Growth Expectations, Dividend Yields, and Future Stock Returns *

Zhi Da[†], Ravi Jagannathan[‡], and Jianfeng Shen[§]

February 22, 2015

Abstract

According to the present value relation, the long-run expected return on stocks, *stock yield*, is the sum of the dividend-to-price ratio and a particular weighted average of expected future dividend growth rates. We develop a proxy for *growth* based on sell-side analysts' near-term earnings forecasts to construct *stock yield*. Our *stock yield* measure predicts monthly stock index returns well, with an out-of-sample R-squared that is consistently above 2% during 1999-2012. The forecast performance considerably worsens when both dividend-to-price ratio and *growth* are used as separate explanatory variables without imposing the present value relation constraint. Incorporating *stock yield* as additional information improves the forecasting performance of the van Binsbergen and Koijen (2010) and Kelly and Pruitt (2013) models.

*We thank Binying Liu for excellent research assistance and valuable comments. We also thank Ben Golez, Shane Johnson, Hwagyun Kim, Raymond Liu, Qinghao Mao, Enrique Sentana, Terry Walter, Yu Yuan, participants at 2014 FIRN conference, and seminar participants at Erasmus University, Indian School of Business, Shanghai Advanced Institute of Finance (SAIF), Texas A&M University, and University of Notre Dame. We are grateful to Ken French, Amit Goyal, Robert Shiller, Bhaskaran Swaminathan, and Hao Zhou for sharing their data with us.

[†]University of Notre Dame, 239 Mendoza College of Business, Notre Dame, IN 46556, USA. Email: zda@nd.edu

[‡]Kellogg School of Management, Northwestern University and NBER, ISB and SAIF. Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, USA. Email: rjaganna@kellogg.northwestern.edu

[§]School of Banking and Finance, UNSW Australia, UNSW Sydney, NSW 2052, Australia. Email: jianfeng.shen@unsw.edu.au

1 Introduction

The predictability of stock market returns is of wide interest to both investors and academics. The common practice is to predict stock returns with various valuation ratios such as dividend to price and earnings to price ratios. When all else remains the same, a higher valuation ratio indicates a higher discount rate, that is, a higher expected return.[1]

In general, all else does not remain the same. Economy-wide events that affect expected future returns on stocks may also affect their expected future cash flows. To see this confounding effect, consider the constant growth model in Gordon (1962): $P = D/(R - G)$, where P is the current price; D is the expected one-period-ahead dividend expected to grow at a constant rate of G; and R is the *stock yield*, defined as the discount rate investors use to compute the present value of expected future dividends, i.e., the long run expected return to buying and holding stocks. Rearranging the equation gives $R = D/P + G$. When G changes across different economic regimes, $D/P + G$ will be a better predictor of R than just the dividend yield D/P. This intuition in a stationary economy is made precise by Campbell and Shiller (1988) who derive a dynamic version of the Gordon growth model. In the dynamic case, stock yield is an affine function of the dividend yield and a weighted average of expected future growth rates in dividends.[2] If the dynamics of the term structure of expected stock returns can be explained to a large extent by a single dominant factor, stock yield will help forecast future stock returns at all horizons.[3]

While it is well recognized that combining dividend yield with a proxy of expected cash flow growth will help in predicting stock returns, what that proxy should be has been the subject of debate. Campbell and Shiller (1988), Bansal and Lundblad (2002), Bakshi and Chen (2005), Lettau and Ludvigson (2005), Binsbergen and Koijen (2010), Lacerda and Santa-Clara (2010), Ferreira and Santa-Clara (2011) all use time series models and historical dividend and earnings data to estimate expected future dividend growth rates. More recently, Golez (2014) uses direct dividend growth rate proxies implied in the derivative markets.

In this paper we combine a time series model for earnings growth that uses analysts' near-

[1]See Basu (1983), Fama and Schwert (1977), Campbell and Shiller (1988), Fama and French (1988) among many others. See also Ball (1978) for a general discussion of yield proxies as predictors of future stock returns.

[2]See Jagannathan, McGrattan, and Scherbina (2000) for an alternative derivation of the continuously compounded analogue of this dynamic version of the Gordon growth Model.

[3]See Bakshi and Chen (2005). As an illustration, we derive the equity term structure in Appendix A for the one-factor case.

term forecasts with the dynamic present-value model in Campbell and Shiller (1988) to develop a measure of stock yield. We focus on expected earnings growth rather than expected dividend growth because earnings better reflect the cash flow prospects of a firm than short-term dividends, which are subject to smoothing and other forms of corporate payout policies. In their seminal work, Miller and Modigliani (1961) argue forcefully that dividend policy is irrelevant: stock prices should be driven by "real" behavior – the earnings power of corporate assets and investment policy – and, crucially, not by how the earnings power is distributed. Similarly, Campbell and Shiller (1988) support the use of earnings "since earnings are constructed by accountants with the objective of helping people to evaluate the fundamental worth of a company." In addition, using direct analyst earnings forecasts avoids several econometric issues associated with modeling dividend growth rates that have become highly persistent since the World War II.

Specifically, we compute the expected earnings growth rate as the ratio between analysts' forecasted earnings in the coming calendar year and the recent realized earnings. We focus on analysts' short-term earnings forecasts since a large finance and accounting literature has found equity analysts' value to mostly come from their short-term earnings forecasts rather than their long-term growth forecasts (see Chan, Karceski, and Lakonishok (2003) and Ivkovic and Jegadeesh (2004) among many others). The stock yield is defined as an affine function of the current log dividend-to-price ratio and our log expected earnings growth rate, where the coefficient on the growth expectation is determined by the present-value model under reasonable assumptions. We find the stock yield to do a very good job in predicting future stock market returns during our sample period 1977 – 2012, both in-sample and out-of-sample.

In our sample period 1977–2012, our stock yield predicts future stock market returns with an adjusted R-squared of 13% at the one-year horizon and up to 54% at the four-year horizon. These R^2s compare favorably with other common stock return predictors proposed in the literature.[4]

Welch and Goyal (2008) show that many popular return predictors in the literature do not consistently outperform a simple historical average in predicting next-month market return out-of-sample. They and Campbell and Thompson (2008) propose an out-of-sample R^2 statistic to gauge

[4] We also confirms that the superior in-sample performance of stock yield is robust to alternative variable definitions and extends to different portfolios of U.S. stocks formed on firm characteristics and the market portfolios of other G7 countries. These in-sample results are not reported given our focus on the more important out-of-sample tests.

such relative out-of-sample performance. We find that stock yield produces an out-of-sample R^2 consistently above 2% for monthly forecasts in our sample period. According to the calculation in Campbell and Thompson (2008), an out-of-sample R^2 of 2% translates to return enhancement of 8% per year for a market timer with a risk aversion of 3 who allocates her investment optimally between the stock market and a risk-free asset. In contrast, the dividend-to-price ratio, by itself, has an out-of-sample R^2 averaged below 1% and often dropping below zero. Stock yield also predicts one-year-ahead returns well, with an out-of-sample R^2 of almost 10%. These results do not seem to be driven by biases contained in the analyst forecasts.

Our findings suggest that analyst earnings forecasts reflect the cash flow expectation of a marginal investor. Jagannathan and Silva (2002) show that analyst earnings forecast does a better job explaining stock market return than a time-series model for expected cash flow. Recently, Chen, Da, and Zhao (2013) find that a significant portion of stock market return variation can be explained by cash flow news measured using analyst earnings forecast revisions. In addition to explaining contemporaneous stock return, analyst earnings forecast is also useful for predicting future stock returns, confirming the fundamental intuition underlying the present value relation.

Our stock yield measure is an affine combination of dividend yield and the growth expectation, with the coefficients determined by the present-value relation under reasonable assumptions. While imposing any coefficient restriction can only hurt in-sample prediction, if the coefficient restriction through the present-value relation is applied appropriately, we should expect the in-sample R-squared of regressing future return on stock yield to be close to that of a unrestricted regression in which both dividend yield and growth expectation are used as separate explanatory variables. More importantly, the out-of-sample return prediction performance of stock yield should beat that of the unrestricted regression.

Indeed we are able to confirm these conjectures. When both dividend yield and growth expectation are used as explanatory variables to predict future stock returns in a multivariate regression without any coefficient restriction , the in-sample adjusted R-squared is almost the same as that for stock yield, ranging from 14.2% at the one-year horizon to 55.7% at the four-year horizon, but the out-of-sample R-squared, 1.38% for monthly forecasts and 5.53% for annual forecasts, is much lower than that for stock yield. These findings suggest that dividend yield, expected dividend growth and expected future returns are tied together through the present-value relation, which is appropriately imposed in the construction of stock yield.

Our work is closely related to the literature on *implied cost of* equity *capital* (ICC), another common measure of stock yield (see Claus and Thomas (2001), Pastor, Sinha, and Swaminathan (2008), and Li, Ng, and Swaminathan (2013), among others). The ICC is computed as the discount rate that equates the present value of future cash flows from holding a stock to the stock's price. The stream of future cash flows is forecasted using a combination of short-term analyst earnings forecasts, long-term growth rate forecasts, and historical payout ratios, and other auxiliary assumptions. In contrast, we use only analysts' forecasts of near-term earnings in conjunction with a time series model for expected earnings growth. Our linear specification allows us to evaluate the relative importance of dividend yield and growth expectation in predicting return.

Empirically, our stock yield measure performs as well as the ICC in predicting future stock market returns. This finding suggests that additional assumptions about long-run cash flows embedded in the ICC appear not critical if the only objective is to forecast future stock returns at the aggregate level. The relevant return-predicting signal in the ICC comes largely from the current dividend-to-price ratio and the expected short-term earnings growth rate, information succinctly summarized in our stock yield measure. Conceptually, when there is more than one factor driving the equity term structure [5], both stock yield and the ICC can be viewed as a first-order approximation of the much richer expected return dynamics.

Interestingly, we find that the forecasting ability of the stock yield is concentrated during bad times when investors' fears are high as measured by the Chicago Fed National Activity Index. This is true both in-sample and out-of-sample, which explains the stock yield's high R^2 during our out-of-sample period with two major recessions. This should not be surprising. When the economy is not doing well, the risk premium tends to be high going forward. At the same time, future growth expectation is high when the economy is bottoming out. This higher growth expectation shows up in our stock yield and enables it to capture the increased expected return.

Suppose that the temporal evolution of the term structure of expected returns on stocks is determined to a large extent by a single unobserved dynamic factor. Then that unobserved single factor can be extracted from the cross section of stock yields on a collection of stock portfolios

[5] See Ai, Croce, Diercks, and Li (2013), Kim (2013), Bansal, Kiku, Shaliastovich, and Yaron (2014), Belo, Collin-Dufresne, and Goldstein (2014).

using sophisticated econometric techniques. We consider two such econometric techniques. The first technique is the three-stage partial least squares regression approach developed by Kelly and Pruitt (2013) in which the single market return predicting factor is extracted from the cross section of portfolio-level valuation ratios. Extracting the forecasting factor from the cross section of stock yields delivers even better performance. The out-of-sample R^2s are 2.71% and 19.49%, respectively, for predicting monthly and annual returns by the forecasting factor extracted from the cross section of stock yields on 25 size- and stock-yield-sorted portfolios. These R^2s are better than those from using just the market-level stock yield, confirming that the Kelly and Pruitt (2013) technique is indeed valuable in extracting useful information from the cross-section. In contrast, when we extract the return forecasting factor from the cross section of book-to-market ratios on 25 size- and BM-sorted portfolios, we find it to predict stock market returns with a slightly lower out-of-sample R^2 of 1.47% (14.11%) when predicting next-month (next-year) returns.

The second technique is the Binsbergen and Koijen (2010) Kalman filter approach in which the expected return on stocks and the expected growth rates on dividends are modeled as latent variables. When stock yield is used as an additional noisy observation of the conditional expected return on stocks, the model's ability to forecast future stock returns in real time improves significantly. The out-of-sample R^2 for predicting next-month return increases from 2.7% to 4.2%. Similarly, the out-of-sample R^2 for predicting next-year return increases from 19.50% to 25.66%.

The rest of the paper is organized as follows. In section 2 we derive the expression for stock yield using a dynamic version of the Gordon growth model. In section 3 we describe the data. We compare the forecasting performance of the stock yield with other popular return predictors in section 4, both in-sample and out-of-sample. Section 5 then applies the Kelly and Pruitt (2013) three-stage regression approach to the cross section of stock yields, and section 6 adds the stock yield to the present-value predictive regression model of Binsbergen and Koijen (2010). We conclude in section 7.

2 Stock Yield

When both expected return and dividend growth rate are time-varying, the static Gordon growth model no longer applies. According to the Campbell and Shiller (1988) return decomposition, a dynamic version can be written as:

$$d_t - p_t = -\frac{\kappa}{1-\rho} - \sum_{j=0}^{\infty} \rho^j E_t(\Delta d_{t+1+j}) + \sum_{j=0}^{\infty} \rho^j E_t(r_{t+1+j}), \tag{1}$$

where ρ and κ are loglinear constants (ρ is often chosen to be 0.95 at annual horizon and $\kappa = -log(\rho) - (1-\rho)log(1/\rho - 1) = 0.2$).

Rearrange the equation, take expectations, and define the expected stock yield (sy) as:

$$sy_t = (1-\rho)(\sum_{j=0}^{\infty} \rho^j E_t(r_{t+1+j})) = \kappa + (1-\rho)(d_t - p_t + \sum_{j=0}^{\infty} \rho^j E_t(\Delta d_{t+1+j})). \tag{2}$$

Equation (2) suggests that the current log dividend-to-price ratio, combined with expected dividend growth forecasts, drives expected stock yield, which algebraically measures expected future expected returns. Jagannathan, McGrattan, and Scherbina (2000) derive a similar equation without loglinearization, and they also show that current dividend-to-price ratio plus weighted-average future expected dividend growth rates measures future expected returns.

In what follows, we focus on earnings growth rates:

$$sy_t = \kappa + (1-\rho)(d_t - p_t + \sum_{j=0}^{\infty} \rho^j E_t(\Delta e_{t+1+j})) + \sum_{j=0}^{\infty} \rho^j E_t(\Delta de_{t+1+j})), \tag{3}$$

where de denotes log cash dividend payout ratio. Miller and Modigliani (1961, p.426) summarize the benefit of using earnings as the meaningful measure of cash flows: "We can follow the standard practice of the security analyst and think in terms of price per share, dividends per share, and the rate of growth of dividends per share; or we can think in terms of the total value of the enterprise, total earnings, and the rate of growth of total earnings. Our own preference happens to be for the second approach primarily because certain additional variables of interest – such as dividend policy, leverage, and size of firm – can be incorporated more easily and meaningfully into test equations in which the growth term is the growth of total earnings."

Over the long run, dividend growth should carry similar information as earnings growth for valuation since dividends have to be paid out from earnings eventually. In the short term, however,dividends can be delinked from earnings due to smoothing and other forms of corporate payout policies. In that case, focusing on earnings growth could better capture long-term cash flow growth prospects of a firm which are what matter for valuation. Algebraically, the difference between the two growth rates reflects the discounted sum of future payout adjustments, or:

$$\sum_{j=0}^{\infty}\rho^j E_t(\Delta de_{t+1+j}). \tag{4}$$

Using data on the S&P500 index and assuming an AR(1) process on the payout rate (de), unreported simulation evidence suggests that the discounted sum of future payout adjustments tends to be stable and contributes little to the time-variation in sy.[6] Not surprisingly, our subsequent empirical results suggest that despite omitting this payout adjustment term, sy still does a good job in predicting future stock returns. Consistent with our finding, Ferreira and Santa-Clara (2011) also report that the payout adjustment adds very little to return predictability.

2.1 Time-varying earnings growth

Equation (3) involves the discounted sum of expected future earnings growth rate. The expression for the discounted sum can be simplified under a reasonable assumption that the conditional expected earnings growth, $g_t \equiv E_t(\Delta e_{t+1})$, follows a stationary AR(1) process with the AR(1) coefficient β:

$$g_{t+1} = \alpha + \beta g_t + u_{t+1},\ 0 < \beta < 1,\ u_{t+1} \sim N(0, \sigma_u^2). \tag{5}$$

Similar assumptions are made by Pastor, Sinha, and Swaminathan (2008), Binsbergen and Koijen (2010), Lacerda and Santa-Clara (2010), Ferreira and Santa-Clara (2011), and Golez (2014), among others.

Taking the expected values of future earnings growth rates based on equation (5), and ignoring long-run payout adjustments, the stock yield (sy) in equation (3) could be simplified as:

$$sy_t = \kappa + \frac{\alpha\rho}{1-\rho\beta} + (1-\rho)(dp_t + \frac{g_t}{1-\rho\beta}). \tag{6}$$

Equation (6) shows that conditional expected earnings growth (g_t) is a sufficient statistic for mea-

[6]The discounted sum of future payout adjustments has a standard deviation of less than 0.05, much lower than that of dp which is 0.34.

suring all future expected earnings growth rates.[7] In addition, the dynamics in stock yield are completely driven by the dividend-to-price ratio augmented by a scaled expected earnings growth rate ($dp_t + \frac{g_t}{1-\rho\beta}$).

We compute g_t as $log[E_t(Earn_{t+1})/Earn_t]$, where $E_t(Earn_{t+1})$ is analysts' forecasted earnings for the next 12 months, and $Earn_t$ is the realized earnings in the most recent fiscal year.[8] The annual autocorrelation in g_t, β, is about 0.3 over our sample period. With a ρ of 0.95, the augmented dividend-to-price ratio ($dp_t + \frac{g_t}{1-\rho\beta}$) becomes $dp_t + 1.4g_t$. Our predictive regression results are driven by this augmented dividend-to-price ratio.

While we estimate β using the full sample period, given the short length of the period (1977 to 2012), the exact value of β is not crucial for our results. When we repeat the analysis for a wide range of β from 0.1 to 0.5, we find very similar predictive results. We confirm that g does a good job of predicting future earnings growth rates at both short horizons (next one year) and long horizons (next five years). While it does not predict the next-year dividend growth rate well, possibly due to dividend smoothing, it does predict long-run dividend growth rates well.

Figure 1 plots the log dividend-to-price ratio (dp) against this augmented dividend-to-price ratio in our sample period 1977 – 2012. The difference between two series is due to the expected growth rate ($1.4g_t$), which is considerable and varies a lot over time.

We focus henceforth on the stock yield (sy_t), defined as an affine transformation of the augmented dividend-to-price ratio: $sy = 0.29 + 0.05(dp_t + 1.4g_t)$. The stock yield has the benefit that it can be interpreted directly as a weighted average long-term expected return as in equation (6). The loglinear constant κ is 0.2. In the calculation of α, we use the average realized earnings growth rate to alleviate the optimism bias associated with analyst forecasts. Note that these constant parameter choices do not affect our predictive regression results at all.

Figure 2 plots the stock yield (sy) against the implied cost of capital (ICC) of Li, Ng, and Swaminathan (2013). The ICC is computed as the discount rate that, when plugged into a present-value model, equates the present value of a stock to its current price. The present value model

[7] In the static case with constant payout ratio, $dp_t = d - p$, $g_t = g$, $r_t = r$, and $\beta = 1$. Equation (6), with the constant added back, reduces to $dp + \kappa/(1-\rho) + g/(1-\rho) = r/(1-\rho)$. After algebra manipulation and noting the fact $\rho = 1/(1 + D/P)$, it is easy to show that $log(1 + D/P) + g = r$, consistent with the Gordon Growth Model.

[8] Implicitly, we are ignoring a convexity adjustment term since $g_t = E_t[log(Earn_{t+1}/Earn_t)] = log[E_t(Earn_{t+1})/Earn_t] - 0.5Var_t[Earn_{t+1}/Earn_t]$. Incorporating the convexity adjustment term by measuring the variance in a rolling window hardly changes our results.

considers the firm's expected future cash flows from next year up to infinity, and these expected cash flows are in turn estimated using a combination of analysts' short-term earnings forecasts, long-term growth rate forecasts, and historical payout ratios, under various assumptions. While the details will be provided later, the ICC can be viewed as a more sophisticated version of the stock yield that incorporates more inputs in the present value model. The figure shows that sy and the ICC are indeed highly correlated (with a correlation coefficient of 0.70). sy, however, seems more volatile, especially following the two recent recessions.

2.2 Equity term structure

Our stock yield (sy) and the implied cost of capital (ICC) are similar to the yield on a bond with an infinite maturity. They are conceptually different from the expected next-period return. In our empirical analysis, however, we will use sy to predict stock return in the next one month, one year, and so on. In Appendix A, we work out the algebra for the equity term structure assuming that the next-period expected return follows a stationary AR(1) process (see Pastor, Sinha, and Swaminathan (2008)). In this simple case, the equity term structure is driven by a single factor (the next-period expected return), and the stock yield is sufficient for predicting all finite-period expected returns.

More generally, the equity term structure can be driven by multiple factors as discussed in Binsbergen, Hueskes, Koijen, and Vrugt (2013) and Kim (2013). Stock yield should still be informative about future finite-period expected returns as it can be viewed as a weighted average of these finite-period expected returns.

3 Data and Variable Descriptions

We obtain stock market data such as share price, share outstanding, and return data from the Center for Research in Security Prices (CRSP), accounting data such as common dividends, net income, and book value for common equity from COMPUSTAT, and analysts' earnings forecasts from Institutional Brokers' Estimate System (I/B/E/S). We measure aggregate market returns by CRSP NYSE/AMEX/NASDAQ value-weighted returns including dividends ($VWRETD$). Our inferences regarding forecasting performance remain unchanged if we use returns on the S&P 500 index as an alternative market return measure.

To compare the forecasting performance of our proposed stock yield measure (sy), we also consider various market return predictors proposed in the literature. We first consider the two component variables in the construction of sy, dividend-to-price ratio (dp) and expected earnings growth (g). We then include other traditional valuation ratios, including book-to-market ratio (BM), earnings-to-price ratio (EP), and Shiller's cyclically adjusted price-to-earnings ratio ($CAPE$), and commonly used business cycle variables, including term spread ($Term$), default spread ($Default$), Treasury bill rate ($T - Bill$), Treasury bond yield ($T - Bond$), and variance risk premium measure in in Bollerslev, Tauchen, and Zhou (2009). Finally, we also compare the predictive performance of sy with another closely related predictor, implied cost of capital (ICC) in Li, Ng, and Swaminathan (2013).

3.1 Stock yield, valuation ratios, and business cycle variables

Dividend-to-price ratio (DP) is the value-weighted average of firm-level DPs of all firms in the CRSP-COMPUSTAT combined sample. We calculate a firm's DP at the end of month t by dividing the total common dividends (DVC, data item 21) in the most recent fiscal year (ending at least three months prior) by market capitalization at month-end. We take the natural log of DP to get dp.[9]

Earnings-to-price ratio (EP) is the value-weighted average of firm-level EPs of all firms in the CRSP-COMPUSTAT combined sample. We calculate a firm's EP at the end of month t as dividing the income before extraordinary items (adjusted for common stock equivalents, $IBADJ$, data item 20) in the most recent fiscal year (ending at least three months prior) by market capitalization at month-end. Following da Silva and Jagannathan (2001) and Abarbanell and Lehavy (2007), we do not use the actual earnings from I/B/E/S. According to da Silva and Jagannathan (2001), "the definition of these actual earnings is not uniform across firms or even across time and the coverage of actual earnings is not so wide." We take the natural log of EP.

Forecasted earnings growth (g) is estimated using the forecasted earnings-to-price ratio (FEP)

[9] An alternative way of computing annual dividends is to use the difference between cum-dividend return (ret) and ex-dividend return ($retx$) from CRSP. We have confirmed that the dividend series computed from the two methods are highly correlated with a correlation of 0.98. We follow Li, Ng, and Swaminathan (2013) and use dividends from CRSP-COMPUSTAT combined sample as they align with the fiscal year and therefore earnings. In subsequent section where monthly dividends are needed, we use the difference between ret and $retx$.

and the most recent realized earnings-to-price ratio (EP) of the aggregate market portfolio. To measure a firm's FEP at the end of month t, we divide the IBES consensus earnings-per-share (EPS) forecast for the next 12-month period by the share price at month-end.[10] We then measure the FEP of the aggregate market portfolio as the value-weighted average of firm-level FEPs.[11] g is calculated as the natural log of the ratio of FEP to EP. Stock yield (sy), as defined previously, is calculated as dp plus 1.4 times g.

Book-to-market ratio (BM) is the value-weighted average of firm-level BMs of all firms in the CRSP-COMPUSTAT combined sample. We calculate a firm's BM at the end of month t by dividing the book equity value, calculated following Daniel and Titman (2006), at the end of the most recent fiscal year (ending at least three months prior) by market capitalization at month-end. We take the natural log of BM.

Shiller's cyclically adjusted price-to-earnings ratio ($CAPE$) is defined as price divided by the average of ten years of earnings, adjusted for inflation, and is obtained from Robert Shiller's website.[12]

Term spread ($Term$) is the AAA-rated corporate bond yield minus the one-month T-bill yield, and default spread ($Default$) is the difference between BAA- and AAA-rated corporate bond yields. The one-month T-bill yield is the average yield on the one-month Treasury bill obtained from WRDS, and the yields of both AAA-rated and BAA-rated corporate bonds are obtained from Amit Goyal's website.[13]

T-bill rate ($T-Bill$) is the one-month T-bill rate obtained from Kenneth French's website.[14] And the treasury bond yield ($T-Bond$) is the yield on 30-year Treasury bond obtained from WRDS.

Variance risk premium (vrp) is computed as the difference between the VIX and realized volatility. Details about vrp construction can be found in Bollerslev, Tauchen, and Zhou (2009).

---

[10] We follow Li, Ng, and Swaminathan (2013) to forecast EPS for the next 12-month period, from month t+1 to t+12, as the weighted average of IBES consensus EPS forecasts for the next two fiscal years ending after month t, where the weight of each fiscal year is its fraction in this 12-month period.

[11] We do not use the stock recommendation data from IBES which are subject to forward-looking bias as documented in Ljungqvist, Malloy, and Marston (2009).

[12] http://aida.wss.yale.edu/ shiller/data.htm.

[13] http://www.hec.unil.ch/agoyal/.

[14] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

3.2 Implied cost of capital in Li, Ng, and Swaminathan (2013)

Li, Ng, and Swaminathan (2013) (LNS hereafter) first estimate firm-level ICCs by solving the following finite-horizon dividend discount model:

$$P_t = \sum_{k=1}^{T}\left(\frac{FE_{t+k} \times (1 - b_{t+k})}{(1 + ICC)^k} + \frac{FE_{t+T+1}}{ICC \times (1 + ICC)^T}\right). \tag{7}$$

where P_t is the stock price at the end of month t, FE_{t+k} and b_{t+k} are the forecasts of earnings per share and plowback rate for the year t+k, and $FE_{t+k} \times (1 - b_{t+k})$ is the free cash flow in year t+k. Equation (7) equates the stock price to the present value of future free cash flows up to a terminal period t+T as captured by the first item on the right-hand side of the equation plus the present value of free cash flows beyond the terminal period as measured by the second item. LNS assume a 15-year terminal horizon (T = 15) to estimate equation (7).

To forecast future earnings, LNS first forecast earnings in year t+1 (FE_1) and earnings growth rate in year t+2 (g_2) based on consensus analyst earnings forecasts for the next two fiscal years in I/B/E/S. They then extrapolate earnings growth rates from year t+3 to t+T+1 by assuming that earnings growth rate mean-reverts exponentially from g_2 in year t+2 to a steady growth rate in year t+T+2 which is assumed as the average value of annual nominal GDP growth rate in the rolling window since 1930. Finally, they forecast earnings from year t+2 to t+T+1 on the basis of the earnings in year t+1 and forecasted earnings growth rates.

To forecast future plowback rates, LNS first explicitly forecast the plowback rate for year t+1 as one minus the dividend payout ratio in the most recent fiscal year, and then extrapolate plowback rates from year t+2 to t+T+1 by assuming that plowback rate reverts linearly from year t+1 to a steady value in year t+T+1. The plowback rate in the steady stage since year t+T+1 is computed as the ratio of the steady earnings growth rate to the ICC, which is implied by the sustainable growth rate formula.

Firm-level ICCs could be estimated from stock prices and forecasts of future earnings and plowback rates by solving equation (7). The ICC for the aggregate market portfolio in a month is calculated as the value-weighted average of firm-level ICCs in that month.[15]

[15] See Li, Ng, and Swaminathan (2013) for more details on forecasting future earnings and plowback rates and estimating ICCs at firm-level. LNS measure monthly ICCs of the aggregate market portfolio based on S&P 500 firms, and indicate that the aggregate ICCs based on all firms in the CRSP-COMPUSTAT-IBES combined sample

3.3 Summary statistics

Table 1 provides the summary statistics of for our main predictive variables. Panel A reports the means, standard deviations, and autocorrelations from one-year to four-year horizons. Panel B reports correlations among the variables.

The expected next-year earnings growth based on analysts' earnings forecasts (g) has a mean of 24%, which is clearly higher than the actual average earnings growth rate, confirming the well-known optimism bias in analyst forecasts. In other words, the analyst-based growth forecast is likely a biased estimator of the true expected cash flow growth rate. However, such a bias is unlikely to predict future market returns. g varies a lot from one month to another, with a standard deviation of 24%, which could indicate useful information for predicting future stock returns. The autocorrelation in g is about 0.3 at an annual frequency, resulting in a scaling factor of 1.4 in the definition of the augmented dividend-to-price ratio. The autocorrelations become negative beyond one year, consistent with mean-reversion.

The stock yield variable, sy, has a mean of 12.85% and a standard deviation of 2.35%. As g is not persistent, neither is sy. Indeed, the autocorrelation in sy is 0.55 at the one-year horizon, and drops to below 0.20 beyond one year.

The correlation between g and the log dividend-to-price ratio (dp) is low (-0.05). As a result, the stock yield, is not perfectly correlated with dp (correlation = 0.69). Among other predictive variables, sy is also highly correlated with the ICC (correlation = 0.70) and BM (correlation = 0.68).

3.4 Cash-flow predictive power of g

The validity of our stock yield measures relies on the analyst-forecast-implied growth rate g being a good proxy for future cash-flow growth expectations. We examine the cash-flow predictive power of g directly in Table 2. When we use g to predict future earnings growth rates, we find strong predictive power at all horizons. While g does not predict next-year dividend growth rate well,

are very similar. Though we use the aggregate ICCs provided by LNS, we also estimate monthly aggregate ICCs by an alternative approach, first forecasting future earnings and plowback rates for the aggregate portfolio of all firms, and then solving equation (7) by equating the aggregate market capitalization to the present value of the aggregate future free cash flows. Over our sample period, the correlation coefficient between the LNS measure and this alternative ICC measure is 0.97 if constructed from S&P 500 firms and 0.95 if constructed from the all-firm sample.

possibly due to dividend smoothing, it does predict long-run dividend growth rates well.

4 Return Predictive Power of Stock Yield

We first examine the return predictive power of the stock yield in itself, both in-sample and more importantly out-of-sample. We also conduct extensive robustness checks to make sure that the predictive power of sy is robust.

4.1 In-sample tests

Following Fama and French (1988) and Li, Ng, and Swaminathan (2013), we run multiperiod predictive regressions of the following form:

$$\sum_{k=1}^{K} \frac{r_{t+k}}{K} = a + b \times X_t + u_{t+K,t}, \tag{8}$$

where r_{t+k} is the continuously compounded monthly return. K is the forecasting horizon which we choose at 12, 24, 36, and 48 months. In other words, we run monthly predictive regressions for average monthly stock returns for the next 1 year, 2 years, and up to 4 years. X_t is a vector of predictive variables, and u is the error term.

Given the overlapping nature of the above regression, we first report Newey-West (1987) corrected t-values where the number of lags is equal to $K-1$. Most predictive variables in X are persistent, which, combined with a small sample, can potentially lead to overestimated t-values. To address all three statistical issues (persistent regressors, small sample, and overlapping regressions) simultaneously, we follow Li, Ng, and Swaminathan (2013) closely and also report p-values from simulation exercises. Appendix B describes the simulation exercises. The adjusted R^2s are from the OLS regressions.

We first consider univariate regressions with one predictor at a time, and report the results in Table 3. Consistent with findings in earlier literature, the dividend-to-price ratio does a reasonable job of forecasting future returns in our sample period. The adjusted R-squared ranges from 11.1% at the one-year horizon to 39% at the four-year horizon. The regression coefficients are also highly significant according to the Newey-West corrected t-values. Once its persistence in a small sample is accounted for, however, dp ceases to be significant, according to our p-values that are all higher than 10%.

The analyst-based expected earnings growth g, in itself, is not a strong return predictor, judging from the adjusted R^2s. Even at the four-year horizon, it is only 14.3%. Its return predictive power, while moderate, does seem to be statistically significant at the 10% confidence level beyond two years, according to the simulated p-values.

While neither dp nor g is a strong and significant return predictor, sy, a simple linear combination of the two, does a striking job in forecasting stock market returns. The adjusted R^2 is as high as 13.0% at the one-year horizon, and increases to 53.5% at the four-year horizon. The regression coefficients are highly significant at all horizons according to either Newey-West corrected t-values or simulated p-values. Interestingly, the adjusted R^2 increases while the p-value declines with the forecasting horizon, suggesting that sy is a stronger predictor of long-horizon returns as in equation (2).

Since sy is a linear combination of dp and g, with the coefficients implied by the present-value relation, its in-sample return predictability will not beat dp and g combined in an unrestricted regression for sure. However, its in-sample regression adjusted R^2 is very close to that of regressing future returns on both dp and g, which ranges from 14.2% at the one-year horizon to 55.7% at the four-year horizon. This finding suggests that dividend yield, expected dividend growth and expected future returns are tied together through the present-value relation, which is appropriately imposed in our construction of stock yield.

We consider three other valuation ratios: BM, EP, and $CAPE$. EP is the weakest return predictor of the three, with low adjusted R^2s below 10% and insignificant regression coefficients at all horizons. $CAPE$ is associated with high adjusted R^2s from 9.3% at the one-year horizon to 46.2% at the four-year horizon. As it is highly persistent, its regression coefficients are not significant, judging from the simulated p-values. Finally, BM is probably the best of the three, with reasonable adjusted R^2s from 10.5% at the one-year horizon to 37.2% at the four-year horizon. Its regression coefficients are marginally significant, according to the simulated p-values. Overall, stock yield outperforms all the three valuation ratios by a clear margin.

We also consider four macroeconomic variables: term spread, default spread, the T-bill rate, and the T-bond rate. In general, they perform poorly in predicting future stock market returns, consistent with the findings in the literature. Their R^2s are low. Their regression coefficients are not significant, with the exception of the T-bond rate whose coefficients are marginally significant for horizons beyond one year.

We further consider *vrp*, a recently developed measure of the aggregate uncertainty in the economy. *vrp* is shown to explain the aggregate stock market returns nontrivially in Bollerslev, Tauchen, and Zhou (2009) and Drechsler and Yaron (2011). More recently, Andersen, Fusari, and Todorov (2013) find a tail factor extracted from equity index option prices to have strong predictive power for market excess returns up to a year. We indeed find that *vrp* is able to explain both next one- and two-year returns. While its R^2s are not particularly high, its coefficient is statistically significant, judging from either Newey-West corrected t-values or simulated p-values, due to its close to zero autocorrelations. Since *vrp* starts in 1990, the regressions cover a shorter sample period, from 1990 to 2012.

Finally, we consider a closely related return predictor, the implied cost of capital (ICC) of Li, Ng, and Swaminathan (2013). As we have shown in section 3.2, ICC can be viewed as a more sophisticated version of stock yield that incorporates more inputs in the present value model. Empirically, however, ICC does a slightly weaker job in predicting future stock market returns than our stock yield measure. While ICC has reasonably high adjusted R^2s from 6.9% at one-year horizon to 40.8% at four-year horizon, its regression coefficients are insignificant or only marginally significant.

We should note that Li, Ng, and Swaminathan (2013) consider a slightly different predictive regression specification that forecasts future market excess returns (over the risk-free rate) using excess ICC (over yields of T-bills or T-bonds). We reproduce their results that excess ICC significantly predicts future excess market returns. We do not consider excess stock yield for the benchmark case because theoretically sy needs to be scaled properly before the T-bill or T-bond yield can be subtracted, but untabulated results confirm that sy does a comparable job predicting returns in-sample relative to excess ICC even if excess returns are used everywhere. [16]

Figure 3 provides a visual representation of the results, plotting both ICC and sy against average annual realized stock market returns in the next four years. It seems that stock yield does

[16] The in-sample adjusted R^2s of excess sy are 7.7%, 15.6%, 21.5%, and 20.1% for one- to four-year horizons respectively, while the corresponding numbers are 6.1%, 17.5%, 28.6%, and 29.0% for excess ICC. And both excess sy and excess ICC outperform other predictors.

a better job than ICC in predicting the high stock returns following the two recent recessions.

The finding that sy can outperform a more complicated ICC is consistent with a large finance and accounting literature documenting that the value of equity analysts' forecasts comes mostly from their short-term earnings forecasts rather than their long-term growth forecasts. Furthermore, the steady-state cash flow growth rate tends to be very slow-moving. Indeed, Li, Ng, and Swaminathan (2013) proxy it with the average historical GDP growth rate. The information signal in the time-varying ICC from month to month largely comes from the current dividend-to-price ratio and the expected short-term earnings growth rate, information succinctly summarized in our stock yield.

Given the strong correlation between sy and other return predictors, it is informative to examine the incremental predictive power of sy. We do this by conducting bivariate regressions with both sy and another predictor simultaneously in equation (8). Overall, we find sy continues to be significant even with the presence of any other return predictor. The p-value of the average slope coefficient on sy is almost always below 10%. It is only slightly above 10% when EP or $CAPE$ is included. Other predictors are never significant alongside sy. These results are not tabulated to save space, but available upon requests.

4.2 Out-of-sample tests

Often the predictability of stock returns is taken out-of-sample. Welch and Goyal (2008) show in an out-of-sample experiment that many well-known predicting variables do not consistently outperform the historical mean. We generate our out-of-sample forecast with a sequence of expanding samples. Specifically, we use information up to month m and estimate the following equation:

$$r_t = \alpha + \beta x_{t-1} + u_t, \forall t <= m \tag{9}$$

where r_t is continuously compounded return in month t, x_{t-1} is a predictive variable at month $t-1$, u_t is the residual, and α and β are coefficients. The coefficients, combined with x_m, are then used to estimate the expected return in month-$m+1$.

We follow this process, adding to the sample by one month each time, thereby generating a series of out-of-sample next-month expected return forecasts, $\hat{r}_{m+1}$, $\hat{r}_{m+2}$, ..., $\hat{r}_T$. The out-of-sample

R^2 then compares the mean-squared errors for a specific predicting variable to those when using historical means:

$$OOS - R^2 = 1 - \frac{\sum_{i=m}^{T-1}(r_{i+1} - \hat{r}_{i+1})^2}{\sum_{i=m}^{T-1}(r_{i+1} - \bar{r}_{i+1})^2}, \tag{10}$$

where $\bar{r}_{i+1}$ is the historical mean of returns up to time i. Month $m+1$ is the starting month of the out-of-sample period, and OOS-R^2 therefore measures the performance of a particular predictor relative to the historical mean during that out-of-sample period.

A positive OOS-R^2 means that this predictor outperforms historical mean. In that case, a market timer optimally allocating her assets between the stock market and a risk-free asset can enhance returns by using this predictor (rather than the historical average) to forecast next-month expected stock return. Campbell and Thompson (2008) show that this return enhancement is approximately OOS-R^2/γ where γ measures relative risk aversion. In other words, if an investor's relative risk aversion coefficient is 3, then the return to her market timing portfolio can be improved by OOS-R^2 per quarter. Of course, such a return improvement can come from taking more risk, so we also compute the implied utility gain for a mean-variance investor with a relative risk aversion coefficient of 3.

Equation (2) suggests that stock yield, sy, is a predictor of long-horizon future returns, and in-sample prediction results also indicate a higher predictive power of sy for longer-horizon returns. Hence it is helpful to include predictors from the past, or sy at different lags, to predict the next month's return. We follow Li, Ng, and Swaminathan (2013) to use a two-year backward moving average of sy and ICC, which indeed helps improve predictive power. Consistent with their results, applying such an average to other predictors does not help improve predictive power much.[17]

For the benchmark case, we choose January 1999 as the start of the out-of-sample period. This allows a minimum of 20 years in the in-sample period. [18] Welch and Goyal (2008) show that well-known return predictors perform poorly starting in the late 1990s, so an out-of-sample period starting in 1999 is of particular relevance to examine.

The results are presented in Panel A of Table 4. Consistent with the prior literature, many

[17]For this reason, the backward moving average calculation is only applied to sy and ICC, and only for predicting next-month returns out-of-sample.

[18]Because of the moving average, the training period for sy and ICC is 1979 – 1998.

predictors such as EP and macroeconomic variables have negative out-of-sample R^2s, and thus underperform the simple average historical return. dp, sy, $CAPE$, and ICC are the only four predictors with significant positive out-of-sample R^2s. Among these four predictors, sy has the highest R^2 of 2.1%. sy is also associated with the highest utility gain among all the predictors, 8.38%, indicating that a risk-averse investor is willing to pay more than 8% in fees per year for access to the information embedded in sy.

We also consider a wide range of alternative starting dates for the out-of-sample period from 1999 to 2007. Figure 4 plots the out-of-sample R^2 for sy and dp. It is clear from the plot that the high out-of-sample R^2 for sy is not particular to the 1999 starting point in the benchmark case. In fact, the R^2 for sy is almost always higher than 2% and can be as high as 3.3%. As a comparison, the R^2 for dp is consistently lower and becomes negative for out-of-sample periods starting after 2002.

To gauge whether sy has incremental predictive power that is not present in other predictors, we conduct Harvey, Leybourne, and Newbold (1998) forecast encompassing test. A forecast encompassing test of predictor i against predictor j tests the null hypothesis that the forecast based on predictor i encompasses the forecast based on predictor j, against the one-sided alternative that the forecast based on predictor i does not encompass the forecast based on predictor j. Untabulated results show that, in the tests of sy against other predictors, the p-values are always higher than 0.3, so we cannot reject the null that sy encompasses other predictors. On the other hand, in the tests of other predictors against sy, the p-values are always lower than 0.1, so we can reject the null that sy is encompassed by other predictors. Overall the results suggest that sy is more informative than other predictors in predicting future stock returns.

We also run similar out-of-sample tests to predict next one-year return. Panel B of Table 4 reports the out-of-sample R^2 and associated p-values computed using Newey-West (1985) correction with the number of lags of 11 (the number of overlapping observations in this exercise). We find that stock yield continues to do a good job. Its out-of-sample R^2 is 9.83% with a p-value of 0.02, compared to 5.17% for the ICC with a p-value of 0.06.

Finally, we also check the out-of-sample predictive power of sy and other predictors for market excess returns (in excess of risk-free rates). For consistency, we subtract the yield on T-bills from sy and ICC for this exercise. Confirming the results in Li, Ng, and Swaminathan (2013), ICC

does the best job when predicting next one-month excess market returns with a OOS-R^2 of 1.43%. sy comes in second with a OOS-R^2 of 1.17%. All other predictors have either negative OOS-R^2s or small positive OOS-R^2s that are insignificant. When predicting next one-year excess returns, sy has the highest OOS-R^2 of 8.96%, followed by ICC with an OOS-R^2 of 6.87%.

To conclude, when the forecasting power is tested out-of-sample, only sy and ICC have consistently positive and significant OOS-R^2s in various specifications.

4.2.1 Robustness checks

sy is computed as $0.29+0.05(dp+\text{beta}\times g)$, where dp is the log value of the dividend-to-price ratio, g is the log value of forecasted earnings growth rate, and beta is a constant. In Table 5, we demonstrate the robustness of return predictive power of our stock yield, sy, by: (1) estimating g in various alternative ways; and (2) combining dp and g in various alternative ways. We focus on the out-of-sample predictions. For consistency, we take the 24-month moving average of all alternative stock yield measures when predicting next-month return.

In our first set of robustness checks, we consider various alternative ways of measuring g. As can be seen in Figure 1, the augmented dividend-to-price ratio, $dp+1.4g$, is much more volatile than dp and has a few spikes, especially following the recent two recessions. To make sure that the results are not driven by the extreme values of g, we consider winsorizing g at the 5 and 95 percentile values across all months in our sample period, -0.05 and 0.67 respectively, and then calculate sy accordingly (sy_{wsg}). Winsorizing g produces a very similar out-of-sample R^2 of sy for predicting next-month return, and actually improves the performance of predicting next one-year return, suggesting that our baseline results are not driven by outliers.

In the benchmark definition of sy, we use earnings forecasts on all stocks covered by analysts in IBES, so the resulting expected earnings growth forecast better proxies the cash flow growth prospects of the overall stock market whose returns are predicted. One potential concern is that analysts' forecasts on small stocks could be highly biased, and thus contaminate the overall g measure. As one alternative definition, we consider calculating sy using g (and dp as well) of the S&P500 stocks only. This alternative stock yield measure, sy_{sp}, produces out-of-sample R^2s very similar to those of our baseline sy.

Another potential concern about our calculation of g is that we are matching actual GAAP realized earnings (before extraordinary items) with IBES forecasted earnings, while the accounting definitions of these two earnings are slightly different. IBES compiles forecasted and reported earnings using its proprietary procedures and definitions, with the aim to exclude from the GAAP-based reported earnings certain nonrecurring items (such as one-time charges or gains associated with acquisitions), other special items, and nonoperating items.[19] To address this concern of a mismatch between expected future earnings and realized past earnings, we consider an alternative definition of sy that uses g computed based on realized and forecasted earnings, both obtained from IBES. This alternative sy, labeled as sy_{ibes}, produces smaller but still significant R^2s, especially when predicting next one-year return.

One important concern with using analyst forecasts is that they are likely to be biased on average. While it is generally believed that the biases associated with short-term earnings forecasts are small, they may still affect our analysis. It is important to note, however, that since we are running predictive regressions, the level of the bias is less relevant. These biases, as long as they persist from one month to another, should not have a strong impact on our monthly predictive regressions. To confirm this, we define two alternative sys in which we compute g not using the IBES median earnings forecasts but rather the lowest or the highest forecasts across different analysts. We label them as sy_{low} and sy_{high}, respectively. If analyst bias has any significant impact on our analysis, we would expect sy_{low} and sy_{high} to produce very different results. We actually find the opposite – the out-of-sample R^2s using sy_{low} and sy_{high} are very similar, and are in fact also similar to those using our benchmark sy.

We also consider two alternative ways to define g using analyst forecasts. First, we compute the expected earnings growth rate as the one implied by analysts' forecasts of earnings for the two fiscal years ending in the next two years. We call the resulting growth rate g_2. Second, we simply use analysts' median long-term growth (ltg) forecasts (available from 1982). sy calculated from these two alternative definitions still predict stock market returns well. However, we find that these two alternative sys perform more poorly than our benchmark sy, especially when predicting next-month return, consistent with the notion that equity analysts are best at forecasting near-term

[19]There are many studies in the accounting literature examining the difference between GAAP-based earnings and "street" earnings produced by commercial forecast data providers and the consequent implications for earnings-related analyses. See Bradshaw and Sloan (2002) and Abarbanell and Lehavy (2007), among others, for related discussions.

earnings.

Finally, recall that g_t is computed as $log[FE_{t+1}/E_t]$, where FE_{t+1} is analyst-forecasted earnings for the next 12 months, and E_t is the realized earnings for the most recent fiscal year. One might be concerned that it is the backward-looking realized earnings (E_t) that drive the return predictive power in g_t. We argue that this is unlikely for two reasons. First, g_2 computed using only forward-looking forecasts still produces significant forecasting results. Second, we decompose $g = log[FE_{t+1}/E_t]$ into two parts: $g = log[FE_{t+1}/E_t^{10}] + log[E_t^{10}/E_t]$, where E_t^{10} measures average earnings in the past 10 years. Untabulated results confirm that the predictive power in g is due mostly to the variation in forecasted earnings in the first term, rather than that in realized earnings in the second term.

In our second set of robustness checks, we consider various alternative ways of combining dp and g to construct sy. Firstly, we take a beta value of 1.4 in the benchmark definition of sy, based on the empirical autocorrelation in g over the full sample period. So we are subject to a criticism of potential forward-looking bias. To ensure that our results are not driven by this particular choice of the scaling factor, we also consider alternative constant values for beta from 1.1 to 1.9, corresponding to an autocorrelation in g from 0.1 to 0.5, respectively. We find that sys computed with these different constant values of beta still predict future stock market returns significantly, with out-of-sample R^2 generally higher for lower values of beta.

Secondly, we construct sy as a linear combination of dp and g, with a constant scaling factor derived from the present-value relation. An alternative way to predict returns with dp and g is to include both of them as separate explanatory variables in the regression. For example, to predict return in month $t + 1$, we first regress monthly returns on dp and g using data up to month t, and then calculate the predicted return with the estimated coefficients on dp and g and their values in month t. This is equivalent to defining sy as dp+beta_ols$\times g$, where beta_ols is the ratio of the coefficient on g to that on dp determined in monthly in-sample OLS regressions and varies across months. While predicting future stock returns with both dp and g still produces significant out-of-sample R^2s, the return predictive power is much weaker than our baseline sy.

We further consider estimating the coefficients on dp and g in the bivariate regressions via restricted least-square method, restricting the ratio of the coefficient on g to that on dp between 1.1 and 1.9, and then predict future returns out of sample with the estimated coefficients. This is equivalent to defining sy as dp+beta_rls$\times g$, where beta_rls is the ratio of the coefficient on g to that on dp as determined in the monthly restricted-least-square regressions. According to the present-value relation, future expected returns, in a weighted-sum form, are a linear function of dp and g, assuming an AR(1) structure of expected earnings growth. The out-of-sample return prediction performance of dp and g should improve by imposing some reasonable restrictions on the ratio of their coefficients, which in term puts restrictions on autocorrelation in expected earnings growth.

Indeed, out-of-sample R^2 improves significantly with these restrictions, increasing from 1.38% (5.53%) to 1.90% (9.85%) for predicting next-month (next-year) return. The finding is consistent with the observation in Cochrane (2008) that imposing economically sensible restrictions help improve out-of-sample prediction of returns. Interestingly, with the aforementioned coefficient restrictions, the out-of-sample return prediction performance of dp and g is very close to that of sy, which combines dp and g linearly with a constant scaling factor derived from the present-value relation. These findings are consistent with the finding in in-sample return predictions that sy performs almost as well as dp and g combined, and reinforce that dividend yield, expected dividend growth and expected future returns are tied together through the present-value relation, which is appropriately imposed in our construction of stock yield.

4.3 Understanding the return predictability of stock yield

What is the source of the return predictability in stock yield? A quick look at Figure 3 seems to suggest that stock yield does a particularly good job in predicting the high stock returns following recessions.

We confirm this observation more formally by conducting our main analysis separately for "good times" and "bad times" categorized according to the Chicago Fed National Activity Index (CFNAI), which measures overall economic activity.We assign all calendar months in our sample period into one of two groups based on the CFNAI in that month: "good times" are those with above-median CFNAIs, and "bad times" are those with below-median CFNAIs.[20]

Figure 5 shows the results from in-sample tests that use stock yield to predict next-year stock returns in good times and bad times separately. We plot realized stock returns against the predicted

[20]We obtain similar results when we use the real-time recession probability as developed by Guan and Parker (2014) to classify "good" and "bad" times.

stock returns for the two groups. It is clear that the predictive power of stock yield is concentrated during bad times when investors' fears are high as measured by a low CFNAI index. During bad times, stock yield has an in-sample R^2 of 26%, much higher than during good times.

This is true for out-of-sample predictions as well. During bad times, the out-of-sample R^2 associated with stock yield is 3.48% (14.69%) when predicting next one-month (one-year) returns, much higher than during good times. This is not surprising, as our out-of-sample period 1999 – 2012 includes two major recessions. Intuitively, when the economy is not doing well, the risk premium tends to be high looking ahead. At the same time, future growth expectation is high as the economy is bottoming out. Such higher growth expectations show up in our stock yield and enables it to capture the increased expected returns. In addition, while analysts' earnings forecasts tend to be optimistic on average, the optimism bias is alleviated during bad times, making the forecasted growth more accurate.[21]

Our finding is also consistent with the intuition in Boyd, Hu, and Jagannathan (2005), who argue that information about future corporate dividends dominates during contraction, and such information needs to be accounted for when predicting future stock returns. The fact that stock yield has stronger predictive power in bad times is also consistent with the recent finding in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014) that skillful mutual fund managers are able to time the market during recessions.

To conclude this section, we find that the stock yield measure does a good job predicting stock returns by itself. Its out-of-sample R^2s are consistently above 2% when predicting monthly returns over various out-of-sample periods. It also predicts next one-year market returns, both raw and excess ones, very well out of sample. In-sample, stock yield predicts future stock market returns with an adjusted R^2 of 13% at the one-year horizon, up to 54% at the four-year horizon. Overall, the performance of stock yield is comparable, if not slightly better than some of the best return predictive variables documented so far in the literature.

In the next two sections, we further show that the stock yield can be easily combined with sophisticated econometric techniques recently developed by Kelly and Pruitt (2013) and Binsbergen and Koijen (2010) to achieve even better return predictive power. For these analyses, we focus on

[21] The correlation between the absolute forecast error in g and CFNAI is 0.28, suggesting forecasted growth is more accurate in bad times.

the out-of-sample prediction.

5 Extracting Factor from the Cross-Section of Stock Yields

Kelly and Pruitt (2013) propose a three-stage regression approach to extract a single market return predictor from the cross-section of portfolio-level predictive variables, such as the book-to-market ratios of 25 Fama-French portfolios. Denoting $r_{t,h}$ the h-period market return after period t and $x_{j,t}$ the predictive variable at period t for the portfolio j, the first step runs a time-series regression of the predictive variable, $x_{j,t}$ on the future market return to forecast, $r_{t,h}$, for each portfolio j separately:

$$x_{j,t} = \hat{\phi}_{j,0} + \hat{\phi}_j r_{t,h} + e_{j,t}. \tag{11}$$

The second step runs a cross-sectional regression of all portfolios' predictive variables on their loadings on the market return estimated in the first-stage regression, for each period t:

$$x_{j,t} = \hat{c}_t + \hat{F}_t \hat{\phi}_j + u_{j,t}. \tag{12}$$

The regression coefficient, $\hat{F}_t$, is the predictive factor for the future market return at period t.

To generate predicted future market returns, the third step regresses future market returns on the lagged predictive factor estimated in the second-stage regression to generate predicted future market returns:

$$r_{t,h} = \hat{\beta}_0 + \hat{\beta}\hat{F}_t + \varepsilon_t. \tag{13}$$

Specifically, to generate the out-of-sample prediction for $r_{t,h}$, the h-period market return over the period t+1 to t+h, we first run the regression in equation (13) with data up to period t, i.e., the last h-period market return entering the regression ends in period t, and the corresponding predictor is in period $t-h$, $\hat{F}_{t-h}$. Then the out-of-sample predicted market return is the product of the regression coefficient, $\hat{\beta}$, and the predictor in period t, $\hat{F}_t$.

We adopt this three-stage regression approach to extract a single market return predictor from the cross-section of portfolio-level stock yields, sy. To ensure sufficient variation in sy, we form size-sy portfolios based on a two-way conditional sort of stocks, first by market capitalization and then by sy. More specifically, to form 6 (25) size-sy portfolios, in each month t we first sort stocks into 2 (5) equally sized groups based on market capitalization at the end of month t, and within each market capitalization group, we further sort stocks into 3 (5) equally sized groups based on sy. We use only stocks with non missing market capitalization and sy. We then calculate the portfolio-level sy and run the three-stage regressions above to generate market return predictor (KP-sy).[22]

For comparison, we apply a similar procedure to form size-dp portfolios based on the two-way conditional sort on market capitalization and log dividend-to-price ratio, dp, and extract a single market return predictor (KP-dp) from the dps of these portfolios. We also replicate the original Kelly and Pruitt (2013) prediction, which uses log book-to-market ratios of Fama-French portfolios since 1930 to extract a single return predictor (KP-BM).[23]

The out-of-sample forecasting performance of these cross-sectionally extracted factors is presented in Table 6. We find that the factor extracted from the cross-section of sy (KP-sy) performs better than the market-level sy. For example, when predicting the next one-month stock return, the out-of-sample R^2 is 2.70% (2.71%) for KP-sy extracted from 6 (25) portfolios. Both R^2s are higher than the 2.10% from using market-level sy (reproduced from Panel A of Table 4). Similarly, when predicting the next-one-year stock return, KP-sy extracted from 6 (25) portfolios has an out-of-sample R^2 of 11.66% (19.49%), higher than the 9.83% from using market-level sy. Likewise, we find KP-dp and KP-BM extracted from the cross-section to outperform the market-level dp and BM, respectively, in general. These results confirm the intuition in Kelly and Pruitt (2013) that information from the cross-section can enhance return predictability.

Finally, we also find KP-sy generally does better than both KP-dp and KP-BM in predicting future stock market returns. For example, when predicting next one-month return by the factors extracted from 25 portfolios, KP-dp and KP-BM both have an out-of-sample R^2 of 1.47%, significantly lower than that of KP-sy. A similar pattern is observed for predicting next one-year return when extracting factors from 25 portfolios.

Overall, these results suggest that the three-stage regression approach developed by Kelly and

[22]For one-month market return predictions, as in the univariate prediction, we use a two-year backward moving average of sy to form portfolios and calculate portfolio-level sy as the value-weighted average of stock-level sy.

[23]We also try to use log book-to-market ratios of Fama-French portfolios from 1977, the starting point of our sample period, to extract a single return predictor. The predictive performance is actually worse. To maintain consistency with the original KP prediction, we use the data from 1930 for this prediction.

Pruitt (2013) can be applied to the cross-section of stock yields to produce an even stronger stock return predictor.

6 Extending the Present Value Approach with Stock Yield

Binsbergen and Koijen (2010) consider a time-series model that satisfies the present value relation in which conditional expected returns and expected dividend growth rates are modeled as latent variables. They show that this approach significantly improves the predictability of both future return and dividend growth. Specifically, they assume AR(1) processes for the two latent state variables: the expected return (μ) and the expected dividend growth (g):

$$\begin{aligned} \mu_{t+1} &= \delta_0 + \delta_1 \left(\mu_t - \delta_0\right) + \epsilon^{\mu}_{t+1} \\ g_{t+1} &= \gamma_0 + \gamma_1 \left(g_t - \gamma_0\right) + \epsilon^{g}_{t+1}. \end{aligned} \tag{14}$$

The de-meaned state variables are:

$$\begin{aligned} \hat{\mu}_t &= \mu_t - \delta_0 \\ \hat{g}_t &= g_t - \gamma_0. \end{aligned}$$

Using the parameters in the above AR(1) processes, equation (1) can be written as:

$$\begin{aligned} dp_t &= -\frac{\kappa}{1-\rho} - \frac{\gamma_0 - \delta_0}{1-\rho} + \frac{\mu_t - \delta_0}{1-\rho\delta_1} - \frac{g_t - \gamma_0}{1-\rho\gamma_1} \\ &= A + B_1 \cdot \hat{\mu}_t - B_2 \cdot \hat{g}_t. \end{aligned} \tag{15}$$

The baseline state-space representation, used by Binsbergen and Koijen (2010), is defined by two measurement equations:

$$dp_t = A + B_1 \cdot \hat{\mu}_t - B_2 \cdot \hat{g}_t,$$

$$\Delta d_{t+1} = \gamma_0 + \hat{g}_t + \epsilon^{d}_{t+1}.$$

The two transition equations are:

$$\hat{\mu}_{t+1} = \delta_1 \cdot \hat{\mu}_t + \epsilon^{\mu}_{t+1},$$

$$\hat{g}_{t+1} = \gamma_1 \cdot \hat{g}_t + \epsilon^g_{t+1}.$$

We have shown the stock yield (sy) to be a good measure of expected stock return, which allows us to add another measurement equation:

$$sy_t = a + b \cdot \hat{\mu}_t.$$

We then evaluate the out-of-sample performance of the original Binsbergen and Koijen (2010) model (original BK) and its extension that includes sy (extended BK) using monthly data.

We use non-overlapping monthly return, dividend growth and dividend-to-price ratio. Monthly dividends are computed from CRSP returns using the difference between the total monthly return and the monthly return without dividends. The dividend-to-price ratio for month t is computed as the ratio between the dividend in that month and the price at the end of the month. We do not apply seasonal adjustment to the monthly dividend data so the present value relation holds every month. sy is computed as the two-year backward moving average as before.

To be consistent with our earlier out-of-sample analysis, we start our out-of-sample period in January 1999. Specifically, we use data before 1999 to estimate both models, and compute expected returns for January 1999. Next, we re-estimate both models using data up to January 1999 and compute expected returns for February 1999. We then repeat this process going forward. This procedure allows us to generate a time series of monthly expected returns under both models from January 1999 through December 2012. We compare them to the realized returns and compute the out-of-sample R^2 as before. Parallel to Figure 4, we also consider a wide range of starting dates for the out-of-sample period from 1999 through 2007, and plot the out-of-sample R^2s for both models in Figure 6.

The original BK model generates an out-of-sample R^2 of 2.71% when we start the out-of-sample period in 1999, even better than the performance of sy as reported in Panel A of Table 4 (2.10%). More interestingly, incorporating stock yield (sy) in the extended BK model improves the out-of-sample R^2 further to 4.22%! Indeed, the extend BK model outperforms the original BK model for a wide range of out-of-sample cutoff points.

Similarly, when predicting next-one year returns, the original BK model generates an out-of-sample R^2 of 19.50% for the out-of-sample period starting in 1999. Incorporating stock yield (sy) in the extended BK model improves the out-of-sample R^2 further to 25.66%![24]

7 Conclusion

According to the Gordon growth model, the long horizon expected return on stock (stock yield) is the sum of the dividend yield and some weighted average of expected future growth rates in dividends. We show how to construct a measure of the weighted average of expected future growth rate in dividends based on sell side analysts' near term earnings forecasts. During 1977-2012, the stock yield measure constructed in this manner predicts future stock returns as well as other variables that have been proposed in the literature for predicting stock returns. The forecasts perform better during times when investors' recession fears are high.

Our work links the insights from earlier literature (Campbell and Shiller (1988), among others) with more recent studies that examine the implied cost of equity capital, ICC (e.g., Li, Ng, and Swaminathan (2013)). Confirming the insight in Campbell and Shiller (1988), we find augmenting the valuation ratio with a measure of growth expectation significantly improves stock return predictability. Consistent with the literature on the ICC, we find that forward-looking growth expectations embedded in analysts' forecasts, especially their near-term earnings forecasts, are useful in predicting stock returns. The additional assumptions in the ICC models, such as those regarding payout ratios, the time for the growth rate to reach a steady-state one, and the way the growth rate fades, appear not critical if the objective is to forecast future stock returns.

Finally, when combined with recently developed econometric techniques, stock yield improves their performance in predicting future stock returns. We give two examples in the paper: (1) the three-stage regression approach of Kelly and Pruitt (2013), and (2) the present-value model of Binsbergen and Koijen (2010).

[24] To predict next-one year returns, we continue to use non-overlapping monthly returns, dividend growth and dividend-to-price ratio so the present value relation holds. The resulting next-one-month return forecasts, after simple annualization, become our next-one-year return forecasts.

References

Abarbanell, J., and R. Lehavy. 2007. Letting the "Tail Wag the Dog": The Debate over GAAP versus Street Earnings Revisited. *Contemporary Accounting Research* 24:675–723.

Ai, H., M. Croce, A. Diercks, and K. Li. 2013. Production-Based Term Structure of Equity Returns. *Working Paper*, UNC Chapel Hill.

Andersen, T., N. Fusari, and V. Todorov. 2013. The Risk Premia Embedded in Index Options. *Working Paper*, Northwestern University.

Bakshi, G., and Z. Chen. 2005. Stock Valuation in Dynamic Economies. *Journal of Financial Markets* 8:111–151.

Ball, R. 1978. Anomalies in Relationships between Securities' Yields and Yield-surrogates. *Journal of Financial Economics* 6:103–126.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron. 2014. Volatility, the Macroeconomy, and Asset Prices. *Journal of Finance* Forthcoming.

Bansal, R., and C. Lundblad. 2002. Market Efficiency, Asset Returns, and the Size of the Risk Premium in Global Equity Markets. *Journal of Econometrics* 109:195–237.

Basu, S. 1983. The Relationship between Earnings' Yield, Market Value and Return for NYSE Common Stocks: Further Evidence. *Journal of Financial Economics* 12:129–156.

Belo, F., P. Collin-Dufresne, and R. Goldstein. 2014. Dividend Dynamics and the Term Structure of Dividend Strips. *Journal of Finance* Forthcoming.

Binsbergen, J.V., W. Hueskes, R. Koijen, and E. Vrugt. 2013. Equity Yields. *Journal of Financial Economics* 110:503–519.

Binsbergen, J.V., and R. Koijen. 2010. Predictive Regressions: A Present-Value Approach. *Journal of Finance* 65:1439–71.

Bollerslev, T., G. Tauchen, and H. Zhou. 2009. Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies* 22:4463–4492.

Boyd, J., J. Hu, and R. Jagannathan. 2005. The Stock Market's Reaction to Unemployment News: Why Bad News Is Usually Good for Stocks. *Journal of Finance* 60:649–672.

Bradshaw, M., and R. Sloan. 2002. GAAP versus the Street: An Empirical Assessment of Two Alternative Definitions of Earnings. *Journal of Accounting Research* 40:41–66.

Campbell, J., and R. Shiller. 1988. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies* 1:195–228.

Campbell, J., and S. Thompson. 2008. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies* 21:1509–1531.

Chan, L., J. Karceski, and J. Lakonishok. 2003. The Level and Persistence of Growth Rates. *Journal of Finance* 58:643–684.

Chen, L., Z. Da, and X. Zhao. 2013. What Drives Stock Price Movement? *Review of Financial Studies* 26:841–876.

Clark, T., and K. West. 2007. Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. *Journal of Econometrics* 138:291–311.

Claus, J., and J. Thomas. 2001. Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets. *Journal of Finance* 56:1629–1666.

Cochrane, J. 2008. The Dog That Did Not Bark: A Defense of Return Predictability. *Review of Financial Studies* 21:1533–1575.

da Silva, A., and R. Jagannathan. 2001. Forecasting Earnings and Dividends on Benchmark Stock Portfolios. *Working paper*, Northwestern University.

Daniel, K., and S. Titman. 2006. Market Reactions to Tangible and Intangible Information. *Journal of Finance* 61:1605–1643.

Diether, K., C. Malloy, and A. Scherbina. 2002. Differences of Opinion and the Cross Section of Stock Returns. *Journal of Finance* 57:2113–2141.

Drechsler, I., and A. Yaron. 2011. What's Vol Got to Do with It. *Review of Financial Studies* 24:1–45.

Fama, E., and K. French. 1988. Dividend Yields and Expected Stock Returns. *Journal of Financial Economics* 22:3–25.

Fama, E., and K. French. 1997. Industry Costs of Equity. *Journal of Financial Economics* 43:153–193.

Fama, E., and G. Schwert. 1977. Asset Returns and Inflation. *Journal of Financial Economics* 5:115–146.

Ferreira, M. A., and P. Santa-Clara. 2011. Forecasting Stock Market Returns: The Sum of the Parts is More than the Whole. *Journal of Financial Economics* 100:514–537.

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

Gordon, M. 1962. *The Investment, Financing, and Valuation of the Corporation* (Homewood, IL: R. D. Irwin).

Guan, J., and J. Parker. 2014. Predicting the Equity Premium using Macroeconomic Data. *Working Paper*, Georgia State University.

Harvey, D.I., S.J. Leybourne, and P. Newbold. 1998. Tests for Forecast Encompassing. *Journal of Business & Economic Statistics* 16:254–259.

Ivkovic, Z., and N. Jegadeesh. 2004. The Timing and Value of Forecast and Recommendation Revisions *Journal of Financial Economics* 73:433–463.

Jagannathan, R., E. McGrattan, and A. Scherbina. 2000. The Declining U.S. Equity Premium. *Quarterly Review, Federal Reserve Bank of Minneapolis* 24:3–19.

Jagannathan, R., and A. B. Silva. 2002. Stock Yields and the Decomposition of Actual Returns on Benchmark Stock Index Portfolios. *Working paper*, Northwestern University.

Kacperczyk M., S. Van Nieuwerburgh, and L. Veldkamp. 2014. Time-Varying Fund Manager Skill. *Journal of Finance* 69:1455–1484.

Kelly, B., and S. Pruitt. 2013. Market Expectations in the Cross-Section of Present Values. *Journal of Finance* 68:1721–1756.

Kim, S. 2013. Asset Prices in Turbulent Markets with Rare Disasters. *Working Paper*, Georgia Institute of Technology.

Lacerda, F., and P. Santa-Clara. 2010. Forecasting Dividend Growth to Better Predict Returns. *Working Paper*, University of Chicago and Universidade Nova de Lisboa.

Lettau, M., and S. Ludvigson. 2005. Expected Returns and Expected Dividend Growth? *Journal of Financial Economics* 76:583–626.

Li, Y., D. Ng, and B. Swaminathan. 2013. Predicting Market Returns Using Aggregate Implied Cost of Capital, *Journal of Financial Economics* 110:419–436.

Ljungqvist, A., C. Malloy, and F. Marston. 2009. Rewriting History. *Journal of Finance* 64:1935–1960.

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

Newey, W., and K. West. 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55:703–708.

Pastor, L., M. Sinha, and B. Swaminathan. 2008. Estimating the Intertemporal Risk-Return Tradeoff Using the Implied Cost of Capital. *Journal of Finance* 63:2859–2897.

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

Appendix A: Term structure of equity returns

Following Pastor, Sinha, and Swaminathan (2008), we assume that the conditional expected return, $\mu_t \equiv E_t(r_{t+1})$, follows a stationary AR(1) process:

$$\mu_{t+1} = a + b\mu_t + u_{t+1},\ 0 < b < 1,\ u_{t+1} \sim N(0, \sigma_u^2).$$

From equation (2), it is easy to show that:

$$sy_t = \frac{a}{(1-b)(1-\rho)} + \left(\mu_t - \frac{a}{1-b}\right)\frac{1}{1-\rho b}.$$

In other words, sy is an affine transformation of the next-period expected return, μ_t. μ turns out to be the one factor that drives the entire equity term structure in this simple case. For example, the next-two-period expected return per period is:

$$\begin{aligned} er_2 &= \frac{E_t(r_{t+1} + r_{t+2})}{2} \\ &= \frac{\mu_t + E_t(\mu_{t+1})}{2} \\ &= \frac{a + (1+b)\mu_t}{2}. \end{aligned}$$

In general, the n-period expected return per period is:

$$er_n = \frac{a}{1-b} + \frac{1}{n}\frac{1-b^n}{1-b}\left(\mu_t - \frac{a}{1-b}\right).$$

When $n \to \infty$, $er_n \to \frac{a}{1-b}$, or the unconditional expected return. Importantly, these future expected returns are all functions of μ and can therefore be predicted by sy.

Appendix B: Simulated p-values

We illustrate our simulation procedure using a bivariate regression where the predictive variables are sy and BM.

Define a 3×1 column vector $Z_t = [r_t, sy_t, BM_t]'$. We first estimate a first-order VAR: $Z_{t+1} = A_0 + A_1 Z_t + u_{t+1}$. We impose the null hypothesis of no return predictability by setting the slope coefficients of the r_t equation to zero and the intercept of the equation to the empirical mean of r_t. The fitted VAR is then used to generate T observations of the simulated variables $[r_t, sy_t, BM_t]'$. The initial observations are drawn from a multivariate normal distribution of the three variables with mean and covariance matrix set to their empirical counterparts. Once the initial observations are chosen, the subsequent T-1 simulated observations are generated from the fitted VAR with the shocks bootstrapped from the actual VAR residuals (sampling without replacement). These simulated data are then used to run a bivariate return predictive regression to produce regression coefficients.

We repeat the process 50,000 times to obtain the empirical distribution of the regression coefficients (under the null of no predictability), which in turn produce the p-values associated with our actual estimated coefficients.

Figure 1: **Dividend-to-Price Ratio (dp) *vs.* Augmented Dividend-to-Price Ratio ($dp + 1.4g$)**

The figure plots the log dividend-to-price ratio (dp) against the augmented ratio ($dp + 1.4g$) at the market level, where g is the log value of forecasted earnings growth rate. The sample period is January 1977 through December 2012.

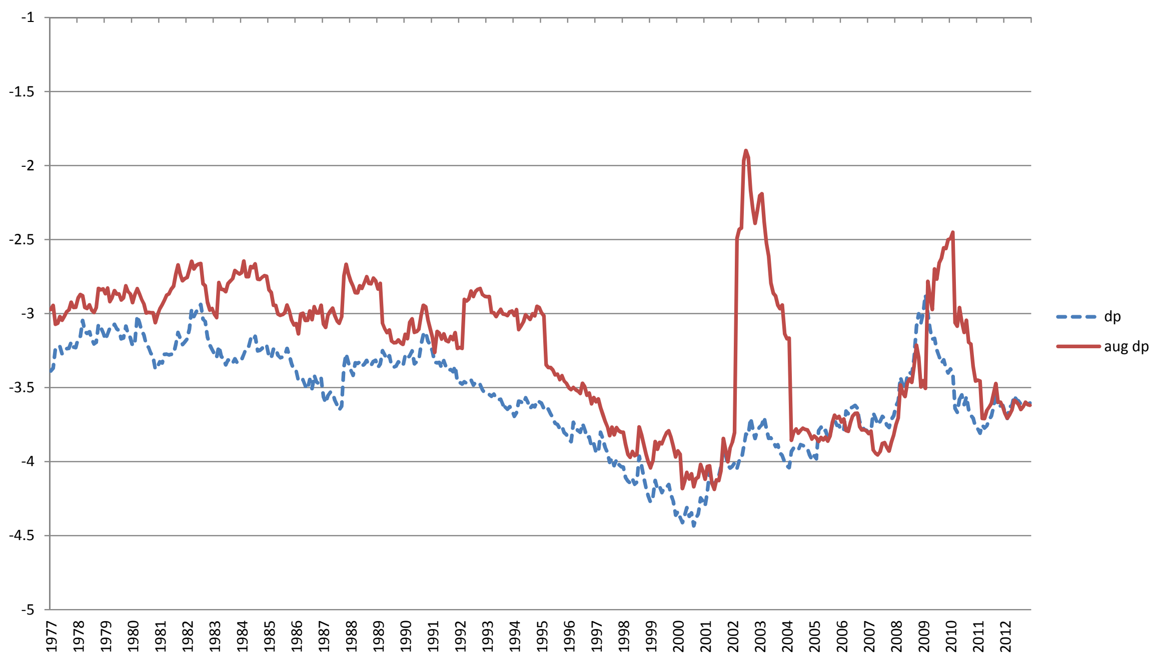

Figure 2: **Implied Cost of Capital (*ICC*) *vs.* Stock Yield (*sy*)**

The figure plots the implied cost of capital (*ICC*) against the stock yield (*sy*) at the market level. *sy* is the stock yield measure, calculated as 0.29+0.05(dp+1.4g), where dp is the log value of the dividend-to-price ratio and g is the log value of forecasted earnings growth rate. *ICC* is the continuously compounded implied cost of capital. The sample period is January 1977 through December 2012.

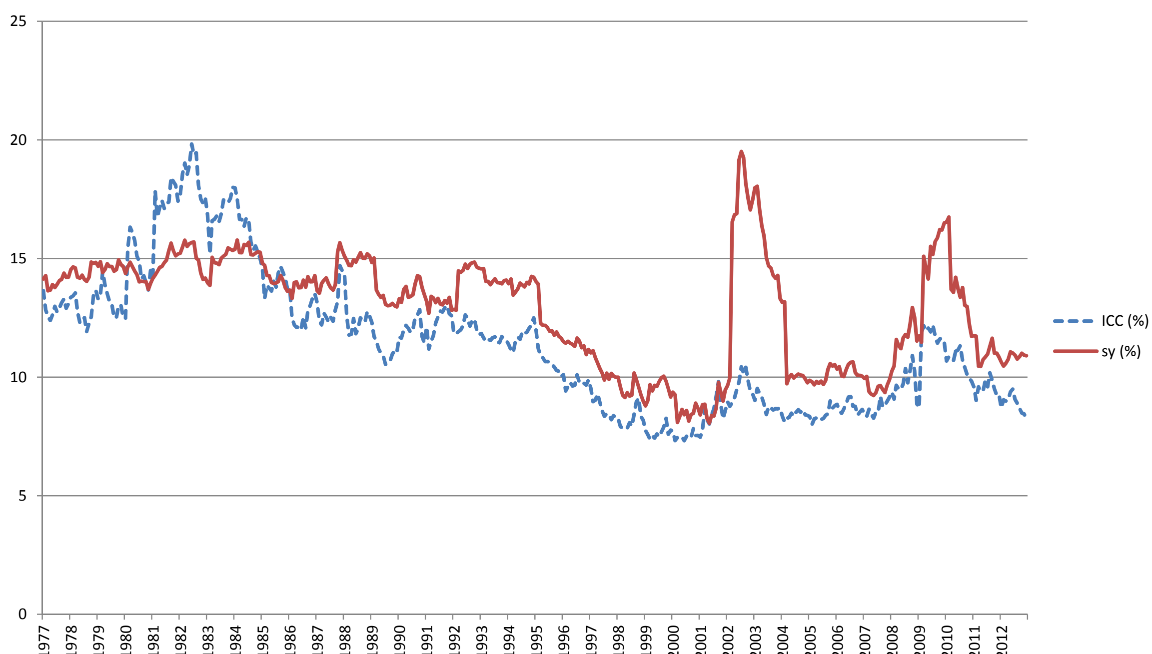

Figure 3: **Stock Yield (sy), Implied Cost of Capital (ICC), and Average Return in the Next Four Years**

The figure plots the stock yield (sy), the implied cost of capital (ICC), and the average monthly continuously compounded return in the next four years at the market level. sy is the stock yield measure, calculated as $0.29{+}0.05(dp{+}1.4g)$, where dp is the log value of the dividend-to-price ratio and g is the log value of forecasted earnings growth rate. ICC is the continuously compounded implied cost of capital. The sample period is January 1977 through December 2012.

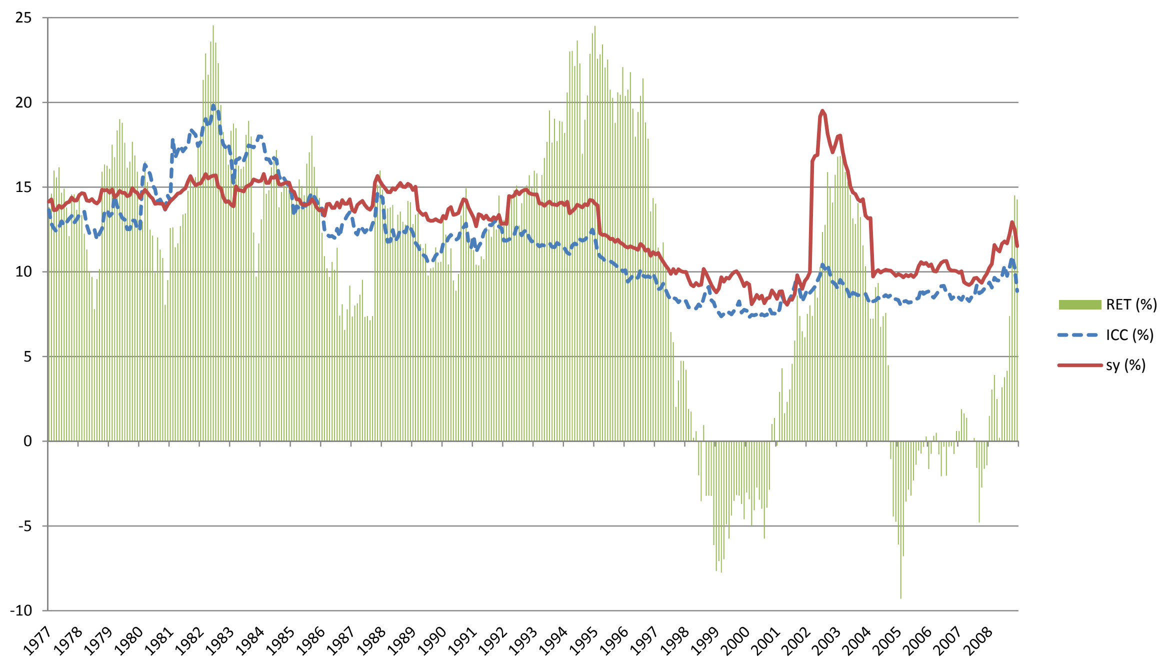

Figure 4: **Out-of-sample R^2s: Stock Yield (sy) *vs.* Dividend-to-Price Ratio (dp)**

The figure plots the out-of-sample R^2s of predicting next one-month returns by the stock yield (sy) and the log dividend-to-price ratio (dp) at the market level for different starting points of the out-of-sample period. sy is calculated as $0.29+0.05(dp+1.4g)$, where dp is the log value of the dividend-to-price ratio and g is the log value of forecasted earnings growth rate. For a specific out-of-sample starting point month t, the in-sample period is January 1977 through month t-1, and the out-of-sample period is month t through December 2012.

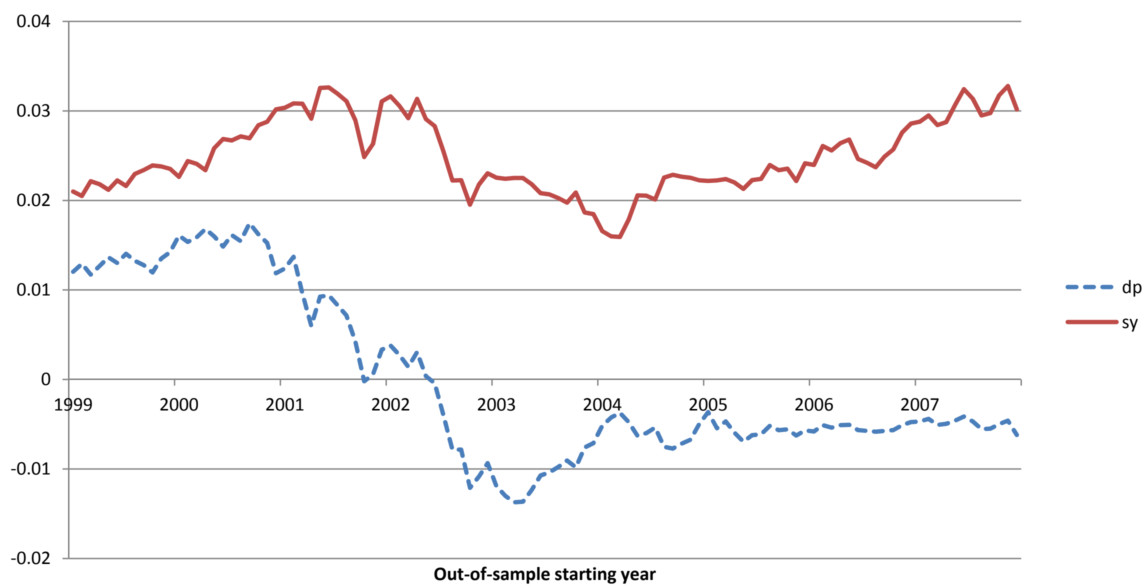

Figure 5: **In-Sample Predictions: Good Times *vs.* Bad Times**

The figure plots realized one-year returns against predicted one-year returns by stock yield for "good times" and "bad times" separately. "Good times" and "bad times" are categorized according to the Chicago Fed National Activity Index (CFNAI), which measures overall economic activity. Each calendar month in the sample period is placed in one of two groups based on the CFNAI in that month: "good times" are those with above-median CFNAIs, and "bad times" are those with below-median CFNAIs.

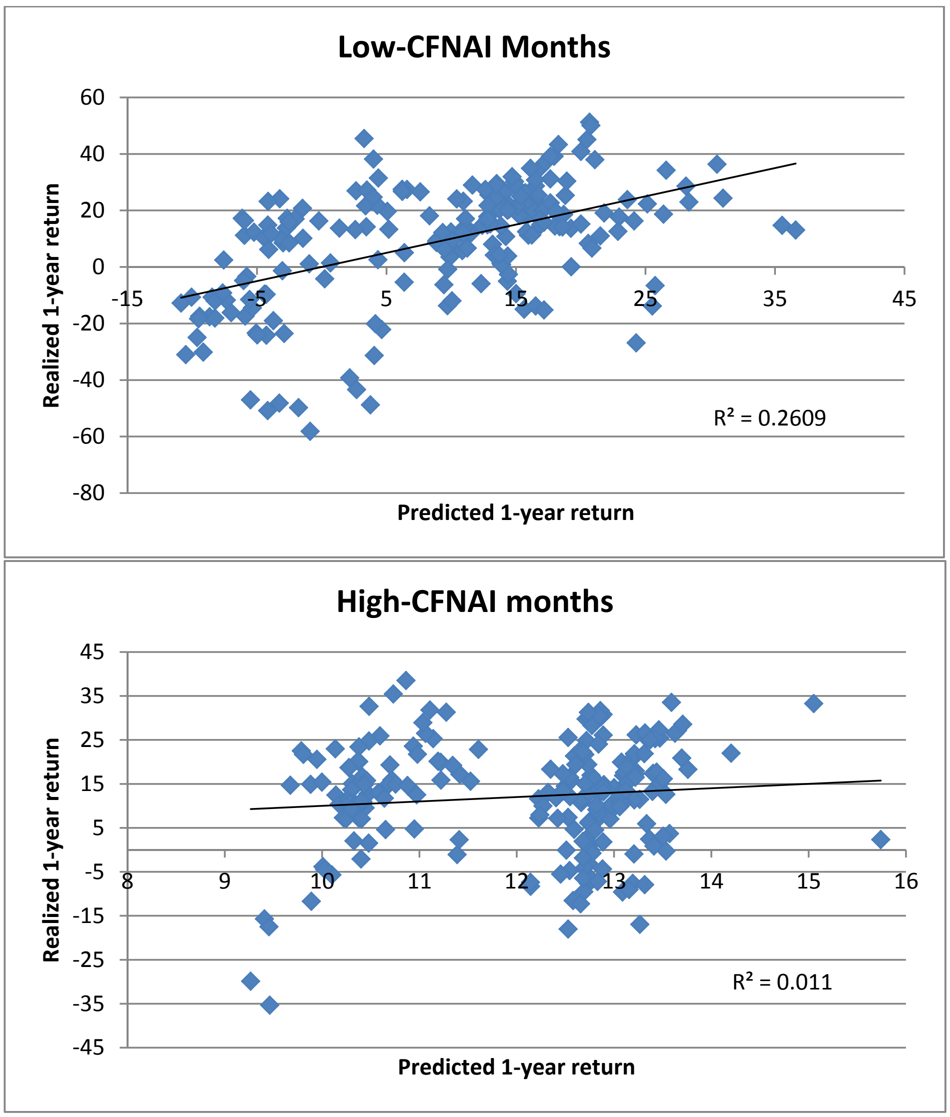

Figure 6: **Out-of-Sample R^2s: Original BK and Extended BK**

The figure plots the out-of-sample R^2s of predicting monthly market returns according to the original Binsbergen and Koijen (2010) model (Original BK) and its extension with stock yield (Extended BK). We vary the starting point (month t) of the out-of-sample period from January 1999 to December 2007, using the period January 1977 to month t-1 as the in-sample period and the period month t to December 2012 as the out-of-sample period.

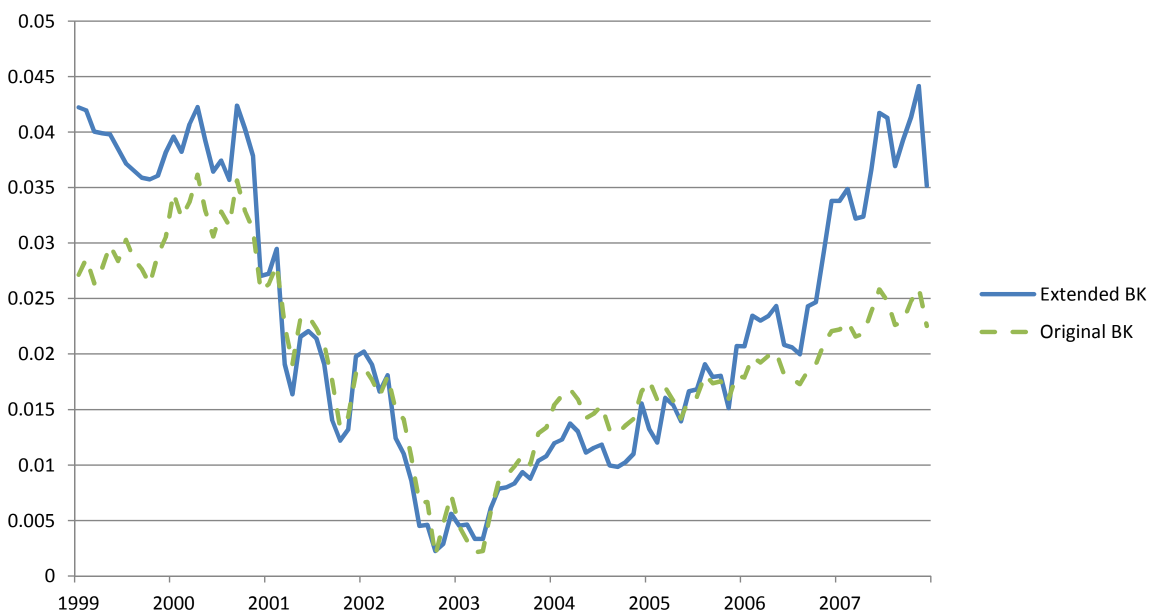

Table 1: **Summary Statistics for Market Return Predictors**

This table reports summary statistics of the main predictive variables used in this paper. dp is the log value of the dividend-to-price ratio; g is the log value of forecasted earnings growth rate; sy is the stock yield measure, calculated as $0.29+0.05(dp+1.4g)$; BM is the log value of the book-to-market ratio; EP is the log value of the earnings-to-price ratio; $CAPE$ is Shiller's cyclically adjusted price-to-earnings ratio; $Term$ is the term spread, calculated as the difference between the AAA-rated corporate bond yield and the one-month T-bill yield; $Default$ is the default spread, calculated as the difference between yields of BAA-rated and AAA-rated corporate bonds; T-$Bill$ is the continuously compounded one-month Treasury bill rate; T-$Bond$ is the continuously compounded yield on 30-year Treasury bonds; VRP is variance risk premium measure in Bollerslev, Tauchen, and Zhou (2009); and ICC is the continuously compounded implied cost of capital of the value-weighted market portfolio. All return, spread, and yield variables are expressed in annualized percentages. The sample period is January 1977 – December 2012, except for VRP which starts from January 1990. We report mean, standard deviation, and autocorrelations of predictive variables in Panel A, and correlations between predictive variables in Panel B.

Panel A: Mean, standard deviation, and autocorrelation

	Mean	Std	AR(1), 1y	AR(1), 2y	AR(1), 3y	AR(1), 4y
dp	-3.57	0.34	0.85	0.72	0.66	0.56
g	0.24	0.24	0.29	-0.18	-0.16	-0.19
sy (%)	12.85	2.35	0.55	0.18	0.16	0.15
BM	-0.45	0.27	0.80	0.67	0.56	0.45
EP	-2.80	0.43	0.72	0.53	0.49	0.39
$CAPE$	20.20	9.11	0.92	0.81	0.69	0.60
$Term$ (%)	3.08	1.61	0.45	0.10	-0.28	-0.42
$Default$ (%)	1.11	0.48	0.47	0.29	0.21	0.10
$T-Bill$ (%)	5.07	3.38	0.82	0.64	0.50	0.42
$T-Bond$ (%)	6.85	2.46	0.90	0.83	0.78	0.70
VRP (%)	18.47	20.35	0.05	0.01	-0.05	-0.02
ICC (%)	11.38	2.87	0.91	0.79	0.67	0.59

Panel B: Correlations

	sy	dp	g	BM	EP	$CAPE$	$Term$	$Default$	$T-Bill$	$T-Bond$	VRP
dp	0.69										
g	0.69	-0.05									
BM	0.68	0.93	0.01								
EP	0.12	0.76	-0.59	0.72							
$CAPE$	-0.75	-0.94	-0.09	-0.93	-0.71						
$Term$	0.24	-0.01	0.35	0.09	-0.34	-0.01					
$Default$	0.39	0.56	-0.03	0.61	0.42	-0.53	0.20				
$T-Bill$	0.28	0.45	-0.07	0.30	0.53	-0.48	-0.58	0.24			
$T-Bond$	0.53	0.59	0.14	0.46	0.49	-0.66	-0.23	0.33	0.88		
VRP	0.05	-0.02	0.07	-0.02	-0.10	0.09	-0.01	0.04	0.03	0.02	
ICC	0.70	0.80	0.17	0.73	0.58	-0.86	0.01	0.55	0.68	0.87	-0.01

Table 2: **Cash Flow Predictive Power of Analyst-Forecast-Implied Growth Rate (g)**

This table examines the predictive power of the analyst-forecast-implied growth rate, g, for log annual earnings growth rates and log dividend growth rates over the next one-, three-, and five-year horizons. The sample period is January 1977 – December 2012. All regressions use overlapping monthly observations. For each regression, we report the slope coefficient, Newey-West HAC t-value, p-value, and adjusted R^2. The p-values of individual regression coefficients and the average coefficient are obtained by comparing the test statistics with their empirical distribution generated under the null of no predictability from five thousand trials of a Monte Carlo simulation as described in the Appendix.

	Earnings growth				Dividend growth			
	Coeff	NW-t	p-value	Adj R^2(%)	Coeff	NW-t	p-value	Adj R^2(%)
1-year	1.17	15.31	0.00	80.24	0.07	1.72	0.16	4.97
3-year	0.45	17.34	0.00	73.52	0.10	4.52	0.05	22.16
5-year	0.26	18.51	0.00	74.98	0.07	4.86	0.03	19.61

Table 3: **In-Sample Regression of Market Returns on Stock Yield and Other Predictors**

This table reports results of in-sample univariate regressions of market returns on stock yield and other predictive variables. The dependent variable is the average continuously compounded market return in the future, ranging from one-year to four-year horizons. dp is the log value of the dividend-to-price ratio; g is the log value of forecasted earnings growth rate; sy is the stock yield measure, calculated as $0.29+0.05(dp+1.4g)$; BM is the log value of the book-to-market ratio; EP is the log value of the earnings-to-price ratio; $CAPE$ is Shiller's cyclically adjusted price-to-earnings ratio; $Term$ is the term spread, calculated as the difference between the AAA-rated corporate bond yield and the one-month T-bill yield; $Default$ is the default spread, calculated as the difference between yields of BAA-rated and AAA-rated corporate bonds; *T-Bill* is the continuously compounded one-month Treasury bill rate; *T-Bond* is the continuously compounded yield on 30-year Treasury bonds; VRP is variance risk premium measure in Bollerslev, Tauchen, and Zhou (2009); and ICC is the continuously compounded implied cost of capital of the value-weighted market portfolio. All return, spread, and yield variables are expressed in annualized percentages. The sample period is January 1977 – December 2012, except for VRP which starts from January 1990. All regressions use overlapping monthly observations. For each regression, we report the slope coefficient, Newey-West HAC t-value, p-value, and adjusted R^2. We also report the average value of regression coefficients across four horizons and its p-value. The p-values of individual regression coefficients and the average coefficient are obtained by comparing the test statistics with their empirical distribution generated under the null of no predictability from five thousand trials of a Monte Carlo simulation as described in the Appendix.

	Coeff	NW-t	p-value	Adj R^2 (%)	Coeff	NW-t	p-value	Adj R^2 (%)	Coeff	NW-t	p-value	Adj R^2 (%)
			dp				g				sy	
1-year	16.43	2.79	0.17	11.14	11.28	1.29	0.12	2.39	2.60	2.87	0.04	12.99
2-year	15.09	2.70	0.18	21.12	11.13	1.61	0.11	5.37	2.46	2.78	0.03	25.91
3-year	14.03	3.44	0.18	29.13	11.23	1.95	0.08	8.70	2.37	3.90	0.02	37.79
4-year	13.82	5.85	0.16	38.98	12.29	2.32	0.05	14.26	2.39	6.07	0.01	53.49
Avg	14.84		0.16		11.48		0.08		2.46		0.02	
			BM				EP				$CAPE$	
1-year	20.26	2.57	0.13	10.47	7.02	1.41	0.27	3.02	-0.56	-2.26	0.53	9.26
2-year	19.07	2.50	0.12	20.84	6.37	1.45	0.29	5.72	-0.60	-2.83	0.56	23.78
3-year	18.31	3.14	0.10	30.57	5.83	1.44	0.30	7.63	-0.59	-4.23	0.57	37.24
4-year	17.23	4.65	0.09	37.17	5.21	1.49	0.32	8.49	-0.55	-7.15	0.55	46.16
Avg	18.72		0.11		6.11		0.29		-0.58		0.55	
			$Term$				$Default$				T-$Bill$	
1-year	0.82	0.73	0.32	0.38	5.35	1.35	0.22	2.09	0.62	0.92	0.21	1.26
2-year	1.23	1.16	0.19	2.81	3.96	1.26	0.28	2.59	0.52	1.75	0.25	1.92
3-year	1.43	1.63	0.12	6.16	3.89	1.29	0.27	4.11	0.52	1.89	0.23	2.89
4-year	1.11	1.35	0.16	4.94	4.44	1.37	0.21	6.47	0.76	2.46	0.13	8.56
Avg	1.15		0.18		4.41		0.23		0.60		0.19	
			T-$Bond$				VRP				ICC	
1-year	1.52	1.50	0.12	4.45	0.16	2.48	0.00	3.03	1.56	1.90	0.18	6.86
2-year	1.66	2.44	0.10	11.52	0.10	1.93	0.01	2.00	1.67	2.46	0.14	17.67
3-year	1.67	3.02	0.10	18.00	0.04	0.66	0.15	0.06	1.67	3.08	0.12	28.42
4-year	1.83	3.96	0.08	29.15	-0.03	-0.46	0.87	0.05	1.67	4.21	0.10	40.76
Avg	1.67		0.09		0.07		0.03		1.64		0.13	

Table 4: **Out-of-Sample Prediction of Market Returns based on Stock Yield and Other Predictors**

This table reports results of out-of-sample prediction of market returns based on stock yield and other predictive variables. In Panel A, the market return to predict is next one-month continuously compounded return. The predictive variables include: stock yield (sy), the log value of the dividend-to-price ratio (dp), the log value of forecasted earnings growth rate (g), the log value of the book-to-market ratio (BM), the log value of the earnings-to-price ratio (EP), Shiller's cyclically adjusted price-to-earnings ratio ($CAPE$), the term spread ($Term$), the default spread ($Default$), the continuously compounded one-month T-bill rate (*T-Bill*), the continuously compounded yield on 30-year Treasury bonds (*T-Bond*), and the continuously compounded aggregate implied cost of capital (ICC). Following Lee, Ng, and Swaminathan (2013), sy and ICC are calculated as their 24-month moving average. All return, spread and yield variables are expressed in annualized percentages. For each predictor, we report out-of-sample R^2 and related statistics. We calculate the p-value of out-of-sample R^2 from the adjusted-MSPE statistic of Clark and West (2007). Following Campbell and Thompson (2008), the utility gain ($Ugain$) is the additional utility to an investor with mean-variance preferences and a risk aversion coefficient of three by forecasting returns using a particular predictive variable rather than the historical average benchmark. The weight on stocks in the investor's portfolio is constrained to lie between 0 and 1.5 (inclusive). In Panel B, the market return to predict is next one-year continuously compounded return. All predictive variables are defined in the same way as in Panel A, except for sy and ICC which take the original value rather than the 24-month moving average. The p-value is corrected for the serial correlation using the Newey-West formula with a lag of 11 (the number of overlapping observations). In each panel, we also report the results of predicting market excess return, which is continuously compounded market return minus the continuously compounded one-month T-bill rate. Following Lee, Ng, and Swaminathan (2013), we subtract the yield of one-month T-bill from sy and ICC. The out-of-sample period is January 1999 – December 2012, and the training period is January 1977 – December 1998.

Panel A: Next one-month market return

	dp	g	sy	BM	EP	$CAPE$	$Term$	$Default$	*T-Bill*	*T-Bond*	ICC
						Raw return					
R^2 (%)	1.20	-3.46	2.10	0.76	0.02	0.91	-0.99	-1.11	-0.43	-0.61	1.01
p-value	0.09		0.03	0.16	0.22	0.03					0.07
$Ugain$	4.86	0.19	8.38	3.16	2.97	2.40	0.38	0.47	0.48	0.67	3.72
						Excess return					
R^2 (%)	0.42	-3.16	1.17	0.46	-0.37	-0.07	-1.05	-0.78	-1.13	-1.63	1.43
p-value	0.21		0.04	0.19							0.05
$Ugain$	0.96	-3.12	4.50	0.72	-0.84	0.90	-2.07	-1.10	-2.33	-1.87	5.42

Panel B: Next one-year market return

	dp	g	sy	BM	EP	$CAPE$	$Term$	$Default$	$T\text{-}Bill$	$T\text{-}Bond$	ICC
						Raw return					
R^2 (%)	9.73	-2.42	9.83	11.13	-2.75	6.07	-1.12	2.47	-1.15	1.20	5.17
p-value	0.01		0.02	0.00		0.05		0.11		0.21	0.06
						Excess return					
R^2 (%)	0.21	-0.58	8.96	5.57	-6.89	-6.06	1.75	-0.22	-4.81	-7.64	6.87
p-value	0.34		0.02	0.03			0.21				0.03

Table 5: **Out-of-Sample Prediction of Market Returns based on Stock Yield: Robustness Checks**

This table reports results of robustness checks of out-of-sample prediction of market returns based on stock yield. The market returns to predict are next one-month and next one-year continuously compounded returns. The stock yield measure, *sy*, is calculated as 0.29+0.05(dp+beta$\times g$), where dp is the log value of the dividend-to-price ratio, g is the log value of forecasted earnings growth rate, and beta is a constant. In the baseline measurement of *sy*, we take the beta value of 1.4, measure dp and g from the value-weighted all-firm portfolio, and calculate g based on IBES median earnings forecasts for the next one-year period and Compustat realized earnings for the most recent fiscal year. In Panel A, we define g in various alternative ways by: (1) winsorizing extreme g at 5 and 95 percentile values across all months (*sy_wsg*); (2) measuring g (and dp) from the S&P500 portfolio (*sy_sp*); (3) using the most recent realized EPS from IBES to calculate g (*sy_ibes*); (4) using IBES lowest or highest EPS forecasts to calculate g (*sy_low* and *sy_high*); (5) using earnings forecasts for the two fiscal years ending in the next two years to compute g (*sy_g*2); and (6) using analysts' median long-term growth (*ltg*) forecast (available since 1982) for g (*sy_ltg*). In Panel B, we define *sy* in various alternative ways by: (1) using different constant values for beta; (2) using time-varying beta values as determined by in-sample OLS regression of market return on dp and g every month (beta_ols); and (3) using time-varying beta values as determined by in-sample restricted-least-square regression of market return on dp and g every month, in which the ratio of the coefficient on g to that on dp is restricted between 1.1 and 1.9 (beta_rls). When predicting next-month return, we take the 24-month moving average of all alternative stock yield measures. For each prediction, we report out-of-sample R^2 and its p-value. We calculate the p-value from the adjusted-MSPE statistic of Clark and West (2007), and correct it for the serial correlation using the Newey-West formula with a lag of 11 (the number of overlapping observations) when predicting next one-year returns. The out-of-sample period is January 1999 – December 2012, and the training period is January 1977 – December 1998.

Panel A: Alternative measures of g

	R^2 (%)	p-value	R^2 (%)	p-value	R^2 (%)	p-value
	sy_wsg		*sy_sp*		*sy_ibes*	
1-month return	2.04	0.03	1.88	0.03	0.95	0.12
1-year return	11.78	0.01	7.59	0.05	8.69	0.03
	sy_low		*sy_high*			
1-month return	2.30	0.02	1.93	0.03		
1-year return	10.45	0.02	9.13	0.02		
	*sy_g*2		*sy_ltg*			
1-month return	1.12	0.08	0.62	0.15		
1-year return	7.49	0.01	8.35	0.03		

Panel B: Alternative values for beta

	R^2 (%)	p-value	R^2 (%)	p-value
	dp+1.1g		dp+1.9g	
1-month return	2.16	0.02	1.88	0.04
1-year return	11.62	0.01	6.69	0.04
	dp+beta_ols×g		dp+beta_rls×g	
1-month return	1.38	0.08	1.90	0.04
1-year return	5.53	0.02	9.85	0.02

Table 6: **Out-of-Sample Regressions: Factors Extracted from the Cross-Section**

This table reports results of out-of-sample prediction of market returns based on a single forecasting factor extracted from the cross-section of stock yield and other predictive variables. For example, to extract a market return predictor from *sy*, we first form size-*sy* portfolios based on a two-way conditional sort on market capitalization and *sy*, then calculate the portfolio-level *sy*, and finally run the three-stage regressions in Kelly and Pruitt (2013) to generate the out-of-sample predictor of market return (KP-*sy*). Similarly, we generate the out-of-sample predictor of market return (KP-*dp*) from the log dividend-to-price ratios (*dp*) of portfolios formed based on a two-way conditional sort on market capitalization and *dp*. We use the data since January 1977 to generate both KP-*sy* and KP-*dp*. For comparison, we also follow Kelly and Pruitt (2013) exactly to extract a return predictor (KP-BM) from the log book-to-market ratios (BM) of Fama-French portfolios, using the data since January 1930. Finally, for ease of comparison, we reproduce the results from Table 4 using the market-level *sy*, *dp*, and BM. For each generated out-of-sample predictor, we calculate the out-of-sample R^2 of predicting market returns and its *p*-value. The out-of-sample period is January 1999 – December 2012. We calculate the *p*-value of out-of-sample R^2 from the adjusted-MSPE statistic of Clark and West (2007), and correct it for the serial correlation using the Newey-West formula with a lag of 11 (the number of overlapping observations) when predicting next one-year returns.

Predictor	Horizon	6 portfolios		25 portfolios		Market	
		R^2 (%)	*p*-value	R^2 (%)	*p*-value	R^2 (%)	*p*-value
KP-*sy*	1-month	2.70	0.03	2.71	0.03	2.10	0.03
	1-year	11.66	0.06	19.49	0.03	9.83	0.02
KP-*dp*	1-month	1.41	0.12	1.47	0.12	1.20	0.09
	1-year	5.06	0.11	9.57	0.07	9.73	0.01
KP-BM	1-month	1.58	0.03	1.47	0.04	0.76	0.16
	1-year	13.15	0.05	14.11	0.03	11.13	0.01