The Disconnect Between Market Capital Gains and the Dividend Yield in Asset Pricing*

Michael Di Carlo **Toronto-Dominion Bank** Michael.DiCarlo@td.com

Jordi Mondria University of Toronto jordi.mondria@utoronto.ca itsiakas@uoguelph.ca

Ilias Tsiakas University of Guelph

September 2024

Abstract

We propose a two-factor Capital Asset Pricing Model (CAPM), which includes two separate factors for the market capital gains and the market dividend yield. We find that the dividend yield factor carries a significant negative premium in the post-1978 period, which coincides with the persistent decline in the number and proportion of US dividend-paying firms. We motivate this finding by proposing a theoretical model, which shows that the predictive information of the dividend yield can be high if capital gains are vastly more volatile than the dividend yield and investors have a behavioural bias against dividends.

Keywords: CAPM; Dividend Yield; Capital Gains; Dividend Disconnect; Factor Models.

JEL classification: G11; G12

^{*}Acknowledgments: The authors are grateful for useful comments to Ronald Balvers, Marie-Claude Beaulieu, Nikola Gradojevic, Delong Li, Fred Liu, Richard Luger, John Maheu, Alex Maynard, Gabriel Power and to seminar participants at the CRETE 2024 conference, the CEA 2022 meetings, the University of Guelph, Université Laval and McMaster University. This research was supported by the Social Sciences and Humanities Research Council of Canada (grant number 511952 for Jordi Mondria and grant number 430418 for Ilias Tsiakas). Part of this work was conducted while Ilias Tsiakas was visiting the University of Toronto. Declarations of interest: none. Corresponding author: Ilias Tsiakas, Department of Economics and Finance, Lang School of Business and Economics, University of Guelph, Guelph, Ontario N1G 2W1, Canada. Tel: 519-824-4120 ext. 53054. Email: itsiakas@uoguelph.ca.

1 Introduction

Financial theory is typically based on the idea that investors care about returns but are indifferent about whether they receive them through capital gains or dividends. This also holds for the Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965) and Black (1972), which makes no distinction about whether the market return is due to capital gains or the dividend yield. In the absence of taxes and other frictions, this idea is economically sound and relates to the dividend irrelevance of Miller and Modigliani (1961) that value-maximizing investors are indifferent about the source of their returns.

In practice, however, Hartzmark and Solomon (2019) find that investors track price changes and dividends separately rather than combine them into total returns. This creates a disconnect between price changes and dividends as investors perceive gains and losses to be primarily driven by price changes and tend to ignore dividends in assessing performance. The disconnect may be exacerbated by the fact that popular stock indices, such as the Dow Jones Industrial Average and the S&P 500, display a biased view of performance because they do not allow for dividend reinvestment (Hartzmark and Solomon, 2022). Accordingly, the dividend disconnect implies that it is sensible to: (1) decompose the performance of portfolios into a capital gains component and a dividend yield component, and (2) allow for different stock betas with market capital gains and the market dividend yield.¹

¹ The idea of the dividend disconnect was first proposed by Shefrin and Statman (1984). See also Baker and Wurgler (2004a, 2004b), Baker, Nagel and Wurgler (2007), and Daniel, Garlappi and Xiao (2021).

In the context of the CAPM model, there is an additional reason why it may be sensible to decompose the market portfolio into its two components. The dividend yield makes a substantial contribution to the mean return: more than 20% of the market return is due to dividends. However, capital gains are vastly more volatile than dividends. The variance of market capital gains can be up to 1000 times higher than the variance of the market dividend yield. This has a profound statistical implication: bundling together two separate components, where one is vastly more volatile than the other one, implies that in regression analysis the volatile component will dominate. As a result, an asset's beta on the market portfolio is effectively the same as the beta on just the market capital gains. In practice, therefore, whether we include or exclude the dividend yield from the market portfolio makes no difference in the estimation of standard CAPM betas, and hence in assessing the risk of financial assets.

We propose a new two-factor CAPM model, where the market capital gains and the market dividend yield comprise two *separate* factors for evaluating the cross-section of expected stock returns. The two-factor CAPM model is designed to give a separate voice to the dividend yield that would otherwise be silenced in estimating the standard CAPM due to the vastly more volatile capital gains component. It also allows for the possibility that the risk associated with capital gains may be distinct from that associated with the dividend yield. To the best of our knowledge, this is the first paper to assess the asset pricing implications of the disconnect between market capital gains and the market dividend yield.

We motivate our empirical analysis by proposing a theoretical model, which shows that the predictive power of a signal about the dividend yield can be substantially higher than the predictive power of an additional signal about capital gains. The intuition of this argument is based on two assumptions. First, following Hartzmark and Solomon (2019), we assume that investors perceive gains and losses to be driven primarily by prices changes and, therefore, tend to ignore information about dividends. Second, we assume that investors process tremendous amounts of information about capital gains because they perceive capital gains as the primary determinant of price uncertainty due to their high variance. After investors take into account information about capital gains, the factor that generates more uncertainty is the dividend yield factor. Consequently, a signal about the dividend yield factor will have higher return predictability and will be more useful in reducing the remaining uncertainty about cash flows than an additional signal about capital gains. In summary, our theoretical model shows that the behavioural bias of investors to ignore dividends combined with the higher variance of capital gains can justify the predictive ability of the market dividend yield in the cross-section of expected stock returns.²

In our empirical setting, we define the new dividend yield factor as the innovation to the 12-month market dividend yield. By construction, therefore, the new dividend yield factor accounts for the strong seasonality and potential non-stationarity of the dividend yield. Specifically, the dividend yield exhibits strong quarterly seasonality. We account for this seasonality by using the 12-month dividend yield,

² Our theoretical model is loosely related to the framework of Peng and Xiong (2006).

which is a standard approach in the literature for avoiding biased results in estimating regression coefficients. Additionally, we account for non-stationarity by using the innovation to the 12-month dividend yield to avoid the effect of the high and long-lived persistence in the 12-month dividend yield.

Consistent with the theoretical model, our main empirical finding is that the new dividend yield factor has strong predictive power for the cross-section of expected stock returns. This is true primarily for the post-1978 sample period. The beginning of this period marks the peak in the number of US dividend-paying firms (Fama and French, 2001). Following this peak, both the number and the proportion of dividend payers declines steadily. This empirical observation is consistent with our theoretical model: the less important are dividends in the US economy, the more likely are investors to ignore them and, hence, the higher the predictive power of dividends in resolving uncertainty about future returns.

To be more specific, we find that, for the post-1978 period, exposure to the new dividend yield factor distinguishes clearly between high-performing stocks and low-performing stocks. The predictive power of the dividend yield factor: (1) is distinct from the market capital gains factor; (2) is also distinct from other standard risk factors; (3) is significant for both dividend-paying and non-dividend paying firms; and (4) is unrelated to *individual* dividend yields.

Consider, for example, the following evidence. For value-weighted quintile portfolios rebalanced monthly by sorting on the beta to the dividend yield factor, the High-minus-Low (H-L) portfolio delivers an expected return of -0.28% per month with a Newey and West (1987) t-statistic of -2.04. The six-factor alpha of the H-L portfolio is equal to -0.27% with a t-statistic of -2.38. The results are stronger for equally-weighted portfolios: the H-L mean return is equal to -0.36% with a t-statistic of -2.16 and the H-L alpha is equal to -0.71% with a t-statistic of -5.62. The significant negative premium of the dividend yield factor is confirmed in Fama and MacBeth (1973) regressions in the presence of standard asset pricing factors.

To understand why this premium is negative, it is essential to note that the dividend yield factor is strongly countercyclical. Consider an asset that has a high positive beta on the dividend yield factor, which implies that the asset performs well when the factor is high. Since the dividend yield factor is high in recessions, the asset performs well in recessions. According to standard asset pricing theory, investors do not require a high expected return to hold an asset that performs well when we need it the most (in the bad states of the world). Consequently, high-beta assets on the dividend yield factor will have low expected returns and vice versa.

The remainder of the paper is organized as follows. In the next section, we present our theoretical model. In Section 3, we describe the data on US stock returns. Our approach to pricing dividend yield risk is described in Section 4. In Section 5, we investigate whether exposure to the market dividend yield factor is related to individual dividend yields. In Section 6, we assess the cross-sectional performance of a version of the two-factor model that we term the "predictive CAPM." Finally, we conclude in Section 7. The Online Appendix provides additional details, robustness tests and results.

2 Theoretical Framework

2.1 A Two-Factor Model

We use a simple theoretical framework to guide our empirical analysis. Let us assume an economy populated by a representative agent. There is one risky asset with payoffs \tilde{v} and the following factor structure:

$$\tilde{\nu} = \beta_g \tilde{g} + \beta_y \tilde{y} + \tilde{f}, \tag{1}$$

where $\tilde{g} \sim N(0, \tau_g^{-1})$ represents the capital gains factor, $\tilde{y} \sim N(0, \tau_y^{-1})$ represents the dividend yield factor, and $\tilde{f} \sim N(0, \tau_f^{-1})$ represents the firm-specific factor.

For simplicity, and without loss of generality, we assume that all factors have a zero mean. The variances V(.) are equal to the inverse of the precisions: $V(\tilde{g}) = \tau_g^{-1}$, $V(\tilde{y}) = \tau_y^{-1}$ and $V(\tilde{f}) = \tau_f^{-1}$. We also assume that all factors are uncorrelated. The parameters β_g and β_y are the factor loadings for the capital gains and dividend yield factors, respectively. Based on the summary statistics of the data (to be discussed in the next section), we assume that the variance of the capital gains factor is much larger than the variance of the dividend yield factor: $V(\tilde{g}) \gg V(\tilde{y})$ or equivalently $\tau_g \ll \tau_y$.

2.2 Estimating the Model Through the CAPM

If econometricians estimate the model in Equation (1) using the CAPM framework, then they would estimate the following regression:

$$\tilde{v} = \beta(\tilde{g} + \tilde{y}) + \tilde{f}, \tag{2}$$

where $\tilde{g} + \tilde{y}$ is the market factor, β is the loading on the market factor and \tilde{f} is the error term in the regression. The estimate of the market factor loading is given by:

$$\hat{\beta} = \frac{Cov(\tilde{v}, \tilde{g} + \tilde{y})}{V(\tilde{g} + \tilde{y})} = \frac{V(\tilde{g})}{V(\tilde{g}) + V(\tilde{y})} \hat{\beta}_g + \frac{V(\tilde{y})}{V(\tilde{g}) + V(\tilde{y})} \hat{\beta}_y, \tag{3}$$

where Cov(.) represents the covariance between two random variables.³

In Equation (3), $\hat{\beta}$ is a weighted average of $\hat{\beta}_g$ and $\hat{\beta}_y$, where the weights depend on the factor variances. Since $V(\tilde{g}) \gg V(\tilde{y})$, $\hat{\beta}$ is very close in value to $\hat{\beta}_g$. Consequently, the capital gains factor dominates the dividend yield factor. In other words, the higher the $V(\tilde{g})$ relative to $V(\tilde{y})$ (or alternatively the lower the τ_g relative to τ_y), the less relevant the information contained in the dividend yield.

If $\hat{\beta}_g = \hat{\beta}_y$, \tilde{g} and \tilde{y} make the same contribution to \tilde{v} . In this case, the CAPM holds regardless of the fact that $V(\tilde{g}) \gg V(\tilde{y})$. If, however, $\hat{\beta}_g \neq \hat{\beta}_y$, \tilde{g} and \tilde{y} make distinct contributions to \tilde{v} . In this case, the higher the $V(\tilde{g})$ relative to $V(\tilde{y})$, the lower the weight on $\hat{\beta}_y$ and the more the CAPM is misspecified by ignoring the distinct information contained in the dividend yield. In short, if $\hat{\beta}_g \neq \hat{\beta}_y$ and $V(\tilde{g}) \gg V(\tilde{y})$, estimating the CAPM of Equation (2) ignores the information contained in the dividend yield, whereas estimating the true model described by Equation (1) does not.⁴

³ Note that $Cov(\tilde{g}, \tilde{y}) = 0$ because the two factors are assumed to be uncorrelated.

⁴ The fact that $V(\tilde{g}) \gg V(\tilde{y})$ is established by the summary statistics of the data discussed in the next section. The hypothesis that $\hat{\beta}_g \neq \hat{\beta}_y$ is assessed throughout our empirical analysis and is also motivated by the empirical findings of Hartzmark and Solomon (2022).

2.3 Predictability of the Dividend Yield Factor

In this section, we build a theoretical model for motivating the predictive ability of the dividend yield factor for stock returns in the context of the two-factor model of Equation (1). We consider a representative agent (investor), who regularly receives signals (i.e., new information) about capital gains and the dividend yield. We assume that the investor regularly pays attention to signals about the capital gains factor but ignores signals about the dividend yield factor. The main capital gains signal they observe is changes in the price. The main dividend yield signal they ignore is changes in the dividend that investors allocate higher attention to capital gains than dividends is consistent with the evidence in Hartzmark and Solomon (2019, 2022), who find that investors perceive gains and losses to be driven largely by price changes and, therefore, trade as if their gains and losses do not include dividends.

Specifically, the representative agent has access to a signal about the capital gains factor $\tilde{s}_g = \tilde{g} + \tilde{\varepsilon}_g$, where $\tilde{\varepsilon}_g \sim N(0, \tau_{\varepsilon g}^{-1})$ and a signal about the firm-specific factor $\tilde{s}_f = \tilde{f} + \tilde{\varepsilon}_f$, where $\tilde{\varepsilon}_f \sim N(0, \tau_{\varepsilon f}^{-1})$. Under these signals, we can calculate the posterior mean $\hat{v} = E[\tilde{v}|\tilde{s}_g, \tilde{s}_f]$ and posterior variance $V[\tilde{v}|\tilde{s}_g, \tilde{s}_f]$ using Bayesian updating.⁵ The posterior mean \hat{v} is given by:

⁵ We could also model the representative agent to observe another signal $\tilde{s}_y = \tilde{y} + \tilde{\varepsilon}_y$, where $\tilde{\varepsilon}_y \sim N(0, \tau_{\varepsilon y}^{-1})$ with $\tau_{\varepsilon y} \to 0$ or with $\tau_{\varepsilon y} \ll \tau_{\varepsilon g}$. In this scenario, the investor processes information about the two factors (\tilde{g}, \tilde{y}) , but information about capital gains is much more precise than information about the dividend yield. We have not added this additional signal because it unnecessarily complicates the model. Instead,

$$\hat{v} = E\left[\tilde{v}\big|\tilde{s}_g, \tilde{s}_f\right] = \frac{\beta_g \tau_{\varepsilon g}}{\tau_g + \tau_{\varepsilon g}} \tilde{s}_g + \frac{\tau_{\varepsilon f}}{\tau_f + \tau_{\varepsilon f}} \tilde{s}_f.$$
(5)

The investor also observes additional signals about the capital gains factor. These additional signals encompass all signals other than changes in the price. For example, the additional signals could be macroeconomic news about employment or interest rates, which indirectly affect stock prices. Although the investor regularly ignores the period-by-period changes in the dividend yield, they may occasionally pay attention to some signals about the dividend yield. An example of such a signal is described in Hartzmark and Solomon (2019): the dividend yield on a given period being such that the return excluding the dividend yield is negative but the return including the dividend yield is positive.

In this context, we measure the return predictive power of an *additional* signal about the capital gains factor $\tilde{m}_g = \tilde{g} + \tilde{\omega}_g$, where $\tilde{\omega}_g \sim N(0, \tau_{\omega}^{-1})$ and the return predictive power of a signal about the dividend yield factor $\tilde{m}_y = \tilde{y} + \tilde{\omega}_y$, where $\tilde{\omega}_y \sim N(0, \tau_{\omega}^{-1})$. Note that $\tilde{\omega}_g$ and $\tilde{\omega}_y$ have the same precision.

Following Peng and Xiong (2006), the return predictive power of an additional signal \tilde{m}_g about the capital gains factor can be measured by the following correlation:

$$\left|Corr(\tilde{v}-\hat{v}),\tilde{m}_{g}\right)\right| = \frac{\left|\beta_{g}\right|\tau_{g}}{\tau_{g}+\tau_{\varepsilon g}}.$$
(6)

we gain tractability without loss of generality by assuming that the representative investor ignores any information about the dividend yield as suggested by Hartzmark and Solomon (2019, 2022).

Similarly, the return predictive power of a signal \tilde{m}_y about the dividend yield factor can be measured by the following correlation:

$$|Corr(\tilde{v} - \hat{v}), \tilde{m}_{v})| = |\beta_{v}|.$$
(7)

Note that $\tilde{v} - \hat{v}$ is the posterior forecast error and hence it is a measure of the uncertainty surrounding \tilde{v} . The higher the correlation (in absolute value) of a signal with $\tilde{v} - \hat{v}$, the more informative the signal is.⁶

This framework allows us to show that a signal \tilde{m}_y about the dividend yield has more predictive power than an additional signal \tilde{m}_q about the capital gains factor when:

$$\frac{|\beta_g|\tau_g}{\tau_g + \tau_{\varepsilon g}} < |\beta_y| \Leftrightarrow (|\beta_y| - |\beta_g|)\tau_g + |\beta_y|\tau_{\varepsilon g} > 0.$$
(8)

Based on stock return data, we know that the precision of capital gains is substantially lower than the precision of the dividend yield, i.e., $\tau_g \ll \tau_y$. We also take the view that investors process tremendous amounts of information about capital gains so that signals about capital gains are highly precise, i.e., $\tau_{\varepsilon g} \gg 0$. The combination of a very low τ_g and a very high $\tau_{\varepsilon g}$ guarantees that $|\beta_y|\tau_{\varepsilon g}$ will be the dominant term and hence the inequality in Equation (8) above holds.

In words, if investors' information about the capital gains factor \tilde{g} is precise, the predictive power of a signal about the dividend yield factor \tilde{y} will be higher than that of an additional signal about \tilde{g} . Intuitively, in an effort to reduce uncertainty about cash flows, investors focus on \tilde{g} because this is the most uncertain (i.e., volatile) factor. As

⁶ Note that \tilde{v} is the realization at time t, and \hat{v} is the posterior forecast for time t conditional on time t-1 information. For notational simplicity, we avoid using time subscripts.

investors spend substantial resources to learn about \tilde{g} , $\tau_{\varepsilon g} \gg 0$. After all the collected information about \tilde{g} is taken into account, the factor that generates more uncertainty is \tilde{y} . Thus, a signal \tilde{m}_y about the dividend yield factor will have higher return predictability and will be more useful in reducing the remaining uncertainty about cash flows than an additional signal about capital gains \tilde{m}_g .

3 Data

3.1 Stock Returns

Our empirical analysis uses the cross-section of US stock returns obtained from the CRSP database. The cross-section includes all common stocks traded on the NYSE, AMEX and NASDAQ exchanges based on the following criteria: (1) the firm must have at least two years of accounting data in COMPUSTAT; (2) the firm must have at least 24 monthly return observations in the past five years; and (3) the book-to-market value (B/M) ratio for the previous fiscal year must be positive.

All data are monthly. For our main analysis, the sample ranges from January 1978 to December 2019. We have chosen 1978 as the year to mark the beginning of the sample period because it coincides with the peak in the number of dividend-paying firms, which occurs at the end of 1977. This is shown in Figure 1, which illustrates that beginning in 1978 both the number and the proportion of dividend payers declines steadily (see also Fama and French, 2001). Hence we argue that in the post-1978 period there is a persistent decline in the importance of dividends in the US equity market.

3.2 The New Dividend Yield Factor

In this section, we describe and motivate the new dividend yield factor. The valueweighted (VW) market return including dividends at time *t* is denoted by mkt_t , while excluding dividends it is denoted by $mktx_t$. The latter is the rate of change in capital gains since $mktx_t = p_t/p_{t-1} - 1$, where p_t is the VW market price. The dividend yield is defined as $dy_t = d_t/p_{t-1}$, where d_t is the VW sum of all dividends paid by all firms at time *t*. We compute the market dividend yield as $dy_t = mkt_t - mktx_t$.⁷

A crucial aspect of dy_t is its strong seasonal behaviour due to the fact that firms pay dividends in different months and at different frequencies. In our sample, 81.9% of the dividend-paying firms pay dividends on a quarterly basis. Some of these firms pay dividends on January-April-July-October, others on February-May-August-November, and yet others on March-June-September-December. In addition, 2.1% of the firms pay monthly dividends, 6.9% of the firms pay semi-annual dividends and, finally, 9% of the firms pay annual dividends. As a result, dy_t displays significant serial correlation at the lags of 1-12 months but very high serial correlation at the quarterly lags of 3, 6, 9 and 12 months.⁸ In estimating regression coefficients, it is essential to account for this seasonality to avoid biased results. The standard approach in the literature to account

⁷ Note that the dividend yield is related to the equity carry, which is defined as the scaled expected dividend yield minus the risk-free rate (see Koijen *et al.*, 2018). For a review of the recent literature on return and dividend growth predictability, see Sabbatucci (2022).

⁸ See Figure 1 of the Online Appendix.

for dividend seasonality is to use the trailing 12-month dividend yield defined as $dy12_t = \Sigma_{j=0}^{11} d_{t-j}/p_{t-1}$. For this reason, our main analysis focuses on $dy12_t$.⁹

In addition to seasonality, the dividend yield may also display non-stationarity due to its high persistence. For example, $dy12_t$ displays an autocorrelation of 0.99 at a 1-month lag, which decays slowly to a value of 0.88 at a 12-month lag.¹⁰ In this context, we compute the Augmented Dickey-Fuller statistic for $dy12_t$, which is equal to -0.78. Hence, the null hypothesis of a unit root (i.e., non-stationarity) cannot be rejected. To account for non-stationarity, we define $\Delta dy12_t$ as the innovation to the 12-month dividend yield, which is calculated as the proportional change in $dy12_t$: $\Delta dy12_t = \frac{dy12_t}{dy12_{t-1}} - 1$. $\Delta dy12_t$ reflects the changes in the market dividend yield at time t and certainly includes changes that have been pre-announced by individual firms.¹¹

3.3 Summary Statistics

Table B1 of the Online Appendix reports summary statistics and Table B2 the crosscorrelations. The main findings can be summarized as follows: (1) the dividend yield accounts for 21% of the market return (0.22% of 1.04% per month); (2) mkt_t has 30

⁹ We have also implemented an alternative way of accounting for seasonality by deseasonalizing the dividend yield using an ARMA model. The results remain qualitatively the same.

¹⁰ Again, see Figure 1 of the Online Appendix.

¹¹ Using the innovation to the dividend yield is consistent with the return decomposition of Campbell and Shiller (1988) into changes in expectations about future cash flows and future discount rates. It is also consistent with the Campbell (1996) implementation of the Merton (1973) ICAPM model.

times higher standard deviation than dy_t (4.40% vs 0.14%). Put differently, mkt_t has 988 times higher variance than dy_t ; (3) whereas mkt_t and $mktx_t$ are strongly cyclical, dy_t and $\Delta dy 12_t$ are strongly countercyclical. Indeed, $\Delta dy 12_t$ is on average *negative* in expansions but strongly *positive* in recessions; and (4) mkt_t and $mktx_t$ exhibit perfect correlation but the correlation between mkt_t and dy_t is equal to 0.10.

4 Dividend Yield Risk in the Cross-Section of Stock Returns

4.1 A Two-Factor CAPM Model

In this section, we propose a simple decomposition of the CAPM with two factors, one based on the capital gains to the market portfolio $(mktx_t)$, and one based on the dividend yield to the market portfolio $(\Delta dy 12_t)$. By explicitly accounting for the market dividend yield as a separate factor in the CAPM regression, we ensure that its contribution to asset pricing is not ignored. This is because in estimating the CAPM, whether we use mkt_t or $mktx_t$, the result is essentially the same since the average betas in the two cases are identical to the second decimal.¹² Furthermore, the quintile portfolios formed by sorting on the betas of either mkt_t or $mktx_t$ are practically identical. Accordingly, we propose the following two-factor model:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,c}(mktx_t - r_{f,t}) + \beta_{i,d}\Delta dy 12_t + \varepsilon_{i,t},$$
(9)

¹² If we estimate the beta on mkt_t for each stock using a five-year rolling window, the average beta across time and across stocks is equal to 1.107. If we replace mkt_t by $mktx_t$, the average beta is equal to 1.109. Hence, the two average betas are identical to the second decimal.

where $r_{i,t}$ is the return of asset *i* at time *t*, $r_{f,t}$ is the riskless rate at time *t*, $r_{i,t} - r_{f,t}$ is the excess return to asset *i* at time *t*, $\beta_{i,c}$ is the loading on the capital gains factor, $\beta_{i,d}$ is the loading on the dividend yield factor, and $\varepsilon_{i,t}$ is the random error term.

4.2 Portfolio Sorts

We begin our empirical analysis with standard portfolio sorts. First, we estimate the coefficient $\beta_{i,d}$ in Equation (9) for each stock in the cross-section using a five-year rolling window. Each month we form quintile portfolios by sorting stocks on $\beta_{i,d}$. We then compute the one-month ahead mean returns of the quintile portfolios and rebalance monthly. We report results for both value-weighted (VW) portfolios based on NYSE weights and for equally-weighted (EW) portfolios (see, e.g., Hou, Xue and Zhang, 2020).

Table 1 reports the performance of quintile portfolios sorted by exposure to the dividend yield factor. Our main finding is that there is a significant *negative* relation between expected returns and exposure to the dividend yield factor. For VW portfolios, the High-minus-Low zero-cost investment portfolio (denoted by H-L), provides a mean return of -0.28% per month, which is statistically significant: the Newey-West (1987) *t*-statistic is equal to -2.04. Furthermore, the H-L portfolio delivers a negative and significant six-factor (FF6) alpha (-0.27% with t-stat=-2.38), where the FF6 model incorporates the five Fama and French (2015) factors plus momentum. The results are stronger for EW portfolios since the H-L mean return is equal to -0.36% with a t-stat=-2.36. The EW six-factor alpha is equal to -0.71% with a t-stat=-5.62. In short, portfolios

with low exposure to the dividend yield factor consistently perform better than portfolios with high exposure.

4.3 Subsample Analysis

Our sample period begins in 1978 to mark the peak in the number of dividend payers and hence the beginning of the declining importance of dividends in the US equity market. In this section, we perform a subsample analysis to shed light on the pre-1978 versus post-1978 performance of the two factor model.

We report the subsample results in Table 2. Our main finding is that for the 1932-1977 period, the VW H-L return spread is equal to 0.08% and insignificant. In contrast, for the 1978-2019 period, the H-L return spread is -0.28% and significant. Therefore, the predictive power of the dividend yield factor is due to the post-1978 sample.

To provide a finer analysis, we also report results for sample periods beginning in 1963, 1968, 1973, 1978, 1983 and 1988. All these subsamples end in 2019. We find that, as the starting date moves forward, the H-L return spread tends to be higher (in absolute value) and more significant. Importantly, the six-factor alphas become significant from 1978 onwards for VW portfolios, which further enhances the importance of using the post-1978 sample period. Overall, these results provide an empirical justification for focusing on the post-1978 sample period.

4.4 Components of the Dividend Yield

The dividend yield itself has two components: dividends (in the numerator) and lagged prices (in the denominator). It is well known that the two components have different behaviour since dividends are issued by corporate management at a low frequency, whereas stock prices are the result of high-frequency trading by market participants. We test whether the dividend yield factor is driven by one or both of its components by decomposing $\Delta dy12$ as follows: $\Delta dy12_t \approx \Delta d12_t - \Delta p_{t-1}$, where $\Delta d12_t = \frac{d12_t}{d12_{t-1}} - 1$ is the year-over-year dividend growth rate, and $\Delta p_{t-1} = \frac{p_{t-1}}{p_{t-2}} - 1$ is the *lagged* mktx.¹³

4.5 Factor-Mimicking Portfolios

The market factor is a tradable portfolio but its two components are not. In this section, we address the non-tradability of mktx and Δ dy12 by constructing factor-mimicking portfolios (FMPs). As the FMP for mktx, we simply use mkt since the correlation between the two variables is equal to one and, as a result, the betas on either mkt or mktx are essentially the same.

For $\Delta dy12$, we generate the FMP by implementing the ordinary least squares (OLS) cross-sectional approach based on Lehmann and Modest (1988). This approach is described in detail in the Online Appendix. The FMP for $\Delta dy12$ is denoted by F $\Delta dy12$.

¹³ The decomposition of Δ dy12 into Δ p and Δ d12 is not exact. However, the two components explain 100% of the variance of Δ dy12: 90.5% for Δ p and 9.5% for Δ d12.

For $\Delta d12$ and Δp , we follow a similar approach to $\Delta dy12$: we form the OLS FMP portfolios, denoted by F $\Delta d12$ and F Δp .¹⁴

4.6 The Price of Dividend Yield Risk

In this section, we estimate two-stage Fama-MacBeth (1973) regressions using the full cross-section of stocks. In the first stage, we estimate the time-series beta coefficients for each stock using the following seven-factor model:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,1} (mktx_t - r_{f,t}) + \beta_{i,2} \Delta dy 12_t + \beta_{i,3} SMB_t + \beta_{i,4} HML_t + \beta_{i,5} RMW_t + \beta_{i,6} CMA_t + \beta_{i,7} MOM_t + \varepsilon_{i,t},$$
(10)

where SMB, HML, RMW and CMA are the Fama and French (1993, 2015) factors, and MOM is the Carhart (1997) momentum factor. Data on SMB, HML, RMW, CMA and MOM are obtained from Ken French's online data library. The betas are estimated using a rolling window of 5 years of monthly data.

In the second stage, we condition on the beta estimates and perform crosssectional estimation at each month *t* as follows:

$$r_{i,t} - r_{f,t} = \gamma_0 + \gamma_1 \hat{\beta}_{i,1,t-1} + \gamma_2 \hat{\beta}_{i,2,t-1} + \gamma_3 \hat{\beta}_{i,3,t-1} + \gamma_4 \hat{\beta}_{i,4,t-1} + \gamma_5 \hat{\beta}_{i,5,t-1} + \gamma_6 \hat{\beta}_{i,6,t-1} + \gamma_7 \hat{\beta}_{i,7,t-1} + \epsilon_{i,t}.$$
(11)

We collect the time-series of gamma estimates and report the mean as well as the Newey-West (1987) *t*-statistic. The mean of each gamma coefficient represents the risk premium associated with each risk factor. The results are reported in Table 3.

¹⁴ The correlations of the variables with their FMPs are reported in Table B3 of the Online Appendix.

Our main finding is that the premium on the dividend yield factor is negative and significant. Specifically, Δ dy12 exhibits a risk premium of -0.26% per month with a t-stat=-2.18. The risk premium is equal to -0.28% for the FMP with a t-stat=-1.84. It is interesting to note that the dividend yield factor has the highest statistical significance among the risk factors. Notably, the capital gains factor ($mktx_t - r_{f,t}$) and its FMP ($mkt_t - r_{f,t}$) display a small and insignificant risk premium. In conclusion, we find clear evidence that the dividend yield factor has a negative and statistically significant price of risk. This finding indicates that the dividend yield is disconnected from capital gains in asset pricing.

To understand why the dividend yield factor carries a negative premium, recall that this factor is strongly countercyclical. If an asset has a high positive beta on the Δ dy12 factor, then it will perform well when Δ dy12 is high. However, Δ dy12 is high in recessions. Therefore, this asset performs well in recessions. Since the asset performs well when we need it the most (in the bad states of the world), investors do not require a high expected return to hold it. As a result, high-beta assets on the Δ dy12 factor will have low expected returns and vice versa. This provides an explanation for the negative premium that is consistent with asset pricing theory.

In Panel B of Table 3, we replace $\Delta dy12$ by its two components: $\Delta d12$ and Δp . We find that the premium for Δp is high (0.26%) and significant (t-stat=2.26) but the premium for $\Delta d12$ is much lower (0.03%) and less significant (t-stat=1.69). These premia remain similar in value for the FMPs but their significance drops slightly. In short, therefore, between the two components of the dividend yield factor, it is the lagged

capital gains factor that remains strong and significant in pricing the cross-section of expected stock returns.

4.7 **Post-Formation Factor Loadings**

In this section, we assess the contemporaneous relation between factor loadings and expected returns. Following a long line of research in asset pricing (see, e.g., Black *et al.*, 1972, Fama and French, 1992, 1993, and Ang *et al.*, 2006), we use pre-formation loadings to form portfolios, and then proceed to examine contemporaneous post-formation loadings. Specifically, we use the FMP for F Δ dy12 to compute post-formation loadings reported on the last column of Table 1. The port-formation loadings are estimated expost for the full data sample using the seven factor model of Equation (10).

The results show that, for all EW portfolios, the quintile portfolio returns load significantly on F Δ dy12 being negative and significant at the 1% level. Importantly, the post-formation loadings for EW portfolios consistently increase (i.e., decrease in absolute value) as we move from the Low to the High portfolio. For VW portfolios, three of the five quintile portfolios exhibit a significant post-formation loading. These results establish that average returns are related to the unconditional covariance between returns and market dividend yield risk for EW portfolios but for VW portfolios the results are slightly weaker.¹⁵

¹⁵ Note that we use individual firms as base assets to construct an FMP that is maximally correlated to the original factor. Consequently, smaller firms may receive FMP weights, which are larger than their

5 The Role of Individual Dividend Yields

In this section, we address two questions: (1) is exposure to the dividend yield factor relevant for both dividend-payers and non-dividend payers?; and (2) for dividend-payers, are portfolios sorted on exposure to the dividend yield factor related to portfolios sorted on the individual dividend yield?

5.1 Dividend Payers vs Non-Dividend Payers

We begin by first separating firms into dividend payers and non-dividend payers, and then re-estimating the Fama-MacBeth regressions for the two separate groups. Dividend payers are identified as firms, which at time t have paid a dividend in any month from time t to time t-11. The remaining firms are labelled as non-payers. On average, 46.5% of firms are identified as dividend payers and 53.5% are identified as non-payers. The results are reported in Table 4.

The results indicate that the predictive power of the F Δ dy12 is significant for both dividend payers and non-dividend payers but it is stronger for dividend payers. Using the original Δ dy12 series, the factor premium is -0.41% for payers (t-stat=2.92) versus -0.25% for non-payers (t-stat=-2.41). The results are similar for FMPs. It is interesting to note that the significance of the dividend yield factor rises when we divide the cross-section in dividend payers and non-payers. In conclusion, the price of

weights in a value-weighted portfolio. This can explain the limited variation in post-formation betas for value-weighted portfolios. It can also explain the large variation for equally-weighted portfolio loadings.

risk for the dividend yield factor is significantly negative for both dividend payers and non-dividend payers but it is more so for dividend payers.

5.2 Portfolio Sorts Based on Individual Dividend Yields

If exposure to the dividend yield factor reflects information on the firm's individual dividend yield, then portfolio sorts based on individual dividend yields should display significant return spreads between high-dividend yield stocks and low-dividend yield stocks. In Table 5, we report results for mean excess returns of VW and EW quintile portfolios sorted on the firms' individual dividend yields. Portfolios for fiscal year *t* are formed using firm dividend yields measured in June of fiscal year *t*-1.

The main finding here is that there is little cross-sectional variation in the performance of stocks according to their dividend yield. For the VW H-L portfolio, the mean return and the six-factor alpha are low (both at 0.12% per month) and insignificant. In short, our evidence indicates that exposure to market dividend yield risk appears to be unrelated to firms' individual dividend yields.

6 A Predictive CAPM Framework

We have previously demonstrated that the factor premium associated with Δp is substantially higher than the factor premium of $\Delta d12$, and Δp is more significant than $\Delta d12$ (see Table 3). For this reason, we examine Δp separately by augmenting the standard CAPM to include the *lagged* excess return to the market. This additional variable essentially captures the effect of Δp since the two variables are perfectly correlated. We refer to this model as the "predictive CAPM," which is described by the following regression:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,1}(mkt_t - r_{f,t}) + \beta_{i,2}(mkt_{t-1} - r_{f,t-1}) + \varepsilon_{i,t}.$$
 (12)

We then consider three cases. First, we sort stocks solely on $\beta_{i,1}$, which is the standard CAPM beta. This approach removes the lagged market return from the analysis. Second, we sort stocks on $\beta_{i,D} = \beta_{i,1} + \beta_{i,2}$, which we refer to as the "Dimson beta." Following Dimson (1979), this is a popular approach that has been implemented by Fama and French (1992) and Liu, Stambaugh and Yuan (2018) among others. Third, we sort stocks solely on $\beta_{i,2}$, which we refer to as the "predictive beta." In all cases, the loadings are estimated over a 5-year rolling window as previously.

The portfolio sorts are reported in Panels A and B of Table 6. Consistent with our previous results, the standard CAPM betas deliver an H-L spread that is low and insignificant. The six-factor alphas are also low and insignificant.

The Dimson betas slightly improve the CAPM performance. The H-L return is higher than for the CAPM-betas but is still low and insignificant. For VW portfolios, the six-factor alphas are also low and insignificant.

Turning to portfolios sorted solely on the predictive betas, the results are striking. There is a positive and almost monotonic relation between the predictive beta and average excess returns. For VW returns, the H-L spread is 0.28% with a t-stat=1.79. For EW returns, the H-L spread is 0.41% with a t-stat=2.32. The six-factor alphas are positive and significant: the VW H-L alpha is 0.33% with a t-stat=2.75 and the EW H-L alpha is 0.78% with a t-stat=5.79. These findings indicate that cross-sectional

predictability lies exclusively in the predictive betas, not in the contemporaneous CAPM betas. Sorting on predictive betas alone (not in conjunction with the contemporaneous betas) delivers a positive beta-return relation. To our knowledge, this is a novel finding in the literature.

We further investigate this finding with Fama-MacBeth (1973) regressions. In Panel C of Table 6, we report the factor premium of the Dimson beta in the presence of the SMB, HML, RMW, CMA and MOM factors. Consistent with the portfolios sorts, we find that the Dimson beta has a positive (0.11) and insignificant market price of risk (*t*stat=0.92).

In Panel D, we use the contemporaneous beta and the predictive beta. The contemporaneous CAPM beta has a low and insignificant price of risk: 0.01% with a *t*-stat=0.11. In contrast, the predictive beta has a positive price of risk (0.25%) and is highly significant (*t*-stat=2.25). Therefore, the lagged market return alone is powerful in predicting the cross-section of expected stock returns.

As a final exercise, we add F Δ d12 to the predictive CAPM framework. In doing so, we are effectively estimating a version of the original three-factor model (mktx, Δ d12 and Δ p) that was initially displayed in Panel B of Table 3. This is because mkt_{t-1} and Δp_{t-1} have a perfect correlation. The results are reported in Panel E of Table 6. We find that F Δ d12 is low (0.03) and marginally significant (t-stat=1.69), while its presence has no effect on the size and significance of the lagged market return. We conclude that the lagged market return, which is the strongest component of the dividend yield factor, is robust and highly significant in predicting the cross-section of expected stock returns.

7 Conclusion

In estimating the CAPM, the beta on the market factor is almost exclusively driven by the capital gains component of the market portfolio. Although the dividend yield makes a substantial contribution to the mean return of the market portfolio, it practically contributes nothing to its variance. As a result, the market dividend yield is effectively ignored in estimating the market beta. This is crucially important if the market dividend yield contains distinct information from market capital gains.

We propose a two-factor CAPM model, which separates the two components of the market portfolio and, therefore, allows the dividend yield to make a distinct contribution to predicting the cross-section of expected stock returns. The results are striking: the capital gains factor performs poorly (same as the market portfolio), but the dividend yield factor performs well in distinguishing between high-performing and low-performing assets. This finding is particularly strong in the post-1978 period that coincides with the persistent decline in the number and proportion of US dividendpaying firms. For this sample period, the high-minus-low VW quintile portfolio delivers a statistically singificant mean return of -0.28% per month, which rises to -0.36% for EW portfolios. The VW and EW six-factor alphas are also significantly negative. Finally, Fama-MacBeth (1973) regressions confirm the presence of a significant negative premium for the dividend yield factor in the presence of other well-known risk factors.

Separating the market factor into a capital gains factor and dividend yield factor is consistent with recent evidence on the dividend disconnect. Hartzmark and Solomon (2019, 2022) find that in practice investors do not treat dividends and capital gains in the same manner and often disregard dividends in making financial decisions. We conjecture that this behavioural bias against dividends became stronger in the post-1978 period as the number and proportion of dividend-paying firms declined significantly in the US. Motivated by this idea, we propose a theoretical model which shows that when investors tend to ignore dividends and the market capital gains have substantially higher variance than the market dividend yield, then the market dividend yield factor has strong predictive ability for stock returns. In this context, our work can be seen as an application of the disconnect between price changes and dividends to asset pricing.

The dividend yield factor has two components: the dividend growth rate and lagged capital gains. The latter is the strongest of the two components since it has a higher risk premium and is consistently significant. This finding is used to motivate a simple extension to the CAPM that we term the "predictive CAPM." The predictive CAPM conditions on both the contemporaneous and the lagged market return. We show that the beta on solely the lagged market return delivers a significant positive factor premium in the cross-section of stock returns.

Overall, our analysis proposes simple extensions of the CAPM that address the dividend disconnect in the market portfolio and the enormous variance differential between capital gains and the dividend yield. The empirical evidence indicates that these extensions to the CAPM can establish a significant beta-return relation, which is the cornerstone of an asset pricing model. Consequently, the two-factor CAPM model can be a useful addition to the toolkit implemented in asset pricing research and financial practice.

References

Ang, A., R.J. Hodrick, Y. Xing, and X. Zhang (2006). "The Cross-Section of Volatility and Expected Returns," *Journal of Finance* **61**, 259-299.

Baker, M., S. Nagel, and J. Wurgler (2007). "The Effect of Dividends on Consumption," *Brookings Papers on Economic Activity* **1**, 277-291.

Baker, M., and J. Wurgler (2004a). "Appearing and Disappearing Dividends: The Link to Catering Incentives," *Journal of Financial Economics* **73**, 271–288.

Baker, M., and J. Wurgler (2004b). "A Catering Theory of Dividends," *Journal of Finance* **59**, 1125–1165.

Black, F. (1972). "Capital Market Equilibrium with Restricted Borrowing," *Journal* of Business **45**, 444–455.

Black, F., M.C. Jensen, and M. Scholes (1972). "The Capital Asset Pricing Model: Some Empirical Tests," *Studies in the Theory of Capital Markets* **81**, 79-121.

Campbell, J.Y. (1996). "Understanding Risk and Return," *Journal of Political Economy* **104**, 298-345.

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**, 195-228.

Carhart, M. (1997). "On Persistence in Mutual Fund Performance," *Journal of Finance* **52**, 57-82.

Daniel, K., L. Garlappi, and K. Xiao (2021). "Monetary Policy and Reaching for Income," *Journal of Finance* **76**, 1145–1192. Dimson, E. (1979). "Risk Measurement when Shares are Subject to Infrequent Trading," *Journal of Financial Economics* **7**, 197-226.

Fama, E.F., and K.R. French (1992). "The Cross-Section of Expected Stock Returns," *Journal of Finance* **47**, 427-465.

Fama, E.F., and K.R. French (1993). "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics* **33**, 3-56.

Fama, E.F., and K.R. French (2001). "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?" *Journal of Financial Economics* **60**, 3-43.

Fama, E.F., and K.R. French (2015). "A Five-Factor Asset Pricing Model," *Journal* of Financial Economics **116**, 1-22.

Fama, E.F., and J.D. MacBeth (1973). "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy* **81**, 607-636.

Hartzmark, S.M., and D.H. Solomon (2019). "The Dividend Disconnect," *Journal* of Finance 74, 2153-2199.

Hartzmark, S.M., and D.H. Solomon (2022). "Reconsidering Returns," *Review of Financial Studies* **35**, 343-393.

Hou, K., C. Xue, and L. Zhang (2020). "Replicating Anomalies," *Review of Financial Studies* **33**, 2019-2133.

Koijen, R.S.J., T.J. Moskowitz, L.H. Pedersen, and E.B. Vrugt (2018). "Carry," Journal of Financial Economics **127**, 197-225.

Lehmann, B.N., and D.M. Modest (1988). "The Empirical Foundations of the Arbitrage Pricing Theory," *Journal of Financial Economics* **21**, 213-254.

Lintner, J. (1965). "Security Prices, Risk, and Maximal Gains from Diversification," *Journal of Finance* **20**, 587-615.

Liu, J., R. Stambaugh, and Y. Yuan (2018). "Absolving Beta of Volatility's Effects," *Journal of Financial Economics* **128**, 1-15.

Merton, R.C. (1973). "An Intertemporal Capital Asset Pricing Model," *Econometrica* **41**, 867-887.

Miller, M.H., and F. Modigliani (1961). "Dividend Policy, Growth, and the Valuation of Shares," *Journal of Business* **34**, 411–433.

Newey, W.K., and K.D. West (1987). "A Simple, Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* **55**, 703–708.

Peng, L., and W. Xiong (2006). "Investor Attention, Overconfidence and Category Learning," *Journal of Financial Economics* **80**, 563-602.

Sharpe, W.F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance* **19**, 425-442.

Sabbatucci, R. (2022). "Are Dividends and Stock Returns Predictable? New Evidence Using M&A Cash Flows," Available at SSRN: https://ssrn.com/abstract=2578393.

Shefrin, H.M., and M. Statman (1984). "Explaining Investor Preference for Cash Dividends," *Journal of Financial Economics* **13**, 253-282.

Table 1: Portfolios Sorted by Exposure to the Dividend Yield Factor

This table presents the performance of portfolios sorted by the exposure (beta) of stock excess returns to the dividend yield factor, Δ dy12. We form value-weighted portfolios based on NYSE breakpoints and equally-weighted portfolios, which are rebalanced monthly. The betas are estimated using Equation (9) based on the most recent five years of monthly data. The mean and standard deviation are for monthly percentage excess returns. Size and B/M report the average log market capitalization and book-to-market ratio for firms in each portfolio. The "H-L" row refers to the difference in monthly excess returns between the High and Low quintile portfolios. The FF6 Alpha column reports the alpha with respect to the Fama–French (2015) five-factor model plus momentum. Post-formation betas are according to Equation (10) using the FMP. Statistical significance is assessed using Newey-West (1987) *t*statistics. The sample period ranges from January 1978 to December 2019.

Panel A: Value-Weighted Portfolios										
	Ret	urns				Factor	or Loadings			
					FF6 Pre-Formation		Post-Formation			
Rank	Mean	St Dev	Size	B/M	Alpha	$\beta_{\Delta dy 12}$	$\beta_{F\Deltady12}$			
High	0.66	4.48	17.34	0.49	0.01	0.33	-0.01			
4	0.63	3.95	17.56	0.51	-0.16***	0.09	0.02*			
3	0.79	4.22	17.50	0.52	0.05	-0.06	0.03**			
2	0.75	4.60	17.42	0.54	0.01	-0.21	0.01			
Low	0.94	5.77	16.28	0.60	0.26***	-0.57	-0.06***			
H-L	-0.28**	2.74			-0.27**		0.05			
(<i>t</i> -stat)	(-2.04)				(-2.38)		(1.40)			

Panel B: Equally-Weighted Portfolios

Returns						Factor Loadings		
				FF6		Pre-Formation	Post-Formation	
Rank	Mean	St Dev	Size	B/M	Alpha	$\beta_{\Delta dy 12}$	$\beta_{F\Deltady12}$	
High	0.70	5.39	14.57	0.73	0.22**	0.47	-0.16***	
4	0.89	4.51	15.00	0.74	0.27***	0.06	-0.13***	
3	0.91	4.69	14.82	0.78	0.34***	-0.13	-0.15***	
2	0.99	5.42	14.04	0.85	0.50***	-0.36	-0.23***	
Low	1.06	7.50	12.79	0.91	0.93***	-0.99	-0.35***	
H-L	-0.36**	3.36			-0.71***		0.19***	
(t-stat)	(-2.16)				(-5.62)		(4.17)	

Table 2: Portfolios Sorted by Exposure to the Dividend Yield Factor Across Subsamples

This table presents the performance of the High-minus-Low (H-L) portfolio across subsamples. The H-L portfolio refers to the difference in monthly excess returns between the High and Low quintile portfolios. The quintile portfolios are generated by sorts on the exposure (beta) of stock excess returns to the dividend yield factor, $\Delta dy12$. We form value-weighted and equally-weighted portfolios with monthly rebalancing. The betas are estimated using Equation (9) based on the most recent five years of monthly data. The FF6 alpha is computed using the Fama–French (2015) five-factor model plus momentum. For the 1932-1977 sample period, we use a four-factor model to compute the alpha because data on RMW and CMA are not available. Statistical significance is assessed using Newey–West (1987) *t*-statistics.

Panel A: Value-Weighted H-L Portfolios									
	1932-1977	1978-2019	1963-2019	1968-2019	1973-2019	1978-2019	1983-2019	1988-2019	
H-L Returns	0.08	-0.28**	-0.18	-0.21	-0.23*	-0.28**	-0.41***	-0.33**	
(t-stat)	(0.48)	(-2.04)	(-1.38)	(-1.59)	(-1.76)	(-2.04)	(-2.80)	(-2.02)	
FF6 Alpha	0.23*	-0.27**	0.01	0.00	-0.03	-0.27**	-0.38***	-0.29**	
(t-stat)	(1.78)	(-2.38)	(0.07)	(-0.03)	(-0.26)	(-2.38)	(-3.09)	(-2.18)	

	Panel B: Equally	v-Weighted H	I-L Portfolios	
070 0010	10(2, 2010	10(0 2010	1072 2010	1070 00

i.

	1932-1977	1978-2019	1963-2019	1968-2019	1973-2019	1978-2019	1983-2019	1988-2019
H-L Returns	-0.06	-0.36**	-0.28**	-0.34**	-0.34**	-0.36**	-0.49***	-0.51**
(<i>t</i> -stat)	(-0.36)	(-2.16)	(-2.00)	(-2.37)	(-2.25)	(-2.16)	(-2.63)	(-2.50)
FF6 Alpha	0.14	-0.71***	-0.42***	-0.48***	-0.50***	-0.71***	-0.83***	-0.81***
(t-stat)	(0.90)	(-5.62)	(-4.16)	(-4.42)	(-4.36)	(-5.62)	(-5.93)	(-5.39)

Table 3: Fama-MacBeth Regressions

This table reports the Fama-MacBeth (1973) factor premiums using the full cross-section of stock returns. The factor premiums are the time-series means of the cross-sectional coefficients γ in Equation (11). The table also reports Newey-West (1987) *t*-statistics. "Original" denotes the original Δ dy12, Δ d12 and Δ p factors and FMP denotes the factor mimicking portfolios. All regressions condition on the five Fama-French (2015) factors plus momentum. The sample period ranges from January 1978 to December 2019.

Panel A: Dividend Yield Factor							
	Orig	ginal	FM	FMP			
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat			
mkt-rf	0.02	0.13	0.01	0.11			
F∆dy12	-0.26	-2.18	-0.28	-1.84			
SMB	-0.03	-0.37	-0.03	-0.49			
HML	0.09	1.13	0.09	1.13			
RMW	-0.01	-0.17	-0.01	-0.18			
CMA	0.02	0.52	0.03	0.58			
MOM	-0.11	-1.53	-0.10	-1.37			

Panel B: Components of the Dividend Yield Factor

	Orig	ginal	FN	FMP	
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat	
mkt-rf	0.01	0.10	0.02	0.17	
F∆d12	0.03	1.69	0.03	1.41	
FΔp	0.26	2.26	0.26	1.89	
SMB	-0.03	-0.41	-0.03	-0.48	
HML	0.08	1.10	0.08	1.07	
RMW	-0.01	-0.11	-0.01	-0.10	
CMA	0.02	0.48	0.02	0.50	
MOM	-0.11	-1.49	-0.09	-1.31	

Table 4: Dividend Payers vs Non-Dividend Payers

This table reports the Fama–MacBeth (1973) factor premiums for the cross-section of two separate groups: dividend payers and non-dividend payers. The specification of the regressions is the same as in Table 3. The sample period ranges from January 1978 to December 2019.

-

Panel A: Dividend Payers								
	Orig	ginal	FMP					
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat				
mkt-rf	0.12	0.82	0.12	0.79				
F∆dy12	-0.41	-2.92	-0.45	-2.57				
SMB	0.06	0.77	0.05	0.62				
HML	0.12	1.57	0.13	1.59				
RMW	-0.07	-1.26	-0.07	-1.32				
CMA	0.03	0.60	0.03	0.70				
MOM	-0.02	-0.26	-0.02	-0.22				

Panel B: Non-Dividend Payers

	Orig	ginal	F	FMP		
	Mean <i>t</i> -stat		Mean	<i>t</i> -stat		
mkt-rf	0.00	-0.04	-0.01	-0.06		
F∆dy12	-0.25	-2.41	-0.27	-2.09		
SMB	-0.02	-0.39	-0.03	-0.52		
HML	0.06	0.85	0.06	0.83		
RMW	0.00	-0.06	0.00	-0.07		
CMA	0.02	0.47	0.02	0.50		
MOM	-0.10	-1.51	-0.09	-1.34		

Table 5: Portfolios Sorted on the Individual Dividend Yield

This table displays the performance of portfolios sorted on the individual dividend yield of each stock. We form value-weighted and equal-weighted portfolios rebalanced monthly. Stocks are sorted into quintiles from lowest dy12 (Low) to highest dy12 (High). The mean and standard deviation are for monthly percentage excess returns. The "H-L" row refers to the difference in monthly excess returns between the High and Low portfolios. Size and B/M report the average log market capitalization and book-to-market ratio for firms in each portfolio. The FF6 alpha is computed using the Fama–French (2015) five-factor model plus momentum. The sample period ranges from January 1978 to December 2019.

Panel A: Value-Weighted Portfolios								
	Ret	urns						
					FF6	Dividend		
Rank	Mean	St Dev	Size	B/M	Alpha	Yield		
High	0.74	4.03	17.39	0.77	-0.07	5.67		
4	0.80	3.94	17.61	0.57	-0.05	3.29		
3	0.76	4.20	17.80	0.45	-0.18***	2.29		
2	0.74	4.63	17.31	0.44	-0.15**	1.50		
Low	0.62	5.08	17.06	0.43	-0.19***	0.63		
H-L	0.12	4.04			0.12			
(t-stat)	(0.61)				(0.90)			

Panel B: Equally-Weighted Portfolios

	Ket	urns				
					FF6	Dividend
Rank	Mean	St Dev	Size	B/M	Alpha	Yield
High	0.87	3.84	15.00	0.95	0.16**	7.45
4	0.96	4.04	15.24	0.80	0.12*	3.28
3	0.91	4.45	15.25	0.74	-0.02	2.30
2	0.93	4.58	15.08	0.69	-0.02	1.49
Low	0.79	5.02	15.03	0.63	-0.14**	0.65
H-L	0.08	2.62			0.30***	
(<i>t</i> -stat)	(0.60)				(3.22)	

Table 6: Predictive CAPM

This table displays the performance of the predictive CAPM. Panels A and B report the performance of value-weighted and equallyweighted quintile portfolios sorted on (1) the standard CAPM betas, (2) the Dimson (1979) CAPM betas, and (3) the predictive betas on the *lagged* market excess return. Panels C, D and E report the Fama–MacBeth (1973) cross-sectional factor premiums for the full cross-section of stock returns. The sample period ranges from January 1978 to December 2019.

Panel A: Value-Weighted Portfolios									
	Standar	d CAPM		Dimso	n (1979)		Pree	dictive CAP	'M beta
	be	eta		CAP	M beta				
Rank	Return	FF6		Return	FF6	-	Return	FF6	Post-
		Alpha			Alpha			Alpha	Formation
TT: 1	0.70	0.07		0.70	0.10	-	0.02	0.00444	$\beta_{FMKT_{t-1}}$
High	0.70	0.06		0.78	0.12		0.92	0.29***	0.11***
4	0.85	-0.05		0.79	-0.04		0.75	0.03	0.01
3	0.85	-0.07		0.85	-0.02		0.81	0.05	-0.02
2	0.76	-0.13**		0.75	-0.11***		0.66	-0.11**	-0.03***
Low	0.64	-0.07		0.66	-0.08		0.64	-0.05	-0.03***
H-L	0.06	0.13		0.12	0.21		0.28*	0.33***	0.14***
(<i>t</i> -stat)	(0.19)	(0.81)		(0.39)	(1.32)		(1.79)	(2.75)	(4.29)
			Panel	B: Equall	y-Weighte	d Portfo	olios		
	Standar	d CAPM		Dimso	n (1979)		Pree	dictive CAF	'M beta
	be	eta		CAP	M beta				
Rank	Return	FF6		Return	FF6		Return	FF6	Post-
		Alpha			Alpha			Alpha	Formation
									$\beta_{FMKT_{t-1}}$
High	0.86	0.45***		0.96	0.58***		1.10	0.97***	0.44***
4	0.92	0.20***		0.99	0.30***		1.00	0.53***	0.29***
3	1.03	0.23***		0.94	0.17***		0.90	0.33***	0.21***
2	0.93	0.19***		0.94	0.17***		0.86	0.24***	0.17***
Low	0.82	0.26***		0.73	0.11		0.69	0.19*	0.18***
H-L	0.05	0.19		0.23	0.47***		0.41**	0.78***	0.25***
(t-stat)	0.15	(1.26)		(0.73)	(2.78)		(2.32)	(5.79)	(4.73)

Panel C: Fama-MacBeth Regressions						
	Ori	ginal	I	FMP		
	Mean <i>t</i> -stat		Mean	<i>t</i> -stat		
Dimson Beta	0.11	0.92	0.11	0.92		
SMB	-0.03	-0.48	-0.06	-0.78		
HML	0.07	1.02	0.07	0.94		
RMW	-0.02	-0.31	-0.02	-0.33		
CMA	0.01	0.24	0.01	0.26		
MOM	-0.12	-1.64	-0.09	-1.27		

Panel D: Fama-MacBeth Regressions

	Original			FMP	
	Mean	<i>t</i> -stat	_	Mean	<i>t</i> -stat
(mkt-rf) _t	0.01	0.11		0.01	0.12
F(mkt-rf) _{t-1}	0.25	2.25		0.25	1.85
SMB	-0.03	-0.39		-0.03	-0.48
HML	0.09	1.15		0.09	1.13
RMW	-0.01	-0.16		-0.01	-0.17
CMA	0.02	0.52		0.03	0.58
MOM	-0.11	-1.53		-0.10	-1.34

Panel E: Fama-MacBeth Regressions

	Original		FMP		
	Mean	<i>t</i> -stat	_	Mean	<i>t</i> -stat
(mkt-rf) _t	0.01	0.10		0.02	0.17
F(mkt-rf) _{t-1}	0.25	2.24		0.26	1.86
F∆d12	0.03	1.69		0.03	1.41
SMB	-0.03	-0.41		-0.03	-0.48
HML	0.08	1.10		0.08	1.07
RMW	-0.01	-0.11		-0.01	-0.10
CMA	0.02	0.48		0.02	0.50
MOM	- 0.11	-1.49		-0.09	-1.30



Figure 1

The top panel displays the number of dividend payers vs non-payers for the sample period of January 1927 to December 2019. The bottom panel displays the proportion of dividend payers for the same sample period.

The Disconnect Between Market Capital Gains and the Dividend Yield in Asset Pricing

Online Appendix

September 2024

A. Factor-Mimicking Portfolio

This section provides details on how we construct the factor-mimicking portfolio (FMP) by implementing the OLS cross-sectional approach based on Lehmann and Modest (1988). The OLS approach involves performing estimation of two-step Fama-MacBeth (1973) regressions. In the first stage time-series regression, loadings are estimated for each firm using the following univariate model:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i \Delta dy 12_t + \varepsilon_{i,t}.$$
 (A1)

The beta estimates from Equation (A1) are used in the second stage crosssectional regressions with β_i as the only explanatory variable. This leads to the following OLS estimates of the factor premium:

$$\hat{\gamma}_t = (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'(r_t - r_{f,t}),$$
(A2)

where $\hat{\beta}$ is the vector of estimated betas using the full sample and $r_t - r_{f,t}$ is the vector of excess stock returns. Then, the factor mimicking portfolio, F Δ dy12, is simply the $\hat{\gamma}_t$ estimate from Equation (A2) above. By design, the OLS approach generates an ex-post full sample FMP, which is maximally correlated with the original Δ dy12 variable. We generate the OLS FMP using data only from dividend-paying firms. This ensures that only firms which contribute to the market dividend are included in mimicking the dividend yield factor. Consequently, both from a methodology point of view and from a data selection point of view, the OLS FMP is maximally correlated with the dividend yield factor. For this reason, the OLS FMP is the factor mimicking portfolio used in our analysis.

B Robustness

B.1 Bad Beta, Good Beta

Our two-factor CAPM has some intuitive similarities with the "bad-beta, good-beta" two-factor model proposed by Campbell and Vuolteenaho (2004). This section evaluates the empirical relation between our dividend yield factor and the bad-beta, good-beta two-factor model, which is based on the Campbell and Shiller (1988) log-linear approximate decomposition of returns in two components:

$$r_{t+1} - E_t[r_{t+1}] = N_{CF,t+1} - N_{DR,t+1},$$
(B1)

where $N_{CF,t+1} = (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}$ denotes news about future cash flows (dividends) and $N_{DR,t+1} = (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j}$ denotes news about future discount rates (expected returns). In these equations, r_{t+1} is the log stock return at time t+1, $E_t[r_{t+1}]$ is the time t expectation of the t+1 return, d_{t+1} is the log dividend paid by the stock at t+1, and $\rho<1$ is a discount coefficient. In words, Equation (B1) suggests that unexpected stock returns are associated with changes in expectations of future cash flows and future discount rates.

Based on this decomposition, Campbell and Vuolteenaho (2004) propose a twofactor model in which they define the cash-flow beta as:

$$\beta_{i,CF} = \frac{Cov(r_{i,t}, N_{CF,t})}{Var(r_{m,t}^{e} - E_{t-1}r_{m,t}^{e})},$$
(B2)

and the discount-rate beta as:

$$\beta_{i,DR} = \frac{Cov(r_{i,t}, N_{DR,t})}{Var(r_{m,t}^e - E_{t-1}r_{m,t}^e)},$$
(B3)

where $r_{m,t}^{e}$ is the excess market return. Since both betas have the same denominator, it is straightforward to show that the market beta ($\beta_{i,m}$) is given by:

$$\beta_{i,m} = \beta_{i,CF} + \beta_{i,DR}.\tag{B4}$$

In this two-beta model, Campbell and Vuolteenaho (2004) refer to $\beta_{i,CF}$ as the "bad beta" and $\beta_{i,DR}$ to as the "good beta." The intuition for this distinction is based on the following argument. The value of the market portfolio may decrease because investors receive bad news about future cashflows (i.e., $N_{CF,t+1}$ decreases) or because investors increase future discount rates (i.e., $N_{DR,t+1}$ increases). The first case corresponds to a bad beta ($\beta_{i,CF}$) because wealth decreases but investment opportunities are unchanged. The second case corresponds to a good beta ($\beta_{i,DR}$) because wealth decreases but now investment opportunities improve.

The decomposition implemented in this paper and the Campbell and Shiller (1988) decomposition are similar but distinct. In this paper, we decompose current market returns into current capital gains and the current dividend yield and then use the latter to form a market dividend yield factor. In contrast, Campbell and Shiller (1988) decompose shocks to future market returns into shocks to future cash flows and shocks to future discount rates. In this section, we rely on an empirical analysis to determine the relation between the two models.

We empirically evaluate the bad-beta, good-beta model using a first-order VAR model to estimate expected returns ($E_t[r_{t+1}]$) and discount rate shocks ($N_{DR,t+1}$). Then, we use the realized return (r_{t+1}) and Equation (B1) to back out the estimated cash flow shocks ($N_{CF,t+1}$). This allows us to compute the cash-flow beta and the discount-rate beta defined above. Our empirical approach follows exactly Campbell and Vuolteenaho (2004). VAR estimation conditions on four state variables: the excess return on the market portfolio, the yield spread between short-term and long-term government bonds, the market smoothed price-earnings ratio, and the small-stock value spread.¹ To be consistent with the ex-ante nature of our analysis, each month we compute ex-ante beta estimates using a 20-year rolling window. Therefore, our bad-beta, good-beta estimates avoid the forward-looking bias.

Armed with the ex-ante estimates of bad-beta and good-beta for each stock, we re-estimate the Fama-MacBeth regressions to include these betas. The results are reported in Table B4. Our main finding is that the factor premiums for bad-beta and

¹ The yield spread is defined as the 10-year treasury bond yield less the 3-month T-bill rate obtained from the FRED database of the Federal Reserve Bank of St. Louis. The smoothed S&P500 price-earnings ratio is obtained from Robert Shiller's website. The small-stock value spread is defined as the difference between the log book-to-market value ratios of small value and small growth stocks obtained from Ken French's online data library.

good-beta are small and statistically insignificant. More importantly, their inclusion in the Fama-MacBeth regressions seems to have no effect on the size and statistical significance of either the dividend yield factor (Δ dy12) or its two components (Δ d12 and Δ p). We conclude, therefore, that empirically the dividend yield factor is unrelated to the bad-beta and good-beta of Campbell and Vuolteenaho (2004).

B.2 Is the Dividend Yield Factor Related to Taxes?

The decomposition of the market return into capital gains and the dividend yield warrants a discussion of taxation since the two return components are subject to a different tax treatment. In this section, we investigate the extent to which the explanatory power of the dividend yield factor can be attributed to taxes.

It is standard in the asset pricing literature to assess the tax effect using the implied tax rate from the municipal bond market (see, e.g., Naranjo, Nimalendran and Ryngaert, 1998). Following Poterba (1986), the implied tax rate is the tax rate that makes an investor indifferent between taxable and non-taxable bonds. Specifically, it is defined as the ratio of the (tax-exempt) municipal yield over the (taxable) treasury yield. This implied tax rate is used as a proxy for the tax differential between dividend income and capital gains income.

We measure the implied tax rate using data from (1) the Standard and Poor's high-grade tax-exempt municipal bond yields with a 20-year maturity, and (2) the 20-year Treasury yield. Both series are obtained from the *LSEG Eikon* database. We form the Δ Tax series, which is the change in the implied tax rate. Using the differenced series

 (ΔTax) ensures that the tax variable is stationary with low persistence. We also generate F ΔTax , which is the factor-mimicking portfolio for the change in the implied tax rate.

We assess the cross-sectional evidence on the tax effect by incorporating F Δ Tax in the Fama-MacBeth regressions, and report the results in Table B5 for dividend-payers and non-dividend payers. Our main finding here is that the tax effect is small and statistically insignificant. Additionally, the inclusion of F Δ Tax does not affect the size and significance of Δ dy12 for either dividend-payers or non-dividend payers. Indeed, the factor premium for Δ dy12 is almost the same in Table 4 (no tax effect) and Table B5 (with tax effect). We conclude, therefore, that taxation is an unlikely explanation for the predictive power of the dividend yield factor.

B.3 Orthogonalizing the Dividend Yield Factor

In this section, we orthogonalize the dividend yield factor to the risk-free rate, the term spread and the default spread. This is motivated by Petkova (2006), who shows that these three variables together with the dividend yield describe well the time-variation in the investment opportunity set. Orthogonalization removes the effect of three prominent variables, which are likely to be correlated with the dividend yield factor. We orthogonalize by estimating a regression of $\Delta dy12$ on the three variables and then using the fitted residuals as our orthogonal $\Delta dy12$. We use a 20-year window and perform ex-ante estimation to avoid a forward looking bias.

The data used in the orthogonalization are obtained as follows. Data on the longterm yields are obtained from Ibbotson's Stocks, Bonds, Bills and Inflation Yearbook. Data on the 3-month treasury bill and the corporate bond yields on AAA-rated and BAA-rated bonds are obtained from the FRED database of the Federal Reserve Bank at St. Louis. The term spread is the difference between the long term yield on government bonds and the Treasury bill rate. The default yield spread is the difference between BAA and AAA-rated corporate bond yields.

The results for portfolio sorts based on the orthogonalized dividend yield factor are in Table B6 and the corresponding Fama-MacBeth regression results are in Table B7. The results indicate that the main findings are similar whether we use the othogonalization or not. Specifically, for VW quintile portfolios rebalanced monthly by sorting on the beta to the orthogonalized Δ dy12 factor, the H-L portfolio delivers a mean return of -0.34% per month with a t-stat=-2.68. The six-factor alpha of the VW H-L portfolio is equal to -0.26% with a t-stat=-2.49. The results are stronger for EW portfolios: the H-L mean return is equal to -0.39% with a t-stat=-2.58 and the H-L alpha is equal to -0.60% with a t-statistic of -5.05. The significant negative premium of the dividend yield factor is also confirmed in Fama and MacBeth (1973) regressions in the presence of standard asset pricing factors. We conclude, therefore, that with or without the orthogonalization, the dividend yield factor has strong predictive power for the cross-section of expected stock returns.

References

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**, 195-228.

Campbell, J.Y., and T. Vuolteenaho (2004). "Bad Beta, Good Beta," *American Economic Review* **94**, 1249-1275.

Fama, E.F., and J.D. MacBeth (1973). "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy* **81**, 607-636.

Lehmann, B.N., and D.M. Modest (1988). "The Empirical Foundations of the

Arbitrage Pricing Theory," Journal of Financial Economics 21, 213-254.

Naranjo, A., M. Nimalendran, and M. Ryngaert (1998). "Stock Returns, Dividend Yields, and Taxes," *Journal of Finance* **53**, 2029-2057.

Petkova, R. (2006). "Do the Fama–French Factors Proxy for Innovations in Predictive Variables?" *Journal of Finance* **61**, 581-612.

Poterba, J.M. (1986). "Explaining the Yield Spread between Taxable and Tax-Exempt Bonds: The Role of Expected Tax Policy," in H. Rosen, ed.: *Studies in State and Local Public Finance*. University of Chicago Press. 5-48.

Table B1: Summary Statistics

This table reports summary statistics for the following monthly variables: mkt is the market excess return, mktx is the market excess return excluding dividends, dy is the market dividend yield, Δ dy12 is the monthly proportional change in the 12-month market dividend yield, Δ d12 is the year-over-year dividend growth rate and Δ p is the lagged rate of capital gains. AR(1) is the degree of serial correlation at a lag of one month. The sample period ranges from January 1978 to December 2019. Expansions and recessions are defined according to the NBER.

			Panel A: F	ull Sample			
	Mean	St. Dev.	Min	Max	Skewness	Kurtosis	AR(1)
mkt	1.04	4.40	-22.64	12.88	-0.75	5.24	0.05
mktx	0.82	4.39	-22.84	12.73	-0.76	5.24	0.05
dy	0.22	0.14	0.06	0.92	2.26	8.81	0.25
∆dy12	-0.09	4.67	-11.42	28.99	1.07	6.86	0.06
Δd12	0.53	1.47	-8.76	18.52	3.33	51.42	0.04
Δp	0.82	4.39	-22.84	12.73	-0.76	5.24	0.05
			Panel B: E	xpansions			
mkt	1.17	4.07	-22.64	12.88	-0.79	5.97	
mktx	0.96	4.05	-22.84	12.73	-0.79	5.99	
dy	0.21	0.13	0.06	0.82	2.26	9.13	
∆dy12	-0.26	4.36	-11.27	28.99	1.21	8.37	
Δd12	0.58	1.50	-8.76	18.52	3.47	52.44	
Δp	1.01	4.05	-22.84	12.73	-0.80	6.03	
			Panel C: F	Recessions			
mkt	0.01	6.44	-17.15	11.90	-0.29	2.56	
mktx	-0.28	6.41	-17.28	11.61	-0.29	2.53	
dy	0.29	0.21	0.06	0.92	1.50	4.34	
Δdy12	1.25	6.55	-11.42	18.72	0.31	2.69	
Δd12	0.10	1.10	-2.52	2.53	-0.13	3.02	
Δp	-0.74	6.31	-17.28	11.61	-0.16	2.58	

Table B2: Cross-Correlations

This table reports the cross-correlations for the variables defined in Table B1. The sample period ranges from January 1978 to December 2019. Expansions and recessions are defined according to the NBER.

		Pane	el A: Full Sam	ple		
	mkt	mktx	dy	∆dy12	Δd12	Δp
mkt	1		-			
mktx	1.00	1				
dy	0.10	0.07	1			
∆dy12	-0.05	-0.05	0.07	1		
Δd12	0.01	0.00	0.17	0.25	1	
Δp	0.05	0.05	0.00	-0.95	0.06	1
		Pane	el B: Expansio	ons		
mkt	1					
mktx	1.00	1				
dy	0.11	0.08	1			
∆dy12	0.04	0.04	0.10	1		
Δd12	0.01	0.00	0.15	0.29	1	
Δp	-0.04	-0.04	-0.04	-0.94	0.05	1
		Pan	el C: Recessio	ons		
mkt	1					
mktx	1.00	1				
dy	0.16	0.13	1			
Δdy12	-0.32	-0.31	-0.13	1		
$\Delta d12$	-0.05	-0.07	0.54	0.09	1	
Δр	0.32	0.31	0.22	-0.99	0.07	1

Table B3: Factor-Mimicking Portfolios

This table reports the cross-correlations between the original factors and the factor-mimicking portfolios (FMPs). The FMPs implement ordinary least squares (OLS) estimation using the cross-section of dividend-paying firms. The sample period ranges from January 1978 to December 2019.

Cross-Correlations	
$Corr(\Delta dv12, F\Delta dv12)$	0.79
$Corr(\Delta d12, F\Delta d12)$	0.85
$Corr(\Delta p, F\Delta p)$	0.79

Table B4: Bad-Beta, Good-Beta

This table reports the Fama–MacBeth (1973) factor premiums for bad-beta, good-beta regressions using the full crosssection of stock returns. We use the Campbell and Vuolteenaho (2004) decomposition of the market beta into a cash flow beta (β_{CF}) and a discount rate beta (β_{DR}). The specification of the regressions is the same as in Table 3 of the paper. The sample period ranges from January 1978 to December 2019.

Panel A: Dividend Yield Factor						
	Orig	ginal	FN	FMP		
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat		
β_{CF}	-0.03	-0.23	-0.04	-0.29		
β_{DR}	0.04	1.29	0.04	1.44		
F∆dy12	-0.26	-2.19	-0.28	-1.83		
SMB	-0.03	-0.42	-0.04	-0.52		
HML	0.08	1.11	0.08	1.11		
RMW	-0.01	-0.10	-0.01	-0.11		
CMA	0.02	0.51	0.03	0.56		
MOM	-0.11	-1.54	-0.10	-1.38		

Panel B: Components of the Dividend Yield Factor

	Original		F	MP
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat
β_{CF}	-0.04	-0.34	-0.04	-0.31
β_{DR}	0.05	1.56	0.05	1.67
F∆d12	0.03	1.70	0.03	1.38
FΔp	0.26	2.27	0.26	1.88
SMB	-0.03	-0.46	-0.04	-0.52
HML	0.08	1.08	0.08	1.06
RMW	0.00	-0.05	0.00	-0.04
CMA	0.02	0.47	0.02	0.49
MOM	-0.11	-1.52	-0.10	-1.33

Table B5: Tax Effects on Dividend Payers and Non-Dividend Payers

This table displays the effect of taxes on the Fama-MacBeth (1973) factor premiums for the cross-section two separate groups: dividend payers and non-dividend payers. F Δ tax is the factor-mimicking portfolio for the change in the implied tax rate. The specification of the regressions is the same as in Table 4 of the paper. The sample period ranges from January 1978 to December 2019.

Panel A: Dividend Payers							
	Orig	ginal	FN	ſP			
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat			
F∆tax	-0.05	-0.67	-0.06	-0.50			
Mkt-rf	0.13	0.87	0.14	0.93			
F∆dy12	-0.41	-2.94	-0.41	-2.66			
SMB	0.06	0.76	0.06	0.75			
HML	0.12	1.55	0.11	1.44			
RMW	-0.07	-1.23	-0.07	-1.28			
CMA	0.03	0.62	0.03	0.54			
MOM	-0.02	-0.27	-0.03	-0.27			

Panel B: Non-Dividend Payers

	Original		FM	IP
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat
F∆tax	-0.01	-0.21	0.01	0.11
Mkt-rf	0.00	0.04	0.01	0.12
F∆dy12	-0.25	-2.40	-0.24	-2.33
SMB	-0.02	-0.38	-0.03	-0.52
HML	0.05	0.79	0.05	0.66
RMW	0.00	-0.06	0.00	-0.04
CMA	0.02	0.33	0.01	0.27
MOM	-0.11	-1.57	-0.10	-1.46

Table B6: Portfolios Sorted by Exposure to the Orthogonalized Dividend Yield Factor

This table presents the performance of portfolios sorted by the exposure (beta) of excess stock returns to the orthogonalized dividend yield factor, Δ dy12. The table reports the same information as Table 1 of the paper but uses a dividend yield factor that is orthogonalized to the risk-free rate, the term spread and the default spread. Statistical significance is assessed using Newey–West (1987) *t*-statistics. The sample period ranges from January 1978 to December 2019.

	Panel A: Value-Weighted Portfolios						
	Retu	ırns			-	Factor	Loadings
			-		FF6	Pre-Formation	Post-Formation
Rank	Mean	St Dev	Size	B/M	Alpha	$\beta_{\Delta dy 12}$	$\beta_{F\Delta dy12}$
High	0.65	4.64	17.28	0.48	0.00	0.35	0.00
4	0.62	4.00	17.51	0.51	-0.17***	0.10	0.01
3	0.78	4.14	17.52	0.53	0.07	-0.05	0.01
2	0.74	4.53	17.40	0.55	-0.02	-0.22	0.02
Low	0.99	5.65	16.54	0.59	0.26***	-0.58	-0.03*
H-L	-0.34***	2.61			-0.26**		0.03
(t-stat)	(-2.68)				(-2.49)		(1.02)
			Panel B: E	qually-We	eighted Port	tfolios	
	Retu	ırns	_		0	Factor	Loadings
					FF6	Pre-Formation	Post-Formation
Rank	Mean	St Dev	Size	B/M	Alpha	$\beta_{\Delta dy 12}$	$\beta_{F\Deltady12}$
High	0.70	5.57	14.48	0.74	0.06	0.51	-0.06***
4	0.88	4.55	14.96	0.75	0.13***	0.08	-0.05***
3	0.93	4.67	14.83	0.78	0.20***	-0.12	-0.07***
2	0.94	5.30	14.16	0.85	0.26***	-0.35	-0.11***
Low	1.10	7.40	12.96	0.89	0.66***	-0.99	-0.23***
H-L	-0.39**	3.11			-0.60***		0.17***
(<i>t</i> -stat)	(-2.58)				(-5.05)		(4.30)

Table B7: Fama-MacBeth Regressions with an Orthogonalized Dividend Yield Factor This table reports the Fama-MacBeth (1973) factor premiums using the orthogonalized dividend yield factor, Δ dy12. The table reports the same information as Table 3 of the paper but uses a dividend yield factor that is orthogonalized to the risk-free rate, the term spread and the default spread. The sample period ranges from January 1978 to December 2019.

Panel A: Dividend Yield Factor							
	Orig	ginal	FN	FMP			
	Mean	<i>t</i> -stat	Mean	<i>t</i> -stat			
mkt-rf	0.02	0.13	0.02	0.15			
F∆dy12	-0.21	-2.35	-0.28	-2.24			
SMB	-0.03	-0.40	-0.03	-0.48			
HML	0.08	1.07	0.08	1.07			
RMW	-0.01	-0.16	-0.01	-0.17			
CMA	0.02	0.50	0.02	0.50			
MOM	-0.11	-1.56	-0.10	-1.41			

Panel B: Components of the Dividend Yield Factor

	Original		FMP	
	Mean	<i>t</i> -stat	Mean	<i>t-</i> stat
mkt-rf	0.01	0.11	0.02	0.20
F∆d12	0.03	1.56	0.03	1.43
FΔp	0.22	2.49	0.27	2.29
SMB	-0.03	-0.46	-0.03	-0.48
HML	0.08	1.02	0.08	1.00
RMW	-0.01	-0.12	-0.01	-0.09
CMA	0.02	0.47	0.02	0.43
MOM	-0.11	-1.50	-0.10	-1.34



Figure B1

This figure displays the autocorrelation function of the monthly market dividend yield (dy, top graph) and the 12-month market dividend yield (dy12, bottom graph) for the sample period of January 1978 to December 2019.



Figure B2

This figure displays the time series of the market dividend yield, dy12, and the dividend yield factor, Δ dy12, for the sample period of January 1978 to December 2019. The shaded areas indicate NBER recessions.