The predictive power of the dividend risk premium

Article

Accepted Version

Avino, D., Stancu, A. and Wese Simen, C. (2020) The predictive power of the dividend risk premium. Journal of Financial and Quantitative Analysis. ISSN 1756-6916 doi: https://doi.org/10.1017/S0022109020000733 Available at https://centaur.reading.ac.uk/83250/

To link to this article DOI: http://dx.doi.org/10.1017/S0022109020000733

Publisher: Cambridge University Press

The Predictive Power of the Dividend Risk Premium

Davide E. Avino, Andrei Stancu and Chardin Wese Simen*

Abstract

We show that the dividend growth rate implied by the options market is informative about (i) the expected dividend growth rate and (ii) the expected dividend risk premium. We model the expected dividend risk premium and explore its implications for the predictability of dividend growth and stock market returns. Correcting for the expected dividend risk premium strengthens the evidence of dividend growth and stock market return predictability both in- and out-of-sample. Economically, a market timing investor who accounts for the time varying expected dividend risk premium realizes an additional utility gain of 2.02 % per year.

JEL classification: C22, C53, G12, G13, G17

Keywords: Dividend risk premium, dividend strip, predictability, present value model

---

*Avino, Management School, University of Liverpool, L69 7ZH, Liverpool, United Kingdom. Contact: d.avino@liverpool.ac.uk; Stancu, Norwich Business School, University of East Anglia, Norwich NR4 7TJ, United Kingdom. Contact: a.stancu@uea.ac.uk; and Wese Simen, ICMA Centre, Henley Business School, University of Reading, Reading, RG6 6BA, United Kingdom. Contact: c.wese-simen@icmacentre.ac.uk.
We are grateful to Jennifer Conrad (the editor) and an anonymous reviewer for constructive comments and suggestions. We are also grateful to Carol Alexander, Nikolaos Antypas, Mascia Bedendo, Maria Boutchkova, Michael Brennan, Chris Brooks, Mike Clements, Olivier Darné, Alfonso Dufour, Carlo Favero, Benjamin Golez, Ranko Jelic, Andreas Kaeck, Andrew Karolyi, Alex Kostakis, Miriam Marra, Michael McKenzie, Alexander Mihailov, Tony Moore, Ogonna Nneji, Ioannis Oikonomou, Carlton-James Osakwe (NFA Discussant), Bradley Paye, Benoît Sévi, Radu Stancu, Simone Varotto, Fang Xu, Yeqin Zheng, Lillian Zhu, seminar participants at the Welsh Accounting & Finance Colloquium, the Midwest Econometrics Group, the ICMA Centre, the University of Edinburgh Business School, the Université de Nantes (LEMNA), the University of Reading (Economics), the University of Sussex, and the Northern Finance Association for helpful comments and discussions. Part of this research project was completed when Chardin Wese Simen visited Leibniz University Hannover.

I. Introduction

The dividend growth forecast implied by the options market is informative about the *risk-adjusted* expectations of future dividend growth. More specifically, the implied dividend growth rate (ig) contains information about (i) the expected dividend growth rate and (ii) the expected dividend risk premium. This insight raises a number of questions. For instance, is ig mainly informative about the expected dividend growth rate or the expected dividend risk premium? What are the theoretical implications of the expected dividend risk premium for the predictability of dividend growth rates and aggregate stock returns?

Addressing these questions is important because a time varying expected dividend risk premium confounds the information content of ig for the expected dividend growth rate. Thus, it might be important to account for these variations when using ig to forecast dividend growth. Furthermore, the logic of present value models suggests that the dividend price (dp) ratio reveals information about the difference between expected stock returns and expected dividend growth rates (Campbell and Shiller, 1988). To the extent that the expected dividend growth rate is time varying, we need to correct the standard dp ratio for these variations in order to strengthen the predictability of stock returns (Campbell, 2008).

This paper makes three contributions to the literature. First, we formally show that ig contains information about the future (i) dividend growth rate and (ii) dividend risk premium. Using a dataset of intraday option prices covering the period 1996–2017, we show that 58 % of the fluctuations in ig are related to the expected dividend risk premium (drp). This leads us to conclude that the expected drp does not only move over time but it is also an important driver of the variations in ig.

Second, we propose a model for the expected drp. In particular, we assume that it depends on the lagged ig and the lagged drp. Although admittedly simple, this 2-factor model achieves a satisfactory empirical performance evidenced by an r-squared (R^2) of more than 60 %. Our model predicts that the implied growth rate corrected for the expected dividend risk premium (ig^{corr}), a linear combination of the lagged ig and the lagged drp, should forecast dividend growth with a slope coefficient equal to 1. We find empirical evidence in support of this prediction. We also examine the out-of-sample performance of ig^{corr} and find that it yields a significantly positive out-of-sample R^2 (R^2_{oos}), while the forecasts implied by ig do not.

Third, we develop a present value model to study the implications of the time varying expected dividend risk premium for the predictability of aggregate stock returns. Our model predicts that the lagged corrected dividend price (dp^{corr}) ratio, an affine function of (i) the lagged standard dp ratio, (ii) the lagged ig, and (iii) the lagged drp, forecasts returns. A regression of the 1-month stock market returns on a constant and the lagged dp^{corr} ratio yields a statistically significant slope estimate. We compare the predictive ability of the standard dp ratio, the dp^{ig} ratio of Golez (2014) (which ignores the variations in the expected dividend risk premium) and the dp^{corr} ratio. Our results reveal that, of all three forecasting variables, the dp^{corr} ratio is the most significant predictor of 1-month returns (t-statistic=2.15 and R^2=1.45 %). Controlling for other predictors of stock market returns does not change this conclusion. Out-of-sample, the dp^{corr} ratio leads to a significantly positive R^2_{oos}. Moreover, this positive R^2_{oos} has economic value. Relative to a strategy based on the recursive mean, an investor with a risk aversion coefficient equal to 4 who follows a timing strategy based on the dp^{corr} ratio achieves a utility gain of 3.07 % per year. In comparison, the strategy based on the dp^{ig} ratio delivers much smaller gains, of 1.05 %, relative to the strategy based on the recursive mean. This finding implies that accounting for the expected drp further elevates the utility gains by 2.02 %. Collectively, these results highlight the relevance of the time varying expected drp.

Our paper is most germane to the innovative work of Golez (2014), who uses ig to correct the standard dp ratio. In a similar vein, Bilson, Kang, and Luo (2015) and Zhong (2016) show that the dividend yield implied by derivatives prices predicts returns. A common feature of these studies is that they assume that dividend risk is not priced. Our main contribution is to provide a formal treatment of the time varying expected dividend risk premium. To do so, we develop a framework that allows us to study its implications for the predictability of dividend growth and stock market returns.

Our paper also relates to the literature on dividend forecasting. Lintner (1956), Marsh and Merton (1987) and Garrett and Priestley (2000) propose to use accounting data, e.g. earnings data, to predict dividend growth rates. We complement this body of works by showing how to obtain dividend growth forecasts from options data. Because option contracts are (i) forward-looking and (ii) observable at higher frequencies than accounting data, our framework could help researchers obtain more timely dividend growth forecasts, e.g. at the daily frequency. This method could also prove very useful when performing event studies.

Furthermore, our work contributes to a broader research agenda emphasizing that derivatives prices are informative about *risk-neutral* expectations whereas, for most practical purposes, one is interested in the physical expectations. The risk premium drives a wedge between these two expectations. Ross (2015) and Borovicka, Hansen, and Scheinkman

(2016), among others, discuss conditions under which it may or may not be possible to recover the physical probability distribution from derivatives prices. Several studies rely on historical data to pin down the dynamics of the risk premium. For instance, Piazzesi and Swanson (2008) focus on the Fed fund futures market and propose a parsimonious time series model for the expected risk premium. They then use their model to correct the forecasts implied by the Fed fund futures market. Chernov (2007) and Prokopczuk and Wese Simen (2014) show how to correct for the variance risk premium when using implied variance to predict realized variance. Our paper is similar in spirit to these works. We posit a time series model for the expected *drp* and analyze the implications of the expected *drp* for the predictability of dividend growth and stock market returns.

The remainder of this paper proceeds as follows. Section II. presents our theory and describes the dataset. Sections III. and IV. discuss our main empirical results on the predictability of dividend growth and stock returns, respectively. Finally, Section V. concludes.

II. Methodology and Data

This section begins by presenting our methodology. We formally show that *ig* contains information about (i) the expected dividend growth rate and (ii) the expected *drp*. We then propose a simple model to capture the dynamics of the expected *drp* and present an empirically testable model for future dividend growth rates and stock market returns. Finally, we introduce the dataset.

A. Methodology

The starting point of our methodology is the put-call parity (Stoll, 1969):

$$p_t(K) + P_t - e^{-rf_t}\mathbb{E}_t^Q(D_{t+1}) = c_t(K) + e^{-rf_t}K \quad (1)$$

where $p_t(K)$ is the price at time t of the put option contract of strike K that expires at the end of the next period, i.e. $t+1$. P_t is the price of the underlying asset at time t. rf_t denotes the 1-period riskless rate observed at t.[1] $\mathbb{E}_t^Q(D_{t+1})$ is the dividend that a risk-neutral (Q) investor expects, at time t, to receive from the underlying security at expiration, i.e. at $t+1$. $c_t(K)$ is the price at time t of the call option contract of strike K that expires at the end of the next period.

In order to clearly show the link between the option prices and the next-period dividend, we introduce the dividend strip. This financial asset entitles the holder to the dividends paid by the underlying security during the life of the strip (van Binsbergen, Brandt, and Koijen, 2012). We can obtain the market price of dividend strips by using two valuation methods: the martingale valuation approach and the standard present value method.

According to the martingale valuation framework of Cox and Ross (1976) and Harrison and Pliska (1981), we can price financial assets as if investors were risk-neutral. A direct implication of this result is that the market price of the dividend strip equals the cash flow

[1] Throughout this paper, we adopt the timing convention that interest rates are given the subscripts for the time when they are observed. As a result, our notation indicates that the interest rate is observed at t even though it is realized at time $t+1$.

that the risk-neutral investor expects to receive discounted to the present at the riskless rate:

$$STRIP_t = e^{-rf_t}\mathbb{E}_t^Q(D_{t+1}) \tag{2}$$

where $STRIP_t$ is the time t market price of the dividend strip expiring at the end of the next period.

Substituting equation (1) into the expression above yields:

$$STRIP_t = p_t(K) + P_t - c_t(K) - e^{-rf_t}K \tag{3}$$

The standard present value approach determines the market price of assets by directly discounting the expected cash flows (under the physical probability measure) at the expected rate of return. The following expression formalizes this idea:

$$STRIP_t = e^{-\mathbb{E}_t(drp_{t+1})}\mathbb{E}_t(D_{t+1}) \tag{4}$$

where $\mathbb{E}_t(drp_{t+1})$ denotes the conditional expectation of the buy-and-hold rate of return on the dividend strip. Throughout this paper, we refer to the return earned by a buy-and-hold investor who opens a long position in the dividend strip as the dividend risk premium (drp).[2,3] $\mathbb{E}_t(D_{t+1})$ is the dividend that the investor expects to receive from the underlying security at

[2]Strictly speaking, the discount rate is the sum of the riskless interest rate and the dividend risk premium. Because interest rates display very little variations in the time series, we commit a slight abuse of terminology and refer to the discount rate as the dividend risk premium. See Cochrane (2011) for a conceptually similar approach. Note also that, in this paper, we take the *drp* to mean the realized (rather than expected) return of the dividend strip. To denote the expected return of the dividend strip, we use the expression "expected *drp*".

[3]It is worth highlighting that, unlike the risk-free rate, the *drp* is only observed ex-post, i.e. at time $t+1$.

$t+1$.

Putting together equations (3) and (4), we derive the following result:

$$\log(\mathbb{E}_t(D_{t+1})) - \mathbb{E}_t(drp_{t+1}) \;=\; \log\left(p_t(K) + P_t - c_t(K) - e^{-rf_t}K\right) \tag{5}$$

Next, we subtract $\log(D_t)$ from both sides of equation (5) and ignore the Jensen inequality term:[4]

$$\mathbb{E}_t(\Delta d_{t+1}) - \mathbb{E}_t(drp_{t+1}) \;\approx\; \underbrace{\log\left(p_t(K) + P_t - c_t(K) - e^{-rf_t}K\right) - \log(D_t)}_{\text{implied growth}} \tag{6}$$

$$\mathbb{E}_t(\Delta d_{t+1}) - \mathbb{E}_t(drp_{t+1}) \;\approx\; ig_t \tag{7}$$

where $\mathbb{E}_t(\Delta d_{t+1})$ denotes the time t expectation of the 1-period dividend growth rate: $\mathbb{E}_t(\Delta d_{t+1}) = \mathbb{E}_t(\log(D_{t+1})) - \log(D_t)$. ig_t denotes the dividend growth rate implied by the options market at time t: $ig_t = \log\left(p_t(K) + P_t - c_t(K) - e^{-rf_t}K\right) - \log(D_t)$.

The expression above reveals that ig is the risk-adjusted expectation of future dividend growth. In particular, ig is positively related to the expected dividend growth and negatively related to the expected drp. An implication of this result is that a time varying expected drp could potentially obscure the information content of ig when predicting the expected dividend growth.

Despite its clear insights, the expression above is merely an accounting identity that is of limited practical use. The reason is that the terms on the left of the equality sign are

[4]It is standard in the literature to ignore the Jensen inequality term, e.g. Golez (2014). We conduct a simple simulation exercise which reveals that the approximation error is small. Most important for our objectives, it displays very little variations. A constant approximation error does not materially affect our results since we include an intercept in all regression models.

conditional expectations which are not directly observable. In order to obtain an empirically testable economic model, one needs to impose a structure on how the conditional expectation of the *drp* is generated.[5] We simply assume that the expected *drp* depends on a constant, the lagged *ig*, and the lagged *drp* (which is included in the information set at time t):

$$drp_{t+1} = \phi_0 + \phi_1 ig_t + \phi_2 drp_t + \epsilon_{t+1}^{drp} \tag{8}$$

where ϕ_0, ϕ_1, and ϕ_2 are the parameters to estimate.

Our assumption that *ig* predicts the *drp* is directly motivated by equation (7), which shows that *ig* is negatively related to the expected *drp*. As a result, we expect ϕ_1 to be negative. The assumption that the *drp* depends on its lagged observation is in keeping with previous works. Because the *drp* is essentially the return to a buy-and-hold trading strategy, our modelling approach is consistent with previous studies which typically assume that returns have an autoregressive component (van Binsbergen and Koijen, 2010; Lacerda and Santa-Clara, 2010; Golez, 2014). Armed with the model above and the identity presented in equation (7), we are now in a position to discuss our first proposition.

Proposition 1: *The lagged corrected implied growth rate (ig^{corr}), an affine function of (i) the lagged implied dividend growth (ig) and (ii) the lagged dividend risk premium (drp),*

[5] One may wonder why we do not use the expected dividend growth, derived from time series models for example, to recover the expected *drp* by manipulating the identity in equation (7). We do not pursue this approach because, if one already has an estimate of the expected dividend growth, then there is no need to use *ig* (and correct for the *drp*). Golez (2014) shows that, in-sample, *ig* outperforms the model of Lacerda and Santa-Clara (2010) which relies on historical dividend growth rates. Our aim is to further improve the forecasting ability of *ig* by explicitly accounting for the expected *drp*.

predicts the next-period dividend growth rate.

$$\Delta d_{t+1} = \underbrace{ig_t + \underbrace{\phi_0 + \phi_1 ig_t + \phi_2 drp_t}_{expected\ drp\ correction}}_{ig_t^{corr}} + \epsilon_{t+1}^{\Delta d} \tag{9}$$

Proof: See Appendix A.1.

This proposition presents our first empirically testable prediction: ig^{corr} predicts dividend growth with a coefficient loading that is not statistically distinguishable from 1.

Next, we build our forecasting model for the next-period return. Campbell and Shiller (1988) derive the following log-linear result:

$$\sum_{j=0}^{+\infty} \overline{\rho}^j \left(\mathbb{E}_t(r_{t+1+j}) - \mathbb{E}_t(\Delta d_{t+1+j})\right) = \frac{k}{1-\overline{\rho}} + dp_t \tag{10}$$

where r_{t+1+j} is the return at time $t+1+j$. k is a constant and $\overline{\rho}$ is the linearization constant computed as follows:

$$\overline{\rho} = \frac{1}{1+e^{\overline{d-p}}} \tag{11}$$

Equation (10) reveals that, to the extent that the expected dividend growth rate is time varying, the standard dividend price ratio is a noisy proxy for the expected return. Thus, it is important to correct the standard dp ratio for the fluctuations in expected dividend growth in order to improve the predictability of returns (Campbell, 2008). Because Proposition 1 shows that the expected dividend growth rate depends not only on ig but also on the expected drp, accounting for the expected drp should therefore strengthen the return predictability results.

We decompose the next-period return (r_{t+1}) into an expected return component (μ_t) and a forecast error (ϵ^r_{t+1}). As is standard in the literature, e.g. Golez (2014), we assume that expected returns and the implied growth rate follow AR(1) processes:

$$r_{t+1} = \mu_t + \epsilon^r_{t+1} \tag{12}$$

$$\mu_{t+1} = \alpha_0 + \alpha_1 \mu_t + \epsilon^\mu_{t+1} \tag{13}$$

$$ig_{t+1} = \delta_0 + \delta_1 ig_t + \epsilon^{ig}_{t+1} \tag{14}$$

where all error terms have zero mean. Armed with these additional assumptions, we can derive Proposition 2.

Proposition 2: *The lagged corrected dividend price (dp^{corr}) ratio, which is an affine function of (i) the lagged standard dividend price (dp) ratio, (ii) the lagged implied dividend growth (ig), and (iii) the lagged dividend risk premium (drp), forecasts the next-period return.*

$$r_{t+1} = \Psi + (1-\bar{\rho}\alpha_1) \underbrace{\left(\underbrace{dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}}_{dp^{ig}} + \underbrace{\frac{\phi_1 ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}}_{expected\ drp\ correction} \right)}_{dp^{corr}_t} + \epsilon^r_{t+1} \tag{15}$$

Proof: See Appendix A.2.

This proposition shows that the standard dividend price ratio alone cannot satisfactorily predict returns. Two adjustments are needed. First, one needs to account for ig to obtain the dp^{ig} ratio (Golez, 2014). Second, one also needs to account for the time variations in the expected drp. By making these two adjustments, we obtain the dp^{corr} ratio. If one assumes a constant expected drp, i.e. $\phi_1 = 0$ and $\phi_2 = 0$, then the dp^{corr} and dp^{ig} ratios are exactly the same. Thus, by comparing the forecasting performance of the dp^{corr} and dp^{ig} ratios, we can shed light on the relevance of the time variations in the expected drp for the predictability of stock market returns. If the expected drp correction plays an important role, then the dp^{corr} ratio should yield significantly better return predictability results than both the dp and dp^{ig} ratios.

B. Data

We obtain intraday quote prices on S&P 500 index option contracts and the underlying spot index from the Chicago Board of Options Exchange (CBOE). Our sample covers the period from January 01, 1996 to March 31, 2017. The S&P 500 index option contracts are of the European type and trade on the CBOE. These options have, among other, monthly expiration dates.

We focus on a sampling frequency of 1-minute and process the dataset as follows. First, we retain observations with non-zero bid and ask prices. Second, we only keep observations with a positive bid–ask spread. Third, we discard observations where the midquote price is lower than five times the minimum tick size of 0.05. Fourth, we expunge observations with no quoted size (either on the bid or ask side). Fifth, we only keep records observed between 10:00 AM and 2:00 PM local time similar to van Binsbergen et al. (2012). We match each option price with the spot index price observed on the same day and at the same time (up to the minute level). It is worth emphasizing that both the underlying price as well as the put and call prices are observed during these trading hours. Thus, our analysis does not suffer from asynchronous closing times induced by the wildcard feature of US derivatives markets.[6]

We proxy the riskless rate with the LIBOR curve, which we obtain from Bloomberg. We then merge together the time series of the interest rate, the spot, and option prices. For each trading day and option maturity, we create a 5-tuple (call option price, put option price, strike price, spot price, and interest rate of corresponding maturity). We plug the relevant values in equation (3) to obtain the dividend strip price. Next, we compute the median of all these dividend strip prices following van Binsbergen et al. (2012). By aggregating across all (i) strike prices and (ii) intraday intervals, we assuage potential concerns related to measurement errors in the dividend strip prices. We repeat the steps above for all maturities observed on each trading day, obtaining the term structure of dividend strips at the daily level. From the term structure, we follow Golez (2014) and linearly interpolate the 6-month dividend strip, which we multiply by a factor of two in order to obtain the annual dividend strip price ($STRIP^A$).

Two caveats are worth discussing. First, our methodology considers all strike prices in the spirit of van Binsbergen et al. (2012). Understandably, one may wonder to what extent are our results affected by deep in-the-money and deep out-of-the-money options. To shed light on this, we repeat our construction of the dividend strips by focusing only on options

[6] As discussed in Harvey and Whaley (1992), the S&P 500 spot market closes at 3:00 PM local time whereas trading in the derivatives market ends at 3:15 PM. This aspect introduces biases in studies that require synchronous observations of spot and derivatives prices. Although consistent with the work of Golez (2014), using asynchronous observations of options and spot data leads to a negative average dividend risk premium, a finding that is difficult to rationalize from an asset pricing perspective. In contrast, by matching the dataset at the intraday level as advocated by van Binsbergen et al. (2012), we obtain a positive dividend risk premium. We refer the interested reader to Boguth, Carlson, Fisher, and Simutin (2012) for a study of the impact of asynchronous observations on the properties of dividend strips.

that are at-the-money, i.e. with a Black and Scholes (1973) delta that is in absolute value between 0.375 and 0.625. Tables A.1 through A.6 of the online appendix present results that are consistent with our benchmark findings. Second, as discussed in Golez (2014), selecting the maturity of the dividend strip involves a tradeoff. On the one hand, options of 6-month maturity are typically more liquid than options of longer maturity, making them particularly useful for predictability purposes. On the other hand, the 12-month maturity could be useful to address concerns related to the seasonality of dividend payments (Fama and French, 1988; Ang and Bekaert, 2007). Because our research builds on that of Golez (2014), we closely follow the author's method of interpolation. To be more specific, we linearly interpolate the dividend strip price of 6-month maturity, which is then multiplied by 2 in order to annualize it. As a robustness check, we depart from the main methodology of Golez (2014) by considering options of longer (and less liquid) maturities when constructing the dividend strips. To this end, we linearly interpolate the 12-month dividend strip and repeat all our main tests. Overall, Tables A.7 to A.12 of the online appendix document that, although the results are slightly weaker than the benchmark findings, our key conclusions are not materially affected.

We obtain the time series of daily dividends and prices related to the S&P 500 index from Bloomberg. We sum all the intra-month dividends to obtain the monthly dividend payments (D^M). The time series of (annualized) monthly returns is computed as:

$$r_{t+1} = 12 \times \log\left(\frac{P_{t+1} + D^M_{t+1}}{P_t}\right) \tag{16}$$

where r_{t+1} is the 1-period annualized return. For the purpose of our empirical analysis, we take 1-period to mean 1-month. P_{t+1} and D_{t+1}^M denote the stock price and the dividend payment related to month $t+1$, respectively. P_t is the stock price observed at the end of month t.

As is standard in the literature, e.g. Ang and Bekaert (2007), we base our analysis on dividends summed over a trailing window of 12 months (D^A). Taking this step ensures we address the issue of seasonality in the dividend series. We then compute the (annualized) 1-month dividend growth rate as follows:

$$\Delta d_{t+1} \quad = \quad 12 \times \log\left(\frac{D_{t+1}^A}{D_t^A}\right) \tag{17}$$

where Δd_{t+1} denotes the monthly growth rate of dividends at $t+1$. D_{t+1}^A and D_t^A are the annual dividends for the periods ending at $t+1$ and t, respectively. Relatedly, we compute the standard dp ratio as:

$$dp_t = \log\left(\frac{D_t^A}{P_t}\right) \tag{18}$$

We then recover the time series of ig by taking the difference between the logarithm of the annual dividend strip and that of the annual dividend:

$$ig_t = \log(STRIP_t^A) - \log(D_t^A) \tag{19}$$

Next, we obtain the time series of the drp:

$$drp_{t+12} \quad = \quad \log(D_{t+12}^A) - \log(STRIP_t^A) \tag{20}$$

Finally, we use all sample information to estimate the parameters δ_1, ϕ_0, ϕ_1, and ϕ_2, necessary to test our two propositions (see equations (9) and (15)).[7] In order to obtain the persistence of ig, i.e. δ_1, we follow Golez (2014) and use successive non-overlapping annual samples. Golez (2014) proposes this approach in order to guard against biases induced by the (i) large overlap between consecutive observations of ig and (ii) potential measurement errors in the implied growth series. To be more specific, we calculate the persistence of ig at the monthly level as follows. We sample all monthly observations of ig recorded on Januaries and estimate the model in equation (14). We repeat these steps for all 12 calendar months and save the corresponding slope estimates and parameter variance-covariance matrices. We then average the slope estimates and parameter variance-covariance matrices across these different non-overlapping samples. Empirically, we find that the average estimate is equal to 0.28 (t-statistic=1.67).[8] Since this average corresponds to the AR(12) persistence estimate, we raise it to the power of 1/12 to recover the AR(1) parameter. In the data, we find $\delta_1 = 0.90$.

The estimation of ϕ_0, ϕ_1, and ϕ_2 is based on equation (8). As before, we use non-overlapping annual samples to estimate the relevant parameters. We average the parameter estimates (and parameter variance-covariance matrices) across all 12 possible samples of annual data. Unlike the estimation of δ_1, we do not convert the annual estimates to the monthly level. This is because, each month, we are interested in the drp expected at the end

[7] When we conduct our analysis out-of-sample, we recursively estimate all parameters to make sure that our results do not suffer from any look-ahead bias.

[8] Because the non-overlapping samples ignore the intermediate data, we are essentially throwing away information. One implication of this is that the derived standard errors are likely too large, making it hard to reject the null hypothesis (Cochrane and Piazzesi, 2005).

of the next year.[9] We find that $\phi_0 = 0.04$ (t-statistic=2.28), $\phi_1 = -0.55$ (t-statistic=-4.19), and $\phi_2 = 0.26$ (t-statistic=1.95). Combining these parameter estimates together with the monthly time series of ig and realized drp, we can recover ig^{corr} (see equation (9)). Next, we compute the linearization constant $\bar{\rho}$ using the whole sample period (see equation (11)). Similar to Golez (2014), we find $\bar{\rho} = 0.98$. Equipped with this information, we then construct the time series of the dp^{ig} and dp^{corr} ratios (see equation (15)).

Figure 1 plots the dividend strip and realized dividends series. For ease of exposition, we align the two time series (van Binsbergen et al., 2012). We can see that the dividend strip price and the realized dividends comove positively. It is also worth noticing that the dividend strip price appears to be more volatile than the realized dividends, indicating the presence of a time varying expected dividend risk premium. This is important because, for the purpose of predictability, the expected drp matters only if it varies over time. The summary statistics reported in Table 1 show that the dividend risk premium displays a volatility of 13 %.

Another interesting observation from Figure 1 is that the realized dividends are generally higher than the corresponding dividend strip prices. An implication of this finding is that, during our sample period, the dividend risk premium is positive as evidenced by its average value of 4 % per year (see Table 1). This observation is in sharp contrast with the puzzling statistics of Golez (2014) that point to a negative average drp of -3.45 % per year.[10] It is possible that the summary statistic of Golez (2014) is affected by measurement errors since the author matches end-of-day S&P 500 derivatives data to the S&P 500 index.

[9] Remember that ig is informative about the growth rate expected over the next year adjusted for the dividend risk premium. Therefore, we need the dividend risk premium expected over the same period.

[10] Tables 1 (Panel B) and 3 of Golez (2014) reveal that the realized and implied growth rates average around 3.86 % and 7.31 %, respectively. Thus, the author's own figures indicate a negative and economically large annualized dividend risk premium of -3.45 %.

This methodology is vulnerable to concerns related to the wildcard feature of US derivative markets, induced by the fact that the derivatives market closes at 3:15 PM local time while the spot market closes at 3:00 PM. To verify this possibility, we obtain OptionMetrics end-of-day options data for our sample period and repeat the analysis. Our untabulated results point to a negative dividend risk premium of around $-5\,\%$ on average. This suggests that asynchronous trading times induce substantial biases in the estimate of the dividend risk premium, a point made by Boguth et al. (2012) among others. Because we (i) match the intraday derivatives and spot data and (ii) aggregate across these matches, our methodology is more robust to concerns about measurement errors (see also van Binsbergen et al. (2012)).

III. Dividend Growth Predictability

The discussion in Section II.A. shows that, if we have a good model for the expected *drp*, we should be able to improve our dividend growth forecasts. Thus, a natural starting point would be to assess the empirical performance of the forecasting model for the *drp* (see equation (8)). If the model does a good job, the expected *drp* should be positively and highly correlated with the subsequently realized *drp*.

Figure 2 displays the dynamics of the realized and expected *drp*. The expected *drp* is the forecast generated by the following equation: $\mathbb{E}_t(drp_{t+12}) = 0.04 - 0.55 ig_t + 0.26 drp_t$. We observe that the two series comove strongly. Indeed, a regression of the realized *drp* on a constant and the expected *drp* yields an insignificant intercept, a slope of 1.08 (t-statistic=7.41) and an R^2 of 63.52 %. We thus conclude that the 2-factor model does a satisfactory job and proceed to analyze its implications for the predictability of dividend growth rates.

A. In-Sample Evidence

We start with the in-sample analysis. This investigation is motivated by Proposition 1, which posits that the lagged ig^{corr}, a linear combination of the lagged ig and the lagged drp, predicts the dividend growth rate. We test this prediction by regressing the time series of 1-month dividend growth rates on a constant and the lagged predictive variable X_t:

$$\Delta d_{t+1} = \gamma_0 + \gamma_1 X_t + \epsilon_{t+1}^{\Delta d} \tag{21}$$

We separately consider the scenarios where $X = ig$ and $X = ig^{corr}$.

Figure 3 displays the dynamics of both forecasting variables. We can see that ig is generally positive and takes negative values ahead of periods of economic crises. It is worth noting that ig^{corr} is on average higher than ig, reflecting the effect of a positive average expected drp. The two series exhibit very similar time series patterns. This could indicate that, of the two factors posited for the expected drp, ig is the main driving force. These results are consistent with the slope estimates ϕ_1 and ϕ_2 as well as the summary statistics in Table 1.

Table 2 summarizes the regression results. The figures in brackets correspond to the

Newey and West (1987) corrected test statistics.[11,12] We test H_0: $\gamma_1 = 0$ against the 1-sided alternative hypothesis H_1: $\gamma_1 > 0$. Throughout this paper, we use a significance level of 5 %. Examining the t-statistics, we can see that the null hypothesis is always rejected, suggesting that each of the two variables predicts the dividend growth rate.

The regression results reveal that ig predicts the dividend growth rate with a slope of 0.42. This slope estimate is important for several reasons. First, it is higher than the slope of 0.19 reported in Golez (2014).[13] The higher estimate presented in this paper is likely due to the fact that the study of Golez (2014) suffers from an attenuation bias induced by noisy estimates of the dividend strip, which in turn affects the accurate measurement of the implied growth rate.

Second, the model of Golez (2014) assumes a constant expected drp, which implies that ig should predict the future dividend growth rate with a slope of 1. Clearly, one can formally reject the null hypothesis that the slope parameter (0.42) is equal to 1. Furthermore, this slope estimate reveals the share of variations in ig that is attributable to the dividend

[11] We follow earlier studies, e.g. Rangvid (2006) and Ang and Bekaert (2007), and set the lag length equal to $h + 1$, where h denotes the forecasting horizon in months. Our results are robust to the choice of the lag length. The simulation results of Ang and Bekaert (2007) show that the Newey and West (1987) standard errors are well-behaved at short horizons while the Hodrick (1992) standard errors perform better than the Newey and West (1987) standard errors at long horizons. Because our study deals with the predictability over the next period, we focus on the Newey and West (1987) standard errors.

[12] To make the statistical inference more robust, we compute the empirical p-values from the wild bootstrap simulation described in the online appendix of Rapach, Strauss, and Zhou (2013). This procedure has a number of desirable features. First, it uses the iterative methodology presented in Amihud, Hurvich, and Wang (2008) to correct for the Stambaugh (1999) bias in a multivariate setting. This correction is important because the Stambaugh (1999) bias is known to generate size distortions. Second, it preserves the contemporaneous correlations across residuals. Third, it allows for general forms of conditional heteroskedasticity. We thank an anonymous reviewer for this very helpful suggestion.

[13] In comparing our results to those of Golez (2014), it is worth keeping in mind that the author regresses the monthly dividend growth rate on implied growth, which is an annualized quantity. Thus, the adapted estimate of the 0.0157 loading on ig at the 1-month horizon shown in Table 4 of Golez (2014) corresponds to $0.0157 \times 12 \approx 0.19$ in our set-up.

growth rate. Exploiting equation (7), we can show that:

$$
\begin{aligned}
var(ig_t) &= cov(\mathbb{E}_t(\Delta d_{t+1}) - \mathbb{E}_t(drp_{t+1}), ig_t) \\
1 &= \frac{cov(\mathbb{E}_t(\Delta d_{t+1}), ig_t)}{var(ig_t)} - \frac{cov(\mathbb{E}_t(drp_{t+1}), ig_t)}{var(ig_t)}
\end{aligned} \tag{22}
$$

The expression above shows that we can decompose the variation in ig into components related to (i) the expected dividend growth and (ii) the expected drp. The first term to the right of the equality sign is essentially the slope coefficient of a regression of the dividend growth rate on a constant and the lagged implied growth rate.[14] The second term to the right of the equality sign is the slope estimate of a regression of the drp on a constant and the lagged ig. Table 2 reveals that only 42 % of variations in ig can be linked to the expected dividend growth rate. One implication of this finding is that the expected drp accounts for the remaining 58 % of variations in ig. In other words, the main driving force of ig is the expected drp, rather than the expected dividend growth rate.

If Proposition 1 holds, then we would expect to find that ig^{corr} predicts the next-period dividend growth with a slope of 1. Table 2 reports that ig^{corr} enters the regression model with a positive and statistically significant slope of 0.92. Using the t-statistic, we can formally test the null hypothesis that this slope equals 1 as predicted by the theory. Our untabulated analysis reveals that the slope estimate is not significantly different from 1, thus supporting the model's prediction.

[14] As Proposition 1 shows, we can express the dividend growth rate as the sum of the expected dividend growth rate and an independent shock. Assuming that the shock is independent of ig, the slope estimate is the same regardless of whether the dependent variable in the regression model is the realized dividend growth or the expected dividend growth.

B. Out-of-Sample Evidence

We now explore the predictability of dividend growth in an out-of-sample setting. We use the last 7 years of the sample period for our out-of-sample test. Accordingly, we use all data up to February 2009 to initially estimate ϕ_0, ϕ_1, and ϕ_2 (see equation (8)). We expand the training window by one month each time, thus recursively estimating the parameters.[15] An upshot of this approach is that there are no look-ahead biases. We consider two distinct forecasting models. Model 1 builds on the work of Golez (2014) to arrive at the forecast ($\hat{y}_t$): $\hat{y}_t = ig_t$. Model 2 uses the insights of Proposition 1 to derive the forecast: $\hat{y}_t = \phi_0 + (1+\phi_1)ig_t + \phi_2 drp_t$. A neat feature of this out-of-sample analysis is that it directly imposes the discipline of the theory and avoids the estimation errors typically associated with dividend growth forecasting regressions.

We compute the out-of-sample R^2 (R^2_{oos}) of each dividend growth forecasting model:

$$R^2_{oos} = 1 - \frac{\sum_{t=1}^{N}(y_{t+1} - \hat{y}_t)^2}{\sum_{t=1}^{N}(y_{t+1} - \bar{y}_t)^2} \tag{23}$$

where y_{t+1} is the realization of the variable of interest at $t+1$. $\bar{y}_t$ is the recursive mean of the variable of interest computed using all observations up to time t. N is the total number of forecasts.

Intuitively, the R^2_{oos} sheds light on the proportional reduction in the mean squared error (MSE) of the forecasting model underpinning $\hat{y}_t$ relative to that of the benchmark

[15] As a robustness check, we also consider a rolling window and reach similar conclusions. These results are not tabulated for brevity.

recursive mean. Next, we compute the $MSE - F$ statistic of McCracken (2007):

$$MSE\text{–}F \quad = \quad N \times \frac{(y_{t+1} - \bar{y}_t)^2 - (y_{t+1} - \hat{y}_t)^2}{(y_{t+1} - \hat{y}_t)^2} \tag{24}$$

McCracken (2007) provides the critical values for this test statistic.

We also compute the $MSE{-}Adj$ statistic of Clark and West (2007). More specifically, we compute the time series of the variable f:

$$f_t = (y_{t+1} - \bar{y}_t)^2 - (y_{t+1} - \hat{y}_t)^2 - (\bar{y}_t - \hat{y}_t)^2 \tag{25}$$

We then regress this time series on a constant and compute the corresponding t-statistic using the Newey and West (1987) standard errors. Clark and West (2007) show that this test statistic has an approximately standard normal distribution.

The two tests enable us to formally examine the null that the mean squared error (MSE) of the benchmark model, i.e. the recursive mean, is smaller than or equal to that of the competing model generating the forecast $\hat{y}_t$. The alternative hypothesis is that the MSE associated with the competing model is lower than that of the recursive mean.

Table 3 reveals that ig^{corr} yields a positive R^2_{oos} whereas ig does not. Moreover, we find that $MSE - F = 26.22$ and $MSE - Adj = 2.20$ for this positive R^2_{oos}. Clearly, the large and positive magnitude of these test statistics indicate that we can reject the null hypothesis. Thus, the improvements in forecast accuracy achieved by ig^{corr} are significant.

IV. Stock Return Predictability

Having established the importance of the expected dividend risk premium correction for the predictability of dividend growth, we now explore the implications for return predictability. We start by examining the predictability of stock market returns in-sample and then we turn our attention to the out-of-sample evidence.

A. In-Sample Evidence

We regress the time series of monthly returns on a constant and the lagged forecasting variable X_t:

$$r_{t+1} = \gamma_0 + \gamma_1 X_t + \epsilon^r_{t+1} \tag{26}$$

We examine the following forecasting variables in turn: dp, dp^{ig}, and dp^{corr}. Comparing the results for the first two forecasting variables sheds light on the importance of accounting for ig. Similarly, by contrasting the results for the last two forecasting variables, we can learn about the relevance of the expected dividend risk premium correction.

Figure 4 shows the dynamics of all 3 forecasting variables. We notice that both dp^{ig} and dp^{corr} are more volatile than the standard dp ratio (see also Table 1). Moreover, they behave in a manner that is reminiscent of ig. This is mainly due to the high magnitude of δ_1 (0.90), which gives more prominence to ig (see Proposition 2). The dp ratio shares a correlation of 0.27 and 0.60 with the dp^{ig} and dp^{corr} ratios, respectively.

Table 4 reports that the slope associated with the dp, dp^{ig}, and dp^{corr} ratios are equal to 0.20, 0.05, and 0.16, respectively. To better understand the slope associated with the dp^{corr} ratio, we implement the following decomposition:

$$(27)\quad \frac{cov(r_{t+1}, dp_t^{corr})}{var(dp_t^{corr})} = \frac{cov(r_{t+1}, dp_t)}{var(dp_t^{corr})} + \frac{cov(r_{t+1}, dp_t^{ig} - dp_t)}{var(dp_t^{corr})} + \frac{cov(r_{t+1}, dp_t^{corr} - dp_t^{ig})}{var(dp_t^{corr})}$$

Empirically, the first, second, and third components to the right of the equality sign amount to 0.05, 0.43, and -0.33, respectively. We thus conclude that adjusting for both ig and the expected drp is necessary when using the dp ratio to forecast stock market returns. Economically, the slope parameter γ_1 in equation (26) is informative about the persistence of expected returns, i.e. α_1 (see equation (13)). As Proposition 2 shows, the slope γ_1 is equal to $1 - \bar{\rho}\alpha_1$. Since $\bar{\rho} = 0.98$, the loadings on dp^{ig} and dp^{corr} imply that the persistence of expected returns is close to 0.97 and 0.86, respectively.[16]

We test H_0: $\gamma_1 = 0$ against the 1-sided alternative hypothesis, i.e. $\gamma_1 > 0$. Sticking to the 5 % significance level, we reject the null hypothesis for dp^{corr} and dp^{ig}. This result holds irrespective of whether we derive the critical values from the asymptotic distribution or the wild bootstrap of Rapach et al. (2013). The finding that dp^{ig} predicts next-month's returns is consistent with the work of Golez (2014). Upon close examination, we notice that the dp^{corr} ratio is the more significant of the two variables. Furthermore, it displays the highest explanatory power (1.45 %) of all three forecasting variables. This suggests that accounting for the expected dividend risk premium helps improve the predictability of stock market returns.

We investigate whether the dp^{corr} ratio contains information which is not included

[16] To get α_1, we look at the slope coefficient of the return forecasting regression. Since theory predicts that the slope equals $1 - \bar{\rho}\alpha_1$, we rearrange the expression to recover α_1.

in variables that have been shown to predict the stock market return at short forecasting horizons. We download the time-series of the book-to-market (bm), the default spread (def), the earnings-to-price ratio (ep), the inflation rate ($infl$), the net equity expansion ($ntis$), the payout ratio (pay), and the level of the Treasury bill ($tbill$) rate from the website of Amit Goyal. We also include the following variables that we describe in Section B of the online appendix: the modified dp ratio (dp^{lac}) of Lacerda and Santa-Clara (2010), the change in the federal fund rate (Δff), the relative interest rate ($rrel$), the implied skewness ($skew$), the sum-of-the-parts forecast (sop), the stock market variance ($svar$), the term spread ($term$), and the variance risk premium (vrp). We sample all the control variables at the end of the month, thus aligning them with the time series of the dp^{corr} ratio. The bivariate regression results of Table 4 show that the dp^{corr} ratio is a robust predictor of stock market returns.

B. Out-of-sample Evidence

We now conduct our analysis out-of-sample. We estimate ϕ_1, ϕ_2, δ_1, and $\bar{\rho}$ recursively. Similar to the out-of-sample setting used for the predictability of the dividend growth rate (see Section III.B.), we use all observations up to February 2009 as our initial training sample period, leaving the last 7 years of data for the out-of-sample test. We exploit all the information in our training sample to estimate the return forecasting regression shown in equation (26). Equipped with the intercept and slope estimates, we use the last observation of the forecasting variable (in the training sample) to predict the next-month's return. If the predicted return is negative, we set the forecast equal to 0 as in Campbell and Thompson (2008). By taking this step, we impose the economic restriction that expected returns are non-negative.[17] We repeat these steps for each month and for each forecasting variable.

Panel A of Table 5 compares the performance of different models to that of the model based on the recursive mean once the economic restriction is implemented. We find that the dp^{corr} ratio yields the highest R^2_{oos} ($R^2_{oos} = 2.18\,\%$). The associated $MSE - F$ (1.92) and $MSE - Adj$ (1.97) statistics suggest that this improvement in forecast accuracy is statistically significant. A similar finding emerges from Panel B of the same table, where we do not impose any economic restrictions. Overall, this finding is consistent with our model's prediction: correcting for the expected drp matters.

C. The Economic Value of Return Predictability

Finally, we explore the implications of the return predictability for the portfolio choice of an investor willing to use the dp^{corr} ratio as a timing signal when implementing a quantitative strategy. In particular, the market timing strategy allocates a fraction of wealth w_t to the risky stock and the remainder to the riskless asset. The risky asset has expected return μ_t and expected volatility $\hat{\sigma}_t$. The riskless asset yields a return rf_t. We assume that the investor has a quadratic utility function with a coefficient of relative risk aversion of γ, thus giving rise to the following optimization problem:[18]

$$\max_{w_t} w_t\mu_t + (1 - w_t)rf_t - \frac{\gamma}{2} w_t^2 \hat{\sigma}_t^2 \tag{28}$$

[17] We thank a referee for suggesting this analysis.

[18] The optimization problem of an investor with quadratic utility is equivalent to maximizing a linear combination of mean and variance. This is true irrespective of the distribution of asset returns. We refer the interested reader to Campbell and Viceira (2002) for an excellent treatment of this topic.

The optimal allocation to the risky asset is given by:

$$w_t = \frac{\mu_t - rf_t}{\gamma \hat{\sigma}_t^2} \tag{29}$$

For each return forecasting model, we compute the expected return on the risky asset and use equation (29) to compute the weights. If the expected return is negative, we set it equal to 0 before computing the weights. By doing so, we align the out-of-sample statistical analysis (R^2_{oos}) with our economic value exercise. We use the 1-month LIBOR rate as our proxy for the riskless rate.[19] We use all monthly returns data available in the recursive window to estimate the variance of the stock returns.[20] Finally, we consider different values for the coefficient of risk aversion, e.g. 2, 4, 6, 8 and 10. Equipped with the portfolio weights and the time series of realized stock returns, we compute the time series of realized portfolio returns.

We then analyze the certainty equivalent rate of return (CE), which is the risk-free rate of return that the investor is willing to accept rather than following a risky market timing strategy:

$$CE = \bar{r}_p - \frac{\gamma}{2}\sigma_p^2 \tag{30}$$

[19]As previously discussed, it is standard in the derivatives pricing community to proxy the interest rate with the LIBOR rate. Consistent with this practice, and thus the earlier part of our study, we use the 1-month LIBOR rate as the risk-free rate proxy. Because the return predictability literature also analyzes the 3-month Treasury bill rate, e.g. Goyal and Welch (2003), one may wonder what impact, if any, does the proxy for the riskless rate have on our portfolio results. To investigate this, we obtain the time series of 3-month Treasury bill rates from the website of the Federal Reserve of St. Louis and repeat our analysis. Untabulated results show that the riskless rate proxy has very little bearing on our results.

[20]In our analysis, we rely on the standard variance estimator. As a robustness check, we also analyze an exponentially weighted moving average (EWMA) model and obtain very similar results.

where $\bar{r}_p$ is the average of the realized portfolio returns. σ_p is the realized volatility of the portfolio returns. All other variables are as previously defined.

Table 6 enables us to answer the following question: How much would an investor pay in order to switch from a quantitative strategy that is based on the recursive mean to a timing strategy that relies on the dp^{corr} ratio? Our results indicate that an investor with a risk aversion coefficient equal to 4 would pay up to 3.07 % per year. This fee speaks directly to the importance of accounting for (i) the implied growth rate and (ii) the expected dividend risk premium. In order to understand the contribution of each component to this result, we also examine the timing strategy based on the dp^{ig} ratio. Computing the difference between the certainty equivalent rate of return of the timing strategy based on the dp^{ig} ratio and that of the strategy based on the recursive mean, we find that, for the same investor, the dp^{ig} ratio leads to a smaller utility gain of 1.05 % per year. This result reveals that accounting for the expected dividend risk premium further elevates the utility gain by 2.02 percentage points from 1.05 % to 3.07 %. We reach qualitatively similar conclusions for other values of the risk aversion coefficient. As an additional check, we also compute the difference between the annualized Sharpe ratio (SR) associated with a given timing strategy and that of the recursive mean. This analysis is interesting because the Sharpe ratio is independent of the risk aversion parameter.[21] The penultimate row of Table 6 shows that the dp^{corr} ratio leads to the highest improvement in SR, confirming that the expected dividend risk premium correction is economically valuable.

[21] We thank an anonymous reviewer for suggesting this analysis.

D. Additional Analyses

1. Annual Data

The theoretical derivation in equation (15) links the return observed over the next period, i.e. $t+1$, with the dp^{corr} ratio constructed using all information from the current period t. The summation of dividends over the past twelve months used to compute the dp ratio (and related quantities) is not well-aligned with the model since all variables should be measured over the same interval. One concern that arises from this is that the magnitude of the predictive slope associated with the dp^{corr} ratio may be the result of mixing annualized dividends with monthly stock prices and returns.[22] To explore this possibility, we repeat our return forecasting analysis using 12 successive non-overlapping annual samples. To be more precise, each non-overlapping annual sample consists of observations recorded during a particular month only, e.g. January. We use each of the 12 samples to estimate the following return forecasting regression:

$$r_{t+12} = \gamma_0 + \gamma_1 X_t + \epsilon^r_{t+12} \tag{31}$$

where r_{t+12} is the annual return realized at $t+12$. For each of the 12 non-overlapping annual samples, we save the parameter estimates and the associated variance-covariance matrices. Next, we average the parameter estimates as well as the variance-covariance matrices across all 12 annual samples in order to obtain the slope parameter estimate and standard error associated with each forecasting variable. The average slope parameters associated with the dp, dp^{ig}, and dp^{corr} ratios are equal to 0.27 (t-statistic=1.76), 0.21 (t-statistic=2.48),

[22]We thank an anonymous reviewer for this valuable comment.

and 0.28 (t-statistic=2.20), respectively. The associated R^2 is equal to 6.64 %, 8.00 %, and 8.22 % for the regression that includes the dp, dp^{ig}, and dp^{corr} ratios, respectively. Given the limited size of each annual sample, we caution that the explanatory power should be interpreted carefully.

We test the null hypothesis that the slope estimate associated with the dp^{corr} ratio obtained using non-overlapping annual data (0.28) is equal to that obtained using monthly data as in our benchmark analysis (0.16). The untabulated analysis reveals that we cannot reject this null hypothesis for the dp^{corr} ratio (t-statistic=0.93). Turning to the slope associated with the dp^{ig} ratio, we observe a marked difference from 0.05 in our benchmark specification to 0.21, when using annual non-overlapping samples. Our untabulated test suggests that we can reject the null hypothesis that the two estimates are equal at the 10 % significance level (t-statistic=1.77). The contrast between the two estimates is reminiscent of the work of Golez (2014) who documents a slope of 0.08 when using monthly returns data and an estimate of around 0.27 when using non-overlapping annual samples.[23]

2. Variance Decomposition

For a better interpretation of our results, we investigate the drivers of the dp ratio variation. Equation (A.11) of the online appendix enables us to decompose the variations in

[23] Table 5 of Golez (2014) reports a slope of 0.0073 for the dp^{ig} at the monthly forecasting horizon. Thus, the annualized slope is equal to $0.0073 \times 12 \approx 0.08$. Column 3 of Table 8 reports that the expected returns have a persistent parameter equal to 0.7487 on average when using non-overlapping annual samples. Since the slope parameter is given as 1 minus the product of the persistence and $\bar{\rho}$, we have $1 - 0.7487 \times 0.98 \approx 0.27$.

the dp ratio into components linked to discount rates (dr) and cash flows (cf):

$$dp_t = k_1 + \underbrace{\frac{\mu_t}{1-\bar{\rho}\alpha_1}}_{dr_t} - \underbrace{\left[\frac{1+\phi_1}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)}\right] ig_t - \frac{\phi_2}{1-\bar{\rho}\phi_2} drp_t}_{cf_t} \tag{32}$$

$$dp_t = k_1 + dr_t - cf_t \tag{33}$$

Using the parameters ϕ_1, $\bar{\rho}$, δ_1, ϕ_2 and the time-series of ig and drp, we can compute the cash flow channel.

Straightforward calculations imply the following variance decomposition of the dp ratio:

$$1 = \frac{cov(dr_t, dp_t)}{var(dp_t)} - \frac{cov(cf_t, dp_t)}{var(dp_t)} \tag{34}$$

Economically, the contribution of the cash flow channel to the variations of the dp ratio is linked to the slope estimate of the regression of the cash flow channel on a constant and the dp ratio. In the data, we find that the covariation of the cash flow channel with the dp ratio accounts for 20.52 % of the variations in the dp ratio. As a result, the covariation of the discount rate channel with the dp ratio accounts for 120.52 % of the variations in the dp ratio. These results are qualitatively similar to those of Golez (2014) who report statistics of 34.22 % and 134.22 % for the covariance of the dp ratio with the cash flow and discount rate channels relative to the variations in the dp ratio, respectively. We thus conclude that most of the variations in prices arise from fluctuations in discount rates.

V. Conclusion

We show that the dividend growth rate implied by the options market contains information about (i) the expected dividend growth rate and (ii) the expected dividend risk premium. We propose a simple model for the expected *drp* and study its implications for the predictability of dividend growth and stock market returns.

Our empirical analysis establishes that accounting for the expected *drp* strengthens the predictability of dividend growth and stock market returns. Our main results hold both in- and out-of-sample. Analyzing the implication of our results for the portfolio choice of an investor, we find that a market timing investor who accounts for the time varying expected dividend risk premium realizes an additional utility gain of 2.02 % per year. Overall, our study highlights, both theoretically and empirically, the importance of the expected dividend risk premium for the predictability of dividend growth and stock market returns.

References

Amihud, Y., C. M. Hurvich; and Y. Wang. "Multiple-predictor regressions: Hypothesis testing." *Review of Financial Studies*, 22 (2008), 413–434.

Ang, A.; and G. Bekaert. "Stock return predictability: Is it there?" *Review of Financial Studies*, 20 (2007), 651–707.

Bilson, J., S. B. Kang; and H. Luo. "The term structure of implied dividend yields and expected returns." *Economics Letters*, 128 (2015), 9–13.

Black, F.; and M. Scholes. "The pricing of options and corporate liabilities." *Journal of Political Economy*, 81 (1973), 637–654.

Boguth, O., M. Carlson, A. J. Fisher; and M. Simutin. "Leverage and the limits of arbitrage pricing: Implications for dividend strips and the term structure of equity risk premia." *Working Paper*, (2012).

Bollerslev, T., G. Tauchen; and H. Zhou. "Expected stock returns and variance risk premia." *Review of Financial Studies*, 22 (2009), 4463–4492.

Borovicka, J., L. P. Hansen; and J. A. Scheinkman. "Misspecified Recovery." *Journal of Finance*, 71 (2016), 2493–2544.

Campbell, J. Y. "Viewpoint: Estimating the equity premium." *Canadian Journal of Economics*, 41 (2008), 1–21.

Campbell, J. Y.; and R. J. Shiller. "The dividend-price ratio and expectations of future dividends and discount factors." *Review of Financial Studies*, 1 (1988), 195–228.

Campbell, J. Y.; and S. B. Thompson. "Predicting excess stock returns out of sample: Can anything beat the historical average?" *Review of Financial Studies*, 21 (2008), 1509–1531.

Campbell, J. Y.; and L. M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors.* Oxford University Press (2002).

Chernov, M. "On the role of risk premia in volatility forecasting." *Journal of Business and Economic Statistics*, 25 (2007), 411–426.

Clark, T. E.; and K. D. West. "Approximately normal tests for equal predictive accuracy in nested models." *Journal of Econometrics*, 138 (2007), 291–311.

Cochrane, J. H. "Presidential address: Discount rates." *Journal of Finance*, 66 (2011), 1047–1108.

Cochrane, J. H.; and M. Piazzesi. "Bond Risk Premia." *American Economic Review*, 95 (2005), 138–160.

Cox, J. C.; and S. A. Ross. "The valuation of options for alternative stochastic processes." *Journal of Financial Economics*, 3 (1976), 145–166.

Drechsler, I.; and A. Yaron. "What's vol got to do with it." *Review of Financial Studies*, 24 (2011), 1–45.

Fama, E. F.; and K. R. French. "Dividend yields and expected stock returns." *Journal of Financial Economics*, 22 (1988), 3–25.

Ferreira, M. A.; and P. Santa-Clara. "Forecasting stock market returns: The sum of the parts is more than the whole." *Journal of Financial Economics*, 100 (2011), 514–537.

Garrett, I.; and R. Priestley. “Dividend behaviour and dividend signaling.” *Journal of Financial and Quantitative Analysis*, 35 (2000), 173–189.

Golez, B. “Expected Returns and Dividend Growth Rates Implied by Derivative Markets.” *Review of Financial Studies*, 27 (2014), 790–822.

Goyal, A.; and I. Welch. “Predicting the equity premium with dividend ratios.” *Management Science*, 49 (2003), 639–654.

Harrison, J. M.; and S. R. Pliska. “Martingales and stochastic integrals in the theory of continuous trading.” *Stochastic Processes and their Applications*, 11 (1981), 215–260.

Harvey, C. R.; and R. E. Whaley. “Market volatility prediction and the efficiency of the S&P 100 index option market.” *Journal of Financial Economics*, 31 (1992), 43–73.

Hodrick, R. J. “Dividend yields and expected stock returns: Alternative procedures for inference and measurement.” *Review of Financial Studies*, 5 (1992), 357–386.

Lacerda, F.; and P. Santa-Clara. “Forecasting dividend growth to better predict returns.” *Chicago Booth Working Paper*, (2010).

Lintner, J. “Distribution of incomes of corporations among dividends, retained earnings, and taxes.” *American Economic Review*, 46 (1956), 97–113.

Maio, P. “The “Fed Model” and the predictability of stock returns.” *Review of Finance*, 17 (2013), 1489–1533.

Maio, P. “Don’t fight the Fed!” *Review of Finance*, 18 (2014), 623–679.

Marsh, T. A.; and R. C. Merton. "Dividend behavior for the aggregate stock market." *Journal of Business*, 60 (1987), 1–40.

McCracken, M. W. "Asymptotics for out of sample tests of Granger causality." *Journal of Econometrics*, 140 (2007), 719–752.

Newey, W. K.; and K. D. West. "A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica*, 55 (1987), 703–708.

Piazzesi, M.; and E. T. Swanson. "Futures prices as risk-adjusted forecasts of monetary policy." *Journal of Monetary Economics*, 55 (2008), 677–691.

Prokopczuk, M.; and C. Wese Simen. "The importance of the volatility risk premium for volatility forecasting." *Journal of Banking & Finance*, 40 (2014), 303–320.

Rangvid, J. "Output and expected returns." *Journal of Financial Economics*, 81 (2006), 595–624.

Rapach, D. E., J. K. Strauss; and G. Zhou. "International stock return predictability: what is the role of the United States?" *Journal of Finance*, 68 (2013), 1633–1662.

Ross, S. "The Recovery Theorem." *Journal of Finance*, 70 (2015), 615–648.

Stambaugh, R. F. "Predictive regressions." *Journal of Financial Economics*, 54 (1999), 375–421.

Stoll, H. R. "The relationship between Put and Call option prices." *Journal of Finance*, 24 (1969), 801–824.

van Binsbergen, J., M. Brandt; and R. Koijen. “On the Timing and Pricing of Dividends.” *American Economic Review*, 102 (2012), 1596–1618.

van Binsbergen, J.; and R. S. Koijen. “Predictive regressions: A present-value approach.” *Journal of Finance*, 65 (2010), 1439–1471.

Zhong, J. “Predictive Regressions Based on Ex Ante Index Futures Market Information.” *Working Paper*, (2016).

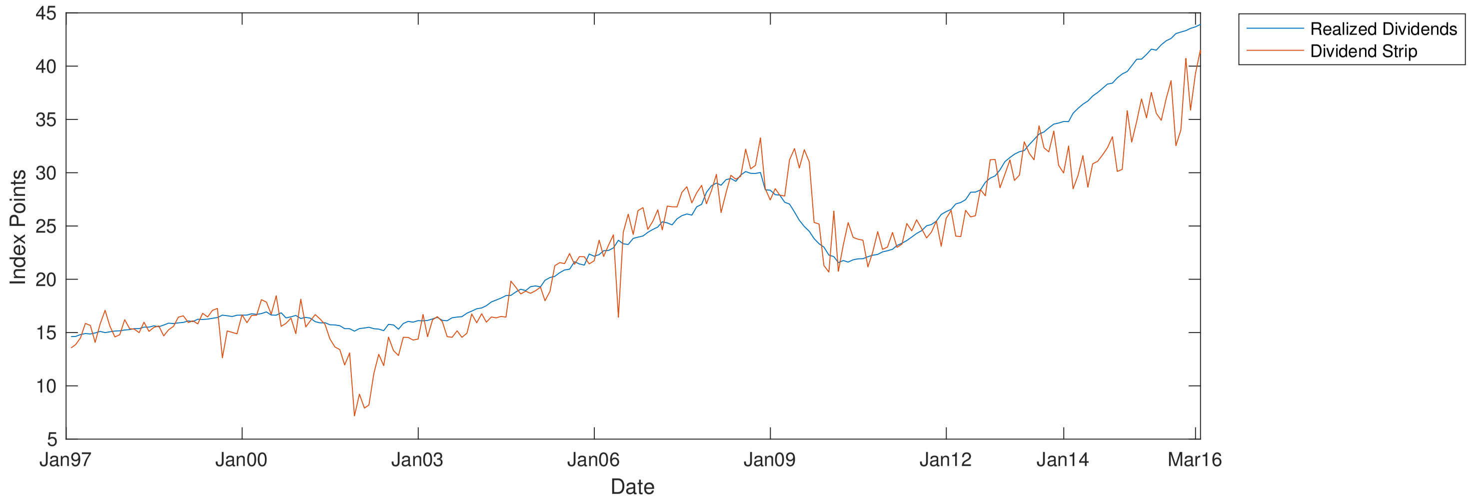

Figure 1: Time Series of Dividend Strips and Realized Dividends

This figure shows the dynamics of the implied dividend strip as well as the realized dividends during our sample period. For ease of exposition, we align the two time series. The horizontal axis displays the observation date. The vertical axis shows the value in index points.

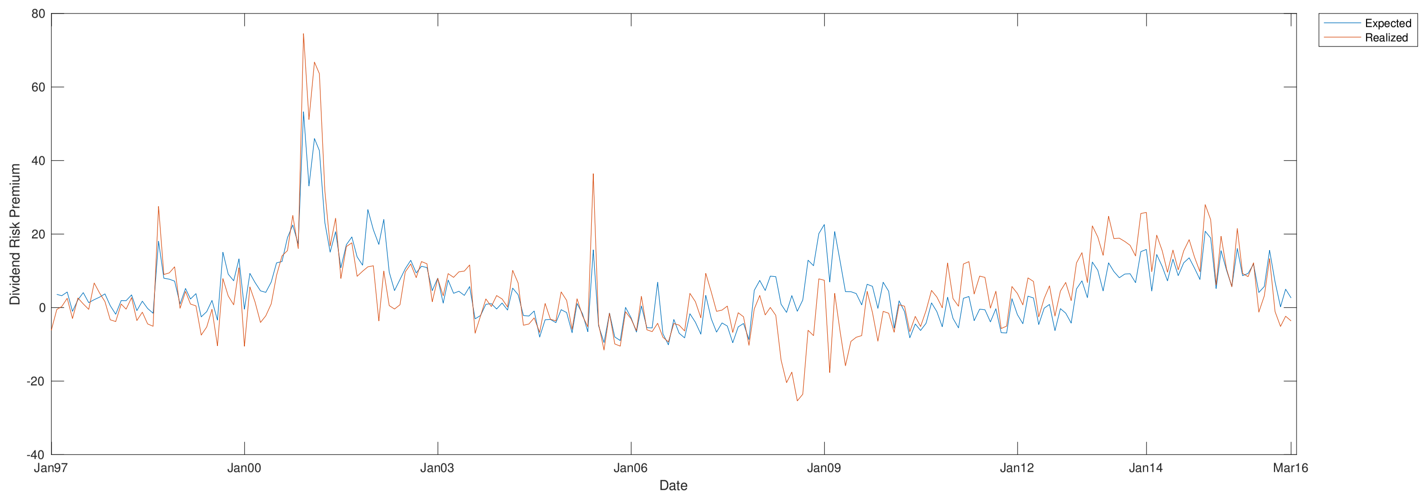

Figure 2: Realized vs. Expected drp

This figure shows the dynamics of the realized and expected drp during our sample period. The expected drp is the forecast generated by the following equation: $\mathbb{E}_t(drp_{t+12}) = 0.04 - 0.55ig_t + 0.26drp_t$*. For ease of exposition, we align the realized and expected drp. The horizontal axis displays the observation date of the expected drp. The vertical axis shows the magnitude of the risk premia. All figures are annualized.*

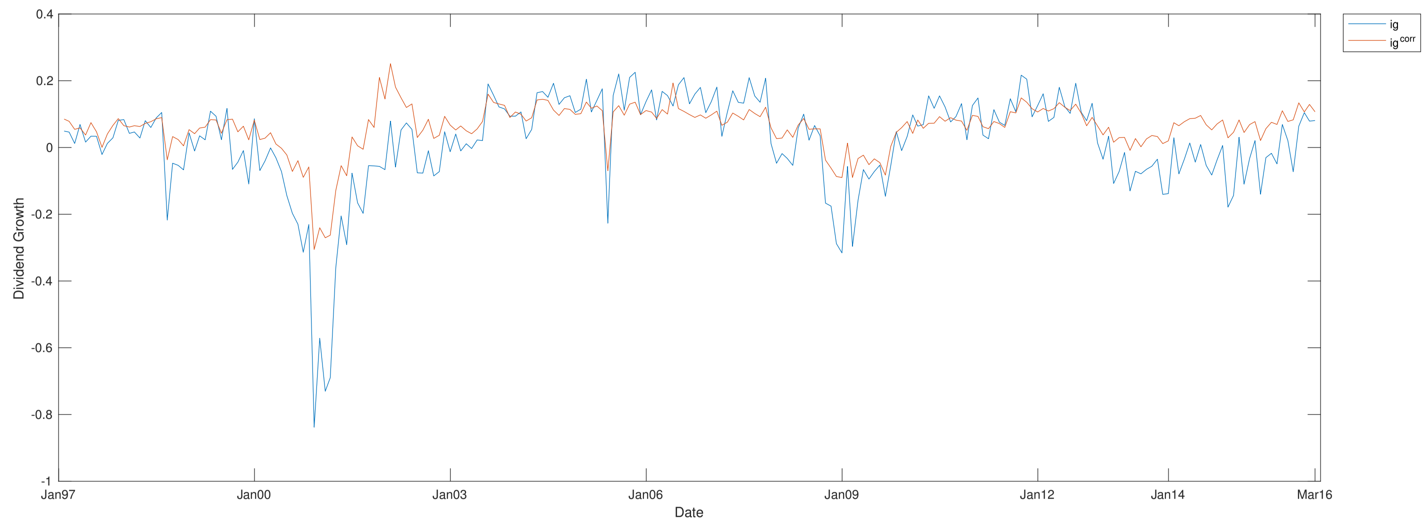

Figure 3: The Dynamics of ig and ig^{corr}

This figure plots the time series dynamics of annualized ig and ig^{corr}, *where ig is the implied dividend growth rate and* $ig_t^{corr} = \phi_0 + (1 + \phi_1)ig_t + \phi_2 drp_t$. *In the data, we find that* $\phi_0 = 0.04$, $\phi_1 = -0.55$, *and* $\phi_2 = 0.26$. *With these parameter values, we construct the relevant time series. The horizontal axis shows the observation date. The vertical axis indicates the (annualized) dividend growth rates.*

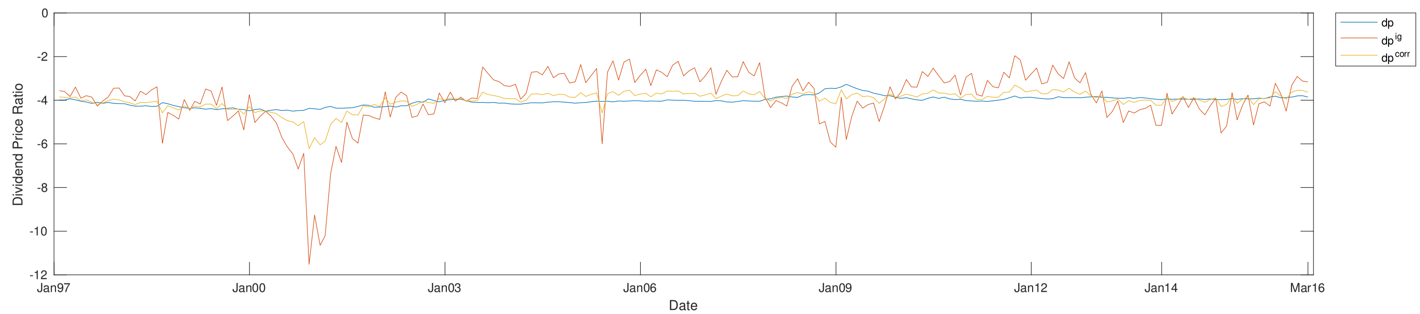

Figure 4: The Dynamics of dp, dp^{ig} and dp^{corr}

This figure plots the time series dynamics of dp, dp^{ig}, and dp^{corr}. dp is the logarithm of the trailing sum of 12-month dividends over the stock index price. $dp^{ig} = dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}$ and $dp^{corr} = dp_t + \frac{(1+\phi_1)ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}$. In the data, we find that $\phi_1 = -0.55$, $\bar{\rho} = 0.98$, $\delta_1 = 0.90$, and $\phi_2 = 0.26$. With these parameter values, we can construct the relevant time series. The horizontal axis shows the observation date. The vertical axis shows the magnitude of the ratios.

Table 1: Summary Statistics

This table reports the summary statistics of several time series. Δd denotes the time series of (annualized) monthly dividend growth. r denotes the time series of (annualized) monthly S&P 500 returns. This corresponds to the return of the trading strategy that buys the index, collects the dividends paid over the next month and sells the index at the end of the following month. ig relates to the implied growth rate. drp refers to the dividend risk premium. dp is the standard dividend price ratio. ig^{corr} is the dividend risk premium corrected implied growth rate. dp^{ig} relates to the growth adjusted dividend price ratio. dp^{corr} denotes the corrected dividend price ratio. The column entitled "Mean" reports the average of the time series [name in row]. Similarly, "Std", "Skew", and "Kurt" relate to the standard deviation, skewness, and kurtosis of the series [name in row]. AR(1) reports the first order autocorrelation. Finally, "Nobs" shows the number of observations.

	Mean	Std	Skew	Kurt	AR(1)	Nobs
Δd	0.06	0.05	-0.60	5.56	0.19	231
r	0.07	0.16	-0.78	4.32	0.08	231
ig	0.01	0.15	-2.17	11.06	0.77	231
drp	0.04	0.13	2.01	11.16	0.69	231
dp	-4.04	0.22	0.22	3.81	0.98	231
ig^{corr}	0.06	0.07	-1.94	9.84	0.81	231
dp^{ig}	-3.93	1.33	-2.29	11.84	0.79	231
dp^{corr}	-4.00	0.42	-2.28	10.58	0.87	231

Table 2: The In-Sample Predictability of Dividend Growth

This table summarizes the results of the predictability of 1-month dividend growth. We regress the time series of dividend growth on a constant and a lagged predictive variable. We consider two distinct predictive variables. The first one, ig, is the implied dividend growth rate. The second predictor, ig^{corr}, is the expected dividend risk premium corrected implied growth rate: $ig_t^{corr} = \phi_0 + (1+\phi_1)ig_t + \phi_2 drp_t$. In the data, we find that $\phi_0 = 0.04$, $\phi_1 = -0.55$, and $\phi_2 = 0.26$. Armed with these parameters, we can construct ig^{corr}. Although all regressions are estimated with an intercept, we report the slope estimates only. The entries in parentheses indicate the Newey–West (1987) adjusted t-statistics computed with 2 lags. The figures in square brackets relate to the bootstrapped p-values computed as in Rapach et al. (2013). R^2 is the r-squared of the regression model.

ig	0.42 (5.09) [0.00]	
ig^{corr}		0.92 (4.47) [0.00]
R^2	16.15%	17.56%

Table 3: The Out-of-Sample Predictability of Dividend Growth

This table presents the out-of-sample R^2 *(*R^2_{oos}*) linked to the predictability of 1-month dividend growth by the variable [name in column]. The benchmark model is the recursive mean. We consider two alternative models. Our first model derives the forecast (*$\hat{y}_t$*) as follows:* $\hat{y}_t = ig_t$*. Our second model derives the forecast as:* $\hat{y}_t = \phi_0 + (1+\phi_1)ig_t + \phi_2 drp_t$*. This forecast corresponds exactly to* ig_t^{corr}*. We use an expanding training window to estimate the parameters* ϕ_0*,* ϕ_1*, and* ϕ_2*.* $MSE-F$ *and* $MSE-Adj$ *denote the McCracken (2007) and Clark and West (2007) test statistics, respectively. The critical values of the* $MSE-F$ *statistic are 3.18, 1.55, and 0.80 at the 1 %, 5 %, and 10 % significance levels, respectively. The critical values for the* $MSE-Adj$ *test statistic are 2.33, 1.65, and 1.28 at the 1 %, 5 %, and 10 % significance levels, respectively. ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % significance levels, respectively.*

	ig	ig^{corr}
R^2_{oos}	-3.20 %	23.57 %
$MSE-F$	-2.67	26.22***
$MSE-Adj$	1.60*	2.20**

Table 4: The In-Sample Predictability of Returns

This table summarizes the results of the predictability of monthly returns. We regress the time series of returns on a constant and the lagged predictive variable. We consider three main predictive variables. The first one, dp, is the standard dividend price ratio. The second predictor, dp^{ig}, is the implied growth augmented dividend price ratio: $dp^{ig} = dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}$. The third predictor, dp^{corr}, is the corrected dividend price ratio: $dp^{corr} = dp_t + \frac{(1+\phi_1)ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}$. Using the following information, $\phi_1 = -0.55$, $\bar{\rho} = 0.98$, $\delta_1 = 0.90$, and $\phi_2 = 0.26$, we compute the relevant forecasting variables. We also consider several control variables discussed in the text: dp^{lac}, bm, def, Δff, ep, $infl$, $ntis$, pay, $rrel$, $skew$, sop, $svar$, $tbill$, $term$, and vrp. Although all regressions are estimated with an intercept, we report the slope estimates only. The entries in parentheses indicate the Newey–West (1987) adjusted t-statistics computed with 2 lags. The figures in square brackets relate to the bootstrapped p-values computed as in Rapach et al. (2013). R^2 is the r-squared of the regression model.

dp	0.20 (0.80) [0.23]																	
dp^{ig}		0.05 (1.96) [0.03]																
dp^{corr}			0.16 (2.15) [0.02]	0.15 (2.02) [0.03]	0.19 (1.99) [0.03]	0.15 (2.09) [0.03]	0.15 (1.94) [0.04]	0.15 (1.76) [0.06]	0.16 (2.11) [0.03]	0.20 (2.80) [0.01]	0.16 (2.17) [0.02]	0.13 (1.80) [0.05]	0.16 (2.13) [0.02]	0.15 (2.07) [0.03]	0.15 (2.02) [0.03]	0.18 (2.38) [0.01]	0.18 (2.47) [0.01]	0.15 (2.06) [0.03]
dp^{lac}				0.00 (0.39) [0.37]														
bm					-0.26 (-0.46) [0.34]													
def						5.60 (0.71) [0.26]												
Δff							13.30 (1.26) [0.12]											
ep								0.02 (0.17) [0.45]										
$infl$									5.79 (0.62) [0.28]									
$ntis$										4.54 (1.52) [0.08]								
pay											-0.01 (-0.06) [0.48]							
$rrel$												9.54 (1.53) [0.08]						
$skew$													0.01 (0.84) [0.22]					
sop														0.33 (0.17) [0.44]				
$svar$															-1.39 (-2.02) [0.03]			
$tbill$																1.21 (0.74) [0.25]		
$term$																	-3.27 (-1.11) [0.15]	
vrp																		0.18 (0.07) [0.48]
R^2	0.64%	1.36%	1.45%	1.60%	1.53%	1.95%	2.39%	1.48%	1.61%	3.74%	1.45%	2.77%	1.61%	1.48%	3.20%	1.63%	1.86%	1.46%

Table 5: The Out-of-Sample Predictability of Returns

This table summarizes the evidence of the predictability of returns out-of-sample. The benchmark model is the recursive mean. We consider the dp, the dp^{ig}, and the dp^{corr} ratios, in turn. The last two forecasting variables are computed using the following formulas: $dp^{ig} = dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}$ *and* $dp^{corr} = dp_t + \frac{(1+\phi_1)ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}$. *We use an expanding training window to recursively estimate the parameters* ϕ_1, $\bar{\rho}$, δ_1, *and* ϕ_2. *We use these parameters to compute the relevant forecasting variables. We also consider several documented predictors of stock market returns discussed in the text:* dp^{lac}, bm, def, Δff, ep, $infl$, $ntis$, pay, $rrel$, $skew$, sop, $svar$, $tbill$, $term$, *and* vrp. *For each of the aforementioned predictive variables, we estimate a return forecasting regression using all information from the training sample. We then use the estimated parameters together with the most recent observation of the forecasting variable to generate the forecast for the next-period, which we subsequently compare to the realized return. Panel A presents the regression results when we constrain the return forecast to be non-negative. Panel B does not impose any restriction on the return forecast. We report the out-of-sample* R^2 (R^2_{oos}) *achieved by each forecasting variable [name in column].* $MSE-F$ *and* $MSE-Adj$ *denote the McCracken (2007) and Clark and West (2007) test statistics, respectively. The critical values of the* $MSE-F$ *statistic are 3.18, 1.55, and 0.80 at the 1 %, 5 %, and 10 % significance levels, respectively. The critical values for the* $MSE-Adj$ *test statistic are 2.33, 1.65, and 1.28 at the 1 %, 5 %, and 10 % significance levels, respectively. ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % significance levels, respectively.*

Panel A: With Restriction

	dp	dp^{ig}	dp^{corr}	dp^{lac}	bm	def	Δff	ep	$infl$	$ntis$	pay	$rrel$	$skew$	sop	$svar$	$tbill$	$term$	vrp
R^2_{oos}	1.25%	0.54%	2.18%	-2.44%	-0.22%	-3.15%	-0.11%	-0.65%	-0.12%	-3.04%	-1.06%	-0.13%	-1.04%	0.02%	-0.27%	-2.29%	-1.91%	-0.69%
$MSE-F$	1.09*	0.46	1.92**	-2.05	-0.19	-2.62	-0.09	-0.55	-0.11	-2.54	-0.90	-0.12	-0.89	0.02	-0.23	-1.93	-1.61	-0.59
$MSE-Adj$	1.78**	0.74	1.97**	-1.12	-0.09	-2.49	-0.17	-0.66	0.14	-1.13	-1.65	0.00	-0.80	0.18	0.13	-2.07	-1.71	-1.37

Panel B: Without Restriction

	dp	dp^{ig}	dp^{corr}	dp^{lac}	bm	def	Δff	ep	$infl$	$ntis$	pay	$rrel$	$skew$	sop	$svar$	$tbill$	$term$	vrp
R^2_{oos}	-3.13%	-0.63%	2.09%	-38.25%	-1.65%	-23.65%	-0.11%	-20.60%	-0.12%	-12.40%	-32.62%	-1.48%	-1.09%	0.02%	-3.90%	-7.49%	-4.63%	-1.96%
$MSE-F$	-2.61	-0.54	1.84**	-23.79	-1.40	-16.45	-0.09	-14.69	-0.11	-9.49	-21.15	-1.26	-0.93	0.02	-3.23	-5.99	-3.80	-1.66
$MSE-Adj$	-0.57	0.00	1.91**	-1.24	-0.92	-1.35	-0.17	-1.15	0.14	-1.46	-1.14	-0.72	-0.83	0.18	-0.68	-1.66	-1.69	-1.31

Table 6: The Economic Value of Return Predictability

This table sheds light on the economic gains of an investor who attempts to exploit the predictability of returns by devising market timing strategies. We assume that the investor has a quadratic utility function with a coefficient of relative risk aversion equal to γ. The first column shows the different values of γ, i.e. $\gamma = 2, 4, 6, 8$ or 10. At the end of each month, we compute the optimal allocation of the investor to the risky stock and the riskless asset. These weights depend, among others, on the forecasting model for expected returns and the risk aversion estimate. The investor considers several forecasting variables: dp, dp^{ig}, dp^{corr}, dp^{lac}, bm, def, Δff, ep, $infl$, $ntis$, pay, $rrel$, $skew$, sop, $svar$, $tbill$, $term$, and vrp. Given these weights, we compute the realized return on the portfolio. We do this for each calendar month and return forecasting variable. The first 5 rows report the difference between the annualized certainty equivalent (ΔCE) of the strategy based on the predictive variable [name in column] and that of the strategy based on the recursive mean. The penultimate row sheds light on the difference between the annualized Sharpe ratio (SR) of the strategy based on the predictive variable [name in column] and that of the strategy based on the recursive mean. The last row presents the annualized Sharpe ratio (SR) of the strategy based on the variable [name in column].

γ	dp	dp^{ig}	dp^{corr}	dp^{lac}	bm	def	Δff	ep	$infl$	$ntis$	pay	$rrel$	$skew$	sop	$svar$	$tbill$	$term$	vrp
2	4.21%	2.10%	6.12%	-4.72%	0.44%	-6.79%	-0.01%	-0.39%	-0.24%	-5.17%	-2.58%	0.85%	-1.84%	0.68%	0.03%	-5.03%	-4.95%	-1.52%
4	2.10%	1.05%	3.07%	-2.37%	0.21%	-3.41%	-0.01%	-0.21%	-0.12%	-2.60%	-1.30%	0.41%	-0.90%	0.33%	0.00%	-2.53%	-2.48%	-0.77%
6	1.39%	0.70%	2.05%	-1.59%	0.14%	-2.28%	0.00%	-0.14%	-0.07%	-1.74%	-0.87%	0.27%	-0.59%	0.22%	-0.01%	-1.69%	-1.66%	-0.51%
8	1.04%	0.52%	1.54%	-1.20%	0.10%	-1.72%	0.00%	-0.11%	-0.05%	-1.31%	-0.66%	0.20%	-0.44%	0.16%	-0.01%	-1.27%	-1.25%	-0.39%
10	0.83%	0.42%	1.23%	-0.96%	0.08%	-1.38%	0.00%	-0.10%	-0.04%	-1.06%	-0.53%	0.16%	-0.34%	0.12%	-0.01%	-1.02%	-1.00%	-0.31%
ΔSR	26.63%	16.08%	34.98%	-20.98%	4.77%	-66.86%	1.06%	-1.49%	-2.45%	-23.90%	-21.07%	7.82%	-12.68%	5.75%	2.52%	-40.96%	-46.92%	-11.69%
SR	44.34%	33.79%	52.69%	-3.28%	22.48%	-49.15%	18.76%	16.22%	15.25%	-6.19%	-3.36%	25.52%	5.03%	23.46%	20.23%	-23.25%	-29.22%	6.02%

Appendix to

"The Predictive Power of the Dividend Risk Premium"

Not Intended for Publication!

Will be Provided as Online Appendix

A Proofs

This appendix presents the detailed proof of the propositions presented in the main text. In order to facilitate the exposition of the derivations, it is useful to re-state our main assumptions:

$$r_{t+1} = \mu_t + \epsilon^r_{t+1} \tag{A.1}$$

$$\mu_{t+1} = \alpha_0 + \alpha_1 \mu_t + \epsilon^\mu_{t+1} \tag{A.2}$$

$$ig_{t+1} = \delta_0 + \delta_1 ig_t + \epsilon^{ig}_{t+1} \tag{A.3}$$

$$drp_{t+1} = \phi_0 + \phi_1 ig_t + \phi_2 drp_t + \epsilon^{drp}_{t+1} \tag{A.4}$$

where all error terms are $i.i.d$ with zero mean.

A.1 Proposition 1

To derive the first proposition of our model, we start from the accounting identity linking together the expected dividend growth rate, the expected drp and the implied growth rate:

$$\mathbb{E}_t(\Delta d_{t+1}) - \mathbb{E}_t(drp_{t+1}) = ig_t$$

This implies that:

$$\begin{aligned} \mathbb{E}_t(\Delta d_{t+1}) &= \mathbb{E}_t(drp_{t+1}) + ig_t \\ &= \mathbb{E}_t(\phi_0 + \phi_1 ig_t + \phi_2 drp_t + \epsilon^{drp}_{t+1}) + ig_t \end{aligned}$$

$$\mathbb{E}_t(\Delta d_{t+1}) = \phi_0 + (1+\phi_1) ig_t + \phi_2 drp_t \tag{A.5}$$

Recall that the realized dividend growth can be decomposed into an expected component and a shock:

$$\Delta d_{t+1} = \mathbb{E}_t(\Delta d_{t+1}) + \epsilon_{t+1}^{\Delta d} \tag{A.6}$$

$$\Delta d_{t+1} = \phi_0 + (1+\phi_1) ig_t + \phi_2 drp_t + \epsilon_{t+1}^{\Delta d} \tag{A.7}$$

This completes the proof of Proposition 1. ■

A.2 Proposition 2:

For ease of exposition, let us restate equation (10) from the manuscript:

$$\sum_{j=0}^{+\infty} \overline{\rho}^j \left(\mathbb{E}_t(r_{t+1+j}) - \mathbb{E}_t(\Delta d_{t+1+j})\right) = \frac{k}{1-\overline{\rho}} + dp_t \tag{A.8}$$

Using equations (A.1) and (A.2), we can compute the first summation term on the left-hand side of equation (A.8):

$$\sum_{j=0}^{+\infty} \overline{\rho}^j \mathbb{E}_t(r_{t+1+j}) \equiv k_r + \sum_{j=0}^{+\infty} \overline{\rho}^j \alpha_1^j \mu_t$$

$$\sum_{j=0}^{+\infty} \overline{\rho}^j \mathbb{E}_t(r_{t+1+j}) \equiv k_r + \frac{\mu_t}{1-\overline{\rho}\alpha_1} \tag{A.9}$$

where k_r is a constant that depends on α_0 and α_1.

Similarly, we combine the result of Proposition 1 (equation (A.7)) together with equations (A.3) and (A.4) to compute the infinite sum of expected dividend growth rates:

$$\sum_{j=0}^{+\infty} \overline{\rho}^j \mathbb{E}_t(\Delta d_{t+1+j}) = k_{\Delta d} + \sum_{j=0}^{+\infty} \overline{\rho}^j \left[\delta_1^j (1+\phi_1) + \phi_1 \phi_2 \frac{\delta_1^j - \phi_2^j}{\delta_1 - \phi_2} \right] ig_t + \sum_{j=0}^{+\infty} \overline{\rho}^j \phi_2^{j+1} drp_t$$

$$\sum_{j=0}^{+\infty} \overline{\rho}^j \mathbb{E}_t(\Delta d_{t+1+j}) \equiv k_{\Delta d} + \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] ig_t + \frac{\phi_2 drp_t}{1-\overline{\rho}\phi_2} \tag{A.10}$$

where $k_{\Delta d}$ is a constant that depends on δ_0, δ_1, ϕ_0, ϕ_1 and ϕ_2.

Substituting equations (A.9) and (A.10) into equation (A.8) yields:

$$\begin{aligned} dp_t &= -\frac{k}{1-\overline{\rho}} + \sum_{j=0}^{+\infty} \overline{\rho}^j \left(\mathbb{E}_t(r_{t+1+j}) - \mathbb{E}_t(\Delta d_{t+1+j}) \right) \\ &= \underbrace{-\frac{k}{1-\overline{\rho}} + k_r - k_{\Delta d}}_{k_1} + \frac{\mu_t}{1-\overline{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] ig_t - \frac{\phi_2 drp_t}{1-\overline{\rho}\phi_2} \end{aligned}$$

$$dp_t \equiv k_1 + \frac{\mu_t}{1-\overline{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] ig_t - \frac{\phi_2}{1-\overline{\rho}\phi_2} drp_t \tag{A.11}$$

Similarly, we can express the next-period dividend price ratio as:

$$dp_{t+1} = k_1 + \frac{\mu_{t+1}}{1-\overline{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] ig_{t+1} - \frac{\phi_2}{1-\overline{\rho}\phi_2} drp_{t+1}$$

Using equations (A.3) and (A.4), we can show that:

$$\begin{aligned} dp_{t+1} &= k_1 + \frac{\mu_{t+1}}{1-\overline{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] ig_{t+1} - \frac{\phi_2}{1-\overline{\rho}\phi_2} drp_{t+1} \\ &= k_1 + \frac{\alpha_0 + \alpha_1 \mu_t + \epsilon_{t+1}^{\mu}}{1-\overline{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] (\delta_0 + \delta_1 ig_t + \epsilon_{t+1}^{ig}) \end{aligned}$$

$$
\begin{aligned}
&- \frac{\phi_2}{1-\bar{\rho}\phi_2}(\phi_0 + \phi_1 ig_t + \phi_2 drp_t + \epsilon^{drp}_{t+1}) \\
&= k_1 + \underbrace{\frac{\alpha_0}{1-\bar{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)}\right]\delta_0 - \frac{\phi_0\phi_2}{1-\bar{\rho}\phi_2}}_{k_2} + \frac{\alpha_1\mu_t + \epsilon^{\mu}_{t+1}}{1-\bar{\rho}\alpha_1} \\
&- \left[\frac{1+\phi_1}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)}\right](\delta_1 ig_t + \epsilon^{ig}_{t+1}) - \frac{\phi_1\phi_2 ig_t}{1-\bar{\rho}\phi_2} - \frac{\phi_2^2 drp_t + \phi_2\epsilon^{drp}_{t+1}}{1-\bar{\rho}\phi_2} \\
&\equiv k_1 + k_2 + \frac{\alpha_1\mu_t}{1-\bar{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)}\right]\delta_1 ig_t - \frac{\phi_1\phi_2 ig_t}{1-\bar{\rho}\phi_2} \\
&- \frac{\phi_2^2 drp_t}{1-\bar{\rho}\phi_2} + \underbrace{\frac{\epsilon^{\mu}_{t+1}}{1-\bar{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)}\right]\epsilon^{ig}_t - \frac{\phi_2\epsilon^{drp}_{t+1}}{1-\bar{\rho}\phi_2}}_{\epsilon^{dp}_{t+1}} \\
&\equiv \alpha_1 k_1 + \frac{\alpha_1\mu_t}{1-\bar{\rho}\alpha_1} - \left[\frac{1+\phi_1}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)}\right](\alpha_1 + \delta_1 - \alpha_1) ig_t \\
&- \frac{\phi_1\phi_2 ig_t}{1-\bar{\rho}\phi_2} - \frac{\alpha_1 + \phi_2 - \alpha_1}{1-\bar{\rho}\phi_2}\phi_2 drp_t + k_2 + (1-\alpha_1)k_1 + \epsilon^{dp}_{t+1}
\end{aligned}
$$

$$
\begin{aligned}
dp_{t+1} &= k_2 + (1-\alpha_1)k_1 + \alpha_1 dp_t - \left[\frac{1+\phi_1}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)}\right](\delta_1 - \alpha_1) ig_t \\
&- \frac{\phi_1\phi_2 ig_t}{1-\bar{\rho}\phi_2} - \frac{\phi_2 - \alpha_1}{1-\bar{\rho}\phi_2}\phi_2 drp_t + \epsilon^{dp}_{t+1}
\end{aligned} \tag{A.12}
$$

Following the steps of Campbell and Shiller (1988), it is straightforward to show that:

$$
r_{t+1} \approx k + \Delta d_{t+1} + dp_t - \bar{\rho} dp_{t+1} \tag{A.13}
$$

The final step of the proof consists in substituting equations (A.7), (A.11) and (A.12) into equation (A.13):

$$
\begin{aligned}
r_{t+1} &= \mathbb{E}_t\left(k + \Delta d_{t+1} + dp_t - \bar{\rho} dp_{t+1}\right) + \epsilon^r_{t+1} \\
&= k + \phi_0 + (1+\phi_1) ig_t + \phi_2 drp_t + dp_t - \bar{\rho}\left(k_2 + (1-\alpha_1)k_1\right) + \epsilon^r_{t+1}
\end{aligned}
$$

$$- \overline{\rho} \left(\alpha_1 dp_t - \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] (\delta_1 - \alpha_1) ig_t - \frac{\phi_1\phi_2 ig_t}{1-\overline{\rho}\phi_2} - \frac{\phi_2 - \alpha_1}{1-\overline{\rho}\phi_2} \phi_2 drp_t \right)$$

$$= \underbrace{k + \phi_0 - \overline{\rho}\left(k_2 + (1-\alpha_1)k_1\right)}_{\Psi} + (1-\overline{\rho}\alpha_1) \left(dp_t + \left[\frac{1+\phi_1}{1-\overline{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} \right] ig_t \right)$$

$$+ (1-\overline{\rho}\alpha_1) \left(\frac{\phi_2 drp_t}{1-\overline{\rho}\phi_2} \right) + \epsilon^r_{t+1}$$

We thus obtain:

$$\text{(A.14)} \qquad r_{t+1} \equiv \Psi + (1-\overline{\rho}\alpha_1) \underbrace{\left(dp_t + \frac{1+\phi_1}{1-\overline{\rho}\delta_1} ig_t + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2}{1-\overline{\rho}\phi_2} drp_t \right)}_{dp^{corr}}$$

$$+ \epsilon^r_{t+1}$$

This completes the proof of Proposition 2. ■

B Additional Control Variables

- dp^{lac}: We compute the dp^{lac} ratio as in Golez (2014):

$$dp^{lac} = dp_t + \frac{\Delta \bar{d}_t}{1 - \overline{\rho}} \tag{B.1}$$

  where $\Delta \bar{d}_t$ is the average dividend growth rate over the past year.

- Δff: We construct Δff as the 1-month change in the federal fund rate as in Maio (2014).

- $rrel$: The $rrel$ measure is the difference between the 3-month Treasury bill rate and its last four-quarter average as in Maio (2013). We obtain all interest rate data from the FRED database of the Federal Reserve Bank of St. Louis.

- $skew$: We download the time-series of the implied skewness from the website of the CBOE.

- sop: We implement the sum-of-part method with no multiple growth. The forecast is given as $\hat{r}_{t+1} = dp_t + \bar{ge}_t$ where $\bar{ge}_t$ is the average earnings growth rate at time t computed using a trailing window of 20 years as in Ferreira and Santa-Clara (2011).

- $svar$: Following Bollerslev, Tauchen, and Zhou (2009) and Drechsler and Yaron (2011), we construct $svar$ as the sum of the squared (1) 5-minute intraday returns and (2) the close-to-open (overnight) returns observed that month. We annualize the monthly $svar$

by multiplying it with 12. All calculations are based on the intraday dataset from the CBOE.

- *term*: Following Maio (2013), we construct *term* as the difference between the yields on the 10-year and the 1-year US Treasury bonds. The data come from the FRED database.

- *vrp*: The *vrp* is defined as the difference between the (1) physical and (2) risk-neutral expectations of next-month's variance. We proxy the risk-neutral expectation of variance with the squared value of the VIX, which we obtain from Bloomberg. In order to compute the physical expectation of the (annualized) realized variance of the index returns, we closely follow the empirical framework of Drechsler and Yaron (2011). Briefly, we regress the monthly time series of *svar* on a constant, the lagged *svar*, and the lagged squared value of the VIX. We use the fitted value from this model as the physical expectation of realized variance.

Table A.1: Summary Statistics: At-the-Money Options

This table reports the summary statistics of several time series. Δd denotes the time series of (annualized) monthly dividend growth. r denotes the time series of (annualized) monthly S&P 500 returns. This corresponds to the return of the trading strategy that buys the index, collects the dividends paid over the next month and sells the index at the end of the following month. ig relates to the implied growth rate. In constructing the dividend strip, we only consider options that are at-the-money, i.e. with a Black and Scholes (1973) delta that is in absolute value between 0.375 and 0.625. drp refers to the dividend risk premium. dp is the standard dividend price ratio. ig^{corr} is the dividend risk premium corrected implied growth rate. dp^{ig} relates to the growth adjusted dividend price ratio. dp^{corr} denotes the corrected dividend price ratio. The column entitled "Mean" reports the average of the time series [name in row]. Similarly, "Std", "Skew", and "Kurt" relate to the standard deviation, skewness, and kurtosis of the series [name in row]. AR(1) reports the first order autocorrelation. Finally, "Nobs" shows the number of observations.

	Mean	Std	Skew	Kurt	AR(1)	Nobs
Δd	0.06	0.05	-0.60	5.56	0.19	231
r	0.07	0.16	-0.78	4.32	0.08	231
ig	0.01	0.15	-2.05	10.40	0.78	231
drp	0.04	0.13	1.92	10.40	0.70	231
dp	-4.04	0.22	0.22	3.81	0.98	231
ig^{corr}	0.06	0.07	-1.92	9.65	0.81	231
dp^{ig}	-3.94	1.37	-2.18	11.22	0.79	231
dp^{corr}	-4.00	0.41	-2.20	10.14	0.87	231

Table A.2: The In-Sample Predictability of Dividend Growth: At-the-Money Options

This table summarizes the results of the predictability of 1-month dividend growth. We regress the time series of dividend growth on a constant and a lagged predictive variable. We consider two distinct predictive variables. The first one, ig, is the implied dividend growth rate. The second predictor, ig^{corr}, is the expected dividend risk premium corrected implied growth rate: $ig_t^{corr} = \phi_0 + (1+\phi_1)ig_t + \phi_2 drp_t$. In constructing the dividend strip, we only consider options that are at-the-money, i.e. with a Black and Scholes (1973) delta that is in absolute value between 0.375 and 0.625. In the data, we find that $\phi_0 = 0.04$, $\phi_1 = -0.58$, and $\phi_2 = 0.24$. Armed with these parameters, we can construct ig^{corr}. Although all regressions are estimated with an intercept, we report the slope estimates only. The entries in parentheses indicate the Newey–West (1987) adjusted t-statistics computed with 2 lags. The figures in square brackets relate to the bootstrapped p-values computed as in Rapach et al. (2013). R^2 is the r-squared of the regression model.

	0.40	
ig	(5.27)	
	[0.00]	
		0.97
ig^{corr}		(4.54)
		[0.00]
R^2	15.56%	17.23%

Table A.3: The Out-of-Sample Predictability of Dividend Growth: At-the-Money Options

*This table presents the out-of-sample R^2 (R^2_{oos}) linked to the predictability of 1-month dividend growth by the variable [name in column]. The benchmark model is the recursive mean. We consider two alternative models. Our first model derives the forecast ($\hat{y}_t$) as follows: $\hat{y}_t = ig_t$. Our second model derives the forecast as: $\hat{y}_t = \phi_0 + (1 + \phi_1)ig_t + \phi_2 drp_t$. This forecast corresponds exactly to ig_t^{corr}. In constructing the dividend strip, we only consider options that are at-the-money, i.e. with a Black and Scholes (1973) delta that is in absolute value between 0.375 and 0.625. We use an expanding training window to estimate the parameters ϕ_0, ϕ_1 and ϕ_2. $MSE - F$ and $MSE - Adj$ denote the McCracken (2007) and Clark and West (2007) test statistics, respectively. The critical values of the $MSE - F$ statistic are 3.18, 1.55, and 0.80 at the 1 %, 5 %, and 10 % significance levels, respectively. The critical values for the $MSE - Adj$ test statistic are 2.33, 1.65, and 1.28 at the 1 %, 5 %, and 10 % significance levels, respectively. ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % significance levels, respectively.*

	ig	ig^{corr}
R^2_{oos}	-12.30%	21.53%
$MSE - F$	-9.42	23.59***
$MSE - Adj$	1.51*	2.12**

Table A.4: The In-Sample Predictability of Returns: At-the-Money Options

This table summarizes the results of the predictability of monthly returns. We regress the time series of returns on a constant and the lagged predictive variable. We consider three main predictive variables. The first one, dp, is the standard dividend price ratio. The second predictor, dp^{ig}, is the implied growth augmented dividend price ratio: $dp^{ig} = dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}$. In constructing the dividend strip, we only consider options that are at-the-money, i.e. with a Black and Scholes (1973) delta that is in absolute value between 0.375 and 0.625. The third predictor, dp^{corr}, is the corrected dividend price ratio: $dp^{corr} = dp_t + \frac{(1+\phi_1)ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}$. Using the following information, $\phi_1 = -0.58$, $\bar{\rho} = 0.98$, $\delta_1 = 0.90$, and $\phi_2 = 0.24$, we compute the relevant forecasting variables. We also consider several control variables discussed in the text: dp^{lac}, bm, def, Δff, ep, $infl$, ntis, pay, rrel, skew, sop, svar, tbill, term, and vrp. Although all regressions are estimated with an intercept, we report the slope estimates only. The entries in parentheses indicate the Newey–West (1987) adjusted t-statistics computed with 2 lags. The figures in square brackets relate to the bootstrapped p-values computed as in Rapach et al. (2013). R^2 is the r-squared of the regression model.

dp	0.20 (0.80) [0.24]																	
dp^{ig}		0.05 (2.02) [0.03]																
dp^{corr}			0.16 (2.24) [0.02]	0.16 (2.11) [0.03]	0.20 (2.12) [0.03]	0.16 (2.19) [0.02]	0.15 (2.04) [0.03]	0.16 (1.85) [0.04]	0.16 (2.20) [0.02]	0.21 (2.89) [0.00]	0.16 (2.27) [0.02]	0.14 (1.89) [0.04]	0.16 (2.23) [0.02]	0.16 (2.17) [0.02]	0.16 (2.14) [0.03]	0.19 (2.49) [0.01]	0.19 (2.57) [0.01]	0.16 (2.16) [0.02]
dp^{lac}				0.00 (0.39) [0.36]														
bm					-0.30 (-0.52) [0.32]													
def						5.63 (0.72) [0.26]												
Δff							13.26 (1.26) [0.13]											
ep								0.02 (0.17) [0.45]										
$infl$									5.70 (0.60) [0.30]									
$ntis$										4.60 (1.54) [0.08]								
pay											-0.01 (-0.07) [0.48]							
$rrel$												9.47 (1.52) [0.08]						
$skew$													0.01 (0.84) [0.22]					
sop														0.34 (0.18) [0.44]				
$svar$															-1.40 (-2.04) [0.03]			
$tbill$																1.23 (0.76) [0.25]		
$term$																	-3.29 (-1.12) [0.16]	
vrp																		0.18 (0.07) [0.47]
R^2	0.64%	1.45%	1.56%	1.71%	1.67%	2.06%	2.50%	1.58%	1.71%	3.91%	1.56%	2.86%	1.73%	1.59%	3.34%	1.74%	1.98%	1.57%

Table A.5: The Out-of-Sample Predictability of Returns: At-the-Money Options

This table summarizes the evidence of the predictability of returns out-of-sample. The benchmark model is the recursive mean. We consider the dp, the dp^{ig}*, and the* dp^{corr} *ratios, in turn. The last two forecasting variables are computed using the following formulas:* $dp^{ig} = dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}$ *and* $dp^{corr} = dp_t + \frac{(1+\phi_1)ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}$*. In constructing the dividend strip, we only consider options that are at-the-money, i.e. with a Black and Scholes (1973) delta that is in absolute value between 0.375 and 0.625. We use an expanding training window to recursively estimate the parameters* ϕ_1*,* $\bar{\rho}$*,* δ_1*, and* ϕ_2*. We use these parameters to compute the relevant forecasting variables. We also consider several documented predictors of stock market returns discussed in the text:* dp^{lac}*, bm,* def*,* Δff*, ep,* $infl$*, ntis, pay, rrel, skew, sop, svar, tbill, term, and vrp. For each of the aforementioned predictive variables, we estimate a return forecasting regression using all information from the training sample. We then use the estimated parameters together with the most recent observation of the forecasting variable to generate the forecast for the next-period, which we subsequently compare to the realized return. We report the out-of-sample* R^2 *(*R^2_{oos}*) achieved by each forecasting variable [name in column].* $MSE-F$ *and* $MSE-Adj$ *denote the McCracken (2007) and Clark and West (2007) test statistics, respectively. The critical values of the* $MSE-F$ *statistic are 3.18, 1.55, and 0.80 at the 1 %, 5 %, and 10 % significance levels, respectively. The critical values for the* $MSE-Adj$ *test statistic are 2.33, 1.65, and 1.28 at the 1 %, 5 %, and 10 % significance levels, respectively. ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % significance levels, respectively.*

	dp	dp^{ig}	dp^{corr}	dp^{lac}	bm	def	Δff	ep	$infl$	$ntis$	pay	$rrel$	$skew$	sop	$svar$	$tbill$	$term$	vrp
R^2_{oos}	1.25%	0.70%	2.38%	-2.44%	-0.22%	-3.15%	-0.11%	-0.65%	-0.12%	-3.04%	-1.06%	-0.13%	-1.04%	0.02%	-0.27%	-2.29%	-1.91%	-0.69%
$MSE-F$	1.09*	0.60	2.10**	-2.05	-0.19	-2.62	-0.09	-0.55	-0.11	-2.54	-0.90	-0.12	-0.89	0.02	-0.23	-1.93	-1.61	-0.59
$MSE-Adj$	1.78**	0.87	2.12**	-1.12	-0.09	-2.49	-0.17	-0.66	0.14	-1.13	-1.65	0.00	-0.80	0.18	0.13	-2.07	-1.71	-1.37

Table A.6: The Economic Value of Return Predictability: At-the-Money Options

This table sheds light on the economic gains of an investor who attempts to exploit the predictability of returns by devising market timing strategies. In constructing the dividend strip, we only consider options that are at-the-money, i.e. with a Black and Scholes (1973) delta that is in absolute value between 0.375 and 0.625. We assume that the investor has a quadratic utility function with a coefficient of relative risk aversion equal to γ. The first column shows the different values of γ, i.e. $\gamma = 2, 4, 6, 8$ or 10. At the end of each month, we compute the optimal allocation of the investor to the risky stock and the riskless asset. These weights depend, among others, on the forecasting model for expected returns and the risk aversion estimate. The investor considers several forecasting variables: dp, dp^{ig}, dp^{corr}, dp^{lac}, bm, def, Δff, ep, $infl$, $ntis$, pay, $rrel$, $skew$, sop, $svar$, $tbill$, $term$, and vrp. Given these weights, we compute the realized return on the portfolio. We do this for each calendar month and return forecasting variable. The first 5 rows report the difference between the annualized certainty equivalent (ΔCE) of the strategy based on the predictive variable [name in column] and that of the strategy based on the recursive mean. The penultimate row sheds light on the difference between the annualized Sharpe ratio (SR) of the strategy based on the predictive variable [name in column] and that of the strategy based on the recursive mean. The last row presents the annualized Sharpe ratio (SR) of the strategy based on the variable [name in column].

γ	dp	dp^{ig}	dp^{corr}	dp^{lac}	bm	def	Δff	ep	$infl$	$ntis$	pay	$rrel$	$skew$	sop	$svar$	$tbill$	$term$	vrp
2	4.21%	2.50%	6.59%	-4.72%	0.44%	-6.79%	-0.01%	-0.39%	-0.24%	-5.17%	-2.58%	0.85%	-1.84%	0.68%	0.03%	-5.03%	-4.95%	-1.52%
4	2.10%	1.25%	3.30%	-2.37%	0.21%	-3.41%	-0.01%	-0.21%	-0.12%	-2.60%	-1.30%	0.41%	-0.90%	0.33%	0.00%	-2.53%	-2.48%	-0.77%
6	1.39%	0.83%	2.20%	-1.59%	0.14%	-2.28%	0.00%	-0.14%	-0.07%	-1.74%	-0.87%	0.27%	-0.59%	0.22%	-0.01%	-1.69%	-1.66%	-0.51%
8	1.04%	0.62%	1.66%	-1.20%	0.10%	-1.72%	0.00%	-0.11%	-0.05%	-1.31%	-0.66%	0.20%	-0.44%	0.16%	-0.01%	-1.27%	-1.25%	-0.39%
10	0.83%	0.50%	1.33%	-0.96%	0.08%	-1.38%	0.00%	-0.10%	-0.04%	-1.06%	-0.53%	0.16%	-0.34%	0.12%	-0.01%	-1.02%	-1.00%	-0.31%
ΔSR	26.63%	17.96%	37.10%	-20.98%	4.77%	-66.86%	1.06%	-1.49%	-2.45%	-23.90%	-21.07%	7.82%	-12.68%	5.75%	2.52%	-40.96%	-46.92%	-11.69%
SR	44.34%	35.67%	54.81%	-3.28%	22.48%	-49.15%	18.76%	16.22%	15.25%	-6.19%	-3.36%	25.52%	5.03%	23.46%	20.23%	-23.25%	-29.22%	6.02%

Table A.7: Summary Statistics: Alternative Interpolation

This table reports the summary statistics of several time series. Δd denotes the time series of (annualized) monthly dividend growth. r denotes the time series of (annualized) monthly S&P 500 returns. This corresponds to the return of the trading strategy that buys the index, collects the dividends paid over the next month and sells the index at the end of the following month. ig relates to the implied growth rate. In constructing the annual dividend strip, we directly interpolate the 12-month maturity. drp refers to the dividend risk premium. dp is the standard dividend price ratio. ig^{corr} is the dividend risk premium corrected implied growth rate. dp^{ig} relates to the growth adjusted dividend price ratio. dp^{corr} denotes the corrected dividend price ratio. The column entitled "Mean" reports the average of the time series [name in row]. Similarly, "Std", "Skew", and "Kurt" relate to the standard deviation, skewness, and kurtosis of the series [name in row]. AR(1) reports the first order autocorrelation. Finally, "Nobs" shows the number of observations.

	Mean	Std	Skew	Kurt	AR(1)	Nobs
Δd	0.06	0.05	-0.60	5.56	0.19	231
r	0.07	0.16	-0.78	4.32	0.08	231
ig	0.00	0.15	-2.29	11.36	0.84	231
drp	0.05	0.11	2.59	14.70	0.75	231
dp	-4.04	0.22	0.22	3.81	0.98	231
ig^{corr}	0.06	0.08	-2.19	10.59	0.86	231
dp^{ig}	-4.00	1.17	-2.46	12.65	0.86	231
dp^{corr}	-4.01	0.47	-2.56	12.49	0.90	231

Table A.8: The In-Sample Predictability of Dividend Growth: Alternative Interpolation

This table summarizes the results of the predictability of 1-month dividend growth. We regress the time series of dividend growth on a constant and a lagged predictive variable. We consider two distinct predictive variables. The first one, ig, is the implied dividend growth rate. The second predictor, ig^{corr}, is the expected dividend risk premium corrected implied growth rate: $ig_t^{corr} = \phi_0 + (1+\phi_1)ig_t + \phi_2 drp_t$. In constructing the annual dividend strip, we directly interpolate the 12-month maturity. In the data, we find that $\phi_0 = 0.04$, $\phi_1 = -0.49$ and $\phi_2 = 0.25$. Armed with these parameters, we can construct ig^{corr}. Although all regressions are estimated with an intercept, we report the slope estimates only. The entries in parentheses indicate the Newey–West (1987) adjusted t-statistics computed with 2 lags. The figures in square brackets relate to the bootstrapped p-values computed as in Rapach et al. (2013). R^2 is the r-squared of the regression model.

	0.45	
ig	(4.86)	
	[0.00]	
		0.86
dp^{corr}		(4.51)
		[0.00]
R^2	18.13%	18.66%

Table A.9: The Out-of-Sample Predictability of Dividend Growth: Alternative Interpolation

*This table presents the out-of-sample R^2 (R^2_{oos}) linked to the predictability of 1-month dividend growth by the variable [name in column]. The benchmark model is the recursive mean. We consider two alternative models. Our first model derives the forecast ($\hat{y}_t$) as follows: $\hat{y}_t = ig_t$. Our second model derives the forecast as: $\hat{y}_t = \phi_0 + (1 + \phi_1)ig_t + \phi_2 drp_t$. This forecast corresponds exactly to ig^{corr}_t. In constructing the annual dividend strip, we directly interpolate the 12-month maturity. We use an expanding training window to estimate the parameters ϕ_0, ϕ_1, and ϕ_2. $MSE - F$ and $MSE - Adj$ denote the McCracken (2007) and Clark and West (2007) test statistics, respectively. The critical values of the $MSE - F$ statistic are 3.18, 1.55, and 0.80 at the 1 %, 5 %, and 10 % significance levels, respectively. The critical values for the $MSE - Adj$ test statistic are 2.33, 1.65, and 1.28 at the 1 %, 5 %, and 10 % significance levels, respectively. ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % significance levels, respectively.*

	ig	ig^{corr}
R^2_{oos}	0.56%	22.52%
$MSE - F$	0.49	25.00***
$MSE - Adj$	1.54*	2.02**

Table A.10: The In-Sample Predictability of Returns: Alternative Interpolation

This table summarizes the results of the predictability of monthly returns. We regress the time series of returns on a constant and the lagged predictive variable. We consider three main predictive variables. The first one, dp, is the standard dividend price ratio. The second predictor, dp^{ig}, is the implied growth augmented dividend price ratio: $dp^{ig} = dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}$. In constructing the annual dividend strip, we directly interpolate the 12-month maturity. The third predictor, dp^{corr}, is the corrected dividend price ratio: $dp^{corr} = dp_t + \frac{(1+\phi_1)ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}$. Using the following information, $\phi_1 = -0.49$, $\bar{\rho} = 0.98$, $\delta_1 = 0.89$, and $\phi_2 = 0.25$, we compute the relevant forecasting variables. We also consider several control variables discussed in the text: dp^{lac}, bm, def, Δff, ep, $infl$, ntis, pay, rrel, skew, sop, svar, tbill, term, and vrp. Although all regressions are estimated with an intercept, we report the slope estimates only. The entries in parentheses indicate the Newey–West (1987) adjusted t-statistics computed with 2 lags. The figures in square brackets relate to the bootstrapped p-values computed as in Rapach et al. (2013). R^2 is the r-squared of the regression model.

dp	0.20 (0.80) [0.23]																	
dp^{ig}		0.05 (1.77) [0.05]																
dp^{corr}			0.14 (2.15) [0.02]	0.13 (2.02) [0.03]	0.16 (1.81) [0.06]	0.13 (1.95) [0.04]	0.13 (1.93) [0.04]	0.14 (1.84) [0.05]	0.14 (2.11) [0.03]	0.16 (2.43) [0.01]	0.14 (2.18) [0.02]	0.11 (1.71) [0.06]	0.14 (2.12) [0.02]	0.14 (2.10) [0.02]	0.12 (1.81) [0.05]	0.16 (2.18) [0.02]	0.16 (2.38) [0.02]	0.14 (2.10) [0.03]
dp^{lac}				0.00 (0.27) [0.41]														
bm					-0.20 (-0.33) [0.39]													
def						4.50 (0.56) [0.31]												
Δff							13.17 (1.25) [0.12]											
ep								0.01 (0.08) [0.47]										
$infl$									5.21 (0.57) [0.31]									
$ntis$										4.02 (1.36) [0.11]								
pay											0.01 (0.07) [0.47]							
$rrel$												9.16 (1.50) [0.08]						
$skew$													0.01 (0.80) [0.23]					
sop														0.09 (0.05) [0.47]				
$svar$															-1.30 (-1.82) [0.05]			
$tbill$																1.00 (0.59) [0.30]		
$term$																	-3.12 (-1.03) [0.18]	
vrp																		0.07 (0.03) [0.49]
R^2	0.64%	1.32%	1.44%	1.51%	1.49%	1.76%	2.37%	1.45%	1.57%	3.30%	1.45%	2.63%	2.63%	1.45%	2.94%	1.57%	1.82%	1.44%

Table A.11: The Out-of-Sample Predictability of Returns: Alternative Interpolation

This table summarizes the evidence of the predictability of returns out-of-sample. The benchmark model is the recursive mean. We consider the dp, the dp^{ig} *and the* dp^{corr} *ratios, in turn. The last two forecasting variables are computed using the following formulas:* $dp^{ig} = dp_t + \frac{ig_t}{1-\bar{\rho}\delta_1}$ *and* $dp^{corr} = dp_t + \frac{(1+\phi_1)ig_t}{1-\bar{\rho}\delta_1} + \frac{\bar{\rho}\phi_1\phi_2 ig_t}{(1-\bar{\rho}\delta_1)(1-\bar{\rho}\phi_2)} + \frac{\phi_2 drp_t}{1-\bar{\rho}\phi_2}$. *In constructing the annual dividend strip, we directly interpolate the 12-month maturity. We use an expanding training window to recursively estimate the parameters* ϕ_1, $\bar{\rho}$, δ_1, *and* ϕ_2. *We use these parameters to compute the relevant forecasting variables. We also consider several documented predictors of stock market returns discussed in the text:* dp^{lac}, bm, def, Δff, ep, $infl$, $ntis$, pay, $rrel$, $skew$, sop, $svar$, $tbill$, $term$, *and* vrp. *For each of the aforementioned predictive variables, we estimate a return forecasting regression using all information from the training sample. We then use the estimated parameters together with the most recent observation of the forecasting variable to generate the forecast for the next-period, which we subsequently compare to the realized return. We report the out-of-sample* R^2 (R^2_{oos}) *achieved by each forecasting variable [name in column].* $MSE-F$ *and* $MSE-Adj$ *denote the McCracken (2007) and Clark and West (2007) test statistics, respectively. The critical values of the* $MSE-F$ *statistic are 3.18, 1.55, and 0.80 at the 1 %, 5 %, and 10 % significance levels, respectively. The critical values for the* $MSE-Adj$ *test statistic are 2.33, 1.65, and 1.28 at the 1 %, 5 %, and 10 % significance levels, respectively. ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % significance levels, respectively.*

	dp	dp^{ig}	dp^{corr}	dp^{lac}	bm	def	Δff	ep	$infl$	$ntis$	pay	$rrel$	$skew$	sop	$svar$	$tbill$	$term$	vrp
R^2_{oos}	1.25%	0.06%	1.58%	-2.44%	-0.22%	-3.15%	-0.11%	-0.65%	-0.12%	-3.04%	-1.06%	-0.13%	-1.04%	0.02%	-0.27%	-2.29%	-1.91%	-0.69%
$MSE-F$	1.09*	0.06	1.38*	-2.05	-0.19	-2.62	-0.09	-0.55	-0.11	-2.54	-0.90	-0.12	-0.89	0.02	-0.23	-1.93	-1.61	-0.59
$MSE-Adj$	1.78**	0.43	1.48*	-1.12	-0.09	-2.49	-0.17	-0.66	0.14	-1.13	-1.65	0.00	-0.80	0.18	0.13	-2.07	-1.71	-1.37

Table A.12: The Economic Value of Return Predictability: Alternative Interpolation

This table sheds light on the economic gains of an investor who attempts to exploit the predictability of returns by devising market timing strategies. In constructing the annual dividend strip, we directly interpolate the 12-month maturity. We assume that the investor has a quadratic utility function with a coefficient of relative risk aversion equal to γ. The first column shows the different values of γ, i.e. $\gamma = 2, 4, 6, 8$ or 10. At the end of each month, we compute the optimal allocation of the investor to the risky stock and the riskless asset. These weights depend, among others, on the forecasting model for expected returns and the risk aversion estimate. The investor considers several forecasting variables: dp, dp^{ig}, dp^{corr}, dp^{lac}, bm, def, Δff, ep, $infl$, $ntis$, pay, $rrel$, $skew$, sop, $svar$, $tbill$, $term$, and vrp. Given these weights, we compute the realized return on the portfolio. We do this for each calendar month and return forecasting variable. The first 5 rows report the difference between the annualized certainty equivalent (ΔCE) of the strategy based on the predictive variable [name in column] and that of the strategy based on the recursive mean. The penultimate row sheds light on the difference between the annualized Sharpe ratio (SR) of the strategy based on the predictive variable [name in column] and that of the strategy based on the recursive mean. The last row presents the annualized Sharpe ratio (SR) of the strategy based on the variable [name in column].

γ	dp	dp^{ig}	dp^{corr}	dp^{lac}	bm	def	Δff	ep	$infl$	$ntis$	pay	$rrel$	$skew$	sop	$svar$	$tbill$	$term$	vrp
2	4.21%	1.22%	4.86%	-4.72%	0.44%	-6.79%	-0.01%	-0.39%	-0.24%	-5.17%	-2.58%	0.85%	-1.84%	0.68%	0.03%	-5.03%	-4.95%	-1.52%
4	2.10%	0.60%	2.43%	-2.37%	0.21%	-3.41%	-0.01%	-0.21%	-0.12%	-2.60%	-1.30%	0.41%	-0.90%	0.33%	0.00%	-2.53%	-2.48%	-0.77%
6	1.39%	0.40%	1.62%	-1.59%	0.14%	-2.28%	0.00%	-0.14%	-0.07%	-1.74%	-0.87%	0.27%	-0.59%	0.22%	-0.01%	-1.69%	-1.66%	-0.51%
8	1.04%	0.30%	1.22%	-1.20%	0.10%	-1.72%	0.00%	-0.11%	-0.05%	-1.31%	-0.66%	0.20%	-0.44%	0.16%	-0.01%	-1.27%	-1.25%	-0.39%
10	0.83%	0.24%	0.98%	-0.96%	0.08%	-1.38%	0.00%	-0.10%	-0.04%	-1.06%	-0.53%	0.16%	-0.34%	0.12%	-0.01%	-1.02%	-1.00%	-0.31%
ΔSR	26.63%	11.49%	29.37%	-20.98%	4.77%	-66.86%	1.06%	-1.49%	-2.45%	-23.90%	-21.07%	7.82%	-12.68%	5.75%	2.52%	-40.96%	-46.92%	-11.69%
SR	44.34%	29.19%	47.08%	-3.28%	22.48%	-49.15%	18.76%	16.22%	15.25%	-6.19%	-3.36%	25.52%	5.03%	23.46%	20.23%	-23.25%	-29.22%	6.02%