



Discussion Paper Series

No. 1205

March 2012

Disappearing Dividends:
Implications for the Dividend-Price Ratio
and Return Predictability

Chang-Jin Kim **and** Cheolbeom Park

The Institute of Economic Research - Korea University

Anam-dong, Sungbuk-ku, Seoul, 136-701, South Korea, Tel: (82-2) 3290-1632, Fax: (82-2) 928-4948

Copyright © 2012 IER.

**Disappearing Dividends:
Implications for the Dividend-Price Ratio and Return Predictability**

Chang-Jin Kim

University of Washington and Korea University

and

Cheolbeom Park¹

Korea University

March, 2012

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 that accounts for the time-varying long-run relationship is stationary with considerably less persistence than the conventional dividend-price ratio. Furthermore, the predictive regression model that employs the adjusted dividend-price ratio as a regressor outperforms the random-walk model in terms of long-horizon out-of-sample predictability. These results are robust with respect to the firm size.

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).

We thank conference and seminar participants at ESWC 2010, KEA 2010, SETA 2010, Korea University, Sogang University, and Sungkyunkwan University for helpful discussions and comments. Usual disclaimers apply.

1. Introduction

Since Campbell and Shiller (1988), a large body of research presents 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 predictive regression models. For example, Goyal and Welch (2008) and Bossaerts and Hillion (1999) suggest that the coefficients of the predictive regression models are unstable, as diagnosed by their poor out-of-sample predictions even in the presence of strong in-sample predictions. Rapach and Wohar (2006) and Paye and Timmerman (2006) directly test for structural breaks in the predictive regressions, and reject the null of no structural break. All these researchers document deterioration in the return predictability in the US stock market, especially since the 1990s.

In the mean time, Fama and French (2001) and Allen and Michaely (2003) document a decreasing dividend-price ratio that results from the changes in the dividend payout policy by firms. They show that the proportion of firms paying cash dividends fell from 66.5% in 1978 to 20.8% in 1999. By employing a mean-adjusted dividend-price ratio in the predictive regression model under this situation, Lettau and van Nieuwerburgh (2007) achieve improved return predictability. They argue that a failure to take into account the decrease or shifts in the mean of the dividend-price ratio is responsible for the evidence of instability in the predictive regressions and the poor out-of-sample predictive power of the dividend-price ratio reported in the literature.

Unlike most researchers who employ the conventional or demeaned dividend-price ratio as a predictor of stock return, Boudoukh *et al.* (2007) and Robertson and Wright (2006) recognize that, while the dividend-price ratio changed remarkably during the 1990's, the total payout ratio defined as the total payout (or dividends plus share repurchases) over price has remained relatively stable. According to them, the total payout has a stable one-to-one long-run relationship with stock price. By employing the total payout ratio in a predictive regression, they achieve improved out-of-sample return predictability.

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

²The 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.

We define the firms with traditional payout policy as the firms that pay out a significant portion of their earnings in the form of dividends, and consider a measure for the proportion of the firms with traditional payout policy in the market (ω_t). We first present a theoretical model in which changes in this proportion results in a time-varying long-run or cointegrating relationship between dividend and price. We then empirically test whether the adjusted dividend-price ratio that reflects a deviation from the time-varying long-run equilibrium has predictive power on stock returns.

The approach in this paper is closely related to that in Boudoukh *et al.* (2007) or Robertson and Wright (2006). The two approaches may be considered two sides of the same coin. Boudoukh *et al.* (2007) or Robertson and Wright (2006) claim that, in the presence of increasing proportion of firms that replace dividends with share repurchases, the one-to-one long-run relationship between dividend and price breaks down. They thus 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, and employ the total payout ratio as a predictor of stock return. On the contrary, we directly estimate a time-varying long-run relationship between dividend and price, and employ the resulting adjusted dividend-price ratio as a predictor of stock return. By doing so, we hope to overcome a potential weakness in Boudoukh *et al.* (2007) and Robertson and Wright's (2006) approach. As noted by Fama and French (2001), a substantial portion of share repurchases 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 presents a theoretical model, with which we discuss the consequences of changes in the payout policy by firms. Section 3 provides empirical evidence on the time-varying long-run relationship between dividend and stock price. Section 4 evaluates the predictive power of the adjusted dividend-price ratio that accounts for the time-varying long-run relationship. Concluding remarks are given in Section 5.

2. Disappearing Dividends and the Time-Varying Long-Run Relationship 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 $D_{2,t}$. $A_{2,t}$ is the hypothetical amount of dividend payment that type-II firms would pay if they adopted the traditional payout policy. The inequality $A_{2,t} > D_{2,t}$ holds, as $A_{2,t}$ includes share repurchases or 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, 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

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

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

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

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

where m_t is a measure of economic conditions (e.g., the aggregate real output) and $E(\varepsilon_{i,t}) = 0$, $i = 1, 2, 3$. As in Fama and French (2001), the two types of firms that we consider have different characteristics such as firm-size, investment opportunity, growth potential, etc. The differences in the characteristics of the two types of firms are responsible for different

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

values for the γ_1 and γ_3 coefficients, even under the hypothetical situation in which both types of firms adopted the traditional payout policy. Different values for the γ_2 and γ_3 coefficients are due to the change in payout policy by type-II firms.

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

$$d_{1,t} - p_{1,t} = -\frac{k_1}{1-\rho_1} + E_t \left[\sum_{j=0}^{\infty} (\rho_1)^j (r_{1,t+1+j} - \Delta d_{1,t+1+j}) \right] \quad (4)$$

$$\lambda_{2,t} - p_{2,t} = -\frac{k_2}{1-\rho_2} + E_t \left[\sum_{j=0}^{\infty} (\rho_2)^j (r_{2,t+1+j} - \Delta \lambda_{2,t+1+j}) \right] \quad (5)$$

where, by defining $r_{1,t} = \ln(P_{1,t} + D_{1,t}) - \ln(P_{1,t})$ is the return for type-I firms at t ; $r_{2,t} = \ln(P_{2,t} + \Lambda_{2,t}) - \ln(P_{2,t})$ is the return for type-II firms at t ; $\rho_1 = \frac{1}{(1+\exp(\overline{d_1 - p_1}))}$, $\overline{d_1 - p_1}$ is the average log dividend-price ratio for type-I firms; $\rho_2 = \frac{1}{(1+\exp(\overline{\lambda_2 - p_2}))}$, $\overline{\lambda_2 - p_2}$ is the average of the hypothetical log dividend-price ratio for type-II firms; k_1 and k_2 are constants that result from the linearization process. The above present value model implies that 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, in the presence of a unit root in $d_{1,t}$, $p_{i,t}$, $i = 1,2$ and $\lambda_{2,t}$. There exists a stable one-to-one long-run relationship between the dividend and the stock price of type-I firms and between the hypothetical dividend and the stock price of type-II firms. In other words, there exists a stable long-run relationship between the aggregate dividend and the aggregate stock price if both types of firms were adopting the traditional payout policy.

⁴Note 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.

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

We present three propositions derived from the model specified above. These propositions describe the key implications of the changes in the market proportion of type-I firms on the aggregate dividend-price ratio and on the return predictability. Proofs of these propositions are given in Appendix 1.

Proposition 1:

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

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

$$d_t - p_t = \delta_t - p_t + (\hat{\omega}_t - \omega_t)d_{1,t} + (1 - \hat{\omega}_t)d_{2,t} - (1 - \omega_t)\lambda_{2,t} \quad (6)$$

where p_t and d_t are the logs of aggregate stock price and aggregate dividend observed in the market; $\delta_t = \omega_t d_{1,t} + (1 - \omega_t)\lambda_{2,t}$ is the log of hypothetical aggregate dividend we would have if type-II firms followed the traditional payout policy (or if all firms in the market were of type-I); ω_t , as defined in (1), is the fraction of type-I firms in terms of the market value of common equities; and $\hat{\omega}_t = \frac{D_{1,t}}{D_{1,t}+D_{2,t}}$ is the fraction of the actual dividends paid out by type-I firms in the market. Note that the last three terms in equation (6) have a unit root. These terms would disappear if all firms in the market were of type-I, where we would have a stationary dividend-price ratio with $d_t - p_t = \delta_t - p_t$. With the presence of both types of firms in the economy, we have a unit root in $d_t - p_t$, due to the last three terms in equation (6). Furthermore, the proportion of the unit root component in $d_t - p_t$ depends upon ω_t .

Proposition 2:

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

$$p_t = \alpha_t d_t + u_t, \quad (7)$$

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

If all the firms in the economy were of type-I ($\omega_t = \widehat{\omega}_t = 1$), we would have $\alpha_t = 1$, and the conventional dividend-price ratio would be stationary as in Proposition 1. Otherwise, the parameter (α_t) describing the long-run relationship between p_t and d_t 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 accounts for the time-varying long-run relationship between dividend and price, is a function of future expected returns, as illustrated below:

$$\begin{aligned} \alpha_t d_t - p_t = & -\frac{k}{1 - \rho} + E_t \left[\sum_{j=0}^{\infty} \rho^j (r_{t+1+j} - \omega_{t+1+j} \Delta d_{1,t+1+j}) \right] \\ & + E_t \left[\sum_{j=0}^{\infty} \rho^j (1 - \bar{\omega}_{t+j}) [(1 - \rho_2)(\lambda_{2,t+1+j} - p_{2,t+1+j}) + \Delta p_{2,t+1+j} - r_{2,t+1+j}] \right] \\ & - E_t \left[\sum_{j=0}^{\infty} \rho^j (1 - \omega_{t+1+j}) (\Delta \lambda_{2,t+1+j}) \right] \end{aligned} \quad (8)$$

where r_t is the aggregate stock return; $\rho = \frac{1}{(1 + \exp(\bar{\delta} - p))}$, $\bar{\delta} - p$ is the average

$\delta_t - p_t$; k is a constant that results from the linearization approximation; and

$\bar{\omega}_{t+j}$ is the average of ω_{t+1+j} and ω_{t+j} .

2.3. Discussion

Propositions 1 and 2 confirm Boudoukh *et al.* (2007) and Robertson and Wright's (2006) argument, which suggests that the changes in firms' payout policy (a shift from dividend payments to share repurchases) results in a breakdown in the one-to-one long-run relationship between dividend and stock price. They thus 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, a considerable portion of small-sized firms have neither paid out dividends nor repurchased shares especially since the 1980s. Besides, as Fama and French (2001) demonstrate, a substantial portion of share repurchase has taken place in consideration of employee stock ownership plans or mergers. These considerations lead us to suspect that even the total payout may not have a one-to-one long-run relationship with stock price. This is why we explicitly and directly consider a time-varying long-run relationship between dividend and stock price in this paper.

Proposition 3 establishes the possibility that the adjusted dividend-price ratio ($\alpha_t d_t - p_t$), or a stationary deviation from the time-varying long-run cointegrating relationship, has predictive power on the aggregate stock return (r_{t+1+j} , $j = 1, 2, 3, \dots$). In the presence of type-II firms which do not pay out dividends, $\alpha_t d_t - p_t$, not $d_t - p_t$, may be employed as a regressor in the predictive regressions. In the case that the fraction of type-I 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 sub-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 the dynamics of ω_t , the fraction of type-I firms in

the market, is given by:

$$\omega_t = (1 - S_t)\omega_0 + S_t\omega_1 \quad (9)$$

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

where τ is the structural break point. Proposition I suggests that, when $\omega_0 \approx 1$ and $\omega_1 < 1$, $d_t - p_t$ is stationary before the structural break and it has a unit root after the structural break. Thus, in a predictive regression that employs 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 $d_t - p_t$. This is because the regressor ($d_t - p_t$) may have 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 $d_t - p_t$ 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 one since the 1990s.

3. Time-Varying Persistence of the Dividend-Price Ratio and the Time-Varying 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 dividends are the average of dividends paid over the previous year, as in the literature to remove any seasonal patterns in dividends payments. 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

⁵ The CRSP data were obtained from <http://wrds.wharton.upenn.edu>.

for the portfolios formed on different firm sizes are obtained from Kenneth French's homepage.⁶ The sample period is 1946.1-2008.12.

We first construct annual data on the fraction (ω_t) of firms that follow the traditional payout policy (type-I firms) in equation (1). For this purpose, we calculate the fraction of firms with share repurchases ($\bar{\omega}_{1t}$) and the fraction of firms that neither pay dividends nor repurchase shares ($\bar{\omega}_{2t}$) in terms of the market values of common equities. These data are based on the year-end market values of the corresponding firms in the CRSP. The annual data on ω_t is constructed by $\omega_t = 1 - \bar{\omega}_{1t} - \bar{\omega}_{2t}$. 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 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- and medium-sized firms most of the period. Besides, it has higher volatility than the other two.⁸

3.2. Time-Varying Persistence of the Dividend-Price Ratio

We examine the implications of Proposition 1, which states that the log of $d_t - p_t$ contains a unit root component, the relative size of which depends upon the fraction of type-I firms (ω_t). 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 unit root null hypothesis for $d_t - p_t$ process, as ω_t decreases. We run the following Augmented Dickey-Fuller (ADF) regression equation recursively to examine this implication:

$$d_t - p_t = \alpha_j + \tau_j(d_{t-1} - p_{t-1}) + \zeta_{1,j}\Delta(d_{t-1} - p_{t-1}) + \cdots + \zeta_{p,j}\Delta(d_{t-p+1} - p_{t-p+1}) + \varepsilon_t,$$

$$t = 1, 2, \dots, T^* + j$$

$$j = 0, 1, 2, \dots, J \quad (10)$$

The first regression is run using a 30-year data set that starts in January 1946 ($j=0$), and the

⁶ The web address for French's homepage is <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>.

⁷The 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/>.

⁸We 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 French (2001)).

subsequent regressions are run recursively by adding one monthly observation at a time ($j=1,2,\dots,J$). 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:

$$\tau_j = c_{\tau,0} + c_{\tau,1}\omega_j + e_{\tau,j} \quad (11)$$

$$adf_j = c_{adf,0} + c_{adf,1}\omega_j + e_{adf,j} \quad (12)$$

where adf_j is the ADF t-statistic for a unit root test and $j = 0,1,2, \dots, J$.

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 $c_{adf,1}$ coefficients are negative and statistically significant. The results are robust with respect to the firm size. Since the estimated ADF test statistics are mostly 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 and the ADF statistic gets closer to zero, making it 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. 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 is the residual from a 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 (A.7) of Appendix 2. 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 0.31, 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 on the time-varying nature of the long-run cointegrating relationship between d_t and p_t , we directly estimate the time-varying cointegration vector based on the Fourier Flexible Form (FFF), as proposed by Park and

⁹We 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 (2010). The null hypothesis is rejected at the 1% significance level for all cases.

Hahn (1999). For detailed discussion of Park and Hahn's (1999) estimation procedure, readers are referred to Appendix 2. Figure 2 depicts the estimates of the time-varying coefficients from a regression of p_t on d_t , along with their 95% confidence bands. As shown in Figure 2, the coefficient on d_t , or the cointegrating vector, gradually declines, with a few large swings during the post World War II period for the aggregate stock market data. We note that the dynamics of the cointegrating vector are closely related to that of the fraction of type-I firms depicted in Figure 1. The tight relation between the time-varying cointegration parameter (α_t) and the fraction of type-I firms (ω_t) can be also ascertained by the regression of α_t on a constant and ω_t . The results are not reported to conserve space but are available upon request.

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 Proposition 3, which suggests that the adjusted dividend-price ratio ($\alpha_t d_t - p_t$), a stationary deviation from the time-varying long-run relationship, has predictive power on stock returns.

We first compare the time-series properties of the conventional dividend-price ratio ($d_t - p_t$) and $\alpha_t d_t - p_t$. Figure 3 plots the two series. Summary statistics for these series are provided in Table 3. As in the literature, the conventional dividend-price ratios 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 reveal much less persistence than $d_t - p_t$. Furthermore, even when the bootstrapped p-values are used to incorporate the pre-testing procedure to identify the time-varying cointegration parameter, the null hypothesis of a unit root is rejected regardless of the firm sizes. The only exception is $\alpha_t d_t - p_t$ for large firms.¹⁰

The predictive regression we employ is given by:

$$r_{t+1,t+k} = \beta_0 + \beta_1 x_t + \varepsilon_{t+1,t+k}, \quad (13)$$

where $r_{t+1,t+i}$ is the cumulative stock return at between time $t+1$ and $t+i$, $i=1,2,\dots,k$; x_t is either $d_t - p_t$ or $\alpha_t d_t - p_t$. The estimates of the β_1 coefficient reported in Table 4 are

¹⁰ Even though the unit root null hypothesis is not rejected by the bootstrapped p-values for large firms, $\alpha_t d_t - p_t$ is still much less persistent than $d_t - p_t$.

based on OLS regressions of equation (13). Based on the p-values from the Hodrick t-statistics¹¹ or the bootstrapped p-values, we find only weak evidence of in-sample predictability of the conventional dividend price ratio, even though the evidence is stronger for large firms and aggregate returns. However, the adjusted dividend-price ratio provides strong evidence of in-sample predictability for all firm sizes and at all horizons. Furthermore, bootstrapped p-values for a joint test of return predictability for one through six-year horizons also suggest strong predictive power of the adjusted dividend-price ratio.

4.2. Out-of-Sample Analysis of Return Predictability

Goyal and Welch (2008) and Bossaerts and Hillion (1999) suggest that the presence of strong in-sample prediction performance does not necessarily lead to significant out-of-sample prediction performance. We thus compare the out-of-sample performances of the predictive regressions that employ $d_t - p_t$ or $\alpha_t d_t - p_t$, relative to that of the random walk with drift model.

Out-of-sample predictions with $\alpha_t d_t - p_t$ are formed based on the following regressions:

$$p_t = \alpha_t d_t + u_t \quad t=1,2,\dots,t_1, \quad (14)$$

$$r_{t-k+1,t} = \beta_{0,t} + \beta_{1,t}(\alpha_{t-k} d_{t-k} - p_{t-k}) + \varepsilon_{t-k+1,t}, \\ t = t_0 + k, t_0 + k + 1, \dots, t_1 \quad (15)$$

Based on equation (14), we first estimate the time-varying vector α_t and estimate the adjusted dividend-price ratio ($\hat{\alpha}_t d_t - p_t$) for $t=1,2,\dots,t_1$. We then estimate equation (15) via the ordinary least squares method using the estimated adjusted dividend-price ratio. Finally, we form the out-of-sample prediction by

$$\hat{r}_{t_1+1,t_1+k} = \hat{\beta}_{0,t} + \hat{\beta}_{1,t}(\alpha_{t_1} d_{t_1} - p_{t_1}). \quad (16)$$

We repeat the above procedure for $t_1 = 1975.12 - 2008.12$. Notice that we run recursive regressions for both regression equations (14) and (15). The first predictive regression is run using 30-year's monthly data that starts in January 1946. That is, the first $t_1 = 1975.12$, adding one month observations at a time for the estimation of parameters and for the prediction of stock returns.

We compare the out-sample performance of the predictive regression with $\alpha_t d_t - p_t$

¹¹ Ang and Bekeart (2007) report that the Hodrick standard errors perform better than the Newey-West standard errors in the long-horizon predictive regressions. With the use of the Hodrick standard errors, p-values based on the asymptotic distribution and bootstrapped distribution are quite close to each other, which is consistent with Ang and Bekeart (2007).

(or $d_t - p_t$) against that of a random walk with drift model. The Clark and West (2007) test is employed for this purpose. The null hypothesis is that two competing forecasting models have equal predictive power. Table 5 reports the test results. We construct the Clark-West test statistic so that it has a significantly positive sign if the regression model with $d_t - p_t$ or $\alpha_t d_t - p_t$ as a regressor has superior predictive power to the random walk model. When $d_t - p_t$ is employed in the predictive regression, the Clark-West test statistics are significant only for medium and small firms, while they are insignificant for large firms and for aggregate returns at almost all horizons. The results with aggregate stock returns are generally consistent with those of Goyal and Welch (2008).

However, when $\alpha_t d_t - p_t$ is employed as a regressor in the predictive regressions, the Clark-West statistics are significant at the 10% level or above at one-year through five-year horizons for large-sized firms and the aggregate returns. For small and medium firms, they are statically significant at all horizons considered in this study. In summary, the model that employs $\alpha_t d_t - p_t$ as a predictor significantly outperforms the random walk model, in terms of out-of-sample predictability. These results are robust with respect to the firm size.

4.3. Comparison of Predictive Power: Adjusted Dividend-Price Ratio and Total Payout Ratio

Boudoukh *et al.* (2007) and Robertson and Wright (2006) document that, as part of the changes in firms' payout policy in the US since the early 1980s, many firms have replaced dividend payments with share repurchases. By defining the payout as the sum of dividends and share repurchases, they show that the payout yield or the payout ratio has strong and stable predictive power for future stock returns. In this section, we examine the predictive power of the payout ratio as employed by Boudoukh *et al.* (2007), and the results are compared to the case of the adjusted dividend price ratio considered in this paper.

By employing the data on net payout yield,¹² which are constructed by Boudoukh *et al.* (2007), we estimate the predictive regressions. The in-sample results are shown in the first panel of Table 6. At horizons longer than two-year, the coefficients on net payout yield (np_y_t) are not significant at the 5% significance level. Furthermore, the joint hypothesis that the net

¹² Net payout yield is defined as the sum of dividend yield and net repurchase yield. Note that net repurchases identify the repurchases that are not related to employees' stock option exercise. Refer to Boudoukh *et al.* (2007) for the exact definition of the net payout yield. The data for the net payout yield is obtained from Michael Roberts' homepage (<http://finance.wharton.upenn.edu/~mrrrobert/>).

payout yield has no predictive power at one-year through six-year horizons is not rejected. These results are consistent with Boudoukh *et al.* (2007). However, note that the coefficients on the adjusted dividend-price ratio ($\alpha_t d_t - p_t$) are significant at the 1% significance level at all horizons considered. The joint hypothesis that the adjusted dividend-price ratio has no predictive power at one year through 6 year horizons is rejected at the 1% significance level. That is, the adjusted dividend-price ratio has better in-sample predictive power than the net payout yield. The comparisons of the out-of-sample predictive performances are shown in the lower panel of Table 6. Note that both the net payout yield and the adjusted dividend-price ratio seem to have comparable predictive power for future stock returns at all horizons considered.

The results in this section suggest that the adjusted dividend-price ratio considered in this paper does as good a job as the net payout yield, in reflecting gradual changes in the payout policy and in forecasting future stock returns. Note that not all firms repurchase shares as dividend payment replacement and some firms skip dividend payment in order to have more retained earnings. Under these situations, the total payout ratio or the net payout ratio is subject to measurement errors. We believe that the approach presented in this paper (based on the adjusted dividend-price ratio) is less subject to these concerns.

5. Summary and Conclusion

We present both the theoretical and the 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. The theoretical model that we develop suggests that the parameter describing the long-run relationship is time-varying and dependent upon the fraction of firms with traditional payout policy.

A failure to consider this time-varying long-run relationship results in highly persistent dynamics in the conventional dividend-price ratio. The adjusted dividend-price ratio that accounts for the time-varying long-run relationship is stationary with much less persistence than the conventional dividend-price ratio, and has predictive power on the stock returns at long horizons. That is, we show that the predictive regression model that employs the adjusted dividend-price ratio outperforms the random walk model or the log net payout yield as in Boudoukh *et al.* (2007) , both in and out of the sample.

References

- Allen, F. and R. Michaely (2003) “Payout Policy” in *North-Holland Handbook of Economics* edited 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.
- Ang, A., and G. Bekeart (2007) “Stock return predictability: Is it there?” *Review of Financial Studies*, 20, 651-707.
- 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 (2010) “Time Varying Cointegration,” *Econometric Theory*, 26, 1453-1490.
- 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.
- 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.
- Clark, T. E., and K. D. West (2007) "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics*, 138, 291-311.
- 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.
- 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.
- Lettau, M., and S. V. Nieuwerburgh (2008) "Reconciling the return predictability evidence," *Review of Financial Studies*, 21, 1607-1652.
- 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.
- Torous, W., R. Valkanov, and S. Yan (2005) "On predicting stock returns with nearly integrated explanatory variables," *Journal of Business*, 77, 937-966.

Appendix 1: Proof of the Propositions

Proof of Proposition 1:

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

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

Since $d_{1,t}$, $d_{2,t}$, and $\lambda_{2,t}$ are all I(1), $d_t - p_t$ contains an I(1) component unless $\omega_t = \widehat{\omega}_t = 1$.¹³ \square

Proof of Proposition 2:

$$\begin{aligned}
 d_t &\approx \widehat{\omega}_t d_{1,t} + (1 - \widehat{\omega}_t) d_{2,t} = \widehat{\omega}_t \gamma_1 \cdot m_t + (1 - \widehat{\omega}_t) \gamma_2 \cdot m_t + \varepsilon_t \\
 &= (\widehat{\omega}_t \gamma_1 + (1 - \widehat{\omega}_t) \gamma_2) m_t + \varepsilon_t \text{ where } \varepsilon_t = \widehat{\omega}_t \varepsilon_{1,t} + (1 - \widehat{\omega}_t) \varepsilon_{2,t}
 \end{aligned}$$

$$\text{Hence, } m_t \approx \frac{d_t}{(\widehat{\omega}_t \gamma_1 + (1 - \widehat{\omega}_t) \gamma_2)} + \tilde{u}_t \text{ where } \tilde{u}_t = \frac{-\varepsilon_t}{(\widehat{\omega}_t \gamma_1 + (1 - \widehat{\omega}_t) \gamma_2)}$$

$$\begin{aligned}
 \text{Then, } \delta_t - p_t &\approx \omega_t (d_{1,t} - p_{1,t}) + (1 - \omega_t) (\lambda_{2,t} - p_{2,t}) \\
 &= \omega_t \gamma_1 m_t + (1 - \omega_t) \gamma_3 m_t - p_t + \tilde{\varepsilon}_t \\
 &= (\omega_t \gamma_1 + (1 - \omega_t) \gamma_3) m_t - p_t + \tilde{\varepsilon}_t \\
 &= \frac{(\omega_t \gamma_1 + (1 - \omega_t) \gamma_3)}{(\widehat{\omega}_t \gamma_1 + (1 - \widehat{\omega}_t) \gamma_2)} d_t - p_t + \hat{u}_t \\
 &= \alpha_t d_t - p_t + \hat{u}_t
 \end{aligned}$$

where $\tilde{\varepsilon}_t = \omega_t \varepsilon_{1,t} + (1 - \omega_t) \varepsilon_{3,t}$ and $\hat{u}_t = \tilde{\varepsilon}_t + \tilde{u}_t$. Since $\delta_t - p_t$ and \hat{u}_t are I(0), $p_t = \alpha_t d_t + u_t$ where $u_t = \hat{u}_t - (\delta_t - p_t)$. \square

Proof of Proposition 3:

¹³ $\widehat{\omega}_t$ equals one when all firms are of type-I.

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

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

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

$$\begin{aligned} r'_{t+1+j} &= \log(P_{1,t+1+j} + P_{2,t+1+j} + D_{1,t+1+j} + \Lambda_{2,t+1+j}) - \log(P_{1,t+j} + P_{2,t+j}) \\ &\approx \bar{\omega}_{t+j} \log(P_{1,t+1+j} + D_{1,t+1+j}) + (1 - \bar{\omega}_{t+j}) \log(P_{2,t+1+j} + \Lambda_{2,t+1+j}) \\ &\quad - \bar{\omega}_{t+j} \log(P_{1,t+j}) - (1 - \bar{\omega}_{t+j}) \log(P_{2,t+j}) \\ &= \bar{\omega}_{t+j} [\log(P_{1,t+1+j} + D_{1,t+1+j}) - \log(P_{1,t+j})] \\ &\quad + (1 - \bar{\omega}_{t+j}) [\log(P_{2,t+1+j} + \Lambda_{2,t+1+j}) - \log(P_{2,t+j})] \\ &\approx \bar{\omega}_{t+j} r_{1,t+1+j} + (1 - \bar{\omega}_{t+j}) [\log(P_{2,t+1+j} + \Lambda_{2,t+1+j}) - \log(P_{2,t+j})] \\ &\quad + (1 - \bar{\omega}_{t+j}) [\log(P_{2,t+1+j} + D_{2,t+1+j}) - \log(P_{2,t+j} + D_{2,t+1+j})] \\ &= \bar{\omega}_{t+j} r_{1,t+1+j} + (1 - \bar{\omega}_{t+j}) [\log(P_{2,t+1+j} + D_{2,t+1+j}) - \log(P_{2,t+j})] \\ &\quad + (1 - \bar{\omega}_{t+j}) [\log(P_{2,t+1+j} + \Lambda_{2,t+1+j}) - \log(P_{2,t+j} + D_{2,t+1+j})] \\ &= \bar{\omega}_{t+j} r_{1,t+1+j} + (1 - \bar{\omega}_{t+j}) r_{2,t+1+j} \\ &\quad + (1 - \bar{\omega}_{t+j}) [\rho p_{2,t+1+j} + (1 - \rho) \lambda_{2,t+1+j} - \log(P_{2,t+1+j} + D_{2,t+1+j})] \\ &= r_{t+1+j} + (1 - \bar{\omega}_{t+j}) [(1 - \rho) (\lambda_{2,t+1+j} - p_{2,t+1+j})] \\ &\quad + (1 - \bar{\omega}_{t+j}) [p_{2,t+1+j} - p_{2,t+j} - \log(P_{2,t+1+j} + D_{2,t+1+j}) + \log(P_{2,t+j})] \\ &= r_{t+1+j} + (1 - \bar{\omega}_{t+j}) [(1 - \rho) (\lambda_{2,t+1+j} - p_{2,t+1+j}) + \Delta p_{2,t+1+j} - r_{2,t+1+j}] \end{aligned} \quad (\text{A.1})$$

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

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

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

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.

$$p_t = \alpha_{0t} + \alpha_{1t}d_t = \alpha_0 + \alpha_{1t}d_t + u_t \quad (\text{A.3})$$

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 $\alpha_{1t} = \alpha\left(\frac{t}{T}\right)$ so that α_{1t} is a smooth function defined on $[0, 1]$.¹⁴ While estimating, 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 $\|\alpha_\kappa(s) - \alpha(s)\| \rightarrow 0$ as $\kappa \rightarrow \infty$, where $\alpha_\kappa(s)$ is an approximation of $\alpha(s)$ given by a combination of a finite series of functions $\varphi_1, \dots, \varphi_\kappa$. Since $\alpha_\kappa(s) = \sum_{i=1}^{\kappa} \theta_i \varphi_i$, the above econometric model can be expressed as:

$$\begin{aligned} p_t &= \alpha_0 + \alpha_{1t}d_t + u_t = \alpha_0 + \alpha\left(\frac{t}{T}\right)d_t + u_t \\ &= \alpha_0 + \left[\sum_{i=1}^{\kappa} \theta_i \varphi_i\left(\frac{t}{T}\right) \right] \cdot d_t + u_{kt} \end{aligned}$$

$$= \alpha_0 + x'_{kt}a_\kappa + u_{kt} \quad (\text{A.4})$$

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

If p_t and d_t are stationary series, then we can establish the asymptotic normality of the 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-square tests. Hence, we have made CCR transformation for p_t and d_t as follows.

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

and

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

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 the Δ matrix. Through the CCR transformation, Equation (2) can be written as

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

where $x_{kt}^* = [\varphi_1\left(\frac{t}{T}\right), \dots, \varphi_\kappa\left(\frac{t}{T}\right)]' d_t^*$. We can derive the asymptotic normality of the LS estimator for a_κ under this transformation. Once the LS estimator for a_κ is obtained, then

¹⁴When 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.

$\alpha(s)$ can be approximated by $\widehat{\alpha}_\kappa = \sum_{i=1}^{\kappa} \widehat{\theta}_i \varphi_i$.

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

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

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 traditional, 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 cointegrating 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:

$$d_t - p_t = \alpha_j + \tau_j(d_{t-1} - p_{t-1}) + \zeta_{1,j}\Delta(d_{t-1} - p_{t-1}) + \dots + \zeta_{p,j}\Delta(d_{t-p+1} - p_{t-p+1}) + \varepsilon_t ,$$

$$t = 1, 2, \dots, T^* + j \text{ and } j = 0, 1, 2, \dots, J$$

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:

$$\tau_j = c_{\tau,0} + c_{\tau,1}\omega_j + e_{\tau,j}$$

$$adf_j = c_{adf,0} + c_{adf,1}\omega_j + e_{adf,j}$$

$$j = 0, 1, 2, \dots, J,$$

where adf_j is the ADF t-statistic. T-statistics are reported in parentheses, and ‘***’, and ‘**’ denote the significance level at the 5%, and 1% level, respectively.

	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 0.31, 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
 Summary Statistics for
 Dividend-Price Ratio ($d_t - p_t$) and Adjusted Dividend-Price Ratio ($\alpha_t d_t - p_t$)

This table shows summary statistics and the ADF test results for $d_t - p_t$ and $\alpha_t d_t - p_t$. The numbers in parentheses show bootstrapped p-values for the augmented Dickey-Fuller test to incorporate the Park and Hahn (1999) pre-testing procedure. 5,000 bootstraps are conducted. ‘*’, ‘**’, and ‘***’ denote the significance level at the 10%, 5%, and 1% level, respectively.

	Mean	Standard Deviation	Autocorrelation	ADF test statistics
Aggregate firms				
$d_t - p_t$	-5.9278	0.4258	0.9931	-1.5387 (0.4040)
$\alpha_t d_t - p_t$	-4.1922	0.1322	0.9373	-5.4449 (0.0596)*
Large firms				
$d_t - p_t$	-5.9068	0.4418	0.9935	-1.4744 (0.5426)
$\alpha_t d_t - p_t$	-3.7812	0.1332	0.9422	-5.1604 (0.1874)
Medium firms				
$d_t - p_t$	-6.0794	0.6061	0.9944	-1.1921 (0.5158)
$\alpha_t d_t - p_t$	-2.3230	0.1578	0.9364	-5.7193 (0.0038)***
Small firms				
$d_t - p_t$	-6.3822	0.6901	0.9944	-1.4138 (0.4366)
$\alpha_t d_t - p_t$	-4.1920	0.1900	0.9410	-5.7262 (0.0008)***

Table 4 Predictive Regressions: In-Sample Analysis

The predictive regression we employ is given by: $r_{t+1,t+k} = \beta_0 + \beta_{1,k}x_t + \varepsilon_{t+1,t+k}$, where $r_{t+1,t+i}$ is the cumulative stock return at between time t+1 and t+i, $i=1,2,\dots,k$; x_t is either the conventional dividend-price ($d_t - p_t$) ratio or the adjusted dividend-price ratio ($\alpha_t d_t - p_t$). The Table reports $\beta_{1,k}$. The numbers in parentheses show p-values for t-tests based on the Hodrick standard errors. Bootstrapped p-values are reported in the brackets. 5,000 bootstraps are conducted. ‘***’, and ‘****’ denote the significance level at the 5%, and 1% level, respectively. The χ^2 test shows bootstrapped p-values for a test that $\beta_{1,12} = \beta_{1,24} = \beta_{1,36} = \beta_{1,48} = \beta_{1,60} = \beta_{1,72} = 0$.

		12 months	24 months	36 months	48 months	60 months	72 months	χ^2 test
Aggregate	$d_t - p_t$	0.0963 (0.0420)** [0.0376]**	0.1741 (0.0677) [0.0700]	0.2456 (0.0864) [0.0876]	0.3131 (0.0989) [0.0930]	0.4100 (0.0890) [0.0842]	0.5075 (0.0835) [0.0796]	0.0780
	$\alpha_t d_t - p_t$	0.6917 (0.0000)*** [0.0000]***	1.0486 (0.0001)*** [0.0000]***	1.1680 (0.0006)*** [0.0000]***	1.1369 (0.0031)*** [0.0050]***	1.2946 (0.0017)*** [0.0034]***	1.3172 (0.0016)*** [0.0024]***	0.0002****
Large	$d_t - p_t$	0.0896 (0.0472)** [0.0438]**	0.1695 (0.0638) [0.0592]	0.2472 (0.0733) [0.0732]	0.3229 (0.0761) [0.0720]	0.4222 (0.0677) [0.0660]	0.5231 (0.0611) [0.0600]	0.0768
	$\alpha_t d_t - p_t$	0.6371 (0.0000)*** [0.0000]***	1.0169 (0.0001)*** [0.0004]***	1.1784 (0.0006)*** [0.0010]***	1.2183 (0.0022)*** [0.0040]***	1.3740 (0.0011)*** [0.0008]***	1.4006 (0.0010)*** [0.0010]***	0.0006****
Medium	$d_t - p_t$	0.0591 (0.1115) [0.1222]	0.0984 (0.1933) [0.2010]	0.1401 (0.2236) [0.2172]	0.1737 (0.2731) [0.2616]	0.2284 (0.2712) [0.2588]	0.2885 (0.2716) [0.2556]	0.1910
	$\alpha_t d_t - p_t$	0.6965 (0.0000)*** [0.0000]***	0.9948 (0.0000)*** [0.0000]***	1.1399 (0.0000)*** [0.0000]***	1.0582 (0.0007)*** [0.0000]***	1.2200 (0.0010)*** [0.0008]***	1.1482 (0.0014)*** [0.0026]***	0.0000****
Small	$d_t - p_t$	0.0521 (0.1374) [0.1382]	0.0972 (0.1641) [0.1658]	0.1328 (0.2123) [0.2146]	0.1570 (0.2782) [0.2712]	0.1955 (0.2980) [0.2866]	0.2316 (0.3234) [0.3138]	0.2482
	$\alpha_t d_t - p_t$	0.6230 (0.0000)*** [0.0192]***	0.9912 (0.0000)*** [0.0000]***	1.1833 (0.0000)*** [0.0196]**	1.1655 (0.0002)*** [0.0000]***	1.3472 (0.0001)*** [0.0206]**	1.3818 (0.0002)*** [0.0000]***	0.0000****

Table 5 Out-of-Sample Tests to Forecast Multiperiod Stock Returns

Out-of-sample predictions are formed based on the following regressions: $p_t = \alpha_t d_t + u_t$ for $t=1,2,\dots, t_1$, and $r_{t-k+1,t} = \beta_{0,t} + \beta_{1,t}(\alpha_{t-k} d_{t-k} - p_{t-k}) + \varepsilon_{t-k+1,t}$, for $t = 1+k, 1+k+1, \dots, t_1$. Based on the first equation, we first estimate the time-varying vector α_t and estimate the adjusted dividend-price ratio ($\hat{\alpha}_t d_t - p_t$) for $t=1,2,\dots,t_1$. We then estimate the second equation via the OLS method using the estimated adjusted dividend-price ratio. Finally, we form the out-of-sample prediction by $\hat{r}_{t_1+1,t_1+k} = \hat{\beta}_{0,t} + \hat{\beta}_{1,t}(\alpha_{t_1} d_{t_1} - p_{t_1})$. We repeat the above procedure for $t_1 = 1975.12 - 2008.12$. Notice that we run recursive regressions for both regression equations. The first predictive regression is run using 30-year's monthly data that starts in January 1946. That is, the first $t_1 = 1975.12$, adding one month observations at a time for the estimation of parameters and for the prediction of stock returns. The Clark-West Test statistics (Clark and West (2007)) is used to compare the forecast ability. A significant positive value of the Clark-West Test statistics indicates that the first model (the predictive regression with either $d_t - p_t$ or $\alpha_t d_t - p_t$) has superior predictive ability to the second model (the random walk model). ‘**’, and ‘***’ denote the significance level at the 5%, and 1% level, respectively.

		12 months	24 months	36 months	48 months	60 months	72 months
Aggregate	$d_t - p_t$ vs random walk	1.6692**	1.5462	1.2262	1.0944	1.1877	1.2170
	$\alpha_t d_t - p_t$ vs random walk	1.7990**	1.9558**	1.7847**	1.5257	1.4974	1.0515
Large	$d_t - p_t$ vs random walk	1.5197	1.3987	1.1852	1.0610	1.1097	1.3174
	$\alpha_t d_t - p_t$ vs random walk	1.7842**	1.8615**	1.7598**	1.5721	1.3972	1.1456
Medium	$d_t - p_t$ vs random walk	3.8114***	3.9541***	2.7721***	1.9786**	1.9548**	1.5282
	$\alpha_t d_t - p_t$ vs random walk	4.0068***	4.9868***	3.6473***	2.0605**	2.3679**	2.2657**
Small	$d_t - p_t$ vs random walk	4.4303***	5.5917***	4.1498***	3.2060***	3.5403***	4.0691***
	$\alpha_t d_t - p_t$ vs random walk	4.0092***	5.7645***	5.2026***	3.1238***	4.1543***	5.2505***

Table 6 Predictability Comparison with the Net Payout Yield

This table compares the predictive power of the adjusted dividend-price ratio with the total net payout yield (npv_t) in Boudoukh *et al.* (2007). The in-sample analysis and out-of-sample analysis are conducted by the same ways in Tables 4 and 5, respectively. The upper panel of the table reports $\beta_{1,k}$. The numbers in parentheses show p-values for t-tests based on the Hodrick standard errors. Bootstrapped p-values are reported in the brackets. 5,000 bootstraps are conducted. The χ^2 test shows bootstrapped p-values for a test that $\beta_{1,12} = \beta_{1,24} = \beta_{1,36} = \beta_{1,48} = \beta_{1,60} = \beta_{1,72} = 0$. The lower panel of the table reports the Clark-West Test statistics. A positive value of the Clark-West Test statistics indicates that the first model (the predictive regression with either npv_t or $\alpha_t d_t - p_t$) has superior predictive ability to the second model (the random walk model). Both analyses are conducted with aggregate stock returns due to the availability of the total net payout yield. ‘***’, and ‘**’ denote the significance level at the 5%, and 1% level, respectively. The results for $\alpha_t d_t - p_t$ are from the results for the aggregate stock market in Tables 5 and 6.

		12 months	24 months	36 months	48 months	60 months	72 months	χ^2 test
In-sample	npv_t	0.1943 (0.0457)** [0.0542]*	0.3775 (0.0415)** [0.0484]**	0.4437 (0.0917) [0.1040]	0.5334 (0.1041) [0.1234]	0.6115 (0.1168) [0.1348]	0.6880 (0.1178) [0.1358]	0.0884
	$\alpha_t d_t - p_t$	0.6917 (0.0000)*** [0.0000]***	1.0486 (0.0001)*** [0.0000]***	1.1680 (0.0006)*** [0.0000]***	1.1369 (0.0031)*** [0.0050]***	1.2946 (0.0017)*** [0.0034]***	1.3172 (0.0016)*** [0.0024]***	0.0002***
Out-of-sample	npv_t	1.7953**	1.8735**	1.6516**	1.5509	1.5392	1.4338	
	$\alpha_t d_t - p_t$	1.7990**	1.9558**	1.7847**	1.5257	1.4974	1.0515	

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

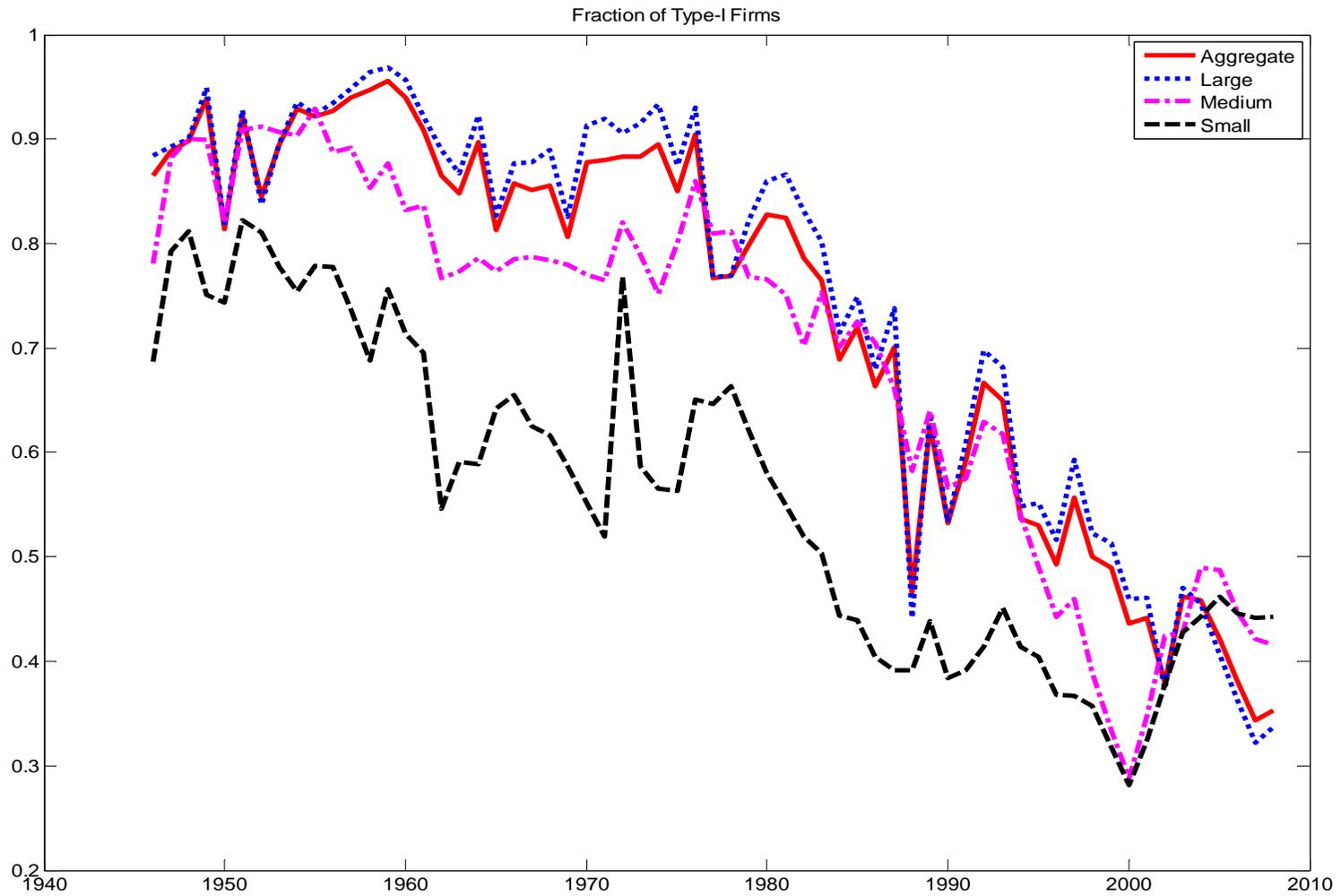


Figure 2. Time-varying Long-Run Relationship (cointegrating vector) between p_t and d_t

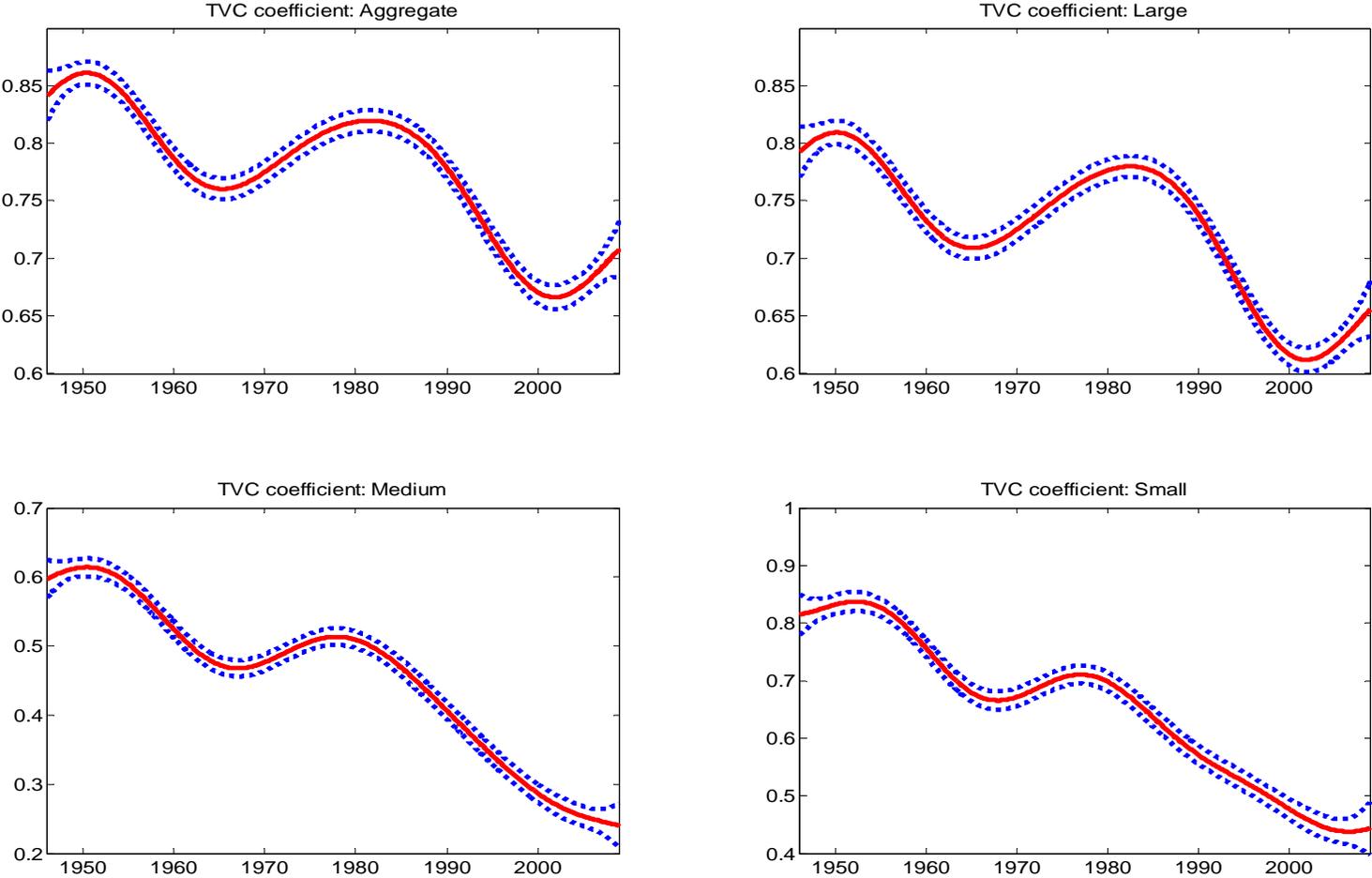


Figure 3. Conventional Dividend-Price Ratio vs. Time-Varying Dividend-Price Ratio

