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

Michael Di Carlo
University of Guelph
mdicarlo@uoguelph.ca

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

Ilias Tsiakas
University of Guelph
itsiakas@uoguelph.ca

December 2022

Abstract

The rate of capital gains of the market portfolio is vastly more volatile than the dividend yield. As a result, standard CAPM betas capture exposure only to market capital gains. We propose a two-factor CAPM that includes a separate market dividend yield factor and find that this factor carries a significant negative premium in the post-1978 period that coincides with the persistent decline in the number 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 when 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, Richard Luger, John Maheu, Alex Maynard, Gabriel Power and to seminar participants at 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). The authors have no conflict of interest to disclose. *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

Most of financial theory is 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, 2022) find that investors track price changes and dividends separately rather than combine them into total returns. The disconnect between price changes and dividends has crucial practical consequences. For example, investors are more likely to sell stocks based on capital gains rather than returns. They also prefer to finance consumption using dividends rather than capital gains. The dividend disconnect reinforces the idea that in practical applications it is sensible to decompose the performance of portfolios into a capital gains component and a dividend yield component.

In the context of the CAPM model, there is an additional reason why it may be sensible to separate the market portfolio into a capital gains component and a dividend yield component. The dividend yield makes a substantial contribution to the total 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 almost 1000 times higher than the variance of the market dividend yield. This has a profound statistical implication: bundling together two separate components (capital gains and the dividend yield), where the first one is vastly more volatile than the second one, implies that in regression analysis the first component will completely dominate the second one. 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.

To address these issues, 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 new dividend yield factor is defined as the innovation to the 12-month market dividend yield. By construction, it accounts for the seasonality and potential non-stationarity of dividends. Following Petkova (2006), we orthogonalize the dividend yield factor to the default spread, the term spread and the risk-free rate. The two-factor CAPM model is explicitly 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 might be distinct from that associated with the dividend yield. In other words, the two-factor model allows for the likelihood of a disconnect between market capital gains and the dividend yield in asset pricing.

We propose a simple theoretical model to motivate our empirical analysis. We assume an economy populated by a representative agent, where asset returns are determined by a two-factor CAPM in which the market factor is decomposed into a capital gains factor and a dividend yield factor. It is straightforward to show that in this case, estimating the standard one-factor CAPM, where the market factor is the sum of the capital gains and dividend yield factors, effectively ignores the information contained in the dividend yield factor. This is because of the enormously higher variance of the capital gains factor relative to the dividend yield factor. A variance decomposition of asset returns further shows that the dividend yield factor is a key component of asset returns and hence should not be overshadowed by the highly volatile capital gains factor.

More importantly, our theoretical model 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 the behavioural bias against dividends explored by Hartzmark and Solomon (2019, 2022). Specifically, in an effort to reduce uncertainty about asset prices, investors process tremendous amounts of information about capital gains because they perceive capital gains as the primary determinant of price uncertainty. As investors spend substantial resources to learn about prices, signals about capital gains are highly precise. In contrast, investors tend to ignore information about dividends. In this context, 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 the market dividend yield in asset pricing can justify the predictive ability of the market dividend yield in the cross-section of expected stock returns.[1]

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 coincides with the peak in the number of dividend-paying firms in the US, which occurred at the end of 1977 (Fama and French, 2001). In other words, the empirical evidence indicates that at the same time that the number of dividend payers began to suffer a sustained decline in the US equity market, the predictive ability of the dividend yield factor increased significantly. 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) remains strong when forming factor-mimicking portfolios; (4) is significant for both

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

dividend-paying and non-dividend paying firms; (5) is unrelated to *individual* dividend yields; and (6) cannot be explained by tax effects. It is important to note that the predictive power of the separate dividend yield factor for our sample period is stronger than that of other well-known factors such as size, value, profitability, investment and momentum.

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 -0.34% per month, which is highly statistically significant. The alphas of the H-L portfolio are also significantly negative. These results hold for value-weighted portfolios and are even stronger for equally-weighted portfolios. 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 since it is substantially higher in business cycle recessions than in expansions. Consider an asset that has a high positive beta on the dividend yield factor. By definition, a high beta implies that this asset performs well when the factor is high. However, the dividend yield factor is high in recessions. Therefore, this is an asset that performs well in recessions. According to standard asset pricing theory, this asset is valuable because it performs well when we need it the most (in the bad states of the world) and hence investors do not require a high expected return to hold it. Consequently, high-beta assets on the dividend yield factor will have low expected returns and vice versa. This argument justifies the negative premium for the dividend yield factor.

Our two-factor CAPM has some intuitive similarities to the "bad-beta, good-beta" two-factor model proposed by Campbell and Vuolteenaho (2004). This model is based on the well-known Campbell and Shiller (1988) decomposition, where unexpected stock returns reflect news about future cash flows or future discount rates. Campbell and Vuolteenaho (2004) propose a two-factor model with a cash-flow beta and a discount-rate beta. The cash-flow beta is the "bad beta" because lower cashflows decrease wealth but leave future investment opportunities unchanged. The discount-rate beta is the "good beta" because higher discount rates decrease wealth but increase future investment opportunities.

We empirically assess the extent to which the bad-beta, good-beta model is related to the dividend yield factor. To do so, we compute ex-ante estimates of the bad beta and good beta for each stock updated every month and incorporate these betas into the Fama-MacBeth regressions. Our empirical findings are unequivocal: the factor premiums for bad-beta and good-beta are small and statistically insignificant. More importantly, their inclusion in the Fama-MacBeth regressions has no effect on the size and statistical significance of the dividend yield factor. We conclude, therefore, that empirically the dividend yield factor is unrelated to the bad-beta, good-beta model of Campbell and Vuolteenaho (2004).

An interesting aspect of the dividend yield factor is that it can be further decomposed into two components: the 12-month market dividend growth rate and the *lagged* market capital gains. We find that both components carry a positive risk premium in the cross-section of expected stock returns. However, only the lagged capital gains component is consistently significant and it carries a substantially higher positive premium than the dividend growth rate.

Based on this decomposition, a simple way to amend the CAPM is to form the "predictive CAPM," which includes the lagged market return in addition to the contemporaneous market factor. This model arises naturally as a special case of the two-factor CAPM since the lagged market return is perfectly correlated with the lagged capital gains component of the dividend yield factor. We find that the predictive CAPM delivers a positive predictive beta-return relation in the cross-section of expected stock returns. Portfolio sorts on the exposure to the lagged market return deliver a value-weighted H-L return spread of 0.36%, which is statistically significant. The factor premium in Fama-MacBeth regressions is also significant. We conclude, therefore, 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.

The remainder of the paper is organized as follows. In the next section, we motivate our empirical analysis by presenting our theoretical model. In Section 3, we describe the cross-sectional 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. Section 6 reports several robustness tests, inlcuding the relation to the bad-beta, good-beta model and taxes. The performance of the predictive CAPM is assessed in Section 7. Finally, we conclude in Section 8.

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:

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

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.[2]

For simplicity, and without loss of generality, we assume that all factors have a zero mean. We express the normal distributions of the factors in terms of their precision, which is the inverse of the variance: $\tau_g^{-1} = V(\tilde{g})$, $\tau_y^{-1} = V(\tilde{y})$ and $\tau_f^{-1} = V(\tilde{f})$, where $V(.)$ represents the variance of a random variable. In addition, we 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

[2] Any variable with a tilde is a random variable, whereas any variable without a tilde is a constant parameter of the model.

gains factor is much larger than the variance of the dividend yield factor: $V(\tilde{g}) \gg V(\tilde{y})$ in terms of variances or $\tau_g \ll \tau_y$ in terms of precisions.

2.2 Estimating the model through CAPM

If an econometrician tries to estimate the model in Equation (1) using the CAPM framework, then she 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), it is clear that $\hat{\beta}$ is a weighted average of $\hat{\beta}_g$ and $\hat{\beta}_y$, where the weights depend on the variance of each factor. Since the variance of the capital gains factor is much larger than the variance of the dividend yield factor, $\hat{\beta}$ is very close in value to $\hat{\beta}_g$. Consequently, the capital gains factor dominates the information contained in the dividend yield factor as long as $\hat{\beta}_g \neq \hat{\beta}_y$. In other words, estimating the CAPM model of Equation (2) effectively ignores most of the information contained in the dividend yield factor because of its small variance. If instead the econometrician were to estimate the true model described by Equation (1), then she would not ignore the information contained in the dividend yield factor.

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

2.3 Variance decomposition

It is straightforward to perform a variance decomposition of the asset payoffs by taking the variance of the payoffs $\tilde{v}$ as described in Equation (1) and plugging in the estimated factor loadings, to obtain:

$$\begin{aligned} V(\tilde{v}) &= \beta_g^2 V(\tilde{g}) + \beta_y^2 V(\tilde{y}) + V(\tilde{f}) \\ &= \frac{Cov(\tilde{v},\tilde{g})^2}{V(\tilde{g})} + \frac{Cov(\tilde{v},\tilde{y})^2}{V(\tilde{y})} + V(\tilde{f}) \\ &= Corr(\tilde{v},\tilde{g})^2 V(\tilde{v}) + Corr(\tilde{v},\tilde{y})^2 V(\tilde{v}) + V(\tilde{f}) \\ &= \frac{V(\tilde{f})}{1 - Corr(\tilde{v},\tilde{g})^2 - Corr(\tilde{v},\tilde{y})^2}, \end{aligned} \tag{4}$$

where $Corr(.)$ represents the correlation between two random variables. The variance decomposition of the asset payoffs shows that the correlation between asset payoffs and the dividend yield factor (i.e., $Corr(\tilde{v},\tilde{y})$) is a key component in explaining the variation of the asset payoffs and should not be overshadowed by the information contained in the capital gains factor.

2.4 Predictability of the dividend yield factor

Let's assume now that 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 $\tilde{\varepsilon}_f \sim N(0, \tau_{\varepsilon f}^{-1})$.[4] Under these signals, we can calculate the posterior

[4] 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 the information about capital gains would be much more precise than the information about

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:

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

In this setup, we are following the premise of Hartzmark and Solomon (2022) that investors do not process any information about the dividend yield factor. Market participants respond to prices (and hence capital gains) but ignore dividend yields.

Next, we measure the return predictive power of an *additional* signal about the capital gains factor $\widetilde{m}_g = \tilde{g} + \widetilde{\omega}_g$, where $\widetilde{\omega}_g \sim N(0, \tau_\omega^{-1})$ and the return predictive power of a signal about the dividend yield factor $\widetilde{m}_y = \tilde{y} + \widetilde{\omega}_y$, where $\widetilde{\omega}_y \sim N(0, \tau_\omega^{-1})$. Note that $\widetilde{\omega}_g$ and $\widetilde{\omega}_y$ have the same precision. Following Peng and Xiong (2006), the return predictive power of an additional signal $\widetilde{m}_g$ about the capital gains factor can be measured by the following correlation:

$$\left|Corr(\tilde{v} - \hat{v}), \widetilde{m}_g)\right| = \frac{\left|\beta_g\right| \tau_g}{\tau_g + \tau_{\varepsilon g}}, \tag{6}$$

---

the dividend yield. We have not added this additional signal in the main text because it unnecessarily complicates the model. Instead, 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 (2022).

[5] The posterior mean and variance are computed according to Hamilton (1994). Let Y_1 be a vector with mean μ_1, and Y_2 be a vector with mean μ_2, where the variance-covariance matrix is given by $\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$. If Y_1 and Y_2 are Gaussian, then: $Y_2 | \, Y_1 \sim N(\mu_2 + \Omega_{21}\Omega_{11}^{-1}(y_1 - \mu_1), \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12})$.

and the return predictive power of a signal $\widetilde{m}_y$ about the dividend yield factor by the following correlation:

$$\left|Corr(\tilde{v}-\hat{v}), \widetilde{m}_y)\right| = \left|\beta_y\right|. \tag{7}$$

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.

This framework allows us to show that a signal $\widetilde{m}_y$ about the dividend yield will have more predictive power than an additional signal $\widetilde{m}_g$ about the capital gains factor when:

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

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 $\left|\beta_y\right|\tau_{\varepsilon g}$ will be the dominant term and hence the inequality in Equation (8) above holds.

In words, if the information about capital gains is precise enough, then the predictive power of a signal about the dividend yield will be higher than the predictive power of additional information about capital gains.[6] Intuitively, the investor would

[6] We would reach a qualitatively similar conclusion if instead of comparing correlations, we compare posterior variances generated by $\widetilde{m}_g$ and $\widetilde{m}_y$.

like to reduce the uncertainty about the factor that is causing more uncertainty in the cash flows. The most uncertain factor is the capital gains factor due to is high variance. This is why investors spend substantial resources to learn about prices (and hence capital gains) so that $\tau_{\varepsilon g} \gg 0$. However, since investors already collect so much information about $\tilde{g}$, the factor that generates more uncertainty after all the collected information is taken into account is the dividend yield factor $\tilde{y}$. Thus, a signal $\widetilde{m}_y$ 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 $\widetilde{m}_g$.

3 Data

Our empirical analysis uses the cross-section of expected stock returns obtained from the CRSP database. The cross-section includes all common stocks traded on the NYSE, AMEX and NASDAQ exchanges (share codes of 10, 11 and exchange codes of 1, 2, 3). Following Fama and French (1993), our analysis uses stocks that satisfy 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 comprises the 42-year period ranging 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 plots the number of dividend payers and non-payers. The post-1978 sample period comprises the period during which the number of dividend payers and hence the importance of dividends declined steadily in the US equity market. This is an important aspect of the analysis because in our theoretical model the declining importance of dividends together with a behavioural bias against dividends is used to justify the high predictive power of the dividend yield factor.

At the beginning of the sample period, the cross-section comprises 3,532 eligible stocks, while at the end of the sample period it comprises 2,832 stocks. The maximum number of stocks across the sample period is 5,396, and the average number is 3,794. The value-weighted market return including dividends at time t is denoted by mkt_t. The value-weighted market return excluding dividends at time t is denoted by $mktx_t$. The value-weighted market price is denoted by p_t so that $mktx_t = p_t/p_{t-1} - 1$. In other words, $mktx_t$ captures the rate of change in capital gains.

Following Lee (1995), we compute the market dividend yield dy_t as follows: $dy_t = mkt_t - mktx_t$. The dividend yield is defined as $dy_t = d_t/p_{t-1}$, where d_t is the value-weighted sum of all dividends paid by all firms at time t. Note that the individual dividend paid by each firm is the US dollar value per share of distributions resulting from cash dividends, spin-offs, mergers, exchanges, reorganizations, liquidations, and rights issues. The definition of the market dividend yield implies that $mkt_t = (p_t + d_t)/\, p_{t-1} - 1$.

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, on a given month, approximately 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, approximately 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. It has become standard in the literature to account for dividend seasonality in a simple way: by using the trailing 12-month dividend yield defined as $dy12_t = \Sigma_{j=0}^{11} d_{t-j}/p_{t-1}$. Our main analysis focuses on $dy12_t$.[7]

In addition to seasonality, the dividend yield may also display non-stationarity (see, e.g., Welch and Goyal, 2008). This is shown in Figure 2, which plots the time-series of $dy12_t$.[8] 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 $dy12_t$: $\Delta dy12_t = \frac{dy12_t}{dy12_{t-1}} - 1$.

The final step in generating the dividend yield factor used in our analysis is that we orthogonalize $\Delta dy12_t$ relative to the 3-month Treasury bill, the term spread and the default spread. According to Petkova (2006), these three variables together with the dividend yield describe well the time-variation in the investment opportunity set.

[7] We have also implemented a more comprehensive way of accounting for seasonality by deseasonalizing the dividend yield using an ARMA model. The results remain qualitatively the same and are available upon request.

[8] For our sample period, the Augmented Dickey-Fuller statistic for the dy12 series is equal to -0.78. Hence the null hypothesis of a unit root cannot be rejected, which implies that the series is non-stationary.

Therefore, orthogonalization removes the effect of variables which are well-known to be correlated with the dividend yield factor. We orthogonalize by estimating a regression of $\Delta dy12_t$ on the three variables and then using the fitted residuals as our orthogonal $\Delta dy12_t$. We do so using a 20-year window and performing ex-ante estimation so that our results do not suffer from a forward looking bias. Henceforth, $\Delta dy12_t$ refers to the orthogonalized $\Delta dy12_t$ and represents the dividend yield factor used in our analysis. The time series of $dy12_t$ and the orthogonalized $\Delta dy12_t$ are displayed in Figure 2.

The data used in the orthogonalization are obtained as follows. Data on the long-term 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. The default yield spread is the difference between BAA and AAA-rated corporate bond yields.

Table 1 reports summary statistics. The main findings can be summarized as follows. As expected, the monthly market dividend yield (dy_t) accounts for a large part of the market return (mkt_t): on average $dy_t = 0.22\%$ per month, whereas $mkt_t = 1.04\%$ per month. Therefore, 21% of the market return is due to the dividend yield.

However, dy_t contributes very little to the variance of mkt_t. Specifically, mkt_t has 30 times higher standard deviation than dy_t (4.40% vs 0.14%). Put differently, mkt_t has 988 times higher variance than dy_t. Therefore, the dividend yield contributes to the mean but not the variance of the market return. Consequently, as seen in Table 1,

mkt_t and $mktx_t$ have almost identical higher moments (volatility, skewness and kurtosis).

Importantly, whereas mkt_t and $mktx_t$ display strong cyclical behaviour, the dividend yield displays strong countercyclical behaviour. The dividend yield is higher during NBER-defined recessions as opposed to expansions.[9] The behaviour of $\Delta dy12_t$ is also strongly countercyclical since on average it is *negative* in expansions but strongly *positive* in recessions.

Table 2 reports the cross-correlations between the variables. The results in this table indicate that mkt_t and $mktx_t$ exhibit perfect correlation. In contrast, the correlation between mkt_t and dy_t is equal to 0.10. We conclude, therefore, that the dividend yield is a low-volatility variable with a low correlation with the market return. When adding the dividend yield (dy_t) plus the high-volatility rate of capital gains ($mktx_t$), this creates a market return variable ($mkt_t = mktx_t + dy_t$), which has a perfect correlation with the rate of capital gains but a very low correlation with the dividend yield. In short, therefore, the contribution of the dividend yield is limited to the mean since it effectively contributes nothing to the variance of the market return.

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

4.1 A Two-Factor CAPM Model

[9] Note that recessions account for 12.8% of our sample period, whereas expansions account for 88.2%.

In this section, we assess the effect of risk due to changes in the market dividend yield on the cross-section of expected stock returns. Our analysis is based on a decomposition of the market return into a capital gains factor and a dividend yield factor in the context of the CAPM model. The standard CAPM model uses the market factor (mkt_t) that incorporates both capital gains ($mktx_t$) and the dividend yield (dy_t). However, as established in the previous section, mkt_t and $mktx_t$ are perfectly correlated. In contrast, the correlation between mkt_t and dy_t is close to zero. Furthermore, $mktx_t$ has enormously higher variance than dy_t. As a result, the variance of mkt_t is practically identical to that of $mktx_t$.

For these reasons, there is a fundamental concern with the CAPM regression: whether we use mkt_t or $mktx_t$, the result will be essentially the same and the contribution of dy_t will be silenced by the vastly more volatile $mktx_t$ component. One way to see this is to estimate the CAPM using the full cross-section of stock returns for two distinct cases: first, using mkt_t as the single factor; and second, using $mktx_t$ as the single factor. In doing so, we find that the average betas in both cases are identical to the second decimal. Furthermore, the quintile portfolios formed by sorting on the betas of either mkt_t or $mktx_t$ are practically identical. We conclude, therefore, that every time we estimate the CAPM, the dividend yield component is effectively silenced and all we observe is capital gains risk.[10]

---

[10] If for each stock we estimate the beta on mkt_t using a five-year rolling window, then the average of these betas across time and across stocks is equal to 1.107. If we repeat the same exercise replacing mkt_t

To address this issue, 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 ($\Delta dy12_t$). 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. We use $\Delta dy12_t$ (as opposed to dy_t) in order to explicitly account for the seasonality, non-stationarity and orthogonalization of the market dividend yield. Consequently, our main analysis is based on the following two-factor model inspired by the CAPM:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,c}(mktx_t - r_{f,t}) + \beta_{i,d}\Delta dy12_t + \varepsilon_{i,t}, \quad (9)$$

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. Consistent with the literature (see, e.g., Ang *et al.*, 2006), our main analysis (pre-formation regression) is based on this two-factor model but in later sections we also discuss the effect of including more factors in the post-formation regression testing.

4.2 Portfolio Sorts

We begin our empirical analysis by testing whether the loading on the dividend yield factor can predict the cross-section of expected stock returns. To do so, we first estimate the coefficient $\beta_{i,d}$ in Equation (9) for each stock in the cross-section using a five-year

by $mktx_t$, then the average beta is equal to 1.109. Therefore, conditioning on either the standard market factor or the capital gains factor leads to identical average betas to the second decimal.

rolling window. At the end of each month, we form quintile portfolios by sorting stocks on the loadings to the dividend yield factor. Stocks in the low (high) quintile have the lowest (highest) loadings on Δdy12 across all stocks in the cross-section. We then compute the one-month ahead mean returns of the quintile portfolios and rebalance monthly. Throughout our analysis, we report results for both value-weighted (VW) portfolios based on NYSE weights and for equally-weighted (EW) portfolios but our discussion will primarily focus on VW portfolios (see, e.g., Hou, Xue and Zhang, 2020).

Table 3 reports the performance of quintile portfolios sorted by exposure to the dividend yield factor. Our main finding in this table 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), which is long on the highest quintile portfolio and short on the lowest quintile portfolio, provides a mean return of -0.34%, which is highly statistically significant: the Newey-West (1987) *t*-statistic is equal to -2.68. Furthermore, the H-L portfolio delivers a negative and significant alpha relative to both the CAPM and the six-factor (FF6) model. The FF6 model incorporates the five Fama and French (2015) factors plus momentum. The results are similar and indeed stronger for EW portfolios. 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 coincide with the peak in the number of dividend payers and hence the beginning of the declining importance of dividends in the US equity market. These empirical facts provide a foundation for our theoretical model in motivating the predictive power of the dividend yield factor. 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 4. Our main finding is that for the period of 1942-1977, the H-L return spread is 0.10 and insignificant. In contrast, for 1978-2019 the H-L return spread is -0.34 and is highly significant. Therefore, the predictive power of the dividend yield factor is due to the post-1978 sample.[11]

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 CAPM and FF6 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 in addition to the conceptual motivation based on our theory for focusing on the post-1978 sample period.

4.4 Components of the Dividend Yield

[11] The earliest start date is 1942 because of the initial data required for orthogonalization.

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. For this reason, it is interesting to test whether the dividend yield factor is driven by one or both of its components. To do so, we decompose Δdy12 into its two components: $\Delta dy12_t \approx \Delta d12_t - \Delta p_{t-1}$, where $\Delta d12_t = \frac{d12_t}{d12_{t-1}} - 1$ is the proportional change in the 12-month trailing dividend series (i.e., the year-over-year dividend growth rate), and $\Delta p_{t-1} = \frac{p_{t-1}}{p_{t-2}} - 1$ is the proportional change in the lagged price series (i.e., the *lagged* mktx). Following Petkova (2006) and to be consistent with our previous analysis, we orthogonalize both components relative to the 3-month Treasury bill, the term spread and the default spread in the same way as the dividend yield factor.[12]

4.5 Factor-Mimicking Portfolios

The market factor is a tradable portfolio but its two components (capital gains and dividend yield) are not. In this section, we address the non-tradability of mktx and Δdy12 by constructing tradable factor-mimicking portfolios (FMPs) that mimic the behaviour of mktx and Δdy12.

[12] Since all three variables, Δdy12, Δp and Δd12 are defined as proportional changes, the decomposition of Δdy12 into Δp and Δd12 is not exact. However, the two components (Δp and Δd12) explain almost 100% of the variation of Δdy12. Specifically, Δp explains 90.5% of the variance of Δdy12, whereas Δd12 explains 9.5% of the variance of Δdy12.

For mktx, there is a straightforward solution: we replace mktx by mkt since the correlation between the two is equal to one. As mentioned earlier, the betas on mktx are essentially the same as the betas on mkt. Therefore, mkt can be thought of as the FMP of mktx.

For Δdy12, we implement two distinct approaches for generating the FMP: (1) the ordinary least squares (OLS) cross-sectional approach based on Lehmann and Modest (1988); and (2) the instrumental variables (IV) approach of Pukthuanthong, Roll, Wang and Zhang (2019). The FMP for Δdy12 is denoted by FΔdy12.

Finally, for Δd12 and Δp, we follow a similar approach to Δdy12: we form the FMP portfolios based on the OLS and IV approaches. These FMPs are denoted by FΔd12 and FΔp. Note that in constructing FΔdy12, FΔd12 and FΔp each underlying series is orthogonalized ex post.

4.5.1 The OLS approach

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:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i \Delta dy12_t + \varepsilon_t. \tag{10}$$

The beta estimates from Equation (10) are used in the second stage cross-sectional regressions with β_i as the only explanatory variable. This leads to the following OLS estimates of the factor premium:

$$\hat{\gamma}_t = (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'\left(r_t - r_{f,t}\right), \tag{11}$$

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, FΔdy12, is simply the $\hat{\gamma}_t$ estimate from Equation (11) above.

Note that the OLS approach generates an ex-post full sample FMP. Pukthuanthong *et al*. (2019) show that by design this ex-post FMP is maximally correlated to the original Δdy12 variable. This is consistent with our empirical evidence since, as we discuss later in detail, the OLS FMP has a high correlation with Δdy12.

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. In short, both from a methodology point of view and from a data selection point of view, the OLS FMP is designed to be maximally correlated to the dividend yield factor. Consequently, the OLS FMP will be the primary factor mimicking portfolio used in our analysis.

4.5.2 The IV approach

The second approach is designed to account for potential errors-in-variables biases and is based on the instrumental variables (IV) approach of Pukthuanthong *et al*. (2019). The IV approach is implemented by dividing the total sample into odd-month and even-month subsamples. For each of the two subsamples, we estimate the beta of each stock using Equations (12) and (13) below so that we have $\hat{\beta}^e$ for the sample of even months and $\hat{\beta}^o$ for the sample of odd months.

On each even month, the risk premium is estimated as follows:

$$\hat{\gamma}_t^e = (\hat{\beta}^{o\prime}\hat{\beta}^e)^{-1}\hat{\beta}^{o\prime}(r_t - r_{f,t}). \tag{12}$$

Correspondingly, on each odd month, the risk premium is estimated as follows:

$$\hat{\gamma}_t^o = (\hat{\beta}^{e\prime}\hat{\beta}^o)^{-1}\hat{\beta}^{e\prime}(r_t - r_{f,t}). \tag{13}$$

Then, the IV FMP is the combination of $\hat{\gamma}_t^o$ and $\hat{\gamma}_t^e$ for the full sample. For the IV FMP, we also use data only from dividend-paying firms.

4.5.3 FMP Correlations

Table 5 reports the full sample correlations between the Δdy12 series and the two FMPs. The table additionally reports the correlations between Δd12 and its FMPs as well as Δp with its FMPs. We find that the OLS FMPs exhibit a high correlation with the original series: 0.80 between OLS and Δdy12, 0.85 between OLS and Δd12, and 0.79 between OLS and Δp. The IV FMPs exhibit slightly lower correlations: 0.68 between IV and Δdy12, 0.71 between IV and Δd12, and 0.65 between IV and Δp. Our main analysis will be based on the OLS method but we will also report the IV results.

4.6 The Price of Dividend Yield Risk

4.6.1 Dividend Yield Risk

In this section, we formalize our analysis of the relation between expected stock returns and the dividend yield factor by estimating two-stage Fama and 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:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,1}(mktx_t - r_{f,t}) + \beta_{i,2}\Delta dy12_t$$
$$+\beta_{i,3}SMB_t + \beta_{i,4}HML_t + \beta_{i,5}RMW_t + \beta_{i,6}CMA_t + \beta_{i,7}MOM_t + \varepsilon_{i,t}, \quad (14)$$

where SMB is the Fama and French (1993) small-minus-big size factor, HML is the high-minus-low value factor, RMW is the Fama and French (2015) robust-minus-weak profitability factor, CMA is the conservative-minus-aggressive investment factor, and MOM is the Carhart (1997) momentum factor. Data on the SMB, HML, RMW, CMA and MOM factors are obtained from Ken French's online data library. The betas are estimated using a rolling window of 5 years of monthly data. In this regression, when using the FMPs, mktx is replaced by mkt, and Δdy12 by FΔdy12.

In the second stage, we condition on the beta estimates available on a given month, and perform estimation of the cross-sectional regression at each month *t* as follows:

$$r_{i,t} - r_{f,t} = a_i + \gamma_1\hat{\beta}_{i,1,t-1} + \gamma_2\hat{\beta}_{i,2,t-1}$$
$$+\gamma_3\hat{\beta}_{i,3,t-1} + \gamma_4\hat{\beta}_{i,4,t-1} + \gamma_5\hat{\beta}_{i,5,t-1} + \gamma_6\hat{\beta}_{i,6,t-1} + \gamma_7\hat{\beta}_{i,7,t-1} + \epsilon_{i,t}. \quad (15)$$

We collect the time-series of gamma estimates and report the mean as well as the Newey and 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 6 for the original factors as well as for the FMPs. Specifically, the column "Original" reports the results using the actual Δdy12 series. The column "OLS FMP" reports the results using FMPs created with the OLS method. Lastly, the column "IV FMP" reports the results using FMPs created with the IV method.

Our main finding is that the premium on the dividend yield factor is consistently negative and significant across all specifications. Specifically, Δdy12 exhibits a risk premium of -0.21% per month with a t-statistic of -2.35. The risk premium is equal to -0.28 for the OLS FMP and -0.26 for the IV FMP and both remain significant. It is interesting to note that the dividend yield factor has the largest risk premium (in absolute value) among all factors. In fact, none of the other risk factors are statistically significant. 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 strong evidence that the dividend yield factor has a negative and statistically significant price of risk in the context of the five-factor Fama-French (2015) model plus momentum. This evidence provides empirical justification for decomposing the market factor into a capital gains and a dividend yield factor and shows that the dividend yield is disconnected from capital gains in asset pricing.

In this context, it is important to understand why the dividend yield factor carries a negative premium in expected stock returns. Recall that the dividend yield factor is strongly countercyclical: Δdy12 is positive in recessions (1.02% per month) and negative in expansions (-0.13% per month). Consider an asset that has a high beta on the Δdy12 factor. By definition, a positive beta implies that this asset performs well when Δdy12 is high. However, Δdy12 is high in recessions. Therefore, this is an asset that performs well in recessions. According to standard asset pricing theory, this asset is valuable because it performs well when we need it the most (in the bad states of the world) and hence investors do not require a high expected return to hold it. As a result, high-beta assets on the Δdy12 factor will have low expected returns and vice versa. In short, the countercyclical behaviour of the dividend yield factor provides an explanation for its negative premium that is consistent with asset pricing theory.

4.6.2 Components of Dividend Yield Risk

In Panel B of Table 6, we replace Δdy12 by its two components: Δd12 and Δp. This allows us to examine which of the two components of Δdy12 is responsible for its cross-sectional predictive power. We find that the premium for Δp is high (0.22) and significant (t-stat=2.49) but the premium for Δd12 is much lower (0.03) and insignificant (t-stat=1.56). These premia remain similar in value and significance for the OLS and IV FMPs. 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.[13]

4.7 Post-Formation Factor Loadings

The performance of quintile portfolios formed on past exposure to the dividend yield factor shows that past loadings on Δdy12 can explain the cross-sectional variation of stock returns. However, a factor explanation requires a contemporaneous pattern between factor loadings and expected returns. In other words, it is important to assess whether the quintile portfolios load significantly on Δdy12 risk over the period used to evaluate portfolio performance. Following a long line of research in asset pricing (see,

[13] In unreported results, we find that Δd12 is also significant before the orthogonalization but it becomes insignificant after the orthogonalization.

among many others, Black *et al.,* 1972, Fama and French, 1992, 1993, and Ang *et al*., 2006), we use past information to form portfolios, and then proceed to examine contemporaneous post-formation loadings. Specifically, we use the FΔdy12 factor based on the OLS FMP approach to compute post-formation loadings reported on the last column of Table 3. The port-formation loadings are estimated ex-post for the full data sample using the seven factor model of Equation (14).

The results show that for all EW portfolios, the quintile portfolio returns load significantly on the FΔdy12 factor mimicking portfolio. Specifically, for EW portfolios, the post-formation loadings on FΔdy12 are negative and statistically 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. However, the results are weaker for VW portfolios since only for the Low quintile portfolio the post-formation loading is significant. In short, 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 weaker.

5 Is Exposure to the Dividend Yield Factor Related to Individual Dividend Yields?

In assessing the role of the dividend yield factor in predicting the cross-section of expected stock returns, two further questions arise: (1) is exposure to the dividend yield factor only relevant for dividend-paying firms or is it also relevant for non-dividend paying firms; and (2) for dividend-paying firms, are portfolios sorted on exposure to the dividend yield factor related to portfolios sorted on the individual dividend yield? Both questions are addressed in this section. First, we examine whether the effect of the dividend yield factor is different for dividend vs non-dividend paying firms. Second, for dividend paying firms only, we evaluate the performance of 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 a given month, an average of 46.5% of firms are identified as dividend payers and 53.5% are identified as non-payers.[14] The results are reported in Table 7.

The results indicate that the predictive power of the FΔdy12 is strong and significant for both dividend payers and non-dividend payers but it is stronger for dividend payers. Specifically, using the original Δdy12 series, the factor premium is -0.35 for payers (t-stat=2.98) and -0.21 for non-payers (t-stat=-2.64). The same is also true when using the OLS FMP approach: the factor premia are -0.43 vs. -0.26 and both are significant. In conclusion, the price of risk for the dividend yield factor is

[14] Note that these proportions change over time. For more details, see Figure 1.

significantly negative for both dividend payers and non-dividend payers but it is more so for dividend payers. Overall, Δdy12 is a systematic risk factor that affects all firms regardless of whether they pay dividends or not.

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. The individual dividend yield for each firm at time t is equal to the sum of dividends from t to t-11 divided by the price at t-1.

In Table 8, we report results for average monthly 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 result here is that there is little cross-sectional variation in the performance of stocks according to their dividend yield. For instance, the VW high-minus-low portfolio (H-L) displays a mean return of 0.12% per month, which is insignificant. Having said that, however, the CAPM-alpha and the six-factor alpha of the H-L portfolio tends to be positive and significant, especially for EW portfolios. For the most important case, however, VW portfolios and FF6 alphas, there is no significant relation. In conclusion, evidence based on returns shows that exposure to market dividend yield risk appears to be unrelated to firms' individual dividend yields.

6 Robustness

6.1 Bad Beta, Good Beta

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

$$r_{t+1} - E_t[r_{t+1}] = N_{CF,t+1} - N_{DR,t+1}, \tag{16}$$

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 ρ<1 is a discount coefficient. In words, Equation (16) 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 two-factor model in which they define the cash-flow beta as:

$$\beta_{i,CF} = \frac{Cov(r_{i,t}, N_{CF,t})}{Var(r_{m,t}^e - E_{t-1}r_{m,t}^e)}, \tag{17}$$

and the discount-rate beta as:

$$\beta_{i,DR} = \frac{Cov(r_{i,t}, N_{DR,t})}{Var(r_{m,t}^e - E_{t-1}r_{m,t}^e)}, \tag{18}$$

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:

$$\beta_{i,m} = \beta_{i,CF} + \beta_{i,DR}. \tag{19}$$

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. The former decomposes current market returns into current capital gains and the current dividend yield and then uses the latter to form a market dividend yield factor. Campbell and Shiller (1988) decompose shocks to future market returns into shocks to future cash flows and shocks to future discount rates. In the VAR model implemented by Campbell and Vuolteenaho (2004), lagged market returns are state variables, which by design are orthogonal to shocks in future market returns. Consequently, N_{CF} and N_{DR} are orthogonal to lagged market returns, which constitute one of the two components of the dividend yield factor. Therefore, theoretically it is unlikely that the Campbell and Shiller (1988) decomposition encompasses the same information as our decomposition. For this reason, we rely on an empirical analysis to determine the relation between the two models.[15]

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 (16) 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.[16] To be consistent with the ex-ante nature of our analysis, we compute ex-ante beta estimates using a 20-year rolling window each month. Therefore, our bad-beta, good-beta estimates avoid the forward-looking bias.

---

[15] Also note that in the Campbell and Vuolteenaho (2004) two-factor extension of the Merton (1973) ICAPM, the price of cash-flow risk and the price of discount risk are linearly related and their variance is the same. Our model differs in that respect since it allows the variance of the dividend yield factor to be different from that of the standard market factor.

[16] Specifically, the yield spread is defined as the ten-year constant maturity treasury bond yield less the annualized three-month constant maturity T-bill. The data are obtained from the FRED database of the Federal Reserve Bank of St. Louis. The smoothed price-earnings ratio of the S&P500 is defined as the log ratio of the index to a ten-year moving average of earnings. The data are obtained from Robert Shiller's website. The small-stock value spread is defined as the difference between the log book-to-market-ratios of small value and small growth stocks. The data are obtained from Ken French's online data library.

Armed with the ex-ante estimates of bad-beta and good-beta for each stock, we reestimate the Fama-MacBeth regressions to include these betas. The results are reported in Table 9. Our main finding is that the factor premiums for bad-beta and 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 (Δdy12) or its two components (Δd12 and Δp). We conclude, therefore, that empirically the dividend yield factor is unrelated to the bad-beta and good-beta of Campbell and Vuolteenaho (2004).

6.2 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 *Refinitiv Eikon* database. We form the ΔTax 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 FΔTax, 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 FΔTax in the Fama-MacBeth regressions, and report the results in Table 10 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 FΔTax does not affect the size and significance of Δdy12 for either dividend-payers or non-dividend payers. Indeed, the factor premium for Δdy12 is almost the same in Table 7 (no tax effect) and Table 10 (with tax effect). We conclude, therefore, that taxation is an unlikely explanation for the explanatory power of the dividend yield factor.

7 A Predictive CAPM Framework

In a previous section, we find that the factor premium associated with Δp is about seven times higher than the factor premium of Δd12, and the latter tends to be insignificant. The high relative size and significance of Δp over Δd12 motivates our examination of Δp in a separate manner. In this section, we isolate the effect of Δp in a simple yet powerful extension to the CAPM. We explicitly augment the standard CAPM to include the *lagged* excess return to the market. This additional variable essentially captures the effect of Δp since the two are perfectly correlated. We refer to this model as the "predictive CAPM," which is described by the following regression:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,1}(mkt_t - r_{f,t}) + \beta_{i,2}(mkt_{t-1} - r_{f,t-1}) + \varepsilon_{i,t}. \quad (20)$$

Based on the predictive CAPM, we 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." Dimson (1979) proposes the use of the sum of the two betas to account for biases arising from non-synchronous trading. This is a popular approach in the literature, which 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 11. Consistent with our previous results, the standard CAPM betas deliver an H-L spread that is low and insignificant. Even worse, the H-L alphas are strongly negative and highly significant, which is aligned with the betting-against-beta factor of Frazzini and Pedersen (2014). In contrast to the CAPM-alphas, the FF6-alphas are low and insignificant.

The Dimson betas slightly improve the CAPM performance: the quintile portfolios tend to exhibit increasing returns with beta. The H-L return is higher than the CAPM-betas but is still low and insignificant. The CAPM H-L alphas are negative and insignificant. The minor improvement of the Dimson betas indicates that accounting for non-synchronous trading by adding the beta of lagged market returns to that of contemporaneous market returns is insufficient in restoring the empirical failure of the CAPM.

Turning to portfolios sorted solely on the predictive betas, the results are striking. There is a positive monotonic relation between the predictive beta and average excess returns. For VW returns, the H-L spread is 0.36% with a *t*-statistic of 2.44. For EW returns, the H-L spread is 0.42% with a *t*-statistic of 2.50. The six-factor alphas are positive and significant. This finding indicates that cross-sectional predictability lies exclusively in the predictive betas, not in the standard 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 11, we report the factor premium of the Dimson beta in the presence of the SMB, HML, RMW, CMA and MOM factors. 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 two components of the Dimson beta, the standard contemporaneous beta and the predictive beta. The standard CAPM beta has a low and insignificant beta: 0.01 with a *t*-stat=0.11. In contrast, the predictive beta has a positive price of risk (0.21%) and highly significant (*t*-stat=2.43). Therefore, consistent with our previous results, the lagged market return alone is powerful in predicting the cross-section of expected stock returns.

As a final exercise, we add FΔd12 to the predictive CAPM framework. In doing so, we are effectively estimating a version of the original three-factor model (mktx, Δd12 and Δp) that was initially displayed in Panel B of Table 6. This is because mkt_{t-1} and Δp_{t-1} have a perfect correlation. The results are reported in Panel E of Table 11. We find that FΔd12 is low (0.03) and insignificant (t-stat=1.56), while its presence has no effect on the size and significance of the lagged market return. We conclude, therefore, 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.

8 Conclusion

The standard CAPM model is well known to empirically fail to distinguish between high-performing and low-performing assets. In other words, the betas on the market risk factor cannot predict the cross-section of expected stock returns. For example, the betting-against-beta factor tends to deliver high and significant excess returns (Frazzini and Pedersen, 2014). An important issue relating to the performance of the CAPM is that 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.

We propose a two-factor model that addresses this issue by separating the two components of the market portfolio. The two-factor model allows capital gains and the dividend yield to make distinct contributions to predicting the cross-section of expected stock returns. In doing so, the results are striking: the separate capital gains factor performs in the same way as the market portfolio, but the separate dividend yield factor performs very 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 of dividend-paying firms in the US. For this sample period, the high-minus-low VW quintile portfolio delivers a statistically singificant mean return of -0.34% per month. The alphas are also significantly negative for both VW and EW portfolios. Finally, Fama and MacBeth (1973) regressions confirm the presence of a significant negative premium for the dividend yield factor, which is unaffected by 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 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, 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.

It is important to note that the predictive power of the dividend yield factor appears to be universal: it is not limited to firms that pay dividends and it does not depend on the individual dividend yield of dividend-paying firms. Therefore, the dividend yield factor captures systematic risk for the entire cross-section of stock returns.

Finally, the dividend yield factor has itself two components: the dividend growth rate and lagged capital gains. The latter is by far the strongest of the two components since it has a much higher risk premium and is consistently significant. The size and significance of the factor price associated with the lagged market capital gains can be 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 enormous variance differential between capital gains and the dividend yield. The evidence indicates that these extensions to the CAPM can establish a significant beta-return relation, which is the cornerstone of an asset pricing model. For these reasons, a simple CAPM model that conditions on the market dividend yield (or its main component, the lagged capital gain) is 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.

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., 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.

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

Dimson, E. (1979). "Risk Measurement when Shares are Subject to Infrequent Trading," *Journal of Financial Economics* **7**, 197-226.

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

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.

Frazzini, A., and L.H. Pedersen (2014). "Betting against Beta," *Journal of Financial Economics* **111**, 1–25.

Hamilton, James D. (1994). *Time series analysis*. Princeton, N.J.: Princeton University Press.

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.

Lee, B.-S. (1995). "The Response of Stock Prices to Permanent and Temporary Shocks to Dividends," *Journal of Financial and Quantitative Analysis* **30**, 1-22.

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, 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.

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

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.

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.

Pukthuanthong, K., R. Roll, J. Wang, and T. Zhang (2019). "A New Method for Factor-Mimicking Portfolio Construction." *Available at SSRN 3341604*.

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

Welch, I., and A. Goyal. (2008). "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies* **21**, 1455-1508.

Table 1: 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, Δdy12 is the orthogonalized monthly proportional change in the 12-month market dividend yield, Δd12 is the orthogonalized monthly growth rate in the 12-month market dividend and Δp is the orthogonalized 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.

	Mean	St. Dev.	Min	Max	Skewness	Kurtosis	AR(1)
Panel A: Full Sample							
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
Δdy12	0.00	4.45	-11.65	28.00	0.93	6.70	0.02
Δd12	0.00	1.46	-9.35	17.95	3.37	52.04	0.04
Δp	0.00	4.21	-22.66	12.01	-0.53	4.93	0.01
Panel B: Expansions							
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	
Δdy12	-0.13	4.26	-11.65	28.00	1.12	8.07	
Δd12	0.05	1.49	-9.35	17.95	3.50	52.91	
Δp	0.15	3.97	-22.66	12.01	-0.66	5.74	
Panel C: 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	1.02	5.71	-11.32	12.72	0.03	2.46	
Δd12	-0.38	1.09	-2.64	2.10	0.00	2.72	
Δp	-1.22	5.69	-12.64	11.46	0.18	2.58	

Table 2: Cross-Correlations

This table reports the cross-correlations for the variables defined in Table 1. The sample period ranges from January 1978 to December 2019. Expansions and recessions are defined according to the NBER.

	mkt	mktx	dy	Δdy12	Δd12	Δp
Panel A: Full Sample						
mkt	1					
mktx	1.00	1				
dy	0.10	0.07	1			
Δdy12	-0.03	-0.03	0.06	1		
Δd12	0.02	0.02	0.17	0.24	1	
Δp	0.03	0.04	0.00	-0.94	0.08	1
Panel B: Expansions						
mkt	1					
mktx	1.00	1				
dy	0.11	0.08	1			
Δdy12	0.05	0.05	0.11	1		
Δd12	0.01	0.01	0.14	0.29	1	
Δp	-0.06	-0.05	-0.05	-0.94	0.05	1
Panel C: Recessions						
mkt	1					
mktx	1.00	1				
dy	0.16	0.13	1			
Δdy12	-0.31	-0.31	-0.18	1		
Δd12	0.06	0.04	0.60	-0.05	1	
Δp	0.32	0.31	0.29	-0.98	0.21	1

Table 3: Portfolios Sorted by Exposure to the Dividend Yield Factor

This table presents the performance of portfolios sorted by the exposure (beta) of individual stock excess returns to the dividend yield factor, Δdy12. We form value-weighted portfolios based on the 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 CAPM and FF6 Alpha columns report Jensen's alpha with respect to the CAPM and the Fama–French (2015) five-factor model plus momentum. Post-formation betas are according to Equation (14) using the OLS 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

	Returns						Factor Loadings	
Rank	Mean	St Dev	Size	B/M	CAPM Alpha	FF6 Alpha	Pre-Formation $\beta_{\Delta dy12}$	Post-Formation $\beta_{F\Delta dy12}$
High	0.65	4.64	17.28	0.48	-0.07	0.00	0.35	0.00
4	0.62	4.00	17.51	0.51	0.00	-0.17***	0.10	0.01
3	0.78	4.14	17.52	0.53	0.13***	0.07	-0.05	0.01
2	0.74	4.53	17.40	0.55	0.02	-0.02	-0.22	0.02
Low	0.99	5.65	16.54	0.59	0.12	0.26***	-0.58	-0.03*
H-L	-0.34***	2.61			-0.19*	-0.26**		0.03
(t-stat)	(-2.68)				(-1.68)	(-2.49)		(1.02)

Panel B: Equally-Weighted Portfolios

	Returns						Factor Loadings	
Rank	Mean	St Dev	Size	B/M	CAPM Alpha	FF6 Alpha	Pre-Formation $\beta_{\Delta dy12}$	Post-Formation $\beta_{F\Delta dy12}$
High	0.70	5.57	14.48	0.74	-0.07	0.06	0.51	-0.06***
4	0.88	4.55	14.96	0.75	0.22**	0.13***	0.08	-0.05***
3	0.93	4.67	14.83	0.78	0.27**	0.20***	-0.12	-0.07***
2	0.94	5.30	14.16	0.85	0.21	0.26***	-0.35	-0.11***
Low	1.10	7.40	12.96	0.89	0.19	0.66***	-0.99	-0.23***
H-L	-0.39**	3.11			-0.26*	-0.60***		0.17***
(t-stat)	(-2.58)				(-1.84)	(-5.05)		(4.30)

Table 4: 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 individual excess stock 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 CAPM and FF6 Alpha rows report Jensen's alpha with respect to the CAPM and the Fama–French (2015) five-factor model plus momentum. Statistical significance is assessed using Newey–West (1987) *t*-statistics.

Panel A: Value-Weighted H-L Portfolios								
	1942-1977	1978-2019	1963-2019	1968-2019	1973-2019	1978-2019	1983-2019	1988-2019
H-L Returns	0.10	-0.34***	-0.20*	-0.25**	-0.29**	-0.34***	-0.45***	-0.36**
(t-stat)	(0.75)	(-2.68)	(-1.66)	(-2.06)	(-2.31)	(-2.68)	(-3.20)	(-2.35)
CAPM Alpha	0.24**	-0.19*	-0.05	-0.11	-0.13	-0.19*	-0.29**	-0.20
(t-stat)	(2.00)	(-1.68)	(-0.51)	(-0.98)	(-1.18)	(-1.68)	(-2.27)	(-1.42)
FF6 Alpha	0.17*	-0.26**	-0.04	-0.07	-0.12	-0.26**	-0.36***	-0.27**
(t-stat)	(1.81)	(-2.49)	(-0.47)	(-0.79)	(-1.23)	(-2.49)	(-3.12)	(-2.20)

Panel B: Equally-Weighted H-L Portfolios								
	1942-1977	1978-2019	1963-2019	1968-2019	1973-2019	1978-2019	1983-2019	1988-2019
H-L Returns	0.00	-0.39**	-0.28**	-0.36***	-0.37***	-0.39**	-0.49***	-0.52***
(t-stat)	(0.00)	(-2.58)	(-2.29)	(-2.72)	(-2.68)	(-2.58)	(-2.88)	(-2.83)
CAPM Alpha	0.06	-0.26*	-0.19	-0.26**	-0.25*	-0.26*	-0.35**	-0.37**
(t-stat)	(0.59)	(-1.84)	(-1.61)	(-2.06)	(-1.92)	(-1.84)	(-2.18)	(-2.09)
FF6 Alpha	0.12	-0.60***	-0.40***	-0.47***	-0.50***	-0.60***	-0.70***	-0.71***
(t-stat)	(1.53)	(-5.05)	(-4.27)	(-4.58)	(-4.61)	(-5.05)	(-5.29)	(-4.92)

Table 5: Factor-Mimicking Portfolios

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

	Cross-Correlations		
	Δdy12	OLS FMP	IV FMP
Δdy12	1		
OLS FMP	0.80	1	
IV FMP	0.68	0.90	1
	Δd12	OLS FMP	IV FMP
Δd12	1		
OLS FMP	0.85	1	
IV FMP	0.71	0.83	1
	Δp	OLS FMP	IV FMP
Δp	1		
OLS FMP	0.79	1	
IV FMP	0.65	0.90	1

Table 6: 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 (21). The table also reports Newey–West (1987) *t*-statistics. The column "Original" uses the original Δdy12, Δd12 and Δp factors. The columns "OLS FMP" and "IV FMP" report results using OLS and IV FMPs for FΔdy12, FΔd12 and FΔp. 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

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
mkt-rf	0.02	0.13	0.02	0.15	0.02	0.15
FΔdy12	-0.21	-2.35	-0.28	-2.24	-0.26	-1.71
SMB	-0.03	-0.40	-0.03	-0.48	-0.03	-0.48
HML	0.08	1.07	0.08	1.07	0.08	1.07
RMW	-0.01	-0.16	-0.01	-0.17	-0.01	-0.16
CMA	0.02	0.50	0.02	0.50	0.03	0.53
MOM	-0.11	-1.56	-0.10	-1.41	-0.11	-1.47

Panel B: Components of the Dividend Yield Factor

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
mkt-rf	0.01	0.11	0.02	0.20	0.03	0.22
FΔd12	0.03	1.56	0.03	1.43	0.03	1.47
FΔp	0.22	2.49	0.27	2.29	0.24	1.73
SMB	-0.03	-0.46	-0.03	-0.48	-0.03	-0.47
HML	0.08	1.02	0.08	1.00	0.08	1.00
RMW	-0.01	-0.12	-0.01	-0.09	-0.01	-0.08
CMA	0.02	0.47	0.02	0.43	0.02	0.46
MOM	-0.11	-1.50	-0.10	-1.34	-0.10	-1.42

Table 7: 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 6. The sample period ranges from January 1978 to December 2019.

Panel A: Dividend Payers

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
mkt-rf	0.12	0.83	0.12	0.82	0.13	0.84
FΔdy12	-0.35	-2.98	-0.43	-2.79	-0.44	-2.20
SMB	0.06	0.76	0.05	0.66	0.05	0.60
HML	0.11	1.47	0.12	1.49	0.12	1.48
RMW	-0.07	-1.27	-0.07	-1.32	-0.07	-1.28
CMA	0.03	0.57	0.03	0.62	0.03	0.68
MOM	-0.02	-0.25	-0.02	-0.25	-0.03	-0.36

Panel B: Non-Dividend Payers

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
mkt-rf	-0.01	-0.06	0.00	-0.02	0.00	-0.03
FΔdy12	-0.21	-2.64	-0.26	-2.55	-0.26	-2.09
SMB	-0.03	-0.46	-0.03	-0.52	-0.03	-0.49
HML	0.05	0.78	0.05	0.77	0.05	0.79
RMW	0.00	-0.04	0.00	-0.05	0.00	-0.05
CMA	0.02	0.45	0.02	0.43	0.02	0.46
MOM	-0.10	-1.53	-0.09	-1.38	-0.10	-1.43

Table 8: 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 CAPM and FF6 Alpha columns report Jensen's alpha with respect to the CAPM and the Fama–French (2015) five-factor model plus momentum. The sample period ranges from January 1978 to December 2019.

Panel A: Value-Weighted Portfolios							
	Returns						
Rank	Mean	St Dev	Size	B/M	CAPM Alpha	FF6 Alpha	Dividend Yield
High	0.74	4.03	17.39	0.77	0.26*	-0.07	5.67
4	0.80	3.94	17.61	0.57	0.25***	-0.05	3.29
3	0.76	4.20	17.80	0.45	0.15*	-0.18***	2.29
2	0.74	4.63	17.31	0.44	0.05	-0.15**	1.50
Low	0.62	5.08	17.06	0.43	-0.17**	-0.19***	0.63
H-L	0.12	4.04			0.43**	0.12	
(*t*-stat)	(0.61)				(2.40)	(0.90)	
Panel B: Equally-Weighted Portfolios							
	Returns						
Rank	Mean	St Dev	Size	B/M	CAPM Alpha	FF6 Alpha	Dividend Yield
High	0.87	3.84	15.00	0.95	0.39***	0.16**	7.45
4	0.96	4.04	15.24	0.80	0.41***	0.12*	3.28
3	0.91	4.45	15.25	0.74	0.29**	-0.02	2.30
2	0.93	4.58	15.08	0.69	0.27**	-0.02	1.49
Low	0.79	5.02	15.03	0.63	0.06	-0.14**	0.65
H-L	0.08	2.62			0.33***	0.30***	
(*t*-stat)	(0.60)				(3.12)	(3.22)	

Table 9: Bad-Beta, Good-Beta

This table reports the Fama–MacBeth (1973) factor premiums for bad-beta, good-beta regressions using the full cross-section 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 6. The sample period ranges from January 1978 to December 2019.

Panel A: Dividend Yield Factor

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
β_{CF}	-0.02	-0.16	-0.03	-0.21	-0.02	-0.16
β_{DR}	0.03	0.94	0.04	1.27	0.03	1.10
FΔdy12	-0.21	-2.36	-0.27	-2.24	-0.26	-1.69
SMB	-0.03	-0.45	-0.04	-0.52	-0.04	-0.51
HML	0.08	1.05	0.08	1.05	0.08	1.05
RMW	-0.01	-0.10	-0.01	-0.11	-0.01	-0.10
CMA	0.02	0.48	0.02	0.49	0.02	0.51
MOM	-0.11	-1.56	-0.10	-1.43	-0.11	-1.50

Panel B: Components of the Dividend Yield Factor

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
β_{CF}	-0.03	-0.25	-0.03	-0.26	-0.02	-0.17
β_{DR}	0.04	1.17	0.05	1.59	0.04	1.35
FΔd12	0.02	1.48	0.03	1.37	0.03	1.41
FΔp	0.22	2.51	0.26	2.29	0.23	1.70
SMB	-0.03	-0.49	-0.04	-0.53	-0.04	-0.53
HML	0.07	0.99	0.07	0.99	0.07	0.98
RMW	0.00	-0.06	0.00	-0.04	0.00	-0.01
CMA	0.02	0.46	0.02	0.43	0.02	0.45
MOM	-0.11	-1.51	-0.10	-1.37	-0.11	-1.45

Table 10: 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. FΔtax is the factor-mimicking portfolio for the change in the implied tax rate. The specification of the regressions is the same as in Table 7. The sample period ranges from January 1978 to December 2019.

Panel A: Dividend Payers

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
FΔtax	-0.05	-0.63	-0.06	-0.50	0.00	-0.01
Mkt-rf	0.13	0.88	0.14	0.93	0.15	1.01
FΔdy12	-0.35	-3.01	-0.41	-2.66	-0.43	-2.15
SMB	0.06	0.80	0.06	0.75	0.06	0.73
HML	0.12	1.52	0.11	1.44	0.12	1.45
RMW	-0.07	-1.24	-0.07	-1.28	-0.07	-1.24
CMA	0.03	0.56	0.03	0.54	0.03	0.53
MOM	-0.03	-0.29	-0.03	-0.27	-0.03	-0.36

Panel B: Non-Dividend Payers

	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
FΔtax	-0.02	-0.25	0.01	0.11	0.10	0.84
Mkt-rf	0.00	0.02	0.01	0.12	0.02	0.14
FΔdy12	-0.20	-2.56	-0.24	-2.33	-0.25	-1.91
SMB	-0.02	-0.39	-0.03	-0.52	-0.02	-0.44
HML	0.05	0.76	0.05	0.66	0.05	0.71
RMW	0.00	-0.06	0.00	-0.04	0.00	-0.07
CMA	0.01	0.32	0.01	0.27	0.01	0.28
MOM	-0.11	-1.60	-0.10	-1.46	-0.10	-1.56

Table 11: Predictive CAPM

This table displays the performance of the predictive CAPM. Panels A and B report the performance of value-weighted and equally-weighted 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. Betas are estimated using the most recent 5 years of monthly data. Statistical significance is assessed using Newey–West (1987) t-statistics. Post-formation betas are computed using a seven-factor model based on the five-factor model of Fama and French (2015) plus momentum plus the OLS factor-mimicking portfolio for 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

	Standard CAPM beta			Dimson (1979) CAPM beta			Predictive CAPM beta			
Rank	Return	CAPM Alpha	FF6 Alpha	Return	CAPM Alpha	FF6 Alpha	Return	CAPM Alpha	FF6 Alpha	Post-Formation $\beta_{FMKT_{t-1}}$
High	0.70	-0.37***	0.06	0.78	-0.30**	0.12	0.99	0.10	0.31***	0.05**
4	0.85	0.04	-0.05	0.79	-0.06	-0.04	0.73	0.00	-0.01	0.02
3	0.85	0.16**	-0.07	0.85	0.12*	-0.02	0.78	0.13***	0.04	-0.02
2	0.76	0.19**	-0.13**	0.75	0.15**	-0.11***	0.67	0.06	-0.09**	0.00
Low	0.64	0.23***	-0.07	0.66	0.24***	-0.08	0.63	-0.07	-0.05	-0.04**
H-L	0.06	-0.60***	0.13	0.12	-0.54***	0.21	0.36**	0.17	0.36***	0.08***
(t-stat)	(0.19)	(-2.85)	(0.81)	(0.39)	(-2.54)	(1.32)	(2.44)	(1.27)	(3.25)	(2.75)

Panel B: Equally-Weighted Portfolios

	Standard CAPM beta			Dimson (1979) CAPM beta			Predictive CAPM beta			
Rank	Return	CAPM Alpha	FF6 Alpha	Return	CAPM Alpha	FF6 Alpha	Return	CAPM Alpha	FF6 Alpha	Post-Formation $\beta_{FMKT_{t-1}}$
High	0.86	-0.27	0.45***	0.96	-0.15	0.58***	1.10	0.19	0.68***	0.28***
4	0.92	0.05	0.20***	0.99	0.12	0.30***	0.96	0.22	0.28***	0.14***
3	1.03	0.31**	0.23***	0.94	0.23*	0.17***	0.94	0.26**	0.21***	0.10***
2	0.93	0.35***	0.19***	0.94	0.36***	0.17***	0.87	0.22**	0.12**	0.06***
Low	0.82	0.36***	0.26***	0.73	0.26**	0.11	0.68	-0.08	0.03	0.06***
H-L	0.05	-0.63***	0.19	0.23	-0.41*	0.47***	0.42**	0.26*	0.65***	0.22***
(t-stat)	0.15	(-2.82)	(1.26)	(0.73)	(-1.66)	(2.78)	(2.50)	(1.67)	(5.00)	(3.97)

Panel C: Fama-MacBeth Regressions						
	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
Dimson Beta	0.11	0.92	0.11	0.92	0.08	0.66
SMB	-0.03	-0.48	-0.06	-0.78	-0.06	-0.84
HML	0.07	1.02	0.07	0.94	0.08	1.02
RMW	-0.02	-0.31	-0.02	-0.33	-0.02	-0.30
CMA	0.01	0.24	0.01	0.26	0.02	0.47
MOM	-0.12	-1.64	-0.09	-1.27	-0.09	-1.30

Panel D: Fama-MacBeth Regressions						
	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
$(mkt\text{-}rf)_t$	0.01	0.11	0.02	0.16	0.02	0.16
$F(mkt\text{-}rf)_{t-1}$	0.21	2.43	0.26	2.24	0.23	1.68
SMB	-0.03	-0.42	-0.03	-0.48	-0.03	-0.47
HML	0.08	1.06	0.08	1.06	0.08	1.08
RMW	-0.01	-0.16	-0.01	-0.15	-0.01	-0.15
CMA	0.02	0.50	0.02	0.49	0.03	0.54
MOM	-0.11	-1.54	-0.10	-1.37	-0.10	-1.44

Panel E: Fama-MacBeth Regressions						
	Original		OLS FMP		IV FMP	
	Mean	NW *t*-stat	Mean	NW *t*-stat	Mean	NW *t*-stat
$(mkt\text{-}rf)_t$	0.01	0.11	0.02	0.20	0.03	0.23
$F(mkt\text{-}rf)_{t-1}$	0.21	2.46	0.26	2.25	0.24	1.73
FΔd12	0.03	1.56	0.03	1.44	0.03	1.46
SMB	-0.03	-0.45	-0.03	-0.48	-0.03	-0.47
HML	0.08	1.01	0.08	1.00	0.08	1.00
RMW	-0.01	-0.12	-0.01	-0.10	-0.01	-0.09
CMA	0.02	0.46	0.02	0.44	0.02	0.45
MOM	-0.11	-1.49	-0.10	-1.33	-0.10	-1.41

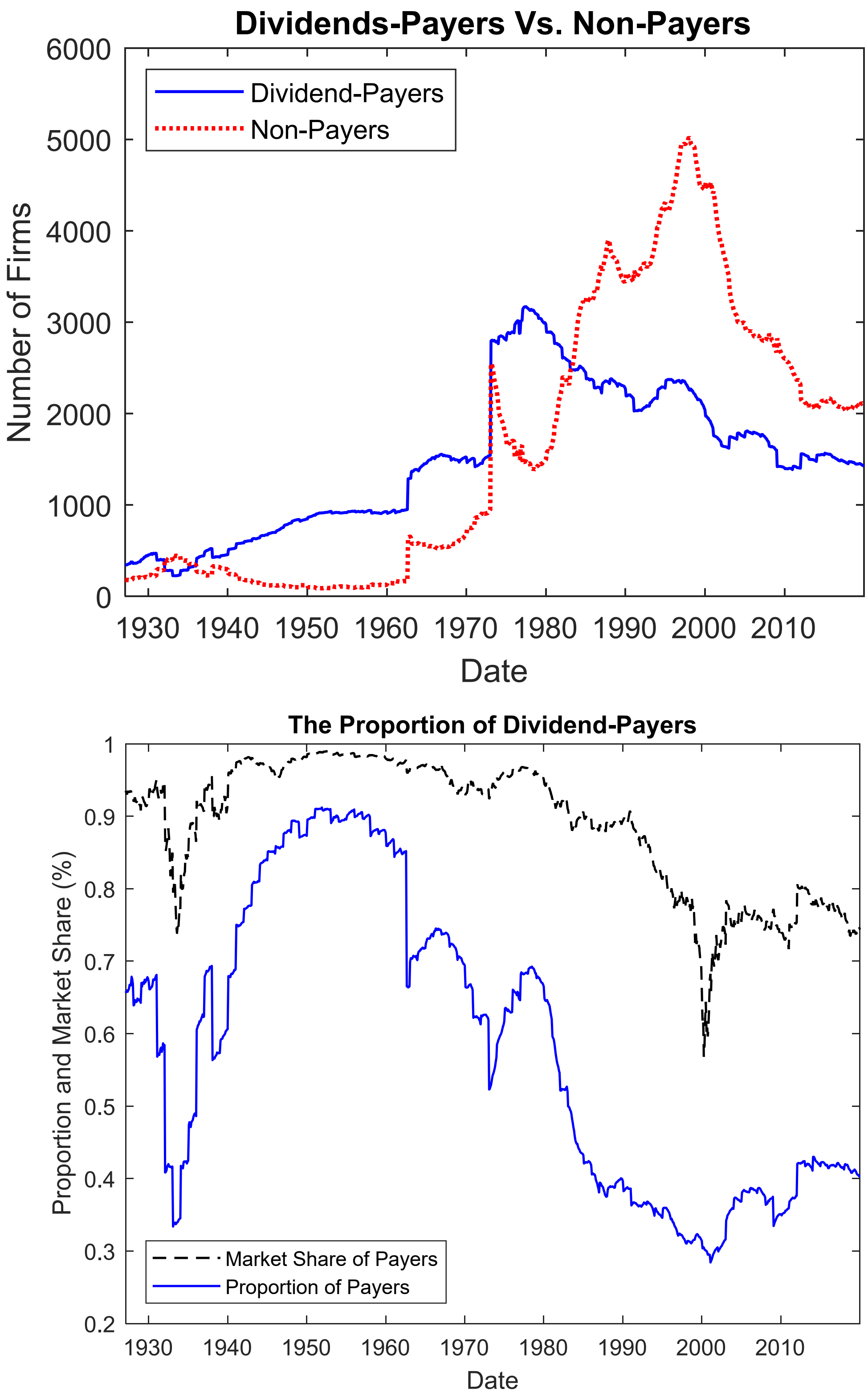

Figure 1

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 and their market share for the same sample period.

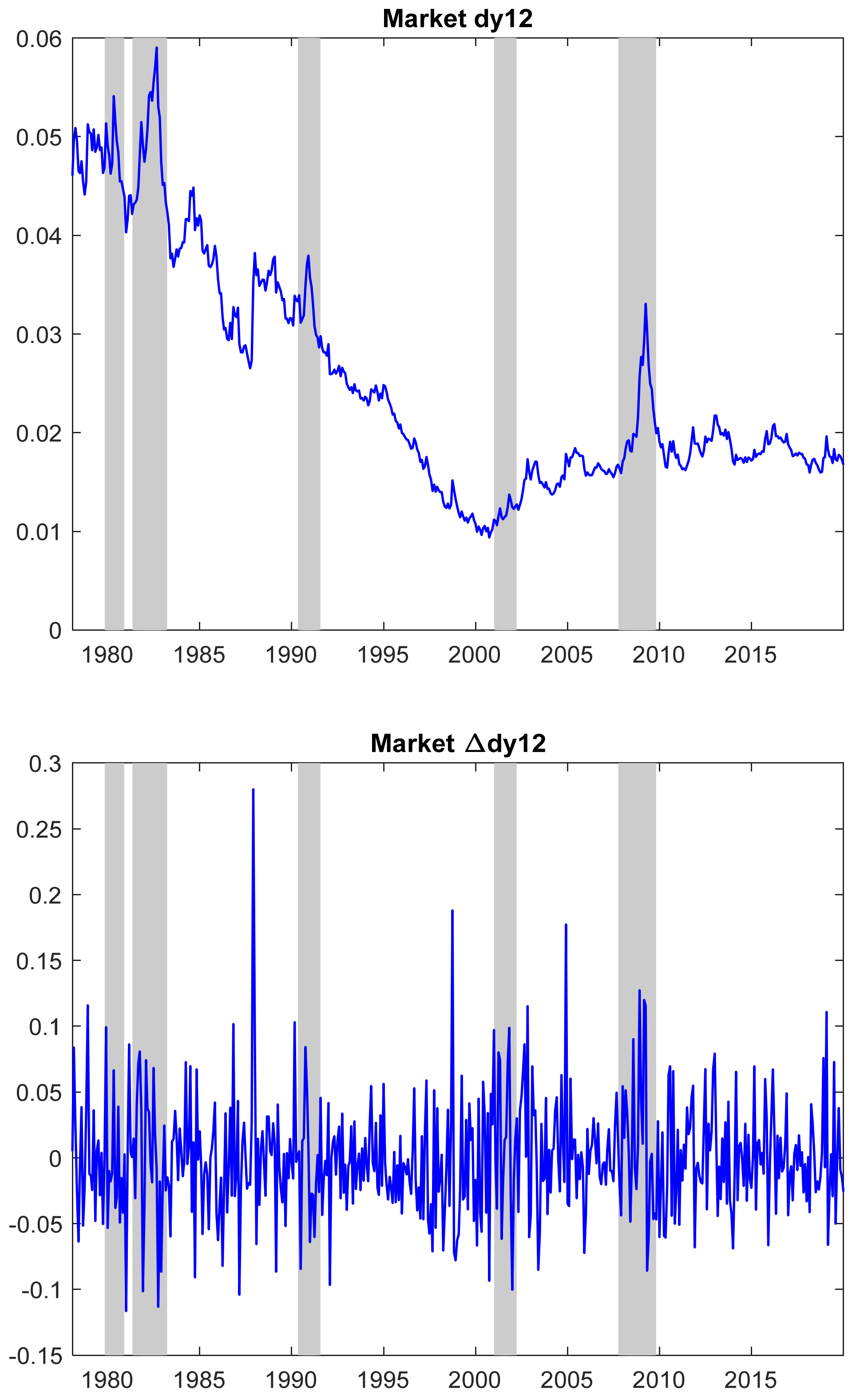

Figure 2

This figure displays the time series of the market dividend yield, dy12, and the orthogonalized dividend yield factor, Δdy12, for the sample period of January 1978 to December 2019. The shaded areas indicate NBER recessions.