanron (2004) claim that stock return predictability arises from a small long-run predictable component and fluctuating uncertainty contained in consumption and dividend growth rates.

predictability. As documented in Fama and French (2001) and Allen and Michaely (2003), we note that there has been changes in the payout policy by the firms in the form of i) many existing firms having stopped or reduced dividend payment; ii) increasing repurchase shares or retained earnings; iii) many newly listed firms usually not paying out dividends since early 1980s. By defining the fraction of firms that follow the traditional payout policy as

$$\omega_t = 1 - \overline{\omega}_{1t} - \overline{\omega}_{2t} \tag{1}$$

where $\overline{\omega}_{1t}$ is the fraction of firms with share repurchases $\overline{\omega}_{2t}$ is the fraction of firms that neither pay dividends nor repurchase shares, we provide a theoretical model in which the evolving payout policy or evolving ω_t results in a time-varying long-run relationship between the dividend and the stock price. We show both theoretically and empirically that the parameter describing the long-run relationship is dependent upon evolving ω_t. Furthermore, we show that the adjusted dividend-price ratio that accounts for a time-varying long-run relationship has predictive power for stock returns both in-sample and out-of-sample. These results suggest that the conventional dividend-price ratio that assumes a stable one-to-one long-run relationship between dividends and stock prices vastly understates the predictive power of the dividend-price ratio for the stock returns.

The approach in this paper and those in Boudoukh *et al*. (2007) and Robertson and Wright (2006) may be equivalent to two sides of the same coin. They employ a measure of total payout (dividends plus repurchases) that is expected to have a stable one-to-one long-run relationship with stock price in the presence of changing payout policy, and construct the total payout ratio defined as dividends plus repurchases over price. We consider a time-varying long-run relationship between the dividends and the stock price and construct the resulting adjusted dividend-price ratio. Our approach is definitely complementary to theirs. However, we hope to overcome potential weaknesses in their approach. As demonstrated by Fama and French (2001), a substantial portion of share repurchase by firms is done in consideration of employee stock ownership plans or mergers, instead of dividend payment replacement. Besides, especially since the 1980s, there have been many firms that neither paid out dividends nor repurchased shares. Under these situations, Boudoukh *et al*. (2007) and Robertson and Wright's (2006) approach may suffer from the problem of measurement error in constructing the total payout ratio.

The rest of this paper is organized as follows. Section 2 discusses the consequence of the gradual change in payout policy by firms. When firms are gradually switching dividend payment to share repurchasing or reducing dividend payment to finance investment opportunities, the conventional dividend-price ratio may become nonstationary, and the cointegration relationship between dividends and stock price varies over time. Section 3 provides empirical evidence on the time-varying cointegration relationship between dividends and stock price, confirming the argument in Section 2. Section 4 evaluates the predictive power of the adjusted dividend-price ratio that accounts for a time-varying cointegrating vector. Concluding remarks are given in Section 5. Appendix contains proofs and a brief presentation about the econometric methodology to estimate the time-varying cointegration parameters between dividends and stock price.

2. Disappearing Dividends and the Time-Varying Long-Run Relation between Dividends and Price: A Theoretical Model

2.1. Model Specification

In this section, we present a simple present-value model of stock returns, from which we can address the issues resulting from the changes in the payout policy by the firms. We assume that there exist two types of firms: i) firms with traditional dividend payout policy (type-I firms) and ii) firms that do not pay dividends or those with reduced dividend payment in an attempt to increase repurchase shares or retained earnings (type-II firms). The traditional payout policy in the present study means that a significant portion of earnings is paid out in the form of dividends. At the end of time t, the total amount of dividends paid by type-I firms is $D_{1,t}$ and that paid by type-II firms is $D_{2,t}$. $\Lambda_{2,t}$ is the hypothetical amount of dividend payment that type-II firms would pay if they adopted the traditional payout policy. The inequality $\Lambda_{2,t} \geq D_{2,t}$ holds, as $\Lambda_{2,t}$ includes share repurchases, a part of retained earnings that has replaced dividends under the traditional payout policy, as well as actual dividend payments by type-II firms. Then, by assuming the Miller-Modigliani theorem holds, the fraction of type I firms (ω_t) in terms of the market value of common equities during the period between $t-1$ and t can be specified by:[3]

[3] The theorem suggests that, as long as investment policy doesn't change for type II firms, altering the mix of retained earnings and payout will not affect firm's value and stock price.

$$\omega_t = \frac{D_{1,t}}{D_{1,t}+\Lambda_{2,t}} \quad (2)$$

The actual dividends ($D_{1,t}$ and $D_{2,t}$) paid out by two types of firms, along with the hypothetical dividends ($\Lambda_{2,t}$) of type II firms under traditional policy, depend upon economic conditions. By letting $d_{i,t} = ln\,(D_{i,t})$, $i = 1,2$ and $\lambda_{i,t} = ln\,(\Lambda_{i,t})$, we assume:

$$d_{i,t} = \gamma_i m_t + \varepsilon_{i,t}, \qquad i = 1,2 \quad (3)$$

$$\lambda_{2,t} = \gamma_3 m_t + \varepsilon_{3,t} \quad (4)$$

where m_t is a measure of economic conditions (e.g., the aggregate output) and $E(\varepsilon_{i,t}) = 0$, $i = 1,2,3$. As documented in Fama and French (2001), the two types of firms considered have different characteristics in terms of firm-size, investment opportunity, growth, etc. These differences in the characteristics for the two types of firms result in different values of γ_1 and γ_3 even under the hypothetical situation in which both types firms adopt the traditional payout policy. In addition, differences in the γ_2 and γ_3 values result from the change in the payout policy by Type-II firms.

Finally, by defining $P_{i,t}$ and $p_{i,t}$, $i = 1,2$, as the stock prices and the logs of stock prices, respectively, we assume that the following present value relation holds, as derived by Campbell and Shiller (1988):[4]

$$d_{1,t} - p_{1,t} = -\frac{k_1}{1-\rho_1} + E_t\left[\sum_{j=0}^{\infty}(\rho_1)^j(r_{1,t+1+j} - \Delta d_{1,t+1+j})\right] \quad (5)$$

$$\lambda_{2,t} - p_{2,t} = -\frac{k_2}{1-\rho_2} + E_t\left[\sum_{j=0}^{\infty}(\rho_2)^j(r_{2,t+1+j} - \Delta \lambda_{2,t+1+j})\right] \quad (6)$$

where, by defining $r_{1,t} = ln\big(P_{1,t} + D_{1,t}\big) - ln\,(P_{1,t})$ is the return for type I firms at t ; $r_{2,t} = ln\big(P_{2,t} + \Lambda_{2,t}\big) - ln\,(P_{2,t})$ is the return for type-II firms at *t;* $\rho_1 = \frac{P_1}{P_1+D_1}$; $\rho_2 = \frac{P_2}{P_2+\Lambda_2}$; k_1 and k_2 are unimportant constants that result from the linearization process. The

---

4 Note that the log present value relation cannot be defined for type II firms, especially for those which are not paying dividends, without considering the hypothetical dividends ($\Lambda_{2,t}$) under traditional payout policy.

above present value model implies that both the dividend price ratio for type I firms $(d_{1,t} - p_{1,t})$ and the hypothetical dividend price ratio for type II firms $(\lambda_{2,t} - p_{2,t})$ are stationary, given that each of $d_{1,t}$, $p_{i,t}$, $i = 1,2$ and $\lambda_{2,t}$ has a unit root. In other words, there exists a long-run or cointegrating relationship between the log of dividend and the log of stock price for each of type-I and type-II firms when both types of firms were adopting the traditional payout policy.

2.2. Implications of Disappearing Dividends: Propositions Derived from the Model

From the model specified above, we derive three propositions describing the implications of the changing payout policy by some of the firms on the aggregate dividend price ratio and the stock return predictability. Proofs are given in the Appendix.

<u>Proposition 1</u>:

The conventional aggregate dividend-price ratio has a unit root ($d_t - p_t$~I(1)) unless all firms in the market are type- I with ω_t=1.

Proposition 1 results from the following representation of the aggregate dividend-price ratio obtained from the model:

$$d_t - p_t = \delta_t - p_t + (\widehat{\omega}_t - \omega_t)d_{1,t} + (1 - \widehat{\omega}_t)d_{2,t} - (1 - \omega_t)\lambda_{2,t} \qquad (7)$$

where p_t and d_t are the actual aggregate stock price and the actual aggregate dividends observed in the market; $\delta_t = \omega_t d_{1,t} + (1 - \omega_t)\lambda_{2,t}$ is the hypothetical aggregate dividends if type-II firms followed the traditional payout policy or if all firms in the market are type I; ω_t, defined in (1), is the fraction of type-I firms in terms of the market value of common equities; and $\widehat{\omega}_t = \frac{D_{1,t}}{D_{1,t}+D_{2,t}}$ is the fraction of the actual dividends paid out by type-I firms in the market. Note that, when all the firms in the market are type I with $\omega_t = \widehat{\omega}_t = 1$, the last three unit root terms equation (7) will disappear and we have a stationary dividend-price ratio with $d_t - p_t = \delta_t - p_t$. With type II firms in the economy, the last three unit root terms in equation (7) would remain, and the persistence of the aggregate dividend price ratio $d_t - p_t$ would depend on ω_t.

Proposition 2:

The aggregate dividends and stock price are cointegrated with a time-varying long-run relationship of the form:

$$p_t = \alpha_t d_t + u_t, \text{ and } u_t \sim \mathrm{I}(0) \tag{8}$$

where $\alpha_t = \frac{(\omega_t \gamma_1 + (1-\omega_t)\gamma_3)}{(\widehat{\omega}_t \gamma_1 + (1-\widehat{\omega}_t)\gamma_2)}$ and u_t is a function of $\varepsilon_{i,t}$, $i = 1,2,3,$ and stationary.

If all the firms in the economy were type I ($\omega_t = \widehat{\omega}_t = 1$), we would have $\alpha_t = 1$, and equation (8) suggests the conventional dividend-price ratio ($d_t - p_t$) would be stationary as in Proposition 1. Otherwise, the parameter (α_t) describing the long-run relationship between the price and the dividend is a function of ω_t, the fraction of type-I firms in terms of the market value of common equities.

Proposition 3:

An adjusted dividend-price ratio, which takes into account the time-varying long-run relationship between dividends and stock price, is a function of future expected returns :

$$\alpha_t d_t - p_t = -\frac{k}{1-\rho} + E_t\left[\sum_{j=0}^{\infty} \rho^j (r_{t+1+j} - \omega_{t+1+j}\Delta d_{1,t+1+j})\right]$$

$$+E_t\left[\sum_{j=0}^{\infty}\rho^j\left(1-\bar{\omega}_{t+j}\right)[(1-\rho_2)(\lambda_{2,t+1+j}-p_{2,t+1+j})+\Delta p_{2,t+1+j}-r_{2,t+1+j}]\right]$$

$$-E_t\left[\sum_{j=0}^{\infty}\rho^j(1-\omega_{t+1+j})(\Delta\lambda_{2,t+1+j})\right] \qquad (9)$$

where r_t is the aggregate stock returns, $\rho=\frac{P_1+P_2}{(P_1+D_1)+(P_2+\Lambda_2)}$, k is an unimportant linearization byproduct constant, and $\bar{\omega}_{t+j}$ is the average of ω_{t+1+j} and ω_{t+j}. Note that all variables except stock returns (r_t) in the right-hand side of equation (9) are unobservable.

2.3. Discussion

Propositions 1 and 2 confirm Boudoukh *et al*. (2007) and Robertson and Wright's (2006) argument, which suggests that the change in firms' payout policy from dividend payments to share repurchases has resulted in a breakdown in the one-to-one long-run relationship between the aggregate dividends and stock price. They consider a measure of total payout (dividends plus repurchases) that is expected to have a stable one-to-one long-run relationship with stock price in the presence of changing payout policy. However, in the presence of type-II firms which neither pay out dividends nor repurchase shares, even the total payout may not have a one-to-one long-run relationship with stock price. In fact, a considerable portion of firms in the portfolio formed on small-sized firms have neither paid out dividends nor repurchased shares especially since the 1980s. Besides, Fama and French (2001) demonstrate that firms may not just replace dividend payments with share repurchases. A substantial portion of share repurchase has taken place in consideration of employee stock ownership plans or mergers. These are the reasons why we explicitly and directly consider a time-varying long-run relationship between dividends and stock price in this paper, unlike the approaches in Boudoukh *et al*. (2007) and Robertson and Wright (2006), who consider a measure of total payout (dividends plus repurchases) that is expected to have a stable one-to-one long-run relationship with stock price.

Proposition 3 establishes the possibility that the adjusted dividend-price ratio ($\alpha_t d_t - p_t$), a stationary deviation from the time-varying long-run or cointegrating relationship, has predictive power for the aggregate stock return ($r_{t+1+j},\ j = 1,2,3, \ldots$). In the presence of type-II firms which do not pay out dividends, the adjusted dividend-price ratio ($\alpha_t d_t - p_t$), and not the conventional dividend price ratio ($d_t - p_t$), may be employed as a regressor in the predictive regression models. In case the fraction of type-II firms in the market changes over time, so does the α_t parameter. Thus, estimation of the α_t parameter is an important empirical issue in this paper.

The propositions in the previous section also have an important implication on the recent findings of the literature, which report sensitivity of stock return predictability to the choice of the sample period. Suppose that ω_t, the fraction of type-I firms (firms that maintain the traditional payout policy) in the market, is given by:

$$\mathrm{\omega_t = (1 - S_t)\omega_0 + S_t\omega_1}$$

$$S_t = 1, \text{if } t \leq \tau; \text{ and } S_t = 0, \text{if } t > \tau$$

where τ is the structural break point. Proposition I suggests that, when $\omega_0 \approx 1$ and $\omega_1 < 1$, the dividend-price ratio ($d_t - p_t$) is stationary before the structural break and it has a unit root after the structural break. Thus, in a predictive regression with the dividend-price ratio as a regressor, the coefficient estimator would always converge in probability to zero for the post-break sample, regardless of the true predictive nature of the dividend price ratio. This is because of the possibility that the regressor (the dividend price ratio) has a unit root in the post-break sample, while the dependent variable (stock return) is always stationary. In fact, Paye and Timmermann (2006) and Rapach and Wohar (2006) document structural breaks in the return prediction models, suggesting that the predictive power of the dividend price ratio has disappeared since the 1990s. In the meantime, Fama and French (2001) and Allen and Michaely (2003) document that a measure of ω_t has declined substantially below 1 since the 1990s.

3. Time-Varying Persistence of Dividend-Price Ratio Long-Run Relationship between Dividends and Stock Price: Empirical Results I

3.1. Data Description

Following Torous, Valkanov, and Yan (2004), the aggregate stock price index is constructed from monthly returns on the CRSP value-weighted market portfolio without dividends, and the aggregate dividend series is constructed from monthly returns on the CRSP value-weighted market portfolio with and without dividends.[5] The aggregate real stock return is constructed by subtracting the CPI inflation rate from the log returns of the CRSP value-weighted market portfolio. Data for portfolios formed on different firm sizes are obtained from Kenneth French's homepage.[6] These data are constructed from CRSP database by French, and hence consistent with our aggregate stock market data. Using the data for monthly returns with and without dividends on portfolios, the stock returns, dividends, stock price indices for portfolios formed on different firm sizes are constructed in the same way as for the aggregate variables. The sample period is 1946.1-2008.12.

In constructing the data on the fraction (ω_t) of firms that follow the traditional payout policy (type-I firms) in equation (1), we calculate the fraction of firms with share repurchases ($\overline{\omega}_{1t}$) and the fraction of firms that neither pay dividends nor repurchase shares ($\overline{\omega}_{2t}$) in terms of the market values of common equities. The data on ω_t is constructed at annual frequency, based on the year-end market values of the corresponding firms in the CRSP. The monthly data on ω_t is then constructed by interpolating the annual data.

Figure 1 depicts the annual measures of ω_t for all the firms in the (CRSP) value-weighted market portfolio as well as those for large-sized, medium-sized, and small-sized

---

[5] The CRSP data were obtained from http://wrds.wharton.upenn.edu.
[6] The web address for French's homepage is http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

firms.[7] All measures of ω_t fluctuate around 0.9 between 1946 and the early 1980s, and they decline sharply since the early 1980s. Note that measure ω_t for small-sized firms is lower than those for large- or medium-sized firms most of the period. Besides, it has higher volatility than the other two.[8]

3.2. Time-Varying Persistence of Dividend-Price Ratio

We examine the implications of Proposition 1, which states that the log of dividend price ratio ($d_t - p_t$) contains a unit root component, the relative size of which depends upon the fraction of type-I firms or firms with traditional payout policy (ω_t). When ω_t is one, there exists no unit root component in $d_t - p_t$. However, as ω_t decreases below 1, the unit root component in $d_t - p_t$ becomes more pronounced, making $d_t - p_t$ a more persistent process. This implies that it will be more difficult to reject the null hypothesis of a unit root for the $d_t - p_t$ process, as ω_t decreases. In order to examine this implication, we run the following Augmented Dickey-Fuller (ADF) regression equation recursively:

$$d_t - p_t = \alpha_j + \tau_j(d_{t-1} - p_{t-1}) + \varsigma_{1,j}\Delta(d_{t-1} - p_{t-1}) + \cdots + \varsigma_{p,j}\Delta\big(d_{t-p+1} - p_{t-p+1}\big) + \varepsilon_t ,$$
$$t = 1,2,\ldots,T^* + j$$
$$j = 0,1,2,\ldots,J \qquad (10)$$

The first regression is run using a 30-year data set that starts in January 1946 (j=0), and the subsequent regressions are run recursively by adding one monthly observation at a time (j=1,2,…,J). Using the estimates of τ_j and the ADF t-statistic for a unit root test obtained from the above recursive regressions, we estimate the following regression equations:

$$\tau_j = c_{\tau,0} + c_{\tau,1}\omega_j + e_{\tau,j} \qquad (11)$$
$$\mathrm{adf}_j = c_{\mathrm{adf},0} + c_{\mathrm{adf},1}\omega_j + e_{\mathrm{adf},j} \qquad (12)$$
$$j = 0,1,2,\ldots,J,$$

where adf_j is the ADF t-statistic.

Table 1 reports the results. $\bar{R}^2$'s from these regressions are 0.51 and 0.31 for the aggregate data, and the estimates of the $c_{\tau,1}$ and $c_{\mathrm{adf},1}$ coefficients are negative and

---

7 The cut-off values of market values for each size (large, medium, and small) are taken from Kenneth French's homepage, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

8 We note that the dynamics of ω_t in terms of market value (Figure 1) is similar to that in terms of the number of firms (depicted in Figure 2 of Fama and Frech (2001)).

statistically significant. The results are robust to the choice of portfolios formed on different firm sizes. Since the ADF test statistics are usually negative, the negative estimates of the $c_{\tau,1}$ and $c_{adf,1}$ coefficients indicate that, as ω_t decreases, τ_j increases toward one, the ADF statistic gets closer to zero and it is more difficult to reject the unit root null hypothesis.

3.3. Time-Varying Long-Run Relationship between Dividends and Stock Price

Proposition 2 states that, even though d_t and p_t move together in the long run, their long-run relationship is not one-to-one (or the cointegrating vector is not fixed at [1, -1]) with disappearing dividends. Rather, the long-run relationship between d_t and p_t is time-varying and dependent upon the fraction of type-I firms in the market (ω_t).

We first test the null hypothesis of constant cointegrating vector, by employing the test procedures proposed by Park and Hahn (1999). We employ two types of test statistics proposed by Park and Hahn (1999): $\tau_1^* = \frac{\sum_{t=1}^{T}(d_t - p_t)^2 - \sum_{t=1}^{T}\hat{s}_t^2}{\hat{\vartheta}_{T\kappa}^2}$ and $\tau_2^* = \frac{\sum_{t=1}^{T}(\sum_{i=1}^{t} d_i - p_i)^2}{T^2\hat{\vartheta}_{T\kappa}^2}$, where $\hat{s}_t$ are the residuals of the regression of $d_t - p_t$ on superfluous regressors such as a constant, $t,$ and t^2, and $\hat{\vartheta}_{T\kappa}^2$ is a long-run variance estimator for where $\hat{s}_t$. Based on the asymptotic distributions derived by Park and Hahn (1999), the 5% critical values for the τ_1^* and τ_2^* statistics under the null hypothesis are given by 7.82 and 0.16, respectively. As shown in Table 2, the null hypothesis of constant long run relationship is rejected regardless of the test statistics employed and regardless of the firm sizes. .[9]

Given the convincing evidence of time-varying cointegration relationship between d_t and p_t, we estimate the time-varying cointegration parameters based on the Fourier Flexible Form (FFF), as proposed by Park and Hahn (1999). For detailed discussion of Park and Hahn's (1999) estimation procedure, readers are referred to the Appendix. Figure 2 depicts the estimates of time-varying coefficients from a regression of p_t on d_t, along with their 95% confidence bands. As shown in Figure 2, the coefficients on d_t or the cointegrating vector gradually declined, with a few large swings during the post World War II period for the aggregate stock market data. We note that dynamics of the cointegrating vector is closely related to that of the fraction of type-I firms in the market, as depicted in

[9] We also consider the likelihood ratio test for the time-varying long-run relationship between d_t and p_t, as proposed by Bierens and Martins (2009). The null hypothesis is rejected at a 1% significance level for all cases.

Figure 1.

In order to see the degree of association between the time-varying cointegrating vectors and the fraction of type-I firms in the market (ω_t), we regress the estimates of time-varying cointegrating vectores on a constant and the monthly data on the fraction of type-I firms. Table 3 reports the estimation results. The coefficients on ω_t are highly significant with $\bar{R}^2$ ranging between as high as 0.69 for the aggregate stock market, suggesting that the long-run relationship between d_t and p_t is time-varying and dependent upon the fraction of type-I firms in the market (ω_t). Again, the results are robust to the choice of portfolios formed on different firm sizes and to the number of series functions in the FFF representation.[10]

4. Adjusted Dividend-Price Ratio and Predictive Regressions: Empirical Results II

4.1. In-Sample Analysis of Return Predictability

In this section, we empirically test the implications of the Proposition 3, which suggests that the adjusted dividend-price ratio ($\alpha_t d_t - p_t = \delta_t - p_t$), a stationary deviation from the time-varying long-run or cointegrating relationship, may have predictive power on the stock returns.

We first compare the time-series properties of the conventional dividend price ratio ($d_t - p_t$) and adjusted dividend-price ratio ($\delta_t - p_t$). Figure 3 plots the $\delta_t - p_t$ and the $d_t - p_t$ series. Summary statistics for these series are provided in Table 4. As in the literature, the conventional dividend-price ratios ($d_t - p_t$) are highly persistent with autocorrelations being close to one. The null hypothesis of a unit root cannot be rejected. In the mean time, the adjusted dividend-price ratios ($\delta_t - p_t$) reveal much less persistence than $d_t - p_t$. Furthermore, the null hypothesis of a unit root is rejected regardless of the firm sizes.

The predictive regression we employ is given by:

$$r_{t+1} = \beta_0 + \beta_1 x_t + \varepsilon_{t+1}, \tag{13}$$

[10] The number of series functions is denoted by κ in Table 3.

$$x_{t+1} = \eta_0 + \eta_1 x_t + e_{t+1} \quad (14)$$

where r_{t+1} is the stock return at time t+1; x_t is either the conventional dividend-price ($\mathrm{d}_t - p_t$) ratio or the adjusted dividend-price ratio ($\delta_t - p_t$); and ε_{t+1} is correlated with e_{t+1}. The estimates of the β_1 coefficient reported in Table 5 are based on OLS regressions of equation (13). As noted by Stambaugh (1999) and Ferson et. al (2003), the persistent nature of x_t in equation (13) and the strong correlation between ε_{t+1} and e_{t+1}, especially when $x_t = d_t - p_t$, may cause distortions in the asymptotic distribution of the estimator for β_1. Hence, we compute the p-values for testing the null hypothesis of $\beta_1 = 0$, based on the bootstrapped distribution of the β_1 estimates. More specifically, we generate pseudo r_t and pseudo x_t under the specification represented by equations (13) and (14) and $\beta_1 = 0$. Then, we regress the pseudo r_t on the pseudo x_t to construct the bootstrapped distribution of the β_1 estimates. For the predictive regression with multi-period stock returns as the dependent variable, multi-period pseudo stock returns are constructed by the sum of one-period pseudo stock returns.

The results are reported in Table 5. While the conventional dividend-price ratio ($\mathrm{d}_t - p_t$) provides little evidence of in-sample return predictability, the adjusted dividend-price ratio ($\delta_t - p_t$) provides evidence of in-sample stock return predictability.

Paye and Timmermann (2006) and Rapach and Wohar (2006) report evidence of structural breaks in the predictive regressions that employ the conventional dividend-price ratio ($d_t - p_t$). In particular, they report that the in-sample predictability of the dividend-price ratio has disappeared since the 1990s. We conjecture that the structural breaks they find may be due to a failure to consider the time-varying long-run relationship between the dividend and the stock price. We thus check whether there exist structural breaks in the predictive regression using the conventional dividend price ratio ($d_t - p_t$) or the adjusted dividend-price ratio ($\delta_t - p_t$) as a regressor. Empirical results on the Bai and Perron test of structural break are reported in Table 6. There exists clear evidence of unstable β_1 coefficient when the conventional dividend price ratio ($d_t - p_t$) is used as a regressor. However, we cannot reject the null hypothesis of no structural break when the adjusted dividend price ratio ($\delta_t - p_t$) is used as a regressor. That is, by accounting for the time-varying long-run relation relationship between the dividend and the stock price, we confirm that the adjusted dividend-price ratio has a stable and statistically significant in-sample predictive power on stock returns.

4.2. Out-Sample Analysis of Return Predictability

For out-sample analysis of return predictability, we first estimate the time-varying cointegrating vector and calculate the resulting adjusted dividend price ratio using data up to time t_0. Second, given the estimates of the adjusted dividend-price ratio $\delta_{t-1} - p_{t-1}$ and data on r_t, for t=2,…,t_0, the predictive regression is run ($r_t = \beta_0 + \beta_1 (\delta_{t-1} - p_{t-1}) + \varepsilon_{t+j-1}$) to get the estimates of β_0 and β_1. Third, using the adjusted dividend-price ratio $\delta_{t_0} - p_{t_0}$, $\hat{\beta}_0$, and $\hat{\beta}_1$, a one-month-ahead out-of-sample return prediction is formed by $E\left(r_{t_0+1} \middle| I_{t_0}\right) = \hat{\beta}_0 + \hat{\beta}_1 (\delta_{t_0} - p_{t_0})$, where I_{t_0} refers to information available up to time t_0. This procedure is repeated recursively. The first predictive regression is run using 30-year's monthly data that starts in January 1946.[11]

In this section, we compare the out-sample performance of the predictive regressions against a random walk null model. The Diebold-Mariano (1995) test is employed for this purpose. The null hypothesis of the Diebold-Mariano test is that both forecasting models have equal predictive power. Table 7 reports the test results. We have constructed the Diebold-Mariano test statistic so that it has negative sign if the first model has inferior predictive power to the second model. We have used absolute forecast error as the loss function in the Diebold-Mariano test statistic to attenuate effects from outliers. When the conventional dividend-price ratio ($d_t - p_t$) is employed in the predictive regression, the Diebold-Marino test statistic is positive and statistically significant, suggesting that a random walk model is superior, as in Goyal and Welch (2008). However, when the adjusted dividend-rice ratio ($\delta_t - p_t$) is used as a regressor in the predictive regressions, the Diebold-Mariano statistics are all negative except the portfolio formed on large size firms. In particular, they are negative and statistically significant for the portfolio formed on small-sized firms. That is, at least for small-sized firms, the model that employs $\delta_t - p_t$ as a predictor beats the random walk model, in terms of the out-of-sample predictability.

[11] For the out-of-sample analysis, the dependent variable is one-month-ahead stock return because of the following two reasons. First, the adjusted dividend-price ratio appears to have the strongest predictive power at one-month horizon in Table 7. Second, we want to avoid econometric problems in the out-of-sample analysis arising from the overlapping observations in the dependent variable for multi-period-ahead stock returns.

5. Summary and Conclusion

We present both the theoretical and empirical frameworks for analyzing the implications of changing dividend payout policy by the firms on the long-run relationship between the dividend and the stock price. Both our theory and empirical results demonstrate that the parameter describing the long-run relationship is time-varying and dependent upon the fraction of firms with traditional payout policy. This time-varying long-run relationship results in highly persistent dynamics in the conventional dividend-price ratio. Furthermore, it explains why the predictive power of dividend-price ratio on stock returns is sensitive to the choice of sample periods, as documented in the literature.

The adjusted dividend-price ratio, which takes into account the time-varying nature of the long-run relationship between dividend and stock price, is stationary with much less persistent than the conventional dividend price ratio. Furthermore, it has a stable and statistically significant in-sample predictive power on stock returns regardless of the firm sizes. The evidence is robust with respect to the firm sizes. For small-sized firms, the predictive regression model that employs the adjusted dividend-price ratio as a regressor beats the random-walk model in terms of the out-sample predictability.

References

Allen, F. and R. Michaely (2003) "Payout Policy" in *North-Hollan Handbook of Economics* editied by G. Constantinides, M. Harris, and R. Stulz, 337-429, Amsterdam: Elsevier – North Holland.

Andrews, D. K. (1991) "Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models" *Econometrica*, 59:2, 307-345.

Bai, J., and P. Perron (1998) "Estimating and testing linear models with multiple structural changes," *Econometrica*, 66, 47-78.

Bansal, R., and A. Yaron (2004) "Risks for the long run: A potential resolution of asset Pricing puzzles," *Journal of Finance*, 59, 1481-1509.

Bierens and Martins (2009) "Time Varying Cointegration," *Econometric Theory*, forthcoming.

Bossaerts, P., and P. Hillion, 1999, "Implementing Statistical Criteria to Select Return Forecasting Models: What do we Learn?," *Review of Financial Studies,* 12, 405-428.

Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts (2007) "On the importance of measuring payout yield: Implications for empirical asset pricing," *Journal of Finance*, 62, 877-915.

Campbell, J. Y., and J. H. Cochrane (1999) "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy*, 107, 205-251.

Campbell, J. Y., and R. J.Shiller (1988) "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, 1, 195-227.

Campbell, J., and M. Yogo (2006) "Efficient tests of stock return predictability," *Journal of Financial Economics* 81, 27-60.

Cecchetti, S. G., P.-S. Lam, and N. C. Mark (2000) “Asset pricing with distorted beliefs: Are equity returns too good to be true?” *American Economic Review*, 90, 787-805.

Diebold, F. X., and R. S. Mariano (1995) “Comparing predictive accuracy,” *Journal of Business and Economics Statistics*, 13, 253-262.

Fama, E. F. and K. R. French (2001) “Disappearing dividends: changing from characteristics or lower propensity to pay?” *Journal of Financial Economics*, 60, 3-42.

Ferson, W. E., S. Sarkissian, and T. T. Simin (2003) “Spurious regressions in financial economics?” Journal of Finance, 58, 1393-1414.

Gallant, A. R., (1981) “On the Basis in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form,” *Journal of Econometrics*, 15, 211-245.

Goyal, A., and I. Welch, (2008) “A comprehensive look at the empirical performance of the equity premium prediction,” *Review of Financial Studies*, 21; 1455-1508.

Inoue, A. and L. Kilian (2004) “In-sample or out-of-sample tests of predictability: Which one should we use?” *Econometric Reviews*, 23, 371-402.

Lanne M. (2002) “Testing the predictability of stock returns,” *Review of Economics and Statistics* 2002; 84; 407-415

Lettau, M., and S. V. Nieuwerburgh (2008) “Reconciling the return predictability evidence,” *Review of Financial Studies*, 21, 1607-1652.

Lo A. (1991) “Long term memory in stock market prices,” *Econometrica* 59; 1279-1313.

Park, J.Y. (1992) “Canonical cointegration regressions,” *Econometrica* 60; 119-143.

Park, J. Y. and S. B. Hahn (1999) “Cointegrating regressions with time varying coefficients,” *Econometric Theory*, 15, 664-703.

Paye, B. S., and A. Timmermann (2006) “Instability of return prediction models,” *Journal of Empirical Finance*, 13, 274-315.

Rapach, D. E., and M. E. Wohar (2006) “Structural breaks and predictive regression models of aggregate US stock returns,” *Journal of Financial Econometrics*, 4, 238-274.

Robertson, D., and S. Wright (2006) “Dividends, total cashflow to shareholders and predictive return regressions,” *Review of Economics and Statistics*, 88, 91-99.

Stambaugh, R. F. (1999) “Predictive regressions,” *Journal of Financial Economics*, 54, 375-421.

Stephens, C., and M. Weisbach (1998) “Actual share reaquisitions in open market repurchases programs,” Journal of Finance, 53 313-333.

Torous, W., R. Valkanov, and S. Yan (2005) “On predicting stock returns with nearly integrated explanatory variables,” *Journal of Business*, 77, 937-966.

Valkanov, R. (2003) “Long-horizon regressions: Theoretical results and applications,” *Journal of Financial Economics*, 68, 201-232.

Wolf, M. (2000) “Stock returns and dividend yields revisited: A new way to look at an old problem,” *Journal of Business and Economic Statistics*, 18, 18-30.

Appendix I: Proof of the Propositions

Proposition 1. The log dividend-price ratio for the aggregate stock market can be written as $d_t - p_t = \delta_t - p_t + (\widehat{\omega}_t - \omega_t)d_{1,t} + (1-\widehat{\omega}_t)d_{2,t} - (1-\omega_t)\lambda_{2,t}$ where d_t and p_t are the aggregate dividends and stock price which are observed in the market, $\delta_t = \omega_t d_{1,t} + (1-\omega_t)\lambda_{2,t}$, and $\widehat{\omega}_t = \frac{D_{1,t}}{D_{1,t}+D_{2,t}}$. As a result, $d_t - p_t$ contains an I(1) component unless $\omega_t = \widehat{\omega}_t = 1$.

[Proof]

Since $d_t \approx \widehat{\omega}_t d_{1,t} + (1-\widehat{\omega}_t)d_{2,t}$ and $p_t \approx \omega p_{1,t} + (1-\omega)p_{2,t}$, $d_t - p_t$ can be written as follows.

$$
\begin{aligned}
d_t - p_t &\approx \widehat{\omega}_t d_{1,t} + (1-\widehat{\omega}_t)d_{2,t} - \omega p_{1,t} - (1-\omega_t)p_{2,t} \\
&= \omega_t d_{1,t} + (\widehat{\omega}_t - \omega_t)d_{1,t} - \omega_t p_{1,t} + (1-\widehat{\omega}_t)d_{2,t} - (1-\omega_t)\lambda_{2,t} \\
&\quad +(1-\omega_t)\lambda_{2,t} - (1-\omega_t)p_{2,t} \\
&= \omega_t\left(d_{1,t} - p_{1,t}\right) + (\widehat{\omega}_t - \omega_t)d_{1,t} + (1-\widehat{\omega}_t)d_{2,t} - (1-\omega_t)\lambda_{2,t} \\
&\quad +(1-\omega_t)(\lambda_{2,t} - p_{2,t}) \\
&= \delta_t - p_t + (\widehat{\omega}_t - \omega_t)d_{1,t} + (1-\widehat{\omega}_t)d_{2,t} - (1-\omega_t)\lambda_{2,t}
\end{aligned}
$$

Proposition 2. The aggregate dividends and the aggregate stock price are cointegrated with time-varying cointegration coefficients. That is, there exist α_{1t} which render $\alpha_{1t}d_t - p_t$ an I(0) variable where $\alpha_{1t} = \frac{(\omega_t\gamma_1 + (1-\omega_t)\gamma_3)}{(\widehat{\omega}_t\gamma_1 + (1-\widehat{\omega}_t)\gamma_2)}$.

(Proof) $d_t \approx \widehat{\omega}_t d_{1,t} + (1-\widehat{\omega}_t)d_{2,t} = \widehat{\omega}_t\gamma_1 \cdot m_t + (1-\widehat{\omega}_t)\gamma_2 \cdot m_t$

$= (\widehat{\omega}_t\gamma_1 + (1-\widehat{\omega}_t)\gamma_2)m_t$.

Hence, $m_t = \frac{d_t}{(\widehat{\omega}_t\gamma_1 + (1-\widehat{\omega}_t)\gamma_2)}$

Then, $\delta_t - p_t \approx \omega_t\left(d_{1,t} - p_{1,t}\right) + (1-\omega_t)(\lambda_{2,t} - p_{2,t})$

$$
\begin{aligned}
&= \omega_t\gamma_1 m_t + (1-\omega_t)\gamma_3 m_t - p_t \\
&= (\omega_t\gamma_1 + (1-\omega_t)\gamma_3)m_t - p_t \\
&= \frac{(\omega_t\gamma_1 + (1-\omega_t)\gamma_3)}{(\widehat{\omega}_t\gamma_1 + (1-\widehat{\omega}_t)\gamma_2)}d_t - p_t \\
&= \alpha_{1t}d_t - p_t
\end{aligned}
$$

Proposition 3. The log present value relation states that $\alpha_{1t}d_t - p_t$ can have the following relation.

$$\alpha_{1t}d_t - p_t = -\frac{k}{1-\rho} + E_t\left[\sum_{j=0}^{\infty}\rho^j(r_{t+1+j} - \omega_{t+1+j}\Delta d_{1,t+1+j})\right]$$

$$+E_t\left[\sum_{j=0}^{\infty}\rho^j\left(1-\bar{\omega}_{t+j}\right)[(1-\rho_2)(\lambda_{2,t+1+j} - p_{2,t+1+j}) + \Delta p_{2,t+1+j} - r_{2,t+1+j}]\right]$$

$$-E_t\left[\sum_{j=0}^{\infty}\rho^j(1-\omega_{t+1+j})(\Delta\lambda_{2,t+1+j})\right]$$

where r_t is the aggregate stock returns observable in the market, $\rho = \frac{P_1+P_2}{(P_1+D_1)+(P_2+\Lambda_2)}$, k is an unimportant linearization constant, and $\bar{\omega}_{t+j}$ is the average of ω_{t+1+j} and ω_{t+j}.

(Proof) $\alpha_{1t}d_t - p_t = \delta_t - p_t = -\frac{k}{1-\rho} + E_t\left[\sum_{j=0}^{\infty}\rho^j(r'_{t+1+j} - \Delta\delta_{t+1+j})\right]$

where $r'_{t+1+j} = \log\left(P_{1,t+1+j} + P_{2,t+1+j} + D_{1,t+1+j} + \Lambda_{2,t+1+j}\right) - \log\left(P_{1,t+j} + P_{2,t+j}\right)$.

Although r'_{t+1+j} is unobservable, it can be expressed as follows.

$$r'_{t+1+j} = \log\left(P_{1,t+1+j} + P_{2,t+1+j} + D_{1,t+1+j} + \Lambda_{2,t+1+j}\right) - \log\left(P_{1,t+j} + P_{2,t+j}\right)$$

$$\approx \bar{\omega}_{t+j}\log\left(P_{1,t+1+j} + D_{1,t+1+j}\right) + \left(1-\bar{\omega}_{t+j}\right)\log\left(P_{2,t+1+j} + \Lambda_{2,t+1+j}\right)$$

$$-\bar{\omega}_{t+j}\log\left(P_{1,t+1+j}\right) - \left(1-\bar{\omega}_{t+j}\right)\log\left(P_{2,t+1+j}\right)$$

$$= \bar{\omega}_{t+j}[\log(P_{1,t+1+j} + D_{1,t+1+j}) - \log\left(P_{1,t+1+j}\right)]$$

$$+\left(1-\bar{\omega}_{t+j}\right)[\log\left(P_{2,t+1+j} + \Lambda_{2,t+1+j}\right) - \log\left(P_{2,t+1+j}\right)]$$

$$\approx \bar{\omega}_{t+j}r_{1,t+1+j} + \left(1-\bar{\omega}_{t+j}\right)[\log\left(P_{2,t+1+j} + \Lambda_{2,t+1+j}\right) - \log\left(P_{2,t+1+j}\right)]$$

$$+\left(1-\bar{\omega}_{t+j}\right)[\log\left(P_{2,t+1+j} + D_{2,t+1+j}\right) - \log\left(P_{2,t+1+j} + D_{2,t+1+j}\right)]$$

$$= \bar{\omega}_{t+j}r_{1,t+1+j} + \left(1-\bar{\omega}_{t+j}\right)[\log\left(P_{2,t+1+j} + D_{2,t+1+j}\right) - \log\left(P_{2,t+1+j}\right)]$$

$$+\left(1-\bar{\omega}_{t+j}\right)[\log\left(P_{2,t+1+j} + \Lambda_{2,t+1+j}\right) - \log\left(P_{2,t+1+j} + D_{2,t+1+j}\right)]$$

$$= \bar{\omega}_{t+j}r_{1,t+1+j} + \left(1-\bar{\omega}_{t+j}\right)r_{2,t+1+j}$$

$$+\left(1-\bar{\omega}_{t+j}\right)[\rho p_{2,t+1+j} + (1-\rho)\lambda_{2,t+1+j} - \log\left(P_{2,t+1+j} + D_{2,t+1+j}\right)]$$

$$= r_{t+1+j} + \left(1-\bar{\omega}_{t+j}\right)\left[(1-\rho)(\lambda_{2,t+1+j} - p_{2,t+1+j})\right]$$

$$+\left(1-\bar{\omega}_{t+j}\right)[p_{2,t+1+j} - p_{2,t+j} - \log\left(P_{2,t+1+j} + D_{2,t+1+j}\right) + \log\left(P_{2,t+j}\right)]$$

$$= r_{t+1+j} + \left(1-\bar{\omega}_{t+j}\right)\left[(1-\rho)(\lambda_{2,t+1+j} - p_{2,t+1+j}) + \Delta p_{2,t+1+j} - r_{2,t+1+j}\right] \quad \text{(A.1)}$$

Also, $\Delta\delta_{t+1+j}$ can be written as follows.

$$\Delta\delta_{t+1+j} = \omega_{t+1+j}\Delta d_{1,t+1+j} + (1-\omega_{t+1+j})\Delta\lambda_{2,t+1+j} \tag{A.2}$$

Using equations (A.1) and (A.2), we can obtain the result in Proposition 3.

Appendix 2: A Cointegrating Regression with a Time-varying Coefficient [Park and Hahn (1999)

In order to estimate the time-varying cointegration relationship and to evaluate whether stationary deviations from this relationship have any predictive power for future stock returns, we have employed the Park and Hahn (1999) approach. Thus, we consider the following econometric model.

$$p_t = \alpha_{0t} + \alpha_{1t}d_t = \alpha_0 + \alpha_{1t}d_t + u_t \tag{A.3}$$

α_{1t} denotes the cointegration coefficient between p_t and d_t, and the gradual changes in ω_t can cause α_{1t} to depend on time as well. We denote the sample size by T and let $\alpha_{1t} = \alpha(\frac{t}{T})$ so that α_{1t} is a smooth function defined on [0, 1].[12] While estimating $\alpha(s)$, no functional form is imposed for $\alpha(s)$. The only assumption required for $\alpha(s)$ is that it is sufficiently smooth to be approximated by a series of polynomials, trigonometric functions, or a mixture of both. That is, we assume that $\|\alpha_\kappa(s) - \alpha(s)\| \to 0$ as $\kappa \to \infty$, where $\alpha_\kappa(s)$ is an approximation of $\alpha(s)$ given by a combination of a finite series of functions $\varphi_1, \ldots, \varphi_\kappa$. Since $\alpha_\kappa(s) = \sum_{i=1}^{\kappa} \theta_i\, \varphi_i$, the above econometric model can be expressed as:

$$\begin{aligned} p_t &= \alpha_0 + \alpha_{1t}d_t + u_t = \alpha_0 + \alpha(\tfrac{t}{T})d_t + u_t \\ &= \alpha_0 + \left[\textstyle\sum_{i=1}^{\kappa} \theta_i\, \varphi_i(\tfrac{t}{T})\right] \cdot d_t + u_{\kappa t} \\ &= \alpha_0 + x'_{\kappa t}a_\kappa + u_{\kappa t} \end{aligned} \tag{A.4}$$

where $u_{\kappa t} = u_t + \left[\alpha\left(\frac{t}{T}\right) - \alpha_\kappa\left(\frac{t}{T}\right)\right] d_t, x'_{\kappa t} = [\varphi_1\left(\frac{t}{T}\right), \ldots, \varphi_\kappa\left(\frac{t}{T}\right)]' d_t$, and $a_\kappa = [\theta_1, \ldots, \theta_\kappa]'$.

If p_t and d_t are stationary series, then we can establish the asymptotic normality of LS estimator for a_κ in equation (2) and chi-square tests (see Andrews (1991)). However, as p_t and d_t are nonstationary, we apply the canonical cointegrating regression (CCR) approach, which was developed by Park (1992), for equation (2) to obtain the asymptotic normality of the LS estimator for a_κ and chi-sqaure tests. Hence, we have made CCR transformation for p_t and d_t as follows.

---

[12] When the time-varying cointegration coefficient is approximated by a series of trigonometric functions, it is desirable to scale the data into the interval [0,1] due to the characteristics of trigonometric functions.

$$p_t^* = p_t - \left[\left[\varphi_1\left(\frac{t}{T}\right), \ldots, \varphi_\kappa\left(\frac{t}{T}\right)\right]' \otimes \Delta_2 \Sigma^{-1} w_t\right]' a_\kappa - (0, \frac{\Omega_{12}}{\Omega_{22}}) w_t \qquad \text{(A.5)}$$

and

$$d_t^* = d_t - \Delta_2 \Sigma^{-1} w_t \qquad \text{(A.6)}$$

where $w_t = (u_t, v_t)'$, $\Delta = E w_t w_t' + \sum_{k=1}^{\infty} E w_t w_{t-k}'$, and $\Omega = \sum_{k=-\infty}^{\infty} E w_t w_{t-k}'$. v_t is the innovation series in the d_t process, Ω_{ij} for $i, j = 1{,}2$ denotes elements of Ω, and Δ_2 denotes the second row of Δ matrix. Through the CCR transformation, Equation (2) can be written as

$$p_t^* = \alpha_0 + x_{\kappa t}^{*\prime} a_\kappa + u_{\kappa t}^* \qquad \text{(A.7)}$$

where $x_{\kappa t}^{*\prime} = [\varphi_1\left(\frac{t}{T}\right), \ldots, \varphi_\kappa\left(\frac{t}{T}\right)]' d_t^*$. We can derive the asymptotic normality of LS estimator for a_κ under this transformation. Once the LS estimator for a_κ is obtained, then $\alpha(s)$ can be approximated by $\widehat{\alpha_\kappa} = \sum_{i=1}^{\kappa} \widehat{\theta_\iota}\, \varphi_i$.

We utilize the Fourier Flexible Form (FFF) to approximate $\alpha_{1t} = \alpha(\frac{t}{T})$ nonparametrically. The FFF, which was introduced by Gallant (1981), extends the traditional Fourier theorem. The FFF expansion of $\alpha(s)$ can be expressed as

$$\alpha_\kappa(s) = \theta_0 + \theta_1 s + \theta_3 s^2 + \sum_{i=1}^{I} [\theta_{3,i} \, cos(\lambda_i s) + \theta_{4,i} \, sin(\lambda_i s)] \qquad \text{(A.8)}$$

where $\lambda_i = 2\pi i$, and $\kappa = 3 + 2I$. It is worth noting the robustness of the FFF approach. Because economic theories provide few guidelines for $\alpha(s)$, except for the conjecture that $\alpha(s)$ might be positive as $\alpha(s)$ is fixed at one when all firms are traditionals, the FFF is ideal, as it approximates $\alpha(s)$ under a flexible representation. If only the first term in equation (6) is considered and set as one, then the time-varying cointegration regression based on FFF becomes a cointegration regression with the usual fixed cointegration coefficient [1, -1]. We choose the number of series functions in the FFF representation as nine ($\kappa = 9$), implying that $\alpha(s)$ is approximated by $\alpha_\kappa(s) = \theta_0 + \theta_1 s + \theta_3 s^2 + \sum_{i=1}^{3} [\theta_{3,i} \, cos(\lambda_i s) + \theta_{4,i} \, sin(\lambda_i s)]$.

Table 1

Persistence of the Log Dividend-Price Ratio and the Fraction of Type-I Firms

The following Augmented Dickey-Fuller (ADF) regression is run recursively:

$$d_t - p_t = \alpha_j + \tau_j(d_{t-1} - p_{t-1}) + \varsigma_{1,j}\Delta(d_{t-1} - p_{t-1}) + \cdots + \varsigma_{p,j}\Delta\left(d_{t-p+1} - p_{t-p+1}\right) + \varepsilon_t \ ,$$
$$t = 1,2,\ldots,T^* + j \ \text{ and } \quad j = 0,1,2,\ldots,J$$

Using the estimates of τ_j and the ADF t-statistic for a unit root test obtained from the above recursive regressions, we estimate the following regression equations:

$$\tau_j = c_{\tau,0} + c_{\tau,1}\omega_j + e_{\tau,j}$$
$$adf_j = c_{adf,0} + c_{adf,1}\omega_j + e_{adf,j}$$
$$j = 0,1,2,\ldots,J,$$

where adf_j is the ADF t-statistic.

	Aggregate firms	Large firms	Medium firms	Small firms
		$\hat{c}_{\tau,1}$		
Coefficient of ω_t	-0.0256*** (-6.6419)	-0.0228*** (-6.0205)	-0.0309*** (-13.1091)	-0.0292*** (-9.6178)
$\bar{R}^2$	0.5316	0.4566	0.8433	0.6503
		$\hat{c}_{adf,1}$		
Coefficient of ω_t	-2.2076*** (-3.6415)	-1.8963*** (-3.0951)	-3.3928*** (-11.4625)	-3.0363*** (-7.5838)
$\bar{R}^2$	0.3182	0.2277	0.7926	0.5655

Table 2

Tests of the Time-varying Long-Run Relation between d_t and p_t

Two types of test statistics given by Park and Hahn (1999) are $\tau_1^* = \frac{\sum_{t=1}^{T}(d_t - p_t)^2 - \sum_{t=1}^{T}\hat{s}_t^2}{\hat{\vartheta}_{T\kappa}^2}$ and $\tau_2^* = \frac{\sum_{t=1}^{T}(\sum_{i=1}^{t} d_i - p_i)^2}{T^2\hat{\vartheta}_{T\kappa}^2}$ where $\hat{s}_t$ are the residuals of the regression of $d_t - p_t$ on superfluous regressors such as a constant, $t,$ and t^2, and $\hat{\vartheta}_{T\kappa}^2$ is a long-run variance estimator of $u_{\kappa t}^*$ in equation (5). The 5% critical values for τ_1^* and τ_2^* are 7.82 and 0.16, respectively.

	Aggregate	Large firms	Medium firms	Small firms
τ_1^*	575.3904	559.6054	1164.0	1054.3
τ_2^*	54.3645	50.7636	111.0943	106.7308

Table 3
The Fraction of Type-I Firms and
Time-varying Long-Run Relation between between d_t and p_t

The estimated time-varying cointegration coefficient is regressed on a constant and ω_t. The numbers in parentheses show t-statistics for the coefficient of ω_t in the regression. The t-statistics are computed from Newey-West standard errors with 24 lags. κ denotes the number of series functions in the FFF representation. The superscripts '*', '**', and '***' denote that the estimates are significant at the 10%, 5%, and 1% significance levels, respectively.

	Aggregate firms	Large firms	Medium firms	Small firms
	$\kappa = 7$			
Coefficient of ω_t	0.3008*** (10.5303)	0.2798*** (8.8406)	0.6219*** (14.4807)	0.7056*** (16.5054)
$\bar{R}^2$	0.6910	0.6191	0.8961	0.8404
	$\kappa = 9$			
Coefficient of ω_t	0.2558*** (8.6674)	0.2437*** (7.5606)	0.6185*** (15.1582)	0.7732*** (16.7022)
$\bar{R}^2$	0.6566	0.5926	0.9015	0.8369
	$\kappa = 11$			
Coefficient of ω_t	0.2534*** (7.9933)	0.2340*** (6.8561)	0.6388*** (16.4890)	0.8584*** (16.0961)
$\bar{R}^2$	0.6354	0.5660	0.9064	0.8314

Table 4
Summary Statistics for
Dividend-Price Ratio ($d_t - p_t$) and Adjusted Dividend-Price Ratio ($\delta_t - p_t$)

	Mean	Standard Deviation	Autocorrelation	ADF test statistics
Aggregate firms				
$d_t - p_t$	-5.9278	0.4258	0.9931	-1.5387
$\delta_t - p_t$	-4.1922	0.1322	0.9373	-4.0933
Large firms				
$d_t - p_t$	-5.9068	0.4418	0.9935	-1.4744
$\delta_t - p_t$	-3.7812	0.1332	0.9422	-3.9450
Medium firms				
$d_t - p_t$	-6.0794	0.6061	0.9944	-1.1921
$\delta_t - p_t$	-2.3230	0.1578	0.9364	-4.2351
Small firms				
$d_t - p_t$	-6.3822	0.6901	0.9944	-1.4138
$\delta_t - p_t$	-4.1920	0.1900	0.9410	-4.1591

Table 5

Predictive Regressions: In-Sample Analysis

$$r_{t+1} = \beta_0 + \beta_1 x_t + \varepsilon_{t+1}, \tag{13}$$

where r_{t+1} is the stock return at time t+1; x_t is either the conventional dividend-price ($d_t - p_t$) ratio or the adjusted dividend-price ratio ($\delta_t - p_t$). The bootstrapped p-values are reported in the brackets and adjusted R^2's from the regressions are reported in the parentheses. '*', '**', and '***' denote the significance level at the 10%, 5%, and 1% level, respectively.

Horizon	Aggregate		Large		Medium		Small	
	$d_t - p_t$	$\delta_t - p_t$	$d_t - p_t$	$\delta_t - p_t$	$d_t - p_t$	$\delta_t - p_t$	$d_t - p_t$	$\delta_t - p_t$
1	0.0075 [0.156] (0.004)	0.0531 [0]*** (0.025)	0.0071 [0.164] (0.004)	0.0500 [0]*** (0.024)	0.0048 [0.309] (0.002)	0.0525 [0.000]*** (0.025)	0.0039 [0.405] (0.000)	0.0448 [0.000]*** (0.020)
3	0.0244 [0.433] (0.016)	0.1812 [0.015]** (0.086)	0.0221 [0.445] (0.015)	0.1617 [0.020]** (0.077)	0.0166 [0.584] (0.009)	0.1948 [0.014]** (0.096)	0.0145 [0.629] (0.006)	0.1764 [0.021]** (0.083)
6	0.0483 [0.565] (0.031)	0.3720 [0.054]* (0.178)	0.0440 [0.580] (0.029)	0.3333 [0.074]* (0.156)	0.0309 [0.684] (0.018)	0.3901 [0.048]** (0.198)	0.0264 [0.738] (0.012)	0.3446 [0.075]* (0.162)

Table 6

Structural Break Tests for the Predictive Regressions

Bai and Perron's structural break tests are conducted for the regression $\tilde{r}_{t+1} = \beta_1 \tilde{y}_t + \varepsilon_{t+1}$ where $\tilde{r}_{t+1}$ is demeaned log stock returns, $\tilde{y}_t$ is demeaned $d_t - p_t$ or $\delta_t - p_t$. Sup-F(*i*,*j*) is the Bai and Perron's (1998) sup-F test statistic where *i* is the number of breaks under the null hypothesis and *j* is the number of breaks under the alternative hypothesis. The test UDmax is defined as the maximum of {Sup-F(0,1),…, Sup-F(0,5)} multiplied by the number of regressors. WDmax is defined in equation (9) of Bai and Perron (1998). The null hypothesis of UDmax and WDmax is that there are 0 breaks and the alternative hypothesis is that there are unknown number of breaks given an upper bound of 5. '*', '**', and '***' denote the significance level at the 10%, 5%, and 1% level, respectively.

	Sup-F(0,1)	Sup-F(0,2)	Sup-F(1,2)	UDmax	WDmax(1%)
demeaned $d_t - p_t$	1.2036	7.7518*	11.8305**	7.7518*	14.8200***
demeaned $\delta_t - p_t$	0.8487	3.0307	2.5684	3.0307	3.9794

Table 7
Out-of-Sample Tests to Forecast One-month Ahead Stock Returns

This table reports the out-of-sample test results to forecast one-month ahead stock returns. The time-varying cointegration parameters for $\delta_t - p_t$ is re-estimated every period using the available data at that period of time. Then, the predictive regression of one-month ahead stock returns is run on the estimated $\delta_t - p_t$ recursively. The first regression is run with the use of 30 year data since January 1946. The Diebold-Mariano test statistics is employed to compare the forecast ability of $\delta_t - p_t$ with that of the random walk model. The loss function of the Diebold-Mariano test is absolute forecast error to mitigate the effect from outliers when comparing the predictive abilities. Negative signs in the Diebold and Mariano test statistics indicate that out-of-sample forecast errors from the predictive regression with the first argument ($\delta_t - p_t$ or $d_t - p_t$) are smaller than those from the second argument (the random walk model or $d_t - p_t$). Results that are significant at the 5% level of the one-side test are shown in boldface.

	Full Sample: 1946.1–2008. 12			1946.1–2005. 12		
	$d_t - p_t$ vs. Random walk	$\delta_t - p_t$ vs. Random walk	$\delta_t - p_t$ vs. $d_t - p_t$	$d_t - p_t$ vs. Random walk	$\delta_t - p_t$ vs. Random walk	$\delta_t - p_t$ vs. $d_t - p_t$
Aggregate	3.0054***	-0.0771	-1.1844	3.0115***	-0.2361	-1.2987*
Large	2.7390***	0.5964	-0.5089	2.8241***	0.4813	-0.6195
Medium	2.9889***	-0.9261	-1.8854**	3.1703***	-0.9915	-1.9940**
Small	2.7650***	-1.8161**	-2.5104***	2.7825***	-1.9327**	-2.6019***

Figure 1. Movements of Market Equity Fraction of Type-I Firms

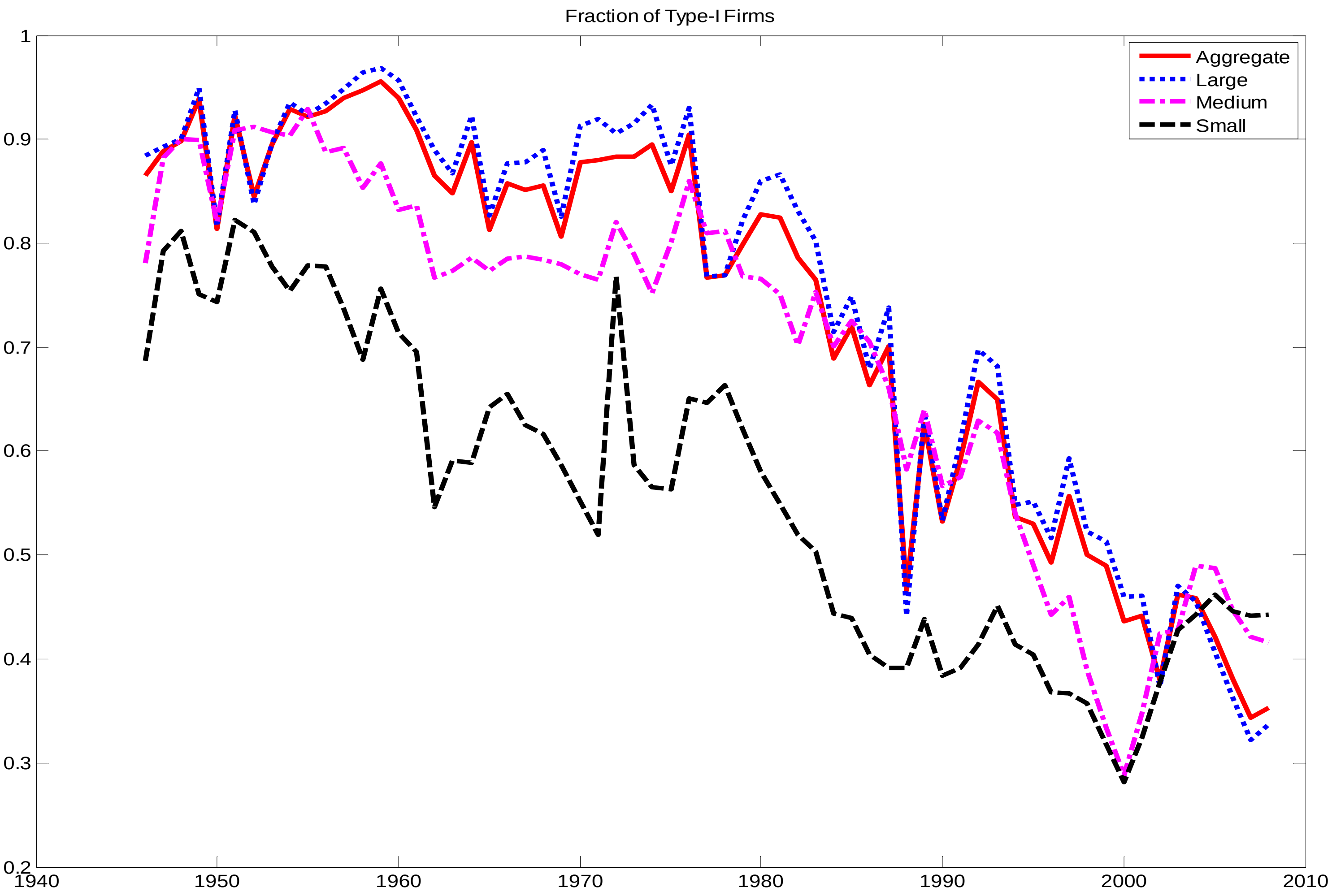

Figure 2. Time-varying Long-Run Relationship (cointegrating vector) between p_t and d_t

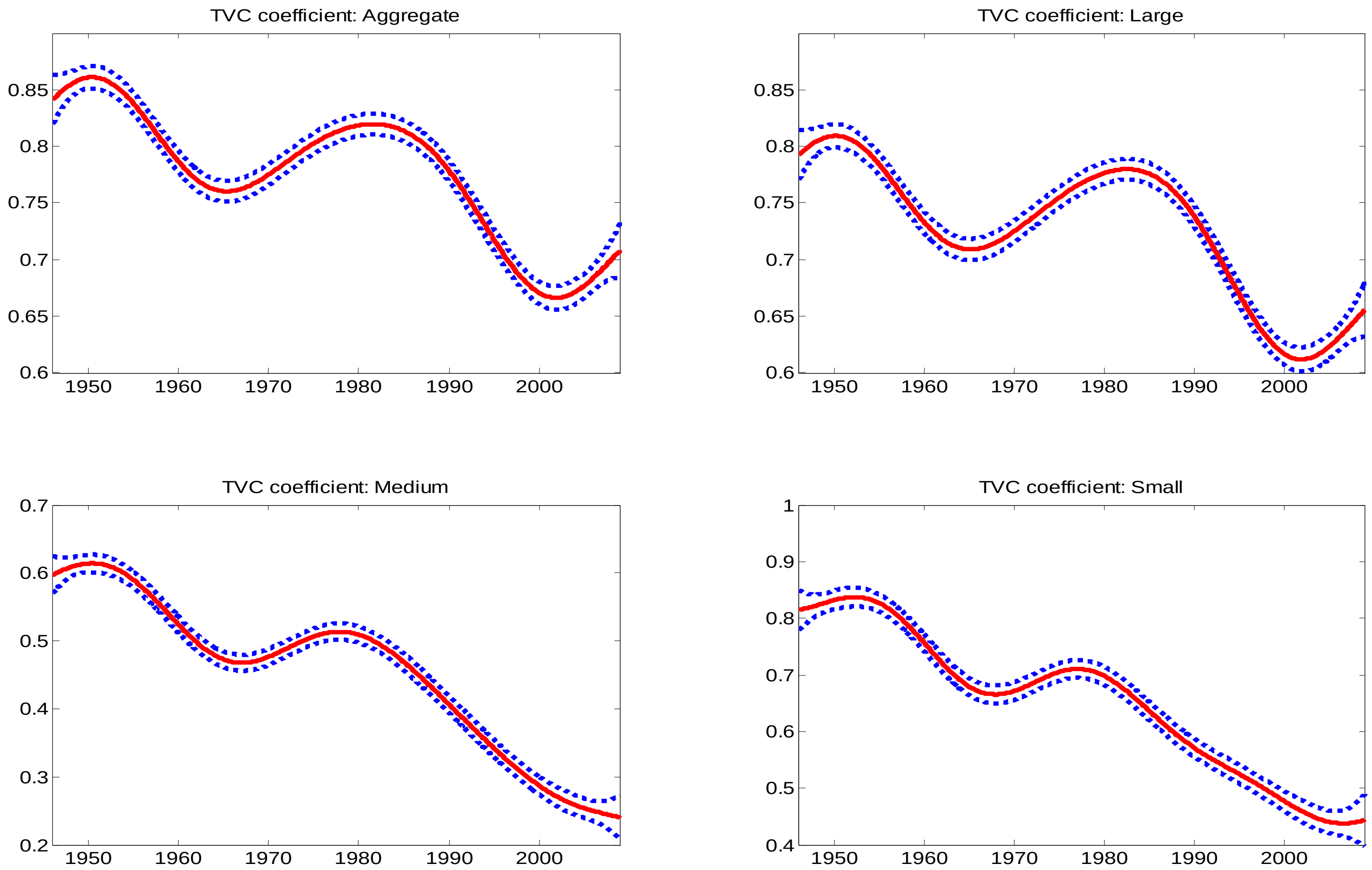

Figure 3. Conventional Dividend-Price Ratio vs. Time-Varying Dividend-Price Ratio